User-generated Student Course Evaluations: (How) Can Key Competencies become Systematic Evaluation Parameters?

Dennis Klinkhammer, Michael Schemmann

Student course evaluations are well known to most students, but their expertise is rarely taken into consideration when it comes to designing them. Since student course evaluations are supposed to provide insights into the quality of heterogeneous courses, which has to be considered challenging from a statistical point of view, students theorized the quality of common student course evaluations within an advanced seminar in educational research. They reflected on different designs, statistical issues as well as imprecise questions and questionnaires before focusing on key competencies as new type of evaluation parameters. Key competencies shall provide a focus on teaching effectiveness and can operate without the need for comparing average scores of heterogeneous courses. By addressing the professional, the methodological, the social as well as the self-competence of each student, the different levels of key competencies within each course can separately be dealt with. Furthermore, a student course evaluation focusing key competencies provides a perfect data basis to be thoroughly tested by factor and reliability analysis in order to highlight the quality of the students' approach and their understanding of evaluation research.

Keywords: User-generated student course evaluation; Key competencies; Factor analysis; Reliability analysis

1 Introduction – students rethinking student course evaluations

Student course evaluations are an ubiquitous phenomenon and most commonly used to evaluate the quality of academic courses (Seldin 1999; Clayson 2009; Davis 2009). They are used within the process of accreditation and institutional evaluation while providing insights for lecturers, research assistants, professors and faculties into the quality and teaching effectiveness of the courses. Most student course evaluations are realized with standardized paper-based questionnaires and take place on a regular basis, commonly at the end of each term. Although student course evaluations are subjective surveys they are supposed to be an objective basis for quality development and are to support the dialogue between lecturers, research assistants, professors and students as well. To most users they appear to be objective because

108

User-generated Student Course Evaluations: (How) Can Key Competencies become Systematic Evaluation Parameters?

they are numerical, but from a statistical point of view that is no conclusive argument (Pounder 2007). Considering student course evaluations as an instrument to measure the teaching effectiveness, this article will sum up some of the most common statistical issues during the evaluation process and unfold the complexity of implementing them.

The article will focus on the process of designing user-generated student course evaluations in order to cope with the complexity as well as the statistical issues. The underlying question is: (How) Can Key Competencies become Systematic Evaluation Parameters?

This user-generated questionnaire is supposed to address the teaching effectiveness by turning away from the usual comparison of average scores and "happy sheets" (Kirkpatrick 1998) and considering key competencies as new type of evaluation parameters. These key parameters have been designed by students who made themselves familiar with the theoretical concept of key competencies as well as the basics of evaluation research and statistics. They have been asked to reflect on their study experience and to apply the general idea of key competencies for designing a new student course evaluation.

Since there are four dimensions of key competencies known in psychological and pedagogical discourse, known as professional, methodological, social and self-competence, the main research questions of this article deal with the quality of the students' approach:

1. How do students operationalize the theoretical framework of key competencies?
2. Can this approach cover the four dimensions of key competencies in practice?

Both research questions provide insights into the students' understanding of statistical methods used in evaluation research and their operationalization skills when it comes to designing a questionnaire. Factor and reliability analysis shall confirm the students' underlying theoretical framework and their techniques in designing the subsequent questions while further insights into the application of the user-generated student course evaluation highlight the potential of the new design.

Therefore, this article provides a synopsis of the complexity of student course evaluations, as taught to the students, focusing the intention of the student course evaluation, common statistical issues as well as the role of the evaluation teams designing them (chapter 2). Subsequently, the theoretical framework of the four dimensions of key competencies and the students' questionnaire shall be highlighted (chapter 3), before a multistage-testing procedure, the application context as well as the used data sets are introduced (chapter 4). The findings focus especially on factor and reliability analysis (chapter 5) before the conclusion provides insights into the application of this user-generated student course evaluation and closes with some recommendations and notes on teaching effectiveness of this approach (chapter 6).

