

Indicators for Research Quality for Evaluation of Humanities Research: Opportunities and Limitations

Michael Ochsner^{1,*}, Sven E. Hug¹, Hans-Dieter Daniel^{1,2}

¹ Professorship for Social Psychology and Research on Higher Education, D-GESS, ETH Zurich, 8001 Zurich, Switzerland

² Evaluation Office, University of Zurich, 8001 Zurich, Switzerland

* E-mail: ochsner@gess.ethz.ch

Abstract

There is currently no agreed upon standard how to evaluate humanities research. Our article addresses one of the several reasons for this: the missing link between indicators and humanities scholars' notions of quality. We present groups of indicators that are tied to quality criteria put forward by humanities scholars of German literature, English literature, and art history. We then outline opportunities and limitations of using indicators to evaluate humanities research.

Introduction

In times of financial cutbacks and increasing autonomy of universities, the interest in quantifying scholarly achievement grows steadily. This process also reached the humanities (see, e.g., Guillory, 2005, p. 28; Royal Netherlands Academy of Arts and Sciences, 2011, p. 5; Zuccala, 2012, p. 229). However, humanities scholars have met the attempt to measure research quality with strong reservations. If indicators are to play a part in evaluations to reduce the problems associated with peer review processes (i. e., subjectivity, low inter-referee reliability, bias, and poor predictive validity; see, e.g., Bornmann & Daniel, 2008; Daniel, Mittag, & Bornmann, 2007; Langfeldt, 2010), humanities scholars should be willing to accept the indicators that are used. This implies that the indicators are connected to quality criteria that humanities scholars apply themselves in their everyday work.

A broad literature on quality criteria and indicators is already in existence. Yet, there is no systematic collection which summarizes the existing indicators, let alone one that points out their value for research assessment in the humanities. In this article we present indicators for research quality in the humanities based on a review of about one hundred documents that contain research indicators that are predominantly linked to humanities research. The collection of indicators is complemented by data of interviews with and a survey of humanities scholars during our project 'Developing and Testing Research Quality Criteria in the Humanities, with an Emphasis on Literature Studies and Art History' (Hug, Ochsner, & Daniel, accepted for publication). The indicators we present are linked to the quality criteria for humanities research developed in this project (Hug, Ochsner, &

Daniel, submitted).

Quality criteria for research in the humanities

The measurement of research quality faces strong opposition in the humanities. In order to develop quality criteria that are adequate for humanities research and consensual in the research community, we conducted extensive literature research. We looked for documents that include criteria or indicators for research in the humanities (and related disciplines) or documents that address criticisms or conceptual aspects of research assessments in the humanities. We included a broad range of documents to find as many arguments, criteria, and indicators as possible: Government or institutional reports on how humanities are evaluated, grey literature on critiques of those procedures by humanities scholars, bibliometric and scientometric literature dealing with social sciences and humanities issues, articles and books by humanities scholars about scientific quality or university policy, and articles on science and technology indicators. This resulted in a database of literature on quality criteria and indicators for humanities research that is continually updated and now contains almost one thousand entries (Peric, Ochsner, Hug, & Daniel, 2012). The database can be accessed on our website (<http://www.psh.ethz.ch/crus/bibliography>).

By analysing this database, we identified the most pertinent objections humanities scholars put forward (Hug et al., accepted for publication). Furthermore, these objections can be traced back to the fact that the assessments are not connected to the humanities scholars' notions of quality. Hence, we designed a framework for developing criteria and indicators for

assessing research quality, which assures a close link to the scholars' own notions of quality. By applying this framework, we developed quality criteria for research in the humanities that adequately assess research quality in the humanities and are consensual in the research community (Hug et al., submitted).

Our framework consists of four pillars. In this article, we focus on the one that is particularly important for the use of indicators in research assessments: the application of a sound measurement approach. Such an approach is necessary because, up to now, indicators frequently have been only loosely tied to definitions of quality (Brooks, 2005; Donovan, 2008). Such weak or missing links between indicators and quality make it difficult for the assessed scholars to understand what is being measured. Therefore, the reluctance of humanities scholars to accept a quantitative representation of research quality is not surprising. A sound measurement approach, such as is typically applied in the social sciences, can replace the missing links. According to social science research methodology, each construct (e. g., 'research quality') has to be defined prior to measuring it. In such a measurement approach, firstly, quality has to be defined by explicating quality criteria. Secondly, every quality criterion is further specified and defined explicitly by one or more aspects (i. e., analytical definition). Thirdly, each aspect is operationalized: Each aspect is tied to one or more indicators that specify how it can be observed, quantified or measured (i. e., operational definition). It is, of course, possible that there will be no suitable quantitative indicators for an aspect and, as a consequence, the aspect cannot be measured. Hence, such an approach allows to quantify the amount of quality criteria or aspects that cannot be quantified.

We identified 19 quality criteria and 70 aspects for humanities research using Repertory Grid interviews (Ochsner, Hug, & Daniel, 2012) and a Delphi survey (Hug et al., submitted). We then extracted the consensual criteria¹ that can be used to assess research quality adequately in these disciplines (Hug et al., submitted). Indicators to evaluate humanities research have to be connected to these consensual criteria (and their respective consensual aspects) if they are to be linked to humanities research quality. The next step, then, is to collect indicators for humanities research.

¹Originally, consensuality is defined for aspects. An aspect is defined as consensual if it received both a high rating (at least 50 % of the scholars rated the aspect with a five in a scale from one to six) and a low disagreement (not more than 10 % of the scholars rated the aspect negatively, i. e., with less than a four). A criterion is then classified as consensual in three disciplines, when in all three disciplines at least one aspect of the criterion has reached consensus (it does not need to be the same aspect in the different disciplines). A criterion is defined as consensual in two disciplines when in two disciplines at least one aspect of the criterion has reached consensus.

