

On including assessments in the calculation of teaching loads

Abstract

Next to courses and seminars, tests and assessments represent the main parameters with which to describe an academic study program independent of its curricular content. Thus, the quality of education depends not only on the quality of the courses taught and how they are interconnected, but also on the quality of testing and the feedback given to students regarding their performance.

Course quality should be ensured through course evaluation. The economic cost of courses is calculated based on the required teaching load. The concept of teaching load stems from the time when program planning was instructor-centered. The main variable in the rules and regulations governing university study was the number of hours per week per semester (or number of course hours). But even in today's student-centered planning, which uses ECTS credits per module as the variable, teaching loads are still used to determine the number of staff necessary to offer an academic study program.

Some universities also include the assessments in the evaluation. Yet the economic costs of testing are de facto ignored almost everywhere, and this does not bode well for the quality of the assessments. Much progress would be made to improve higher education if assessments counted as part of the teaching loads and the curricular norm values. This paper identifies which requirements must be considered in order to include assessments in teaching loads.

Keywords: assessments, teaching load, examination parameters, curricular standard value

Introduction

Assessments are an important component of university study programs and can lead to long legal battles. Two rulings handed down by the Federal Administrative Court (Bundesverwaltungsgericht) [1], [2] have resulted in a more precise wording of the relevant passages in the regulations governing the study and testing of midwifery (*Studien- und Prüfungsverordnung für Hebammen*) [<https://www.gesetze-im-internet.de/hebstprv/BJNR003900020.html>], although these rulings referred to other courses of study in Germany. Precisely because these court decisions address the basic framework for objective testing, it is surprising that such improvements and revisions are still needed. It is not clear why this is the case since this issue is clearly a topic of wide discussion: There are

- international standards for tests [3], [4], [5], [6],
- recommendations for practitioners [7], [8], and
- literature on the laws governing academic testing ([9] or a new edition, [10]).

On the other hand, tests and assessments are traditionally part of the pre- and post-work done for courses. And only the courses themselves are factored in to the calcu-

Volkhard Fischer¹

1 Medizinische Hochschule
Hannover, Studiendekanat
Bereich Evaluation &
Kapazität, Hannover,
Germany

lation of curricular norm values (*Curricularnormwerte* or CNW) and counted toward the individual teaching load (or teaching duties). If it is understood that a university study program is comprised of (classroom-based and/or online) courses, different types of feedback, self-guided study and extra-curricular activities, it is clear that feedback, together with the courses, belongs to the responsibilities owed by the university and its teachers to the students.

Several state-level (Land) rules and regulations regarding the assignment of teaching duties in higher education (*Lehrverpflichtungsverordnung*, or LVVO) contain statements phrased to allow faculties the option of paying tests and assessments the attention due by treating them as a subgroup of the feedback students receive regarding their academic progress. Passages from four LVVOs are presented in table 1 as examples.

The requirements which should be met so that a university faculty can take advantage of this option are covered in the following. The background for this entails the following considerations:

Table 1: Excerpts from different state regulations governing the assignment of university teaching duties that allow the possibility of counting courses or lectures towards the calculation of the teaching load (translated)

Bavaria	LUFV of 2019	§ 3 (1), first sentence	Courses not provided for in examination regulations, study regulations or curricula are taken into account if all courses of a subject provided for in these regulations are offered by full-time or part-time teachers working at the university; ...
Berlin	LVVO of 2021	§ 3 (3)	Lectures, tutorials, seminars, colloquia, revision courses, individual and group artistic tuition and equivalent courses in terms of preparation and follow-up time, as well as seminar teaching and practical courses at universities of applied sciences, are fully credited towards the teaching obligation.
Lower Saxony	LVVO of 2021	§ 13 (1), first sentence	When fulfilling the teaching obligation, the courses, including those outside the lecture period, which are provided for in the study or examination regulations or curricula are taken into account.
Saxony-Anhalt	LVVO of 2018	§ 5 (1)	In order to fulfil the teaching obligation, those courses must be taken into account which are mandatory for a proper course of study in the respective semester according to examination or study regulations

- P1) Assessments should be sufficiently defined so that the resulting interference with students' rights is still legally compliant.
- P2) Assessments do not belong to conventional seminars and courses and generally are not used in the calculation of teaching load, but rather count as part of the work done before and after teaching courses.
- P3) If an LVVO does not permit the inclusion of courses not originally provided for, then these courses must be calculated toward the teaching load in a manner similar to conventional courses.
- S) Thus, in order to include assessments in the teaching load, it is necessary to establish which of the assessment features that must be sufficiently defined to ensure specificity correspond to the factors used to calculate teaching load and if these features can be considered separately from the course.

In its decision regarding a digitally administered module assessment, the Administrative Court in Hannover formulated a list of requirements for the Hannover Medical School that its examination regulations should satisfy [11]. Even if the judicial system is now more familiar with assessments using electronic devices [10], the requirements stipulated by the court at that time (see table 2) represent not only important guidance to ensure the legality of the tests and assessments administered over the course of an academic study program, but also for including them in the calculation of teaching loads. The starting point for the legal considerations is that the basic right to freely choose an occupation, as provided for in § 12 (1) of the Basic Law of the Federal Republic of Germany (Grundgesetz), is restricted since a final failure on an assessment permanently hinders future professional practice. Even in the very general case of a test or assessment, the materiality principle must be satisfied, and this fundamental right may only be restricted on legal grounds or within the scope granted by law under a statute. To comply with this, assessment procedures and require-

ments should be based on the principle of equal opportunity and it should be ensured, to the greatest extent possible, that the methods for grading assessments will stand up to judicial review. This aligns very well with the international standards for objective, valid and reliable assessments [3], [4], [6].

The first column of table 2 specifies the variables used to calculate the curricular percentage (CAP) of a conventional course. The second column lists features that need to be sufficiently defined for a test, according to the court decisions cited above. It does not, however, present an exhaustive list. All of the components of these features are then listed in the third column. The variables needed to include assessments in the teaching load are what enables quantification analogous to that for the conventional courses.

The features listed in the second column of table 2 all affect the objectivity of assessments, whereby it is already clear that they are in no way sufficient to ensure "good" tests. Nonetheless, it is (usually) the objectivity of an assessment that is the subject of legal scrutiny. Exceptions confirm this rule, such as the judgment handed down in the "numerus clausus case" by the Federal Constitutional Court (Bundesverfassungsgericht) [12]. For, at the time, when clarifying the constitutionality of multiple-choice testing, the fundamental question was not about the reliability of state medical exams using the multiple-choice format, but rather about their level of difficulty across different cycles.

Understood metaphorically, this means that the rules and regulations required by law affect, above all, the foundation, not the walls nor the roof, of the structure "housing" the tests and examinations connected with a particular study program and that they should be self-evident. For this reason, they are referred to, for instance, in the "recommendations" [8]. The focus of the recommendations is based, however, primarily on the threats to the validity and reliability of assessments because it is thought that

Table 2: Necessary features of an assessment, as defined in the examination regulations, to ascertain sufficient specificity, to include it in the calculation of teaching load, and to determine its curricular percentage

	Variables determining the curricular percentage of a course	Assessment features that must be sufficiently defined to ensure specificity	Variables determining the curricular percentage of an assessment
1		Prerequisites for taking the assessment	
2	Course content	Assessment content	Assessment content
3	Course duration	Assessment duration / scope	Assessment duration
4	Participating instructors	Number of examiners	Participating instructors
5	Number of students	Number of examinees	Number of examinees
6		Criteria for passing	
7		Consequences	
8		Documentation	

the greatest number of weak points can be found there. To this extent, the characteristics in the second column of table 2 only represent a partial list of the criteria for “good” or legally compatible assessment, but they do make it feasible to fulfill the demand contained in table 1 to define the other courses relevant to computing teaching load.

Assessments

In its examination regulations, the Medical University of Vienna differentiates in a very helpful manner between courses in which ongoing assessment is an intrinsic part, course-based tests, and comprehensive assessments [13]. There is also, in addition to the bachelor thesis, a final exam at the end of the study program. These differences clearly show the range that assessments can have and why it is not unreasonable to consider separately counting at least certain forms of testing toward the teaching load. The order in which these assessment formats are addressed in the following is based according to the severity of the consequences generally associated with them [14].