2   Considering the complexity

In the following we will discuss the complexity of student course evaluation. As simple as the usually generated means and standard deviations of student course evaluations seem to be, as complex is the way to an adequate questionnaire for student course evaluations. Authors like Marsh and Roche (1997), Perry and Smart (1997) and Diehl (2001), for example, highlight the complexity of student course evaluations according to their variety of approaches and intentions. Therefore, Rindermann (2001; 2003) uses a multidimensional model in order to structure this variety among measurable outcomes (chapter 2.1). Furthermore, Marsh (2007), as well as Cashin and Clegg (1987) and McKeachie (1997), focus on the statistical issues in performing student course evaluations (chapter 2.2.), which may lead to difficulties in terms of clarity and interpretation. In order to counteract these issues, they need to be addressed by the evaluating teams with reasonable care (chapter 2.3).

2.1   Approach and intention

Depending on approach and intention, student course evaluations may differ in terms of their content as well as in quality. A variety of variables is at hand and leads to different evaluation designs between universities and even between different faculties within one university. As a result, it is challenging to identify design standards when talking about student course evaluations in general. However, in terms of approach and intention Rindermann (2003, p. 235 f.) differentiates four main dimensions of measurable outcomes for student course evaluations:

1. First of all, they may address the behaviour, knowledge, working materials and engagement of the lecturer, research assistant or professor.
2. Secondly, they may focus on determinants concerning mainly the students, e. g. their background knowledge and diligence.
3. Contextual factors like requirement level and type of course may also be dealt with in student course evaluations.
4. Finally, they may capture the educational success of the students.

Given that kind of differentiation, Rindermann (2001) develops a *Multidimensional Condition Model for Educational Success,* which summarizes different aspects within educational contexts. Similar references to the complexity of educational contexts and a variety of aspects can be found in the scientific debate since 1997 (Marsh & Roche 1997; Perry & Smart 1997; Diehl 2001; Greimel 2002, Clayson 2009). Evaluation teams can try to cover all four dimensions, but most of them are likely to make a selection according to their intentions.

A focus on key competencies, as presented in chapter 3, might address the intersection of these main dimensions. Hereby, measuring key competencies might take the individual background of the students into account, by accepting different outcomes of educational success as perceived by the students themselves, as well as ac-

knowledging a broad spectrum of influencing factors and possibilities of interpretation.

2.2   Common statistical issues

Discussing student course evaluations there is much that can be criticized: First of all, most higher education courses tend to get a distinctly good evaluation result with a statistical bias in terms of interpretation (Daniel 1996; Marsh 2007). Of course, there is a variety of explanations for that kind of behaviour at hand:

1. A quite obvious one is, that some students might want to give their lecturers good credits despite his or her teaching effectiveness.
2. Following that clue, some of the variance seems only to be related to the personal influence or charisma of the lecturer or the subject of the course, instead of its relevance and general significance. This issue is especially addressed by the work of Marsh (2007).
3. Furthermore, students from different academic fields tend to rate higher education courses differently, as Cashin and Clegg (1987) point out.
4. McKeachie (1997) reminds evaluators to recall the fact that some of the courses are mandatory while others are freely chosen and that this might have an impact on the outcome of student course evaluations as well.

It should also be noted that not all students participate in the student course evaluation. Some of them are missing, because of their absence on evaluation day or their general drop out. These kinds of nonresponses produce uncertainties, because it is unlikely that the nonresponders would have acted the same way than the responders did; whether they liked the course or left it because they didn't like it at all. Remaining students might be influenced by the size of the course. Smaller courses tend to decrease the level of anonymity and that might reduce the willingness to respond truthfully.

In summary, a lot of issues can influence the response behaviour, starting at the micro-level of the students, covering various aspects at the level of the educational context and going up to the level of the evaluated course itself. That said, none of these levels can be considered as static. In summary, students are not alike: Some of them have a better comprehension of things, a different social behaviour towards their lecturers or even a different background knowledge. Furthermore, courses and their contexts are also not alike (Beleche, Fairris & Marks 2012), which requires a more flexible and yet comparable evaluation instrument. Again, a focus on measuring key competencies might address these issues, by accepting different outcomes at an individual level, according to the differences of each students' context. The flexible evaluation goals of measuring key competencies and their counteracting potential in terms of these statistical issues will be highlighted in chapter 3.