Collecting indicators for research in the humanities

There is a wealth of literature on research assessment and indicators of research quality. However, its focus lies on the natural and life sciences (see, e. g., Hemlin, 1996, p. 53). Hence, there is no canon on evaluation procedures of and indicators for research quality in the humanities. As a first step, we collected indicators from those documents of the above-mentioned database that propose indicators. Quite to the contrary of our expectations, we found an abundance of indicators, some very specific, some more general. Because of the high number of indicators, we grouped the indicators in clusters of related topics, obtaining a workable number of groups that suits a questionnaire for a rating of the indicators. After scanning the most pertinent documents for indicators, no new indicator groups were defined, resulting in a systematic analysis of the one hundred documents containing indicators for research quality. These documents are listed in the references section indicated by an asterisk. In a second step, we complemented our list of groups of indicators by indicators gathered directly from humanities scholars themselves to obtain a list of indicators tailored to humanities research. On the one hand we analysed literature on research quality by humanities scholars and transferred some suggestions into indicators. On the other hand, during our Repertory Grid interviews and the first round of our Delphi survey, we asked the participating humanities scholars to name indicators they deemed adequate.

The grouping of the indicators followed two principles: Firstly, the indicators in one group should be of a similar kind. Secondly, according to our measurement model, it should be possible to connect the group to a specific quality criterion or aspect. The grouping resulted in 62 groups of indicators for research quality in the humanities. The groups are listed in table A.1 in the appendix.

The measurement of research quality in the humanities

Each group of indicators can be assigned to one or more aspects of the quality criteria. A group is assigned to an aspect if the occurrence of the aspect can be deduced from the indicator(s) of the group. For example, the occurrence of the aspect *interdisci-*

iplinary exchange can be deduced from the existence of interdisciplinary personal contacts. However, the occurrence of the aspect *insights are recognized by the research community* cannot be deduced directly from *personal contacts*: While it is likely that *personal contacts* suggest that the insights of the scholar find recognition by his or her personal contacts, it is not a sufficient condition for recognition since a personal contact can be established without recognition of the insights of the scholar. Hence, the group *personal contacts* was not assigned to the aspect *insights are recognized by the research community*. Table A.2 in the appendix shows the groups of indicators along with the aspects they were assigned to, i. e., the aspects they potentially measure. Due to limitations of space, we cannot present all consensual aspects in detail. The interested reader is referred to the forthcoming publication by Hug et al. (submitted). Instead, for each group of indicators the measured aspects are reported in abbreviated form grouped by the criteria to which they belong.

The clear assignment of indicators to the aspects and criteria enables us to quantify the amount of quality criteria and aspects that can be measured quantitatively. We were only able to identify indicators for 23 of the 42 consensual aspects. Thus, only 55% of the relevant aspects can be quantified. The other

45% of the aspects cannot be measured with indicators and are only accessible to the judgement of peers. When each discipline is analysed separately, the portion of measurable aspects is even lower: In German literature studies, we identified indicators for 19 of the 36 consensual aspects (53%), in English literature studies, 15 of 29 aspects can be measured (52%), and in art history, for 15 of 31 aspects (48%) indicators were found. This leads us to the conclusion that only about 50% of the consensual aspects can be operationalized quantitatively, i. e., with indicators. In other words, indicators can only capture 50% of the humanities scholars' notions of quality.

The analysis works in the reverse direction as well: Since we know the quality criteria and aspects that humanities scholars put forward, we can track down what the most commonly used indicators can measure. This is achieved by linking them to the aspects of the quality criteria. The indicators that appear in most documents we scanned (e. g., Cunningham, 2008; Geschwind & Larsson, 2008; Hinze, Schelling, & Ulrich, 2010; Palomares-Montero & Garcia-Aracil, 2011; Zwaan & Nederhof, 1989) and in renowned rankings of higher education institutions (e. g., Centrum für Hochschulentwicklung, 2012; Times Higher Education, 2012) are shown in table 1 along with the criteria they potentially measure.

Table 1: Frequently used indicators and criteria they can potentially measure

Indicators	Criterion
Citations	Recognition; impact on research community; relevance
Prizes	Recognition; impact on research community; relevance
Third party funding	Recognition; impact on research community; relevance; relation to and impact on society
Collaborations	Scholarly exchange; recognition
Transfers to society and economy	Relation to and impact on society
Publications	Scholarly exchange; productivity
Board memberships	Scholarly exchange; recognition; impact on research community
Recruitment	Continuity, continuation

Needless to say, there are also more sophisticated lists of indicators and procedures to assess research quality (e. g., Montada, Krampen, & Burkard, 1999; Spaapen, Dijkstra, & Wamelink, 2007; Wissenschaftsrat, 2004) and procedures that were adapted to the humanities (e. g., Royal Netherlands Academy of Arts and Sciences, 2011; Wissenschaftliche Kommission Niedersachsen, 2003, 2004a, 2004b; Wissenschaftsrat, 2010b). They include some of the indicators listed in our collection of

indicators as well, such as *editorship; presentations; documentation activities; usage statistics; strategies; started initiatives; reviews of the researcher's work; appointments to professorship; attractiveness to, qualification of, and success of junior researchers; or a preference for monographs.*

Do those indicators measure the relevant criteria to assess research quality in the humanities? Table 2 shows the list of quality criteria for research in the humanities along with information on consensuality in

the three disciplines covered in our study and their measurability by indicators. The criteria that are consensual in all three disciplines are marked with two asterisks (**), the criteria that are consensual in two disciplines are marked with one asterisk (*) (for detailed information about consensuality of the quality criteria and their aspects, see Hug et al., submitted). The measurability of the criteria with indicators is in-

dicated by three daggers (†††) if the criterion is measurable with the frequently used indicators, with two daggers (††) if the criterion is measurable using the more sophisticated lists of indicators and procedures, and one dagger (†) if it is measurable with the indicators we present in this article (note, however, that in this article we do not present indicators for the criteria that are not consensual in any discipline).