The question as to whether or not feedback is graded is only of secondary importance here. What is relevant is if the feedback brings consequences with it; if the assessment can only be repeated so many times if it has not been passed. Therefore, as a first step there needs to be a delineation between feedback with no consequences (course attendance certificates, formative assessments) and feedback with consequences (summative assessments). This would mean that much of the ungraded feedback designated in the study and examination regulations as academic coursework is summative assessment because failing interferes with the free choice of occupation by extending the length of study time or even leading to disenrollment from the study program.

Student theses and state exams

Regulations governing academic study define the different types of courses offered in a study program. In them there

is usually also at least one section on the requirements for approval and completion of a student thesis. In regard to the inclusion of student theses in teaching loads, separate sections were added to the LVVOs only after the Bologna Process (see table 3), although student theses are much older than these reforms.

This makes it clear that this traditional assessment format can be counted toward the teaching load. The difference here is that in the case of thesis advising, the number of hours spent advising cannot be determined objectively, as is the case with conventional classroom teaching, but rather only indirectly based on the student workload.

The state exams in dentistry and medicine are not currently included in the calculation of the curricular norm values for these study programs. No teaching effort is assigned at all to these exams in the overall calculation of the curricular norm value for each program. This is surprising, considering that, for example, in the dental licensing regulations (ZApprO) the definitions of the different sections of the state exam go on for 22 pages [<https://www.gesetze-im-internet.de/zappro/BJNR093310019.html>] and a total of 12 pages are needed in the current version of the medical licensing regulations (ÄApprO) [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html]. This is surprising because these characteristics – mandated by the administrative courts precisely for these state exams to ensure sufficient definition of a test – are duly defined in the relevant university rules and regulations, and they are normally documented so well when administering assessments that it would be thoroughly feasible to include this in the calculation of an individual teaching load and the curricular norm value.

In contrast to bachelor and master theses, the institution responsible for assessment is not the university but a state-level examination office. Carrying out the actual assessment, however, lies with the university faculty members. Thus, factoring the assessment times into teaching loads would be entirely justified; the more so as no one doubts that the oral assessment formats, unlike the written state exams consisting of nationally standardized questions, require unique preparation by teachers at specific universities. Reversely, the administrative

Table 3: Excerpts from different regulations governing teaching duties on the inclusion of student theses in the calculation of teaching load (translated)

Bavaria	LUFV of 2019	§ 3 (8)	Advising activities for "Diplom," bachelor or other degree program theses and comparable student theses may only be counted once per student with consideration for the necessary time and effort, but only up to an amount of two course hours at universities and art academies and only up to an amount of three course hours at universities of applied sciences.
Berlin	LVVO of 2021	§ 3 (6)	If university instructors with regular teaching duties pursuant to § 5 (1), sentence 1, number 1(b) and § 5 (3), sentence 1, number 1 can demonstrate an above-average workload due to student thesis advising or comparable academic advising, the university administration or human resources office may offset this against the regular teaching duties based on the budget if the particular situation in the subject area permits.
Lower Saxony	LVVO of 2021	§ 15	The time required for advising activities for written theses and comparable theses may be taken into account by the university in an amount up to 2 course hours per semester if a particular burden is incurred by the advising activities.
Saxony-Anhalt	LVVO of 2018	§ 3a	Increased workloads resulting from thesis advising may be counted toward the performance of teaching duties, taking into account the required effort and workload of the affected advising instructor.

courts do not recognize the administering of these state exams as a reason to reduce teaching load because they are clearly part of the academic teaching duties.

Comprehensive assessments

Comprehensive assessments cover the content of all courses offered in a semester or academic year, regardless if the comprehensive assessment is designed to be formative or summative [13]. Like the final exams at the end of a study program or the state exams, summative comprehensive assessments are in no way conceived of as preparatory or follow-up work for specific courses. It is very difficult to view creating, administering and grading them as part of the work done before and after a course or module. Nonetheless, they are an integral component of the study program, for instance, when they determine a student's progression to the next phase of study.

A current example for such a summative comprehensive assessment is found in the Objective Clinical Examination (OSCE) that is required under § 37 of the revised draft legislation for medical licensing (ÄApprO) [15]. According to § 37 of the revised licensing regulations, the future OSCE would meet all of the requirements laid out by the LVVOs to count it as part of the teaching load. Also, the duration of the assessment time would even be set down in statute.

One example of a formative comprehensive assessment is voluntarily taking the progress tests offered at many universities [16] if they are irrelevant to a student's passage on to the next level. These types of assessments represent a special form of self-guided study.

Module tests

Course-specific tests come at the end of a course concluding it in the form of a singular occasion on one day, or at a maximum on two. The Viennese regulations [13] leaves open whether these cover only the block's lectures or a combination of lecture, seminar and practicum, as would be assumed for a module test. In general, it is module tests that are meant here.

At first glance, there are many reasons to continue viewing the assessments that give students feedback on their academic performance at the end of a module as simply part of the pre- and post-work associated with teaching. This traditional approach also makes sense for unstandardized assessments of unclear objectivity which do not (and are not meant to) fulfill the minimum requirements for documented assessment. However, as a result, they are on par with a testing practice that has been repeatedly rejected by the administrative courts as unlawful [8], [9]. Since modules are often a combination of different teaching and course formats, it is impossible to clearly assign module testing to the preparatory or follow-up work done to teach a particular course. Moreover, a standardized test requires a greater level of effort if it is to be objective, reliable and valid. A written multiple-choice test – to which the administrative courts, in a misjudgment of the literature on assessment quality [17], generally show a particular distrust – with 60 or more questions cannot be created in the same amount of time needed to write a two-hour lecture. Furthermore, a 12-station OSCE also requires more time and effort than teaching a cohort of medical students in small groups in the Skill-Lab for four hours each [18].

In addition to the effort required for assessments based on table 2, purely pragmatic reasons also speak for inclusion in the teaching load. It does not take special genius to list the lecture and the assessment separately in a module description, to specify the testing format in the examination regulations, or to list explicitly all of the module tests. Even if the lecture itself should remain unchanged for several years, it would be possible, over the same course of time, to change the written test with open-ended questions into a multiple-choice test taken on electronic end devices, and this can then be documented by means of a minor revision to the examination regulations.

In regard to differentiating between module tests with and without consequences, the same applies as for the final exams taken at the end of a study program: formative module tests pose a special form of self-guided study.

Courses with inherent examination character

If grades relevant to passing are given during a course, then, according to the Vienna Medical University's examination regulations [13], this involves a course with ongoing assessment as an inherent part. The model examination regulations for the Leibniz University in Hannover [19] mentions course-related assessments. Both terms express the fact that this does not involve one test occasion, but rather a repetition of several occasions of assessment that are then evaluated as an overall grade.

Classic examples of such graded work are seminar presentations, making plaster casts and models in dental medicine, and writing lab reports for practicums in the natural sciences. These student assignments cannot be separated from the classroom sessions precisely because they are completed during the normal weekly seminar or practicum hours. They also cannot be counted separately in the teaching load as assessments, since an individual student's presentation during a seminar session is coursework for the listeners and a test of ability for the presenter, whereby the roles are switched from session to session. Furthermore, in the student-to-teacher ratio, "student" refers to both the presenter (examinee) and the listeners.

The case is different for a seminar in which different topics are covered during the sessions and which are then gone into more depth in seminar papers that are turned in at the end of the semester. A practicum in which clinical examination techniques are practiced is not a course with ongoing inherent examinations if the summative feedback is given via an OSCE at the very end. Because these assessments are chronologically separate from the course, they have a unique character and fall into the category described in the previous section.