2.3 Comprehension of the evaluation teams

It should also be noted that the statistical and methodological comprehension of the evaluation team is vital to the whole process of designing and implementing a student course evaluation. Otherwise, a lack of statistical and methodological knowledge may lead to more or less sophisticated evaluation designs (Marsh 2007; McCollough & Radson 2011). Therefore, Marsh, as well as McCollough and Radson, define four necessary key aspects of statistical and methodological knowledge:

1. The evaluation teams need to be familiar with applicable standards in terms of evaluation designs.
2. They need to know how to formulate precise questions in terms of reliability and validity.
3. Furthermore, they need to anticipate the possible response behaviour of the students in order to prevent biases.
4. And they should use objective methods of interpretation.

Without these four key aspects of statistical and methodological knowledge imprecise questions and misleading evaluation designs could differ from the genuine intention of the student course evaluation or falsify the results, as the chapter on common statistical issues pointed out (chapter 2.2). Therefore, the students involved in setting up a user-generated student course evaluation (chapter 3.2), have previously discussed all of these issues and familiarized themselves with necessary knowledge before constructing, testing and implementing a new student course evaluation. The theoretical framework for constructing a student course evaluation that will cover up the complexity is presented in the next chapter.

3 Designing the questionnaire: Key competencies as theoretical framework

The European Commission (2018) considers specific knowledge, skills and attitudes as key competencies for personal fulfilment, participation on the labour market and social participation. All of them can be considered content of academic seminars. The European Commission is quite concrete in describing some of these key competencies in detail, e. g. communication in one's mother language and foreign languages as well, scientific community takes a more general approach. Our approach refers to four acknowledged key competencies: Roth (1971), one of the central actors in the field of key competencies, stated that professional competence, social competence and self-competence are the basic competencies that would have to be learned for educational and work-related success. This approach refers mainly to White's (1959) psychological concept of competencies that seem to be necessary to interact effectively with the environment. Since then, all three key competencies have served as reference for further adaptions of the original concept. Nowadays experts differ in four dimensions of key competencies and their subdivision by professional, meth-

odological, social and self-competence, as summarized, for example, by Maurer (2006) in his article on competencies and educational standards as well as the items of the International Centre for Higher Education Research (INCHER) Kassel (2014).

3.1   Synopsis of the four dimensions of key competencies

One advantage of measuring key competencies is the clear intention (chapter 2.1.) of the theoretical framework taking into account that each student is unique and might have different evaluation outcomes. The students also assumed that academic courses differ in respect of their content and so should the key competencies of the students within. Hence, academic courses are rarely comparable figure by figure, but could be categorized according to their content and in respect to the key competencies achieved by the students. As a result, lecturers, research assistants and professors shall have a better insight into their teaching effectiveness without the need for statistically invalid comparisons with all too different courses, but in accordance with the accreditation and evaluation goals of their universities.

Following the psychological and pedagogical discourse on key competencies, the students designed a set of questions (chapter 3.2) in respect to the known statistical issues (chapter 2.2) and in accordance with the four dimensions of key competencies as follows:

1. The professional competence addresses a broad range of knowledge, theories, attitudes and skills required in order to work in a specialized area or profession (Klippert 1994; Maurer 2006).
2. Within the scientific discourse methods and methodological competence are referred to as tools for almost every educational and work-related success (Klippert 1994; Trautwein 2011). A correct and situation-specific application of these methods can be critical in most processes.
3. The concept of social competence refers to the ability to act appropriately within social interactions and to get along with others (Roth 1971; Maurer 2006). Thus, social competence is related to interpersonal communication, the perception of others and the self-perception in respect to others.
4. Self-competence can be described as someone's attitudes and abilities to reflect upon himself and his own strengths and weaknesses. It is sometimes referred to be the basis for developing other competencies (Maurer 2006).

In summary, a reference to Klipperts (1994) model of extended learning concept can be observed within the four dimensions of key competencies. First research projects operate explicitly on basis of this fourfold division (e. g. Trautwein 2011) by turning away from measuring items that reflect solely on the lecturer, his materials or the external situation. Since meanwhile universities use these four dimensions of key competencies, which can be referred to as classification of learning outcomes, in order to design the contents of their courses (Zentrum für Qualitätssicherung und -entwicklung der Johannes Gutenberg Universität Mainz 2014), a focus in terms of evaluation seems appropriate.