Table 2: Quality criteria for humanities research: consensuality and measurability

1. Scholarly exchange ^{**†††}	8. Continuity, continuation ^{†††}	15. Scholarship, erudition ^{**††}
2. Innovation, originality ^{**}	9. Impact on research community ^{**†††}	16. Passion, enthusiasm ^{*,††}
3. Productivity ^{†††}	10. Relation to and impact on society ^{†††}	17. Vision of future research ^{**††}
4. Rigour ^{**}	11. Variety of research ^{*,††}	18. Connection between research and teaching, scholarship of teaching ^{**†}
5. Fostering cultural memory ^{**††}	12. Connection to other research ^{**††}	19. Relevance ^{†††}
6. Recognition ^{†††}	13. Openness to ideas and persons ^{**†}	
7. Reflection, criticism ^{*,††}	14. Selfmanagement, independence ^{*,†}	

Table 2 illustrates that the frequently used indicators (e. g., the indicators used in rankings) are not representing research quality in the humanities at all. Only two criteria that are consensual in at least two disciplines (i. e., the criteria marked with one or two asterisks) are measured by those indicators (as indicated by three daggers). The other criteria measured by the frequently used indicators are exactly those criteria that are *not* consensual in more than one discipline or even not consensual in any discipline (i. e., criteria without an asterisk). The table also shows that the initiatives targeted at the assessment of humanities research are heading in the right direction: Many of the criteria that are consensual in more than one discipline show two daggers. Yet, those initiatives also fail to take some relevant criteria into account, which are highly consensual and measurable – as shown, e. g., by the indicators we present – namely *openness to ideas and persons*; *selfmanagement, independence*; and *connection between research and teaching, scholarship of teaching*. And two of the highly consensual criteria, namely *innovation, originality* and *rigour* elude direct measurement entirely. However, it must be noted that, due to limitations of space and for simplicity, we reported consensuality for the quality criteria. If the consensuality of the as-

pects, for which consensuality originally was defined, is taken into account, an even smaller fraction of the consensual aspects can be measured quantitatively, since indicators can only operationalize about 50 % of the consensual aspects.

Conclusion

Research assessments face strong opposition from humanities scholars. In our project ‘Developing and Testing Research Quality Criteria in the Humanities, with an Emphasis on Literature Studies and Art History’ (Hug et al., accepted for publication), we present adequate quality criteria for humanities research that are linked to the notions of quality that humanities scholars apply themselves. In this article, we have presented indicators for research quality in the humanities that are tied to quality criteria put forward by humanities scholars of German literature, English literature, and art history. We have shown that the most commonly applied indicators do not measure the relevant quality criteria for humanities research and that only about 50 % of the relevant aspects of the quality criteria can be measured with indicators. Hence, an evaluation of research quality in the humanities

should not rely exclusively on indicators. While this is a limitation for the use of indicators in an evaluation of humanities research, the project has shown some opportunities as well. Our project has shown that when, in a bottom-up process, the scholars themselves are involved to define quality criteria, humanities scholars are interested in the discussion about and definition of quality criteria for research (Hug et al., submitted). Humanities researchers' refusal of evaluations could be alleviated if the assessment is based on peer review using consensual criteria and aspects for research quality and if the scholars are involved in the process early on (i. e., definition of research quality). To reduce possible shortcomings of peer review, like subjectivity, low inter-referee reliability, bias, and poor predictive validity (see, e. g., Bornmann & Daniel, 2008; Daniel et al., 2007; Langfeldt, 2010), indicators attached to the relevant criteria can be used to inform peers (informed peer review). While the link of indicators to quality criteria and aspects ensures that the right aspects are measured and the peers are informed about the most important criteria, it still might be the case that humanities scholars will not accept some indicators (or even indicators in general). Therefore, a rating of the indicators by the scholars is required. Furthermore, more research in other humanities disciplines is needed to fine-tune quality criteria and indicators for other humanities disciplines.

Furthermore, the present study pinpoints validity issues concerning the measurement of research quality in the humanities. We have shown that, concerning German literature studies, English literature studies, and art history, the commonly applied indicators for assessing research quality do not cover the relevant quality criteria for research in the respective disciplines. They might cover, however, expectations of the administration and the public towards humanities research. Yet, these indicators are not valid instruments to measure *research quality* in the humanities. This raises the question of validity issues for the measurement of research quality in the social sciences or the technical, natural, or life sciences as well. Hence, more research is needed on what exactly research quality means in different fields and disciplines if a valid measurement of research quality is to be achieved.

Acknowledgements

This work was conducted within the cooperative project of the Universities of Zurich and Basel entitled 'Developing and Testing Research Quality Criteria in the Humanities, with an emphasis on Literature Studies and Art History' and supported by the Rectors' Conference of the Swiss Universities (CRUS) within the framework of the SUK B-05 innovation and

cooperation project 'Mesurer les performances de la recherche' [Measuring research performance]. The authors would like to thank Esther Germann and Julia Wolf for proofreading.

References

References marked with an asterisk indicate studies included in the literature review.