Discussion

If feedback is meant to give students a differentiated evaluation of their learning progress, then how it is designed, delivered and evaluated entails time and effort. The justification given by the German Association of University Professors and Lecturers (Deutsche Hochschulverband) for its push in 2019 to reduce the teaching load of university teachers is based, among other things, on the increased requirements placed on assessments [20]. However, such a solution does not promote the creation of high-quality assessments, but rather is just less encumbering than the existing process. Moreover, since the feedback from summative assessments is connected with consequences for a student's future career path, this feedback should be sufficiently defined so that it is not only objective, valid and reliable [21], but also compatible with the law. The higher an assessment's stakes are, the more transparent the conditions under which it is taken should be and the more detailed (legally speaking) the assessment's specification should be [9]. The proposal made in this paper to count assessments separately as part of the teaching load would work to facilitate this.

Only the relevant examination regulations containing explicit definitions according to the characteristics listed in table 2 can, as the lowest bar, apply to all of the summative assessment formats described in the previous section. Due to their lack of consequences, formative assessments are not required to be explicitly defined or specified in the examination regulations for reasons of legal conformity. Meeting this requirement, however, would be necessary in order to count this type of assessment toward the teaching load, as well as to improve the quality of higher education overall.

Figure 1 illustrates the delineations suggested in this paper between what is already counted toward teaching load (green), what could potentially be counted (yellow), and which assessments cannot be counted (red). If this classification is deemed acceptable, the question then arises as to why the criteria viewed by the administrative courts as necessary to sufficiently define assessments are, in turn, necessary only in part when calculating the teaching load.

To factor a course into the teaching load, its topic, duration in hours, and the standard teacher-to-student ratio or group size must be known. It is generally assumed that a teacher is constantly present in the classroom. These variables are also defined for assessments in legally compliant examination regulations and are indispensable for inclusion in the teaching load. The fact that there are still many study and examination regulations which do not implement this only says something about the affected study programs' lack of transparency.

Under which conditions students are admitted to take a course or sit for an exam, how course attendance or passing is certified, what the consequences are for not attending or failing a course, and how all of this is documented are usually explicitly laid down in good study and

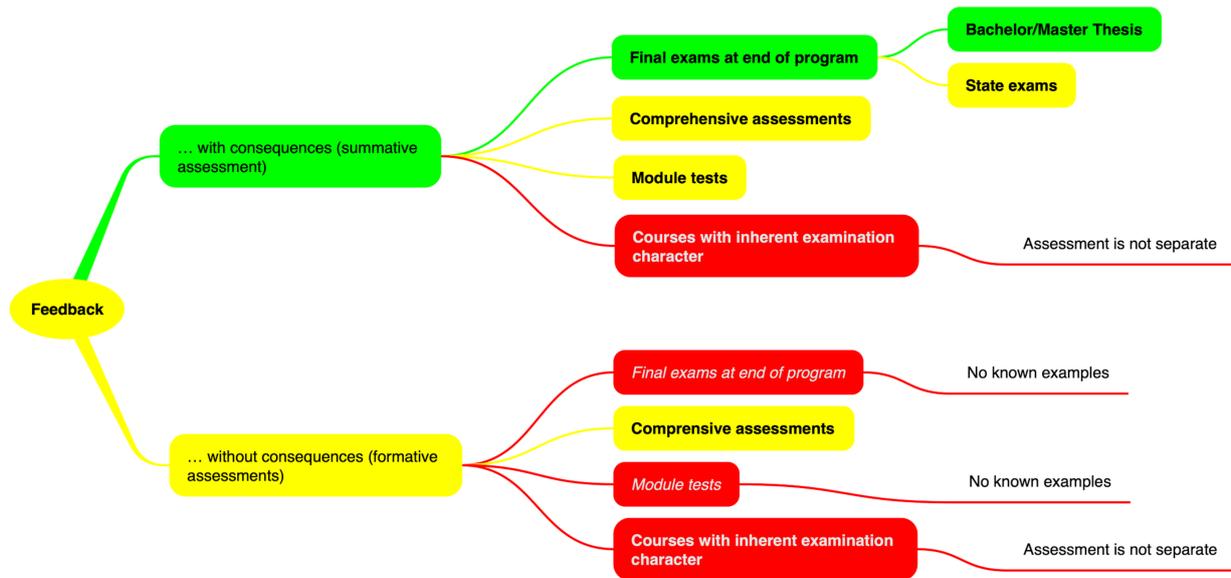


Figure 1: Delineation between feedback which is already counted in teaching loads (green), which could potentially count toward teaching loads (yellow), and feedback that cannot be included in the teaching load (red).

examination regulations. Yet this information is irrelevant for both courses and assessments when calculating their curricular percentage (CAp).

The question remains as to why the basic standards specified by the Federal Constitutional Court for legally sound assessments are not explicitly implemented in all examination regulations, especially since these are also basic criteria for assessment objectivity. The answer may have to do with the unwillingness to depart from long-established practices and the accounting for costs [14]. However, the cost for one individual assessment generally does not change significantly if the test is based on international standards [3], [22] and/or its design is theory-based [23]. Perhaps the working hours expended would change if the assessments could be counted as part of the teaching load. However, working hours already now involve preparing, giving and evaluating assessments. What would definitely see an increase are the estimated costs of a study program or its curricular norm value. Whether the actual cost increases depends on the present quality of precisely these assessments (or the effort invested in them). But if the notion is taken seriously that assessments could constitute an interference with an individual's basic rights, then they should not be hidden in the budget under "miscellaneous". Perhaps then, the absolute number of graduates will increase.

Author's ORCID

Volkhard Fischer: 0000-0001-8499-9437

Competing interests

The author declares that he has no competing interests.

References

1. Bundesverwaltungsgericht. Überdenkungsverfahren bei mündlichen berufsbezogenen Abschlussprüfungen. Urteil vom 10.04.2019. BVerwG 6 C 19.18. Leipzig: Bundesverwaltungsgericht; 2019. Zugänglich unter/available from: <https://www.bverwg.de/100419U6C19.18.0>
2. Bundesverwaltungsgericht. Staatliche Ergänzungsprüfung für den Beruf "Notfallsanitäter". Urteil vom 28.10.2020. BVerwG 6 C 8.19. Leipzig: Bundesverwaltungsgericht; 2020. Zugänglich unter/available from: <https://www.bverwg.de/281020U6C8.19.0>
3. American Educational Research Association, American Psychological Association, National Council on Measurement in education. Standards for Educational and Psychological Testing. Washington, DC: AERA, APA, NCME; 2014.
4. Bandaranayake RC. Setting and maintaining standards in multiple choice examinations: AMEE Guide NO.37. Med Teach. 2008;30(9-10):836-845. DOI: 10.1080/01421590802402247
5. Messick S. Validity. In: Linn RL, editor. Educational Measurement. New York: American Council on Education, Macmillan; 1989. p.13-103.
6. Kane MT. Validating the Interpretations and Uses of Test Scores. J Educ Meas. 2013;50(1):1-73. DOI: 10.1111/jedm.12000
7. Gesellschaft für Medizinische Ausbildung, Kompetenzzentrum Prüfungen Baden-Württemberg, Fischer M. Leitlinie für Fakultätsinterne Leistungsnachweise während des Medizinstudiums: Ein Positionspapier des GMA-Ausschusses Prüfungen und des Kompetenzzentrums Prüfungen Baden-Württemberg. GMS Z Med Ausbild. 2008;25(1):Doc74. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2008-25/zma000558.shtml>
8. Jünger J, Just I. Empfehlungen der Gesellschaft für Medizinische Ausbildung und des Medizinischen Fakultätentags für fakultätsinterne Leistungsnachweise während des Studiums der Human-, Zahn- und Tiermedizin. GMS Z Med Ausbild. 2014;31(3):Doc34. DOI: 10.3205/zma000926
9. Niehues N, Fischer E. Prüfungsrecht. 5. ed. München: C.H. Beck; 2010. p.360.
10. Morgenroth C. Hochschulstudienrecht und Hochschulprüfungsrecht. 3. ed. Baden-Baden: Nomos Verlagsgesellschaft; 2021. p.229.