Nevertheless, this fourfold division is rarely tested. Therefore, the following questions designed by students themselves not only provide first insights on how students perceive their courses and study experience in respect to the key competencies, but are also a basis for multistage-testing of the theoretical framework. The user-generated questions and items will be presented in the next chapter.

3.2 User-generated item structure

This section introduces the user-generated items. As for the designing procedure, students in the field of adult education who have been familiar with the complexity of student course evaluations, statistics and the theoretical framework of four dimensions of key competencies have been asked to work in groups, in order to work out relevant questions for each competence. One group addressed professional competence, another methodological competence, while the two other groups were dealing with social competence and self-competence. Since all groups have been familiar with all four competencies, they could discuss their set of questions with other students and with the accompanying research team. A selection of the most precise and purposeful questions during pretest (chapter 4.2) served as basis for the new questionnaire. In total, the students developed 16 different questions on basis of a five-level scale and in accordance with the four dimensions of key competencies (chapter 3.1):

(1) I have acquired and expanded professional knowledge. (pc_knowledge)

(2) The course was a useful addition to my field of studies. (pc_study)

(3) The course provided relevant concepts and theories. (pc_theory)

(4) I know methods that refer to my professional competence. (pc_methods)

(5) I can practically apply the theoretical knowledge. (mc_use)

(6) I could try out techniques and methods. (mc_test)

(7) I can objectively reflect techniques and methods. (mc_reflect)

(8) I can purposefully work towards a result. (mc_results)

(9) I know how to deal with conflict situations. (soc_conflict)

(10) I was able to take responsibilities. (soc_responsibility)

(11) I was able to reach out and ask questions. (soc_questions)

(12) I'm sure to gain acceptance. (soc_enforce)

(13) I can deal with complex subjects. (sec_facts)

(14) The course encourages my organizational skills. (sec_organize)

(15) I can work in a structured way. (sec_structure)

(16) I was able to set my own priorities within the course. (sec_own)

Associated items are expelled in addition to the questions, indicating one of the four key competencies by prefix. Labelled items make it easier to identify each question in the continuing analysis (chapter 5). The methodological approach and application context are summarized within the next chapter.

4   Multistage-testing procedure and application context

This chapter refers to the methodological approach for testing the students' concept of a student course evaluation focusing key competencies as evaluation parameters. Up to four steps of analysis will be introduced, before the application of the new student course evaluation at the University of Cologne and the generated data sets are presented in detail.

4.1   Multistage-testing of the user-generated student course evaluation

To confirm the consistency of the four dimensions of key competencies within the questionnaire, a thorough check of all user-generated items was performed by using exploratory factor analysis within the pretest and confirmatory factor analysis, reliability analysis and correlations with larger data sets.

Exploratory factor analysis was used during pretest to identify the structure of the user-generated items according to each of the four dimensions of key competencies as presented in chapter 3.2. It is a statistical approach to uncover the underlying structure of a large set of items. When student course evaluation data is normally distributed one can focus the statistical significance of each factor loading by using maximum likelihood as extraction method. Later on, after pretesting and with a larger data set from other terms, a confirmatory factor analysis can verify the identified structure of all student course evaluation items (chapter 5.1.). The internal consistency of user-generated items can be verified via reliability analysis, when each item is grouped according to the previously identified structure (chapter 5.2.). Correlations are to identify the connectedness of each key competence to its counterparts. A relation between professional and methodological competence seems more likely than a relation between professional and social competence, since professional and methodological competence work often consecutive in real life (chapter 5.3.).

In summary, this multistage-testing procedure identifies and counts the dimensions of key competencies within the student course evaluation, so that each item can be reconciled with the theoretical framework on dimensions of four key competencies. The underlying data sets will be highlighted within the next chapter.

4.2   Application context at the University of Cologne and datasets

The application context for the user-generated student course evaluation is the Professional Center at the University of Cologne. It offers interdisciplinary courses to students from all faculties. Up to 60 courses can be realised each term, in which up to 700 students can participate. The evaluation is voluntary for all the students. As a

result, a consistent student course evaluation, applicable to most students from different faculties and with strong focus on teaching effectiveness is required for evaluation and accreditation.