- *Ad-hoc-Kommission der phil.-hist. Klasse zur Bewertung geisteswissenschaftlicher Forschungsleistungen der Österreichischen Akademie der Wissenschaften. (2009). *Kommissionsbericht: Überlegungen zur Qualitätssicherung geisteswissenschaftlicher Forschungsleistungen vom 11. Juni 2008*. Österreichische Akademie der Wissenschaften. Retrieved 2012-10-15, from http://www.oeaw.ac.at/imafo/fileadmin/bilder/Research_Policy_Material/BewertungGeisteswissAdhoc08.pdf
- *Arts and Humanities Research Council. (2009). *Research funding guide (version 1.7)*. Arts and Humanities Research Council.
- *Australian Research Council. (2009a). *ERA indicator benchmark methodology. Physical, Chemical and Earth Sciences (PCE) and Humanities and Creative Arts (HCA) clusters*. Australian Research Council.
- *Australian Research Council. (2009b). *ERA indicators consultation paper*. Australian Research Council.
- *Australian Research Council. (2009c). *Evaluation guidelines for the 2009 ERA trial. Physical, Chemical and Earth Sciences (PCE) and Humanities and Creative Arts (HCA) clusters*. Australian Research Council.
- *Bauerlein, M. (2008). The future of humanities labor. Let's change the standards for tenure in the humanities before it's too late. *Academe Online, September-October*. Retrieved 2012-10-15, from <http://www.aaup.org/AAUP/pubsres/academe/2008/S0/Feat/baue.htm>
- *Bollen, K., van de Sompel, & Rodriguez, M. A. (2008). Towards usage-based impact metrics: first results from the MESUR project. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries (JCDL)* (pp. 231–240). Association for Computing Machinery.
- Bornmann, L., & Daniel, H.-D. (2008). The effectiveness of the peer review process: inter-referee agreement and predictive validity of manuscript refereeing at *Angewandte Chemie*. *Angewandte Chemie International Edition*, 47(38), 7173–7178. doi: 10.1002/anie.200800513

- *Broek, S., & Nijssen, A. (2009). *Impact Assessment Geesteswetenschappen. Onderzoek uitgevoerd door Research voor Beleide in opdracht van Gebiedsbestuur Geesteswetenschappen*. Zoetermeer: Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).
- Brooks, R. L. (2005). Measuring university quality. *The Review of Higher Education*, 29(1), 1-21. doi: 10.1353/rhe.2005.0061
- *Butler, L., & Visser, M. S. (2006). Extending citation analysis to non-source items. *Scientometrics*, 66(2), 327-343.
- *Campaign for Science & Engineering in the UK. (2009). *Impacts of investment in the science & engineering research base (CaSE Policy Report No. 10)*. Campaign for Science & Engineering in the UK (CaSE).
- *Centra, J. A. (1977). *How universities evaluate faculty performance: a survey of department heads*. Graduate Record Examinations Program, Educational Testing Service.
- Centrum für Hochschulentwicklung. (2012). *ForschungsRanking Indikatoren*. Retrieved 2012-10-15, from http://www.che-ranking.de/methodenwiki/index.php/ForschungsRanking_Indikatoren
- *CHERPA-Network. (2010a). *U-Multirank: interim progress report: design phase of the project 'Design and Testing the feasibility of a Multi-dimensional Global University Ranking'*. CHERPA-Network.
- *CHERPA-Network. (2010b). *U-Multirank: project 'Design and Testing the Feasibility of a Multi-dimensional Global University Ranking'. Interim progress report: preparation of the pilot phase*. CHERPA-Network.
- *Cunningham, S. (2008). University and discipline cluster ranking systems and the humanities, arts, and social sciences. *Higher Education in Europe*, 33(2/3), 245-258.
- Daniel, H.-D., Mittag, S., & Bornmann, L. (2007). The potential and problems of peer evaluation in higher education and research. In A. Cavalli (Ed.), *Quality assessment for higher education in Europe* (pp. 71-82). London: Portland Press.
- *de Jong, S., van Arensbergen, P., Daemen, F., van der Meulen, B., & van den Besselaar, P. (2011). Evaluation of research in context: an approach and two cases. *Research Evaluation*, 20(1), 61-72.
- *de Jong, S., Wamelink, F., & van der Meulen, B. (2009). *Evaluating research in context: pilot study at the faculty of architecture – TU Delft*. Evaluating Research in Context Rathenau Instituut.
- *Dinges, M., & Hofer, R. (2008). *Der Erfolg von Forschungsprojekten*. Joanneum Research Forschungsgesellschaft mbH.
- *Dolan, C. (2006). *Impact and quality assessment practices in the humanities*. Humanities in the European Research Area.
- *Dolan, C. (2008). *Feasibility study: the evaluation and benchmarking of humanities research in Europe*. Humanities in the European Research Area (HERA).
- *Donovan, C. (2005). *A review of current Australian and international practice in measuring the quality and impact of publicly funded research in the humanities, arts, and social sciences*. Research Evaluation and Policy Project, Australian National University.
- Donovan, C. (2008). Das zweiköpfige Lama zähmen: Die australische Suche nach den besten Evaluierungsmethoden für die Geisteswissenschaften. In E. Lack & C. Marksches (Eds.), *What the hell is quality? Qualitätsstandards in den Geisteswissenschaften* (p. 74-98). Frankfurt a. M.: Campus.
- *Donovan, C., & Butler, L. (2007). Testing novel quantitative indicators of research 'quality', esteem and 'user engagement': an economics pilot study. *Research Evaluation*, 16(4), 231-232.
- *Drooge van, L., Besselaar van den, P., Elsen, G., Haas de, M., Heuvel van den, J., Maassen van den Brink, H., ... Westenbrink, R. (2010). *Evaluating the societal relevance of academic research: a guide*. Den Haag: Netherlands Association of Universities of Applied Sciences (HBO-raad), the Royal Netherlands Academy of Arts and Sciences (KNAW), the Netherlands Organisation for Scientific Research (NWO), the Association of Universities in the Netherlands (VSNU) and the Rathenau Institute.
- *East, J. W. (2006). Ranking journals in the humanities: an Australian case study. *Australian Academic and Research Libraries*, 37(1), 3-16.
- *Finkenstaedt, T. (1990). Measuring research performance in the humanities. *Scientometrics*, 19(5/6), 409-417.
- *Finkenstaedt, T., & Fries, M. (1988). Forschungsmessung in der Anglistik. Eine Längsschnittuntersuchung. In H. D. Daniel & R. Fisch (Eds.), *Evaluation von Forschung: Methoden – Ergebnisse – Stellungnahmen* (pp. 151-176). Konstanz: Universitätsverlag Konstanz.
- *Fisher, D., Rubenson, K., Rockwell, K., Grosjean, G., & Atkinson-Grosjean, J. (2000). *Performance indicators and the humanities and social sciences*. Centre for Policy Studies in Higher Education and Training.
- *Frost, C. O. (1979). The use of citations in literary research: a preliminary classification of citation