11. Verwaltungsgericht Hannover. Modulprüfung in elektronischer Form. Beschluss vom 10.12.2008. 6 B 5583/08. Hannover: Verwaltungsgericht Hannover; 2008. Zugänglich unter/available from: <https://openjur.de/u/323500.html>
12. Bundesverfassungsgericht. Verfassungsrechtliche Prüfung des Antwort-Wahl-Verfahrens (Multiple-Choice-Verfahren) als Form der Ärztlichen Prüfung. BVerfG vom 14.03.1989. 1 BvR 1033/82, 1 BvR 174/84. Karlsruhe: Bundesverfassungsgericht; 1989.
13. Medizinische Universität Wien. Curriculum für das Diplomstudium Humanmedizin. Konsolidierte Fassung; Juni 2022. Wien: Medizinische Universität Wien; 2022.
14. Fischer V. Gütekriterien bei universitären Prüfungen im Lichte von Kanes Rahmenwerk [Quality criteria for academic assessments according to Kane's framework]. *Wien Med Wochenschr.* 2019;169(5-6):110-118. DOI: 10.1007/s10354-018-0661-z
15. Bundesministerium für Gesundheit. Überarbeiteter Referentenentwurf der Approbationsordnung für Ärzte (ÄApprO). Berlin: Bundesministerium für Gesundheit; 2023.
16. Nouns ZM, Brauns K. Das Prinzip des Progress-Testing - Detaillierte Leistungsdarstellung und Lehrevaluation auf Basis der kontinuierlichen Wissensentwicklung der Studierenden. In: Dany S, Szczyrba B, Wildt J, editors. *Prüfungen auf die Agenda! Blickpunkt Hochschuldidaktik.* Bielefeld: W. Bertelsmann Verlag; 2008. p.114-128.
17. Schuwirth LW, van der Vleuten CP. Programmatic assessment and Kane's validity perspective. *Med Educ.* 2012;46(1):38-48. DOI: 10.1111/j.1365-2923.2011.04098.x
18. Boulet JR, De Champlain AF, McKinley DW. Setting defensible performance standards on OSCEs and standardized patient examinations. *Med Teach.* 2003;25(3):245-249. DOI: 10.1080/0142159031000100274
19. Leibniz Universität Hannover. Musterprüfungsopdrnung 2021 für Bachelorstudiengänge mit Anhörungsmodell. Hannover: Leibniz Universität Hannover; 2022.
20. Deutscher Hochschulverband. Zur Lehrverpflichtung von Wissenschaftlerinnen und Wissenschaftlern an Universitäten. Resolution. Berlin: Deutscher Hochschulverband; 2019.
21. Kane MT. Validating high-stakes testing programs. *Educ Meas.* 2002;21(1):31-41. DOI: 10.1111/j.1745-3992.2002.tb00083.x
22. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ.* 2015;49(6):560-575. DOI: 10.1111/medu.12678
23. Schuwirth LW, van der Vleuten CP. General overview of the theories used in assessment. Dundee: Association for Medical Education in Europe (AMEE); 2008.

Corresponding author:

PD Dr. Dipl.-Psych. Volkhard Fischer
 Medizinische Hochschule Hannover, Studiendekanat
 Bereich Evaluation & Kapazität, Carl-Neuberg-Str. 1,
 D-30625 Hannover, Germany
fischer.volkhard@mh-hannover.de

Please cite as

Fischer V. On including assessments in the calculation of teaching loads. *GMS J Med Educ.* 2024;41(1):Doc3.
 DOI: 10.3205/zma001658, URN: urn:nbn:de:0183-zma0016584

This article is freely available from

<https://doi.org/10.3205/zma001658>

Received: 2023-08-03

Revised: 2023-11-05

Accepted: 2023-11-15

Published: 2024-02-15

Copyright

©2024 Fischer. 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/>.

Über die Anrechnung von Prüfungen auf das Lehrdeputat

Zusammenfassung

Neben Lehrveranstaltungen stellen Prüfungen die zentralen Bestimmungsgrößen zur Beschreibung eines Studiengangs unabhängig von seinen Inhalten dar. So hängt die Ausbildungsqualität einerseits von der Qualität der Lehrveranstaltungen und ihrer Verzahnung untereinander, andererseits aber auch von der Qualität der Prüfungen und der mit ihnen erfolgenden Leistungsrückmeldungen an die Studierenden ab. Die Qualität der Lehrveranstaltungen soll durch deren Evaluation sichergestellt werden. Der ökonomische Aufwand für die Lehrveranstaltungen wird über das benötigte Lehrdeputat ermittelt. Lehrdeputate stammen aus einer Zeit als dozierendenzentriert akademische Lehre geplant wurde. Zentrale Kenngröße in den Studienordnungen waren die Semesterwochenstunden (bzw. Lehrveranstaltungsstunden). Aber auch in der aktuellen studierendenzentrierten Studiengangsplanung mit ECTS-Punkten pro Modul als Kenngröße, haben Lehrdeputate ihren Nutzen zur Bestimmung des für die Studiengangsdurchführung notwendigen Personalbedarfs nicht verloren.

An manchen Standorten gehen die Prüfungen auch mit in die Evaluation ein. Der mit Prüfungen verbundene ökonomische Aufwand wird de facto aber fast überall ignoriert. Für die Qualität von Prüfungen verheißt dies nichts Gutes. Würde man die Prüfungen in die Berechnung der Lehrdeputate und der Curricularnormwerte einbeziehen, wäre für die Verbesserung der studentischen Ausbildung schon viel gewonnen.

Welche Voraussetzungen beachtet werden müssen, um Prüfungen in die Erfassung der Lehrdeputate mit einzubeziehen, behandelt dieser Artikel.

Schlüsselwörter: Prüfungen, Lehrdeputat, Prüfungsparameter, Curricularnormwert

Einleitung

Prüfungen sind ein wichtiger Bestandteil eines Studiums und können lange juristische Streitigkeiten auslösen. Zwei Urteile des Bundesverwaltungsgerichts [1], [2] haben zu einer Präzisierung der einschlägigen Paragraphen in der Studien- und Prüfungsverordnung für Hebammen [<https://www.gesetze-im-internet.de/hebstprv/BJNR003900020.html>] geführt, obwohl sie sich auf andere Studiengänge bezogen. Gerade weil es in den Urteilen um Grundbedingungen für objektive Prüfungen geht, ist es überraschend, dass es solcher Nachbesserungen immer noch bedarf. An der Aufbereitung des Themas kann es nicht liegen. Es gibt

- internationale Standards für Prüfungen [3], [4], [5], [6],
- Empfehlungen für Praktiker [7], [8] und
- Literatur zum Prüfungsrecht ([9] oder eine neuere Auflage, [10]).

Andererseits zählen Prüfungen im deutschen Hochschulalltag traditionell zur Vor- und Nachbereitung von Lehrver-

Volkhard Fischer¹

1 Medizinische Hochschule
Hannover, Studiendekanat
Bereich Evaluation &
Kapazität, Hannover,
Deutschland

anstaltungen (LV). Und nur LV gehen in die Berechnung der Curricularnormwerte (CNW) ein und werden auf die individuellen Lehrdeputate (bzw. Lehrverpflichtungen) angerechnet. Wenn man sich dann vergegenwärtigt, dass sich ein Hochschulstudium aus (Präsenz- und/oder Online-)LV, verschiedenen Arten von Leistungsrückmeldungen, Eigenstudium und extracurricularen Aktivitäten zusammensetzt, ist klar, dass Leistungsrückmeldungen zusammen mit den LV zu von der Fakultät bzw. den Dozierenden zu erbringenden Leistungen gegenüber den Studierenden gehören.