As stated before, all 16 questions are items that have been pretested on basis of five different courses with a total of 71 students in winterterm 2014/2015. Despite this first exploratory factor analysis, all students of the pretest have been asked to give a feedback in terms of comprehensibility and general design. There was no critical feedback that marked the initial concept of the questionnaire as incomprehensible. The students participating in the pretest had no affiliation to the students designing the student course evaluation.

Subsequent to this first alignment and in order to increase the number of participants, the questionnaire has been used to evaluate Professional Center courses at the University of Cologne. A total of 481 students from 40 different courses participated within summer term 2015 – called Round 1. Six months later, in winter term 2015/2016 a total of 521 students from 43 different courses participated – called Round 2. Both, Round 1 and Round 2, have been used for thorough multistage-testing, as introduced in chapter 3. Since there are no significant differences between the findings of Round 1 and Round 2, the latest data set will be referred to in the upcoming analysis.

5 Findings

This chapter presents the confirmatory factor analysis, reliability analysis and interrelatedness of all four dimensions of key competencies within the largest data set, Round 2. If the students design matches the theoretical concept of four dimensions of key competencies, all items should be differentiable accordingly and there should be sufficient reliability scores.

5.1 Factor analysis

Confirmatory factor analysis underlines the expected fourfold division of key competencies (table 1), covered up by the 16 user-generated items. There is only a slight overlap with one item of professional competence (0,396), that also seems to load on methodological competence (0,694). Thus, *pc_methods* is the weakest item of professional competence and this had to be expected, since its original question refers to the link between professional competence and necessary tools in order to apply this competence.

Table 1: Factor analysis with four key competencies (Round 2) (Source: Own calculation (2016); N = 521;
Varimax, 6 Iterations.).

	1	2	3	4
pc_knowledge	,774			
pc_study	,675			
pc_theory	,768			
pc_methods	,396	,694		
mc_use		,826		
mc_test		,836		
mc_reflect		,757		
mc_results		,766		
soc_conflict			,774	
soc_responsibility			,799	
soc_questions			,805	
soc_enforce			,791	
sec_facts				,677
sec_organize				,520
sec_structure				,782
sec_own				,822

According to similar findings in Round 1 and Round 2, it can be assumed, that the
four key competencies can be clearly identified within this type of student course
evaluation. Nevertheless, the strong link between professional competence and
methodological competence deserves a more detailed consideration, such as a check
for internal consistency.

5.2   Reliability analysis

The reliability analysis confirms the internal consistency of the four key competen-
cies (table 2), which supports the previous factor analysis findings. There is only a
small difference between non-standardized and standardized Cronbachs Alpha val-
ues and all of them meet the criteria that values should be above ,700. These values
indicate that each of the four items, that propose to measure the same general con-
struct of competence, produce similar scores within their key competence.

Table 2: Reliability analysis (Round 2) (Source: Own calculation (2016); N = 521.).

	Cronbachs Alpha (not standardized)	Cronbachs Alpha (standardized)
professional	,744	,756
methodological	,860	,871
social	,834	,836
self	,750	,767

However, professional competence generates the lowest Cronbachs Alpha values, probably because of the mentioned connection between professional competence and methodological competence. This connection can be further analysed by focusing the inter-relatedness of each key competence via correlations.

5.3 Inter-relatedness of all four key competencies

Arithmetic scales out of the four items of each key competence can be analysed in detail by the use of a correlation matrix. As expected and stated before, the strongest correlation ($r = 0,664$; $p = 0,000$) can be found between professional competence and methodological competence. It can be assumed, that they depend on each other in higher education. There are also remarkable correlations between self-competence with methodological competence ($r = 0,512$; $p = 0,000$) and social competence ($r = 0,522$; $p = 0,000$). The weakest correlations can be found between social competence with professional competence ($r = 0,316$; $p = 0,000$) and methodological competence ($r = 0,356$; $p = 0,000$). These results are also not unexpected. Thus, the correlation matrix highlights the connection between professional competence and methodological competence on the one hand, while, on the other hand, it sets the social competence apart of that connection. Self-competence is connected to all other competencies.