- functions. *Library Quarterly*, 49(4), 399–414.
- *Furlong, J., & Oancea, A. (2005). *Assessing quality in applied and practice-based educational research: a framework for discussion*. Oxford University Department of Educational Studies. Retrieved 2012-10-16, from http://mt.educarchile.cl/MT/jjbrunner/archives/assessing_quality_shortreport_tcm6-8232.pdf
- *Geschwind, L., & Larsson, K. (2008). *Getting Pole Position. Pre reform research strategies in the humanities at Swedish universities* (Working Paper Series in Economics and Institutions of Innovation No. 140). Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies. Retrieved 2012-10-15, from <http://ideas.repec.org/p/hhs/cesisp/0140.html>
- *Giménez-Toledo, E., & Román-Román, A. (2009). Assessment of humanities and social sciences monographs through their publishers: a review and a study towards a model of evaluation. *Research Evaluation*, 18(3), 201–213.
- *Ginevicius, R., & Gineviene, V. B. (2009). The compliance of master's degree studies with the economic needs of the country. *Technological and Economic Development of Economy*, 15(1), 136–153.
- *Guetzkow, J., Lamont, M., & Mallard, G. (2004). What is originality in the humanities and the social sciences? *American Sociological Review*, 69(2), 190–212.
- Guillory, J. (2005). Valuing the humanities, evaluating scholarship. *Profession (MLA: Modern Language Association)*, 28-38. doi: 10.1632/074069505X79071
- *Hart, D., Lang, D., Watson, N., & Oldfield, M. (1999). *Performance indicators in the humanities and social sciences: report to the humanities and social sciences federation of Canada*. Ontario Institute for Studies in Education (University of Toronto).
- *Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, 5(3), 446–457.
- *Hemlin, S. (1996). Social studies of the humanities: a case study of research conditions and performance in ancient history and classical archaeology and English. *Research Evaluation*, 6(1), 53–61.
- *Hemlin, S., & Gustafsson, M. (1996). Research production in the arts and humanities. A questionnaire study of factors influencing research performance. *Scientometrics*, 37(3), 417–432.
- *Higher Education Funding Council for England. (2009). *Research excellence framework – second consultation on the assessment and funding of research*. Higher Education Funding Council for England.
- *Hinze, S., Schelling, A., & Ulrich, S. (2010). *Social Sciences and Humanities in Germany. 2010 Country Report* (Report to the European Commission by the Metris Network). European Commission.
- *Hornbostel, S. (2008). Gesucht: Aussagekräftige Indikatoren und belastbare Datenkollektionen. Desiderate geisteswissenschaftlicher Evaluierung in Deutschland. In E. Lack & C. Marksches (Eds.), *What the hell is quality? Qualitätsstandards in den Geisteswissenschaften* (pp. 55–73). Frankfurt/New York: Campus Verlag.
- *Hornbostel, S., Klingsporn, B., & von Ins, M. (2008). Messung von Forschungsleistungen – eine Vermessenheit? In Alexander von Humboldt-Stiftung (Ed.), *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen. Beiträge zur Beurteilung von Forschungsleistungen* (pp. 11–32). Bonn: Alexander von Humboldt-Stiftung.
- Hug, S. E., Ochsner, M., & Daniel, H.-D. (accepted for publication). A framework to explore and develop criteria for assessing research quality in the humanities. *International Journal for Education Law and Policy*.
- Hug, S. E., Ochsner, M., & Daniel, H.-D. (submitted). Criteria for assessing research quality in the humanities – a Delphi study among scholars (of English literature, German literature and art history).
- *Hutchison, H. (2009). *Humanities and creative arts: recognising esteem factors and non-traditional publication in Excellence in Research for Australia (ERA) Initiative*. Council for the Humanities, Arts and Social Sciences. Retrieved 2012-10-16, from <http://www.chass.org.au/papers/PAP20090820HH.php>
- *International Peer Review Committee on the Research in History, Arts and Culture. (2006). *Research in history, arts and culture: evaluation report of the committee on the research in history, arts and culture 1999-2004 of the faculty of history and arts*. Erasmus University.
- *Kern, H. (2008). Kreativität in den Geisteswissenschaften. *Arbeits- und Industriesoziologische Studien*, 1, 45–57.
- *Kern, H. (2009). Kreativität in den Wissenschaften. Wovon hängt sie ab? In N. Elsner & N. A. Rupke (Eds.), *Wissenswelten – Bildungswelten* (pp. 35–54). Göttingen: Wallstein Verlag.
- *Kladakis, G., & Olesen, L. (2004). *European research funding systems in the humanities. Report 2: best practice*. Copenhagen: European Network of Research Councils in the Human-