In einigen Lehrverpflichtungsverordnungen (LVVO) der Bundesländer gibt es Formulierungen, die es Fakultäten ermöglichen könnten, Prüfungen als Untergruppe von Leistungsrückmeldungen die diesen gebührende Aufmerksamkeit zu schenken. Beispielhaft seien in Tabelle 1 die entsprechenden Textstellen aus vier LVVOs aufgeführt. Im Folgenden soll erläutert werden, welche Voraussetzungen erfüllt sein sollten, damit eine Fakultät von diesen Regelungen Gebrauch machen kann. Hintergrund ist folgende Überlegung:

Tabelle 1: Auszüge aus verschiedenen Lehrverpflichtungsverordnungen zur Berücksichtigung von weiteren Lehrveranstaltungen bei der Ermittlung der Lehrleistung

Bayern	LUFV von 2019	§ 3 Absatz 1 Satz 1	Nach Prüfungsordnungen, Studienordnungen oder Studienplänen nicht vorgesehene Lehrveranstaltungen werden berücksichtigt, wenn alle nach diesen Vorschriften vorgesehenen Lehrveranstaltungen eines Fachs durch hauptberuflich oder nebenberuflich an der Hochschule tätige Lehrpersonen angeboten werden; ...
Berlin	LVVO von 2021	§ 3 Absatz 3	Vorlesungen, Übungen, Seminare, Kolloquien, Repetitorien, künstlerischer Einzel- und Gruppenunterricht sowie hinsichtlich der Vor- und Nachbereitungszeit gleichwertige Lehrveranstaltungen, an Fachhochschulen auch seminaristischer Unterricht und Praktika, werden auf die Lehrverpflichtung voll angerechnet.
Niedersachsen	LVVO von 2021	§ 13 Absatz 1, Satz 1	Bei der Erfüllung der Lehrverpflichtung werden die Lehrveranstaltungen einschließlich solcher außerhalb der Vorlesungszeit berücksichtigt, die nach den Studien- oder Prüfungsordnungen oder Studienplänen vorgesehen sind.
Sachsen-Anhalt	LVVO von 2018	§ 5 Absatz 1	Für die Erfüllung der Lehrverpflichtung sind diejenigen Lehrveranstaltungen zu berücksichtigen, die jeweils im Semester nach Prüfungs- oder Studienordnungen für ein ordnungsgemäßes Studium verbindlich sind

- P1) Prüfungen sollten hinreichend bestimmt sein, damit der durch sie erfolgende Eingriff in die Rechte Studierender rechtskonform ist.
- P2) Prüfungen gehören nicht zu den klassischen LV und werden in der Regel nicht auf das Lehrdeputat angerechnet, sondern zählen zur Vor- und Nachbereitung von LV.
- P3) Lässt die LVVO die Anrechnung originär nicht vorgesehener LV zu, dann müssen diese analog zu klassischen LV berechenbar sein.
- S) Um also Prüfungen auf das Lehrdeputat anrechnen zu können, muss man klären, welche Aspekte für die hinreichende Bestimmtheit einer Prüfung den Parametern für die Lehrdeputatsberechnung entsprechen und ob sie losgelöst von LV betrachtet werden können.

Für die Medizinische Hochschule Hannover hatte das Verwaltungsgericht Hannover in einem Beschluss zu einer digital durchgeführten Modulprüfung eine Reihe von Bedingungen formuliert, die Prüfungsordnungen erfüllen sollten [11]. Auch wenn die Rechtsprechung mit Prüfungen an elektronischen Eingabegeräten inzwischen vertrauter ist [10], stellen die damals aufgezeigten Punkte (siehe Tabelle 2) nicht nur wichtige Pfeiler für die Rechtssicherheit von Prüfungen in Studiengängen dar, sondern auch für die Anrechenbarkeit von Prüfungen auf das Lehrdeputat. Ausgangspunkt der juristischen Überlegungen ist, dass durch die Prüfungen das Grundrecht auf Freiheit der Berufswahl nach Artikel 12 Absatz 1 GG eingeschränkt wird, weil ihr endgültiges Nichtbestehen eine spätere Berufsausübung dauerhaft verhindert. Aber auch im allgemeinen Fall einer Prüfung muss diese dem Wesentlichkeitsgrundsatz genügen und darf nur auf Grund eines Gesetzes oder eines durch ein Gesetz eingeräumten Spielraums im Rahmen einer Satzung dieses Grundrecht einschränken. Dafür sollten sich die Abläufe und Anforderun-

gen am Grundsatz der Chancengleichheit orientieren und die gerichtliche Kontrolle des Bewertungsvorgangs weitestgehend sichergestellt werden. Dies passt sehr gut zu den internationalen Standards für objektive, valide und reliable Prüfungen [3], [4], [6].

In der ersten Spalte von Tabelle 2 sind die Parameter für die Berechnung des Curricularanteils (CAp) von klassischen LV aufgelistet. In der zweiten Spalte sind die Merkmale für eine hinreichende Bestimmtheit von Prüfungen ohne Anspruch auf Vollständigkeit aus den genannten Gerichtsurteilen zusammengestellt. Die in der dritten Spalte von Tabelle 2 aufgeführte Teilmenge dieser Aspekte sollte dagegen vollständig sein. Es sind die für die Anrechnung von Prüfungen auf das Lehrdeputat notwendigen Parameter, die eine Quantifizierung analog zu den klassischen LV ermöglichen.

Die in der zweiten Spalte von Tabelle 2 aufgeführten Aspekte betreffen alle die Objektivität von Prüfungen, womit schon deutlich wird, dass sie keineswegs hinreichend sind, um „gute“ Prüfungen sicher zu stellen. Aber es ist die Objektivität einer Prüfung, die (meist) Gegenstand ihrer juristischen Überprüfung ist. Scheinbare Ausnahmen wie das NC-Urteil des Bundesverfassungsgerichts [12] bestätigen dies. Denn damals ging es bei der Klärung der Verfassungsmäßigkeit des Antwort-Wahl-Verfahrens im Kern nicht um die Reliabilität von staatlichen Prüfungen mit Fragen im Mehrfach-Wahl-Format, sondern um deren Schwierigkeit über verschiedene Zyklen hinweg.

D.h. die juristisch geforderten Regelungen betreffen vor allem das Fundament, nicht die Wände oder das Dach des Prüfungsgebäudes eines Studiengangs und sollten selbstverständlich sein. Deshalb werden sie z.B. in den „Empfehlungen“ [8] natürlich auch angesprochen. Der Fokus der Empfehlungen richtete sich aber vor allem auf

Tabelle 2: Auszüge aus verschiedenen Lehrverpflichtungsverordnungen zur Berücksichtigung von Studienabschlussarbeiten bei der Ermittlung der Lehrleistung

	Parameter zur Ermittlung des Curricularanteils von Lehrveranstaltungen	Aspekte von Prüfungen, die für ihre hinreichende Bestimmtheit definiert sein müssen	Parameter zur Ermittlung des Curricularanteils von Prüfungen
1		Teilnahmevoraussetzungen	
2	LV-Inhalt	Prüfungsinhalt	Prüfungsinhalt
3	LV-Dauer	Prüfungsdauer/-umfang	Prüfungsdauer
4	Beteiligte Dozierende	Anzahl der Prüfenden	Beteiligte Prüfende
5	Anzahl der Studierenden	Anzahl der Prüflinge	Anzahl der Prüflinge
6		Bestehensregeln	
7		Konsequenzen	
8		Dokumentation	

die Bedrohungen der Validität und der Reliabilität von Prüfungen, weil man meinte, dass es dort die meisten Schwachstellen gäbe. Insoweit stellen die in der zweiten Spalte von Tabelle 2 aufgeführten Merkmale nur eine Teilmenge der Kriterien für „gute“ bzw. rechtssichere Prüfungen dar. Aber sie machen die in Tabelle 1 enthaltene Forderung, nach einer Definition der weiteren für das Lehrdeputat zu berücksichtigen LV, umsetzbar.

Prüfungen

Die Medizinische Universität Wien unterscheidet in ihrer Prüfungsordnung in einer sehr hilfreichen Weise zwischen LV mit immanentem Prüfungscharakter, Lehrveranstaltungsprüfungen und Gesamtprüfungen [13]. Außerdem gibt es mit der Diplomarbeit noch eine Studienabschlussprüfung. Diese Unterscheidung verdeutlicht die Bandbreite, die Prüfungen haben können und warum es nicht abwegig ist, über die getrennte Anrechnung zumindest bestimmter Prüfungsformen auf das Lehrdeputat nachzudenken. Die Reihenfolge der Erörterung richtet sich dabei nach der Schwere, der mit ihnen in der Regel verbundenen Konsequenzen [14].