6 Summary and recommendations

Students at the University of Cologne designed their own questionnaire in order to address some of the statistical issues within common student course evaluations and to focus on skills and competencies in respect to the difficult measurable teaching effectiveness. They have been aware of the necessary techniques to design a proper questionnaire and the theoretical background on skills and competencies. A research team took care of their approach.

Although there is quite some theoretical reference on competence-based evaluations, this is the first time it is used within a student course evaluation entirely designed by students themselves. As a result, the students came up with their own questions for professional competence, methodological competence, social compe-

tence and self-competence. All user-generated questions have been a match according to the four dimensions of key competencies and can be identified via exploratory factor analysis within the new questionnaire. Confirmatory factor analysis, reliability analysis and correlations confirm the concept of four key competencies on basis of two different data sets (Round 1 and Round 2), reflecting different terms and seminars with a total of 1.002 students participating.

However, a strong correlative connection between methodological and professional competence can be found. Reliability analysis as well as factor analysis indicate the statistical challenges in designing an own scale or dimension that fits the theoretical concept of both competencies, since there is a theoretical and practical link between professional competence and methodological competence. This makes it statistically hard to differentiate between these two competencies. Despite one linking-item, factor analysis provided a very consistent alignment of items for professional, methodological, social and self-competence. That covers perfectly the academic discourse on that broad topic and stands in line with the field-tested design of INCHER.

As a result, a competence-based student course evaluation is a different approach than the current comparison of means and standard deviations. It works on the assumption that each course is different and sets different teaching goals. While one course may focus on professional skills like theoretical terms and concepts, another may focus on the practical application, self-experience and social interaction with others. Furthermore, students are not alike. Measuring four key competencies allows for considering their different background knowledge and learning behaviour. This can be clearly identified within both datasets, Round 1 and Round 2.

In respect to the initial research questions it can be stated that the students' approach has proven valid during a multistage-testing procedure. The students have been able to address common statistical issues and designed adequate questions during the process of operationalization. Their set of questions underlines the fourfold division of key competencies and differentiates adequately between different contents of different courses. Thus, their approach is nowadays the standard student course evaluation at the Professional Center and this is a remarkable achievement, considering all different types of courses offered by the Professional Center. Some of them are centred on professional and methodological aspects, while others focus the students and their interactions with others. Therefore, program managers can use student course evaluations that focus on key competencies, in order to systematise their courses and check for differences according to the intended goals of the courses. The students involved in designing the student course evaluation, for example, indicated clear effects on their professional as well as methodological competencies during the course.

A Scientific-Use-File (SUF) on term 2016 has been created, covering 786 students that can be analysed in detail, focusing different faculties, a broad range of different seminars, the progress of the students as well as their sociodemographic factors. The SUF will be provided by the authors on request.

References

Beleche, T., Fairris, D. & Marks, M. (2012). Do course evaluations truly reflect student learning? Evidence from an objectively graded post-test. *Economics of Education Review, 31 (5)*, pp. 709–719.

Cashin, W. E. & Clegg, V. L. (1987). *Are student ratings of different academic fields different? Paper presented at the annual meeting of the American Educational Research Association.* Washington, DC.

Clayson, D. E. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of literature. *Journal of Marketing Education, 31 (1)*, pp. 16–30.

Daniel, H. D. (1996). Evaluierung der universitären Lehre durch Studenten und Absolventen. *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie, 16 (2)*, pp. 149–164.

Davis, B. G. (2009). *Tools for Teaching. 2nd edition.* San Francisco, CA: John Wiley & Sons.

Diehl, J. M. (2001). Studentische Lehrevaluation in den Sozialwissenschaften. Fragebögen, Normen, Probleme. In: Keiner, E. (Ed.): *Evaluation (in) der Erziehungswissenschaft.* Weinheim: Beltz, pp. 63–90.

European Commission (2018). *Key competences.* Available at: http://ec.europa.eu/education/policy/school/competences_en (Access on: August 27th 2018).