- ties.
- *Krull, W. (2010). *Vom Nutzen und Nachteil der Qualitätsbewertung für die Geisteswissenschaften* (Vortrag an der Universität Zürich im Rahmen der Kolloquiumsreihe 'Qualitätskriterien für die Forschung in den Geisteswissenschaften'). Retrieved 2012-10-16, from <http://www.psh.ethz.ch/crus/kolloquium/Krull2010.pdf>
- *Kuitunen, S., & Haila, K. (2007). *Impact cycle model as an analytical for assessing the societal relevance and impacts of science* (Presentation at the International Conference on 'Science Impact', 10-11 May 2007, Vienna).
- *Kyvik, S. (1989). Productivity differences fields of learning, and Lotka's law. *Scientometrics*, 15(3/4), 205–214.
- *Lahtinen, E., Koskinen-Ollonqvist, P., Rouvinen-Wilenius, P., Tuominen, P., & Mittelmark, M. B. (2005). The development of quality criteria for research: a Finnish approach. *Health Promotion International*, 20(3), 306–315.
- *Lamont, M. (2009). *How professors think: inside the curious world of academic judgment*. Cambridge, MA: Harvard University Press.
- Langfeldt, L. (2010). Expert panels evaluating research: decision-making and sources of bias. *Research Evaluation*, 13(1), 51–62.
- *López, G. M. A. (2008). Principios teóricos para la evaluación del documento filosófico [Theoretical principles to evaluate a philosophical document]. *Investigacion Bibliotecologica*, 22(45), 41–62.
- *Luwel, M., Moed, H., Nederhof, A. J., De Samblanx, V., Verbrugghen, K., & van der Wurff, L. J. (1999). *Towards indicators of research performance in the social sciences and humanities. An exploratory study in the fields of law and linguistics at Flemish universities*. Leiden: Centre for Science and Technology Studies (CWTS).
- *Meagher, L., Lyall, C., & Nutley, S. (2008). Flows of knowledge, expertise and influence: a method for assessing policy and practice impacts from social science research. *Research Evaluation*, 17(3), 163–173.
- *Moed, H., Luwel, M., & Nederhof, A. J. (2002). Towards research performance in the humanities. *Library Trends*, 50(3), 498–520.
- *Montada, L., Krampen, G., & Burkard, P. (1999). Personal and social orientations of psychology college teachers on evaluative criteria for own job performances: results of an expert survey in German psychology college teachers. *Psychologische Rundschau*, 50(2), 69–89.
- *Moses, I. (1990). Teaching, research and scholarship in different disciplines. *Higher Education*, 19(3), 351–375.
- *Nederhof, A. J., Meijer, R. F., Moed, H., & van Raan, A. F. J. (1993). Research performance indicators for university departments – a study of an agricultural university. *Scientometrics*, 27(2), 157–178.
- *Nederhof, A. J., & Noyons, E. C. M. (1992). International comparison of departments' research performance in the humanities. *Journal of the American Society for Information Science*, 43(3), 249–256.
- *Nederhof, A. J., & Zwaan, R. A. (1991). Quality judgements of journals as indicators of research performance in the humanities and the social and behavioral sciences. *Journal of the American Society for Information Science*, 42(5), 332–340.
- *Nederhof, A. J., Zwaan, R. A., De Bruin, R. E., & Dekker, P. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social sciences – a comparative study. *Scientometrics*, 15(5/6), 423–435.
- *Oancea, A., & Furlong, J. (2007). Expressions of excellence and the assessment of applied and practice-based research. *Research Papers in Education*, 22(2), 119–137.
- Ochsner, M., Hug, S. E., & Daniel, H.-D. (2012). Four types of research in the humanities: setting the stage for research quality criteria in the humanities. *Research Evaluation*. doi: 10.1093/reseval/rvs039
- Palomares-Montero, D., & Garcia-Aracil, A. (2011). What are the key indicators for evaluating the activities of universities? *Research Evaluation*, 20(5), 353–363. doi: 10.3152/095820211X13176484436096
- Peric, B., Ochsner, M., Hug, S. E., & Daniel, H.-D. (2012). *AHRABi. Arts and Humanities Research Assessment Bibliography*. ETH Zurich. Retrieved 2012-10-16, from <http://www.psh.ethz.ch/crus/bibliography>
- *Peyraube, A. (2008). *Evaluation of the humanities at the ERC* (Presentation at the Conference 'Relevance and Impact of the Humanities', 15-16 December 2008, Vienna). European Research Council.
- *Poole, M. E. (1993). Reviewing for research excellence – expectations, procedures and outcomes. *Australian Journal of Education*, 37(3), 219–230.
- *Powles, J. (2005). *Measures of quality and impact of publicly funded research in the humanities, arts and social sciences* (CHASS Occasional Papers No. 2). Canberra: Council for the Humanities, Arts and Social Sciences.
- *Research Assessment Exercise. (2006a). *RAE panel criteria and working methods – Panel M: UOA 53 – German, Dutch and Scandinavian languages* (RAE Panel criteria and work-