Der Frage, ob eine Leistungsrückmeldung benotet oder unbenotet erfolgt, kommt dabei nur eine sekundäre Bedeutung zu. Relevant ist, ob die Leistungsrückmeldung Konsequenzen hat, also nur begrenzt wiederholt werden kann, wenn sie nicht bestanden wurde. Deshalb soll hier als erstes zwischen Leistungsrückmeldungen ohne Konsequenzen (Testate oder formative Prüfungen) und Leistungsrückmeldungen mit Konsequenzen (summative Prüfungen) unterscheiden werden. In diesem Sinne sind dann viele der in Studien- und Prüfungsordnungen als Studienleistung bezeichneten unbenoteten Leistungsrückmeldungen summative Prüfungen, weil ihr Nichtbestehen über Studienzeitverlängerungen oder die Exmatrikulation der Studierenden in die Freiheit der Berufswahl eingreift.

Studienabschlussarbeiten und Staatsexamina

Studienordnungen definieren die verschiedenen Arten von LV, die im Rahmen eines Studiums angeboten wer-

den. Außerdem findet sich in ihnen mindestens ein Paragraph zu den Zulassungs- und Durchführungsbedingungen einer Studienabschlussarbeit. Bzgl. der Anrechnung der Studienabschlussarbeiten sind erst im Nachgang zu den Bologna-Reformen separate Paragraphen in die LV-VOs aufgenommen worden (siehe Tabelle 3), obwohl Studienabschlussarbeiten viel älter als die Bologna-Reform sind.

Damit ist für dieses traditionelle Prüfungsformat klar, dass es auf das Lehrdeputat angerechnet werden kann. Das Besondere dabei ist, dass bei Studienabschlussarbeiten die Anzahl der Betreuungsstunden, anders als beim klassischen Präsenzunterricht, nicht objektiv ermittelt, sondern nur indirekt aus der Workload der Studierenden abgeleitet wird.

Die Staatsexamina in Zahn- und Humanmedizin gehen aktuell nicht in die Berechnung der CNW für diese Studiengänge mit ein. Diesen Prüfungen ist nicht einmal in der zusammenfassenden Berechnung des jeweiligen CNW ein Lehraufwand zugeordnet. Wenn man bedenkt, dass z.B. in der ZApprO die Definition der verschiedenen Abschnitte des Staatsexamens 22 Seiten einnimmt [<https://www.gesetze-im-internet.de/zappro/BJNR093310019.html>] und in der aktuellen Fassung der ÄApprO [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html] immerhin auch schon 12 Seiten umfasst, verwundert dies. Denn gerade für diese Staatsexamina sind in den einschlägigen Verordnungen die von den Verwaltungsgerichten eingeforderten Merkmale für eine hinreichende Bestimmtheit einer Prüfung gut geregelt und sie werden normalerweise auch bei der Prüfungsdurchführung so gut dokumentiert, dass sowohl eine Anrechnung auf das individuelle Lehrdeputat als auch auf den CNW durchaus möglich wäre.

Anders als bei Bachelor- oder Masterarbeiten, ist die die Prüfung verantwortende Institution zwar nicht die Hochschule, sondern ein Landesprüfungsamt. Die Durchführung der eigentlichen Prüfung obliegt aber Mitgliedern einer Hochschule. Eine Anrechnung der Prüfungszeiten auf das Lehrdeputat wäre also durchaus gerechtfertigt, zumal auch niemand bezweifelt, dass die mündlichen Prüfungsformate, im Unterschied zu den schriftlichen Staatsexamina mit bundeseinheitlichen Prüfungsfragen, eine Vorbereitung durch die Dozierenden vor Ort erfor-

Tabelle 3: Auszüge aus verschiedenen Lehrverpflichtungsverordnungen zur Berücksichtigung von Studienabschlussarbeiten bei der Ermittlung der Lehrleistung

Bayern	LUFV von 2019	§ 3 Absatz 8 Satz 1	Betreuungstätigkeiten für Diplom-, Bachelor- und andere Studienabschlussarbeiten sowie vergleichbare Studienarbeiten (Abschlussarbeiten) können nur einmal je Studierenden unter Berücksichtigung des notwendigen Aufwands, insgesamt aber an Universitäten und Kunsthochschulen nur bis zu einem Umfang von zwei und an Fachhochschulen bis zu einem Umfang von drei Lehrveranstaltungsstunden angerechnet werden.
Berlin	LVVO von 2021	§ 3 Absatz 6 Satz 1	Weisen Hochschullehrer mit einer Regellehrerverpflichtung nach § 5 Abs. 1 Satz 1 Nr. 1 Buchstabe b und Abs. 3 Satz 1 Nr. 1 eine überdurchschnittliche Belastung durch die Betreuung von Studienabschlussarbeiten oder vergleichbaren Studienarbeiten nach, so kann die Dienstbehörde oder Personalstelle diese nach Maßgabe des Haushalts auf die Regellehrerverpflichtung anrechnen, wenn es die Situation im jeweiligen Fach zulässt.
Niedersachsen	LVVO von 2021	§ 15	Betreuungstätigkeiten für Studienabschlussarbeiten und vergleichbare Studienarbeiten können durch die Hochschule in einem dem Zeitaufwand entsprechenden Umfang mit bis zu 2 LVS je Semester berücksichtigt werden, wenn eine besondere Belastung durch die Betreuungstätigkeit vorliegt.
Sachsen-Anhalt	LVVO von 2018	§ 3a Satz 1	Besondere Belastungen durch die Betreuung von Studienabschlussarbeiten können unter Berücksichtigung des notwendigen Aufwands und der Auslastung der jeweiligen Lehreinheit auf die Erfüllung der Lehrverpflichtung angerechnet werden.

dem. Umgekehrt erkennen die Verwaltungsgerichte die Durchführung dieser Staatsexamina aber nicht als Gründe für eine Deputatsreduktion an, weil sie ja Teil der akademischen Lehrverpflichtung seien.

Gesamtprüfungen

Gesamtprüfungen fassen die Inhalte aller LV eines Semesters oder eines Studienjahres zusammen, unabhängig davon, ob die Gesamtprüfung formativ oder summativ angelegt ist [13]. Ähnlich wie Studienabschlussprüfungen oder Staatsexamina sind die summativen Gesamtprüfungen keineswegs als Vor- und Nachbereitung einzelner LV konzipiert. Ihre Erstellung, Durchführung und Auswertung kann also nur schwer als Teil der Vor- und Nachbereitung einer LV oder eines Moduls verstanden werden. Gleichwohl sind sie integraler Bestandteil des Studiengangs, wenn sie z.B. über das Fortschreiten in den nächsten Studienabschnitt entscheiden.

Ein aktuelles Beispiel für eine solche summative Gesamtprüfung stellt die im überarbeiteten Referentenentwurf einer neuen Approbationsordnung Humanmedizin [15] in § 37 geforderte klinisch-praktische Prüfung (also eine Objective Clinical Examination: OSCE) dar. Die zukünftige OSCE nach § 37 ÄApprO würde alle Voraussetzungen der LVVOs für eine Anrechnung auf das Lehrdeputat erfüllen. Und die Festlegung der Prüfungszeit wäre sogar in einem materiellen Gesetz geregelt.

Ein Beispiel für formative Gesamtprüfungen wären die an vielen Standorten angebotenen freiwilligen Teilnahmen an einem Progress-Test [16], wenn sie für den formalen Studienfortschritt irrelevant sind. Sie stellen eine besondere Form des Eigenstudiums dar.

Modulprüfungen

Lehrveranstaltungsprüfungen schließen eine LV durch eine Prüfung als singulären Akt an einem, maximal zwei Tagen ab. Die Wiener Ordnung [13] lässt dabei offen, ob sich diese nur auf die Vorlesung eines Blockes oder auf die Kombination aus Vorlesung, Seminar und Praktikum bezieht, wie man es für eine Modulprüfung unterstellen würde. Verallgemeinernd soll hier von Modulprüfungen gesprochen werden.

Auf den ersten Blick spricht viel dafür, dass man Prüfungen, die den Studierenden eine Leistungsrückmeldung zum gerade besuchten Modul geben, wie bisher einfach als Teil der Vor- und Nachbereitung betrachtet. Diese traditionelle Herangehensweise macht auch Sinn bei unstandardisierten Prüfungen mit einer unklaren Objektivität, die nicht die Minimalbedingungen an eine nachvollziehbare Prüfung erfüllen (sollen). Sie entsprechen damit aber einer Prüfungspraxis, die in den letzten Jahren immer wieder von den Verwaltungsgerichten als rechtswidrig verworfen wurde [8], [9].