Greimel, B. (2002). Lehrevaluation durch Beurteilung der Lernenden. Eine Analyse des Standes der Evaluationsforschung. *Zeitschrift für Berufs- und Wirtschaftspädagogik, 98,* pp. 197–224.

International Centre for Higher Education Research (INCHER) Kassel (2014). *Fragebogen der KOAB-Absolventenbefragung 2014.* Available at: https://www.uni-goettingen.de/de/fragebogen-pj-2012/497972.html (Access on: August 27th 2018).

Kirkpatrick, D. (1998). *Evaluating Training Programs: The Four Levels.* San Francisco, CA: Berrett-Koehler Publishers Inc.

Klippert, H. (1994). *Methoden-Training.* Weinheim, Basel: Beltz.

Marsh, H. W. & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. In: *American Psychologist, 52 (11)*, pp. 1187–1197.

Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In: Perry, R. P. & Smart, J. C. (Eds.): *The scholarship of teaching and learning in higher education: An evidence-based perspective.* Dordrecht: Springer, pp. 319–383.

Maurer, K. M. (2006). Kompetenzbeschreibung und Bildungsstandards. In: *Erziehungskunst, 11/2006,* pp. 1165–1174.

McKeachie, W. J. (1997). Student ratings. The validity of use. *American Psychologist, 52,* pp. 1218–1225.

McCollough, B. D. & Radson, D. (2011). Analysing student evaluations of teaching: Comparing means and proportions. *Evaluation & Research in Education, 24 (3)*, pp. 183–202.

120

User-generated Student Course Evaluations: (How) Can Key Competencies become Systematic
Evaluation Parameters?

Perry, R. P. & Smart, J. C. (1997). *Effective teaching in higher education: Research and practice*. New York: Agathon Press.

Pounder, J. S. (2007). Is student evaluation of teaching worthwhile?: An analytical framework for answering the question. *Quality Assurance in Education, 15 (2)*, pp. 178–191-.

Rindermann, H. (2001). *Lehrevaluation. Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Mit einem Beitrag zur Evaluation computerbasierten Unterrichts*. Landau: Empirische Pädagogik.

Rindermann, H. (2003). Lehrevaluation an Hochschulen: Schlussfolgerungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. *Zeitschrift für Evaluation, 2/2003,* pp. 233–256.

Roth, H. (1971). *Pädagogische Anthropologie. Band II. Entwicklung und Erziehung*. Hannover: Schroedel.

Seldin, P. (1999). Building successful teaching evaluation programs. In: Seldin, P. (Ed.): *Current practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions*. Bolton, MA: Anker.

Trautwein, C. (2011). *Unternehmensplanspiele im industriebetrieblichen Hochschulstudium*. Wiesbaden: Gabler.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66 (5)*, pp. 297–333.

Zentrum für Qualitätssicherung und -entwicklung der Johannes Gutenberg Universität Mainz (2014). *Handreichung zur Formulierung von Lernergebnissen (Learning Outcomes)*. Available at: https://www.blogs.uni-mainz.de/verwaltung-sl/files/2018/07/Leitfaden_Lernergebnisse.pdf (Acces on: August 27th 2018).

Authors

Since 2018 Prof. Dr. Dennis Klinkhammer is professor for health and social management with focus on empirical social research and quantitative research methods. His lectures and research projects address social inequalities within society as well as Big Data analyses in the healthcare sector. In addition, he is engaged as module leader at the FOM University of Applied Sciences for Economics and Management when it comes to research methods for social work and related areas of study.

Contact
Professur für Gesundheits- und Sozialmanagement, insbesondere empirische Sozialforschung, FOM Hochschule für Oekonomie & Management, Hochschulzentrum Köln
Agrippinawerft 4, 50678 Köln, Germany
dennis.klinkhammer@fom.de

Prof. Dr. Michael Schemmann is Professor of Adult and Continuing Education at the University of Cologne. His research interests concern structural developments in continuing education, research on organisations of adult education and internationally comparative adult education research.

Contact
Universität zu Köln, Humanwissenschaftliche Fakultät, Department Erziehungs- und Sozialwissenschaften, Professur für Erwachsenenbildung/Weiterbildung
Innere Kanalstraße 15, 50823 Köln, Germany
michael.schemmann@uni-koeln.de