- ing methods). Higher Education Funding Council for Wales, Higher Education Funding Council for England, Scottish Funding Council, Department for Employment and Learning.
- *Research Assessment Exercise. (2006b). *RAE panel criteria and working methods – Panel M: UOA 57 – English Language and Literature* (RAE panel criteria and working methods). Higher Education Funding Council for Wales, Higher Education Funding Council for England, Scottish Funding Council, Department for Employment and Learning.
- *Research Evaluation and Policy Project. (2005). *Quantitative indicators for research assessment – a literature review* (REPP Discussion Paper No. 5). Research Evaluation and Policy Project.
- *Román, A., Alcain, M. D., & Giménez, E. (2007). Evaluation of Scientific Publications in the Humanities. In D. Torres-Salinas & H. Moed (Eds.), *Proceedings of ISSI 2007. 11th International Conference of the International Society for Scientometrics and Informetrics, 25-27 June 2007, Madrid* (Vol. 2, pp. 672–676). Madrid: CSIC.
- *Rons, N., De Bruyn, A., & Cornelis, J. (2008). Research evaluation per discipline: a peer-review method and its outcomes. *Research Evaluation*, 17(1), 45–57.
- *Royal Irish Academy. (2009). *Developing key performance indicators for the humanities*. Dublin: Royal Irish Academy.
- *Royal Netherlands Academy of Arts and Sciences. (2005). *Judging research on its merits. An advisory report by the Council for the Humanities and the Social Sciences Council*. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- Royal Netherlands Academy of Arts and Sciences. (2011). *Quality indicators for research in the humanities*. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- *Royal Netherlands Academy of Arts and Sciences, Association of Universities in the Netherlands & Netherlands Organisation for Scientific Research. (2009). *Standard Evaluation Protocol 2009-2015: protocol for research assessment in the Netherlands*. Royal Netherlands Academy of Arts and Sciences (KNAW), Association of Universities in the Netherlands (VSNU), Netherlands Organisation for Scientific Research (NWO). Retrieved 2012-10-16, from http://www.know.nl/content/Internet_KNAW/publicaties/pdf/20091052.pdf
- *Scheidegger, F. (2007). *Darstellung, Vergleich und Bewertung von Forschungsleistungen in den Geistes- und Sozialwissenschaften. Bestandesaufnahme der Literatur und von Beispielen aus dem In- und Ausland*. Bern: Zentrum für Wissenschafts- und Technologiestudien (CEST).
- *Schultheis, K. (2005). Leistungsparameter zur Steuerung von Hochschulen und Forschungseinrichtungen. In A. Hopbach & V. Chalvet (Eds.), *Qualität messen - Qualität managen. Leistungsparameter in der Hochschulentwicklung* (pp. 84–89). Bonn: Hochschulrektorenkonferenz.
- *Schütte, G., & Schuh, C. (2009). *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen. Beiträge zur Beurteilung von Forschungsleistungen* (second extended ed.). Bonn: Stiftung Alexander von Humboldt.
- Spaapen, J., Dijstelbloem, H., & Wamelink, F. (2007). *Evaluating Research in Context. A method for comprehensive assessment* (Second ed.). The Hague: Consultative Committee of Sector Councils for Research and Development.
- *Spanish Foundation for Science and Technology & National Evaluation and Foresight Agency. (2007). *Quality criteria for research in the humanities*. Spanish Foundation for Science and Technology (FECYT) and National Evaluation and Foresight Agency (ANEP). Retrieved 2012-10-16, from www.micinn.es/stfls/MICINN/Investigacion/.../2007-criterios-hh3.pdf
- *Strand, D. (1998). *Research in the creative arts* (Evaluations and Investigations Program report No. 98/6). Canberra: Department of Employment, Education, Training and Youth Affairs.
- *Suárez-Balseiro, C. A., Iribarren-Maestro, I., & Sanz-Casado, E. (2003). A study of the use of the Carlos III University of Madrid library's online database service in scientific endeavor. *Information Technology and Libraries*, 22(4), 179–183.
- *Teichert, D. (2006). Zwischen Vorurteilen und Missverständnissen. Zur Situation der Geisteswissenschaften. In K. Kempter & P. Meusbürger (Eds.), *Bildung und Wissensgesellschaft* (pp. 127–150). Berlin/Heidelberg: Springer.
- *The Modern Language Association of America. (2007). Report of the MLA task force on evaluating scholarship for tenure and promotion. *Profession*, 9, 9–71.
- Times Higher Education. (2012). *Times Higher Education World University Rankings. Methodology*. Retrieved 2012-10-15, from <http://www.timeshighereducation.co.uk/world-university-rankings/2012-13/world-ranking/methodology>
- *Tippelt, R. (2004). Standards und Kriterien von Forschungsevaluationen in der Erziehungswissenschaft. In H. Merckens (Ed.), *Evaluation in der Erziehungswissenschaft* (pp. 89–100).

- Wiesbaden: VS.
- *Tsay, A., Lamont, M., Abbott, A., & Guetzkow, J. (2003). From character to intellect: changing conceptions of merit in the social sciences and humanities, 1951-1971. *Poetics*, 31(1), 23–49.
- *Wanner, R. A., Lewis, L. S., & Gregorio, D. I. (1981). Research productivity in academia: a comparative study of the sciences, social sciences and humanities. *Sociology of Education*, 54(4), 238–253.
- *Webler, W. D. (2009). Resümee: Zum Stand der Bewertung geisteswissenschaftlicher Forschungsleistungen – zwei Konflikte im Vergleich. *Qualität in der Wissenschaft*, 3(3/4), 95–104.
- *Webler, W. D., & Pätzold, M. (2009). Gespräch mit Dr. Mathias Pätzold, Generalsekretär der Wissenschaftlichen Kommission Niedersachsen (WKN) über niedersächsische Erfahrungen mit der Bewertung geistes- und sozialwissenschaftlicher Forschung. *Qualität in der Wissenschaft*, 3(3/4), 78–82.
- *Weingart, P., Prinz, W., Kastner, M., Maasen, S., & Walter, W. (1991). *Die sogenannten Geisteswissenschaften: Aussenansichten: Die Entwicklung der Geisteswissenschaften in der BRD, 1954-1987*. Frankfurt a. M.: Suhrkamp.
- *Williams, R., & van Dyke, N. (2006). *Rating major disciplines in Australian universities: perceptions and reality*. Melbourne: Melbourne Institute of Applied Economic and Social Research. Retrieved 2012-10-16, from http://melbourneinstitute.com/downloads/reports/DR-Paper_Rating.pdf
- *Wissenschaftliche Kommission Niedersachsen. (2003). *Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen: Kunstgeschichte / Kunstwissenschaft. Ergebnisse und Empfehlungen*. Hannover: Wissenschaftliche Kommission Niedersachsen.
- *Wissenschaftliche Kommission Niedersachsen. (2004a). *Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen: Anglistik und Amerikanistik. Ergebnisse und Empfehlungen*. Hannover: Wissenschaftliche Kommission Niedersachsen.
- *Wissenschaftliche Kommission Niedersachsen. (2004b). *Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen: Germanistik mit Niederdeutsch, Niederlandistik und Skandinavistik, Kulturanthropologie, europäische Ethnologie, Medienwissenschaften. Ergebnisse und Empfehlungen*. Hannover: Wissenschaftliche Kommission Niedersachsen.
- *Wissenschaftliche Kommission Niedersachsen. (2009). *Von Agrar- bis Wirtschaftswissenschaften: Alle über einen Leisten? Zu Forschungsevaluationen in Niedersachsen und ihren fachspezifischen Kriterien*. Hannover: Wissenschaftliche Kommission Niedersachsen. Retrieved 2012-10-16, from <http://www.wk.niedersachsen.de/servlets/download?C=60199907&L=20>
- *Wissenschaftsrat. (2004). *Empfehlungen zu Rankings im Wissenschaftssystem. Teil 1: Forschung*. Hamburg: Wissenschaftsrat.
- *Wissenschaftsrat. (2008a). *Bericht der Steuerungsgruppe zur Pilotstudie Forschungsrating Chemie und Soziologie*. Köln: Wissenschaftsrat. Retrieved 2012-10-16, from <http://www.wissenschaftsrat.de/download/archiv/8453-08.pdf>
- *Wissenschaftsrat. (2008b). *Pilotstudie Forschungsrating Soziologie Abschlussbericht der Bewertungsgruppe*. Köln: Wissenschaftsrat. Retrieved 2012-10-16, from <http://www.wissenschaftsrat.de/download/Forschungsrating/Dokumente/Grundlegende%20Dokumente%20zum%20Forschungsrating/8422-08.pdf>
- *Wissenschaftsrat. (2010a). *Aufgaben, Kriterien und Verfahren des Evaluationsausschusses des Wissenschaftsrates*. Lübeck: Wissenschaftsrat. Retrieved 2012-10-16, from <http://www.wissenschaftsrat.de/download/archiv/10296-10.pdf>
- *Wissenschaftsrat. (2010b). *Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften*. Köln: Wissenschaftsrat.
- *Wooding, S., & Grant, J. (2003). *Assessing research: the researchers view* (RAND Europe Report for HEFCE). Bristol: Higher Education Funding Council for England (HEFCE). Retrieved 2012-10-16, from <http://www.ra-review.ac.uk/reports/assess/AssessResearchReport.pdf>
- Zuccala, A. (2012). Quality and influence in literary work: evaluating the 'educated imagination'. *Research Evaluation*, 21(3), 229–241. doi: 10.1093/reseval/rvs017
- Zwaan, R. A., & Nederhof, A. J. (1989). *Performance indicators in the humanities and the social sciences* (Research Report for the Netherlands Advisory Council for Science Policy (RAWB)). Leiden: Center for Studies of Science and Technology (CWTS).