Weil sich Module häufig aus verschiedenen Lehrveranstaltungsformaten zusammensetzen, ist die Modulprüfung nicht mehr uneindeutig der Vor- und -nachbereitung einer

LV zuzuordnen. Außerdem erfordert eine standardisierte Prüfung einen hohen Aufwand, wenn sie objektiv, reliabel und valide sein soll. So ist eine schriftliche Prüfung im Antwort-Wahl-Verfahren, denen Verwaltungsgerichte generell in Verkennung der Literatur zur Prüfungsqualität [17] ja ein besonderes Misstrauen entgegenbringen, mit 60 oder mehr Fragen nicht in der Zeit erstellt, die man für die Erstellung einer zweistündigen Vorlesung benötigt. Und eine 12-Stationen OSCE erfordert auch mehr Aufwand, als wenn man einen Jahrgang Medizinstudierende in Kleingruppen je vier Stunden lang im SkillsLab unterrichtet [18].

Für die Anrechenbarkeit auf das Lehrdeputat sprechen neben dem mit Tabelle 2 genügenden Prüfungen verbundenem Aufwand auch rein pragmatische Gründe. Es ist kein Hexenwerk in einer Modulbeschreibung neben der Vorlesung auch die Prüfung separat auszuweisen und die Prüfungsform darüber hinaus in der Prüfungsordnung zu definieren und alle Modulprüfungen explizit aufzulisten. Selbst wenn die Vorlesung über die Jahre unverändert bestehen bleibt, kann dann aus der rein schriftlichen Prüfung mit Freitextfragen im Laufe der Jahre eine Prüfung nach dem Antwort-Wahl-Verfahren an elektronischen Endgeräten werden und dies durch eine kleine Änderung der Prüfungsordnung dokumentiert werden.

Bzgl. der Differenzierung zwischen Modulprüfungen mit und ohne Konsequenzen gilt das Gleiche wie schon für die Studienabschlussprüfungen gesagt wurde. Formative Modulprüfungen würden eine besondere Form des Eigenstudiums darstellen.

Lehrveranstaltungen mit immanentem Prüfungscharakter

Wenn innerhalb einer LV bestehensrelevante Leistungsrückmeldungen gegeben werden, handelt es sich nach der Wiener Prüfungsordnung [13] um eine LV mit immanentem Prüfungscharakter. Die Musterprüfungsordnung der Leibniz-Universität Hannover [19] spricht von Veranstaltungsbegleitenden Prüfungen. In beiden Namen kommt die Tatsache zum Ausdruck, dass es sich nicht um einen singulären Prüfungsakt, sondern um wiederholte, mehrfache Prüfungsakte handelt, die aber als Gesamtleistung bewertet werden.

Klassische Beispiele für solche Leistungsrückmeldungen sind Referate in Seminaren, das Anfertigen von Gipsabdrücken und Modellen in der Zahnmedizin oder auch die Erstellung von Laborprotokollen in naturwissenschaftlichen Praktika. Eben weil diese studentischen Leistungen während der normalen Seminar- oder Praktikumstermine erbracht werden, sind sie aber nicht von den Präsenzterminen zu trennen. Sie können auch nicht separat als Prüfung auf das Lehrdeputat angerechnet werden, denn das Referat einzelner Studierender in den Seminarveranstaltungen ist für die Zuhörenden eine LV und für die Vortragenden eine Prüfungsleistung, wobei die Rollen von Termin zu Termin wechseln. Und das Betreuungsverhältnis Lehrende: Studierende bezieht sich bei den Stu-

dierenden auf die Vortragenden (Prüflinge) und die Zuhörenden zusammen.

Anders sieht es bei einem Seminar aus, in dem verschiedene Themen an den Terminen besprochen werden, die dann in zum Semesterende abzugebenden Hausarbeiten vertieft werden sollen. Auch ein Praktikum, in dem klinische Untersuchungstechniken trainiert werden, ist keine LV mit immanentem Prüfungscharakter, wenn die summative Leistungsrückmeldung über eine abschließende OSCE erfolgt. Diese Prüfungen haben einen singulären Charakter und fallen unter die im vorherigen Abschnitt dargestellte Kategorie, weil sie von den LV zeitlich getrennt sind.

Diskussion

Wenn Leistungsrückmeldungen Studierenden eine didaktisch differenzierte Rückmeldung geben sollen, ist ihre Erstellung, Durchführung und Auswertung mit Aufwand verbunden. Deshalb hatte der Deutsche Hochschulverband 2019 seinen Vorstoß zu einer Reduktion des Lehrdeputates für Hochschullehrer u.a. mit den erhöhten Anforderungen an Prüfungen begründet [20]. Eine solche Lösung fördert aber nicht die Erstellung qualitativ hochwertiger Prüfungen, sondern behindert sie nur weniger als die bestehende Praxis. Und da Leistungsrückmeldungen mit Konsequenzen für den weiteren Lebensweg der Studierenden verbunden sind, wenn es sich um summative Prüfungen handelt, sollten sie hinreichend bestimmt sein, damit sie nicht nur objektiv, valide und reliabel [21], sondern auch rechtskonform sind. Je weitreichender die Konsequenzen von Prüfungen sind, desto transparenter sollten die Bedingungen sein, unter denen eine Prüfung durchgeführt wird und desto höher (in juristischen Sinne) muss die Prüfung verankert sein [9]. Die hier vorgeschlagene separate Anrechnung von Prüfungen auf das Lehrdeputat würde dies begünstigen.

Für alle im vorhergehenden Abschnitt diskutierten summativen Prüfungstypen kann dies als minimales Niveau nur die jeweilige Prüfungsordnung mit expliziten Festlegungen nach den in Tabelle 2 genannten Aspekten sein. Wegen der fehlenden Konsequenzen müssten formative Prüfungen aus Gründen der Rechtskonformität nicht in einer Prüfungsordnung explizit geregelt sein. Für ihre Anrechenbarkeit auf das Lehrdeputat wäre es aber eine notwendige Voraussetzung. Für die Verbesserung der Studienqualität ebenfalls.

Abbildung 1 verdeutlicht die hier vorgeschlagene Unterscheidung von schon jetzt auf das Lehrdeputat anrechenbare (grün), potenziell auf das Lehrdeputat anrechenbare (gelb) und nicht anrechenbare Prüfungen (rot). Akzeptiert man diese Unterteilung so stellt sich die Frage, warum von den Verwaltungsgerichten als notwendig angesehene Kriterien für eine hinreichende Bestimmtheit von Prüfungen nur zum Teil für die Anrechnung auf das Lehrdeputat notwendig sind.

Für die Anrechnung von LV muss man den Inhalt der LV in Form des Themas, die Veranstaltungsdauer in Form



Abbildung 1: Unterteilung verschiedener Arten von Leistungsrückmeldungen in schon jetzt auf das Lehrdeputat anrechenbare (grün), potenziell auf das Lehrdeputat anrechenbare (gelb) und nicht anrechenbare Prüfungen (rot).

von Stunden und die normative Betreuungsrelation bzw. Gruppengröße kennen. In der Regel wird dabei davon ausgegangen, dass eine lehrende Person permanent im Raum ist. Diese Parameter sind auch für Prüfungen in einer rechtskonformen Prüfungsordnung festgelegt und für eine Anrechnung unverzichtbar. Die Tatsache, dass es trotzdem viele Studien- und Prüfungsordnungen gibt, die dies nicht umsetzen, sagt nur etwas über das Ausmaß an Intransparenz der jeweiligen Studiengänge aus.

Unter welchen Bedingungen Studierende für eine LV oder eine Prüfung zugelassen werden, die Teilnahme bzw. das Bestehen bescheinigt wird, welche Konsequenzen eine Nicht-Teilnahme bzw. ein Nicht-Bestehen hat und wie dies alles dokumentiert wird, sind in guten Studien- bzw. Prüfungsordnungen in der Regel auch explizit geregelt. Für die Berechnung des jeweiligen Curricularanteils sind diese Daten aber weder bei LV noch bei Prüfungen relevant.

Es bleibt die Frage, warum nicht in allen Prüfungsordnungen die vom BVerwG festgestellten grundlegenden Standards für rechtssichere Prüfungen explizit umgesetzt werden, zumal diese auch grundlegende Kriterien für die Objektivität von Prüfungen sind. Die Antwort hat vielleicht mit der Bereitschaft eingefahrene Wege zu verlassen und Kosten auszuweisen zu tun [14]. Die Kosten einer einzelnen Prüfung verändern sich jedenfalls in der Regel nicht signifikant, wenn man sich für eine Prüfung an internationalen Standards orientiert [3], [22] und/oder sie theoriegeleitet konstruiert [23]. Vielleicht würden sich die aufgewendeten Arbeitsstunden verändern, wenn man Prüfungen auf das Lehrdeputat anrechnen kann. Arbeitsstunden sind Vorbereitung, Durchführung und Auswertung von Prüfungen allerdings auch heute schon. Definitiv steigen würden die zu veranschlagenden Kosten für die Durchführung eines Studiengangs bzw. dessen CNW. Ob die tatsächlichen Kosten steigen, hängt von der aktuellen Qualität eben dieser Prüfungen (bzw. dem in sie investierten Arbeitsaufwand) ab. Aber wenn man den Gedanken

erst nimmt, dass Prüfungen eine Grundrechtseinschränkung darstellen können, sollte man dies nicht versteckt unter „Sonstiges“ verbuchen. Und vielleicht steigen dann die Absolventenzahlen absolut betrachtet an.

ORCID des Autors

Volkhard Fischer: 0000-0001-8499-9437

Interessenkonflikt

Der Autor erklärt, dass er keine Interessenkonflikte im Zusammenhang mit diesem Artikel hat.

Literatur

1. Bundesverwaltungsgericht. Überdenkungsverfahren bei mündlichen berufsbezogenen Abschlussprüfungen. Urteil vom 10.04.2019. BVerwG 6 C 19.18. Leipzig: Bundesverwaltungsgericht; 2019. Zugänglich unter/available from: <https://www.bverwg.de/100419U6C19.18.0>
2. Bundesverwaltungsgericht. Staatliche Ergänzungsprüfung für den Beruf "Notfallsanitäter". Urteil vom 28.10.2020. BVerwG 6 C 8.19. Leipzig: Bundesverwaltungsgericht; 2020. Zugänglich unter/available from: <https://www.bverwg.de/281020U6C8.19.0>
3. American Educational Research Association, American Psychological Association, National Council on Measurement in education. Standards for Educational and Psychological Testing. Washington, DC: AERA, APA, NCME; 2014.
4. Bandaranayake RC. Setting and maintaining standards in multiple choice examinations: AMEE Guide NO.37. Med Teach. 2008;30(9-10):836-845. DOI: 10.1080/01421590802402247
5. Messick S. Validity. In: Linn RL, editor. Educational Measurement. New York: American Council on Education, Macmillan; 1989. p.13-103.
6. Kane MT. Validating the Interpretations and Uses of Test Scores. J Educ Meas. 2013;50(1):1-73. DOI: 10.1111/jedm.12000

7. Gesellschaft für Medizinische Ausbildung, Kompetenzzentrum Prüfungen Baden-Württemberg, Fischer M. Leitlinie für Fakultätsinterne Leistungsnachweise während des Medizinstudiums: Ein Positionspapier des GMA-Ausschusses Prüfungen und des Kompetenzzentrums Prüfungen Baden-Württemberg. *GMS Z Med Ausbild.* 2008;25(1):Doc74. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2008-25/zma000558.shtml>
8. Jünger J, Just I. Empfehlungen der Gesellschaft für Medizinische Ausbildung und des Medizinischen Fakultätentags für fakultätsinterne Leistungsnachweise während des Studiums der Human-, Zahn- und Tiermedizin. *GMS Z Med Ausbild.* 2014;31(3):Doc34. DOI: 10.3205/zma000926
9. Niehues N, Fischer E. Prüfungsrecht. 5. ed. München: C.H. Beck; 2010. p.360.
10. Morgenroth C. Hochschulstudienrecht und Hochschulprüfungsrecht. 3. ed. Baden-Baden: Nomos Verlagsgesellschaft; 2021. p.229.
11. Verwaltungsgericht Hannover. Modulprüfung in elektronischer Form. Beschluss vom 10.12.2008. 6 B 5583/08. Hannover: Verwaltungsgericht Hannover; 2008. Zugänglich unter/available from: <https://openjur.de/u/323500.html>
12. Bundesverfassungsgericht. Verfassungsrechtliche Prüfung des Antwort-Wahl-Verfahrens (Multiple-Choice-Verfahren) als Form der Ärztlichen Prüfung. BVerfG vom 14.03.1989. 1 BvR 1033/82, 1 BvR 174/84. Karlsruhe: Bundesverfassungsgericht; 1989.
13. Medizinische Universität Wien. Curriculum für das Diplomstudium Humanmedizin. Konsolidierte Fassung: Juni 2022. Wien: Medizinische Universität Wien; 2022.
14. Fischer V. Gütekriterien bei universitären Prüfungen im Lichte von Kanes Rahmenwerk [Quality criteria for academic assessments according to Kane's framework]. *Wien Med Wochenschr.* 2019;169(5-6):110-118. DOI: 10.1007/s10354-018-0661-z
15. Bundesministerium für Gesundheit. Überarbeiteter Referentenentwurf der Approbationsordnung für Ärzte (ÄApprO). Berlin: Bundesministerium für Gesundheit; 2023.
16. Nouns ZM, Brauns K. Das Prinzip des Progress-Testing - Detaillierte Leistungsdarstellung und Lehrevaluation auf Basis der kontinuierlichen Wissensentwicklung der Studierenden. In: Dany S, Szczyrba B, Wildt J, editors. Prüfungen auf die Agenda! Blickpunkt Hochschuldidaktik. Bielefeld: W. Bertelsmann Verlag; 2008. p.114-128.
17. Schuwirth LW, van der Vleuten CP. Programmatic assessment and Kane's validity perspective. *Med Educ.* 2012;46(1):38-48. DOI: 10.1111/j.1365-2923.2011.04098.x
18. Boulet JR, De Champlain AF, McKinley DW. Setting defensible performance standards on OSCEs and standardized patient examinations. *Med Teach.* 2003;25(3):245-249. DOI: 10.1080/0142159031000100274
19. Leibniz Universität Hannover. Musterprüfungsopdrnung 2021 für Bachelorstudiengänge mit Anhörungsmodell. Hannover: Leibniz Universität Hannover; 2022.
20. Deutscher Hochschulverband. Zur Lehrverpflichtung von Wissenschaftlerinnen und Wissenschaftlern an Universitäten. Resolution. Berlin: Deutscher Hochschulverband; 2019.
21. Kane MT. Validating high-stakes testing programs. *Educ Meas.* 2002;21(1):31-41. DOI: 10.1111/j.1745-3992.2002.tb00083.x
22. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ.* 2015;49(6):560-575. DOI: 10.1111/medu.12678
23. Schuwirth LW, van der Vleuten CP. General overview of the theories used in assessment. Dundee: Association for Medical Education in Europe (AMEE); 2008.

Korrespondenzadresse:

PD Dr. Dipl.-Psych. Volkhard Fischer
 Medizinische Hochschule Hannover, Studiendekanat
 Bereich Evaluation & Kapazität, Carl-Neuberg-Str. 1,
 30625 Hannover, Deutschland
fischer.volkhard@mh-hannover.de

Bitte zitieren als

Fischer V. On including assessments in the calculation of teaching loads. *GMS J Med Educ.* 2024;41(1):Doc3.
 DOI: 10.3205/zma001658, URN: urn:nbn:de:0183-zma0016584

Artikel online frei zugänglich unter

<https://doi.org/10.3205/zma001658>

Eingereicht: 03.08.2023

Überarbeitet: 05.11.2023

Angenommen: 15.11.2023

Veröffentlicht: 15.02.2024

Copyright

©2024 Fischer. 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/>.