Appendix

Table A.1: Groups of indicators for research quality in the humanities

Name of group	Specification ^a
Publications	Number and weighting of publications (e.g., monographs, article in edited volume, exhibition catalogue, historical critical edition, art work, documentary film)
References	Number of references (i.e., number of sources I quote in my publications)
Presentations	Number and weighting of presentations
Editorship	Number, weighting and duration of editorships (e.g., edited volume, historical critical edition, art work, documentary film)
Organized events	Number and weighting of organized events (e.g., colloquium, series of lectures, exhibition, conference)
Collaborations	Number, weighting and duration of collaborations (e.g., joint research projects with other institutions, co-authorship with peer, membership in a research network)
Personal contacts	Number and weighting of personal contacts with researchers (e.g., visits, email correspondence, meetings)
Review activities	Number, weighting and duration of review activities (e.g., grants, journal manuscripts, contributions to congresses, project proposals)
Academic associations	Number, weighting and duration of activities in learned societies, professional societies or associations
Panels	Number, weighting and duration of service on professional committees, expert panels, academic boards or advisory boards
Documentation activities	Number, weighting and duration of documentation or preservation activities
Output of documentation activities	Number and weighting of outputs reflecting documentation and preservation activities
Activities for the public	Number and weighting of activities for the public (e.g., guided tours, public lectures, readings, media appearances, performances)
Outputs for the public	Number and weighting of outputs for the public (e.g., popular books or articles, exhibitions, documentary films)
Survey: renewal of interpretations	Survey of students, alumni and the public about whether I renewed their understanding and interpretation of aspects of the past
Reviews of my work	Number and weighting of review articles of my work
Citations	Number and weighting of citations (i.e., frequency with which other researchers quote my work)
Acknowledgments	Number of times my name is mentioned in acknowledgments
Peer-reviewed channels	Number and weighting of publications in peer-reviewed channels (e.g., essay in an edited volume reviewed by the editor, article in peer-reviewed journal)
Usage statistics	Usage statistics (e.g., number of libraries holding my books, number of editions, number of translations, number of downloads of my publications, number of visits of my research website or to a databank I administer, sales figures of my publications)
Third party funding	Number, weighting and duration of third party funded projects
Prizes	Number and weighting of honours, awards and prizes
Appointments to professorship	Number and weighting of appointments to a professorship, visiting or guest or replacement or interim professorships, fellowships, research stays abroad, research grants

(continued)

Table A.1: Groups of indicators for research quality in the humanities (continued)

Name of group	Specification ^a
Attractivity to junior researchers	Attractivity to junior researchers (e.g., number of Ph.D. students I have, postdoctoral researchers and researchers from abroad I have; number of participants in my courses)
Monographs	Number of my monographs
Monographs relative to articles	Number of my monographs related to the number of my articles
Qualification of junior researchers	Qualification of students and junior researchers (e.g., number of bachelor/master/doctoral degrees; success rate (appointments to a professorship) of former students; drop-out rate of students and junior researchers; survey of alumni about the skills/competencies/qualifications they acquired)
Success of junior researchers	Success of junior researchers (e.g., number and weighting of publications; honours, awards and prizes of my students and junior researchers; number of citations of junior researchers; stay in research of junior researchers; teaching quality among junior researchers I have trained)
Teaching	What I offer in teaching (e.g., teaching hours, the time that I spend in helping and guiding junior researchers; my participation in a graduate program, graduate school or comparable program; the number and quality of further training courses I offer)
External education	External education of junior researchers (e.g., research stays of junior researchers at other institutions; number of external further training these junior researchers have attended; the financial resources I make available to them for attending congresses or receiving additional training)
Collaboration with junior researchers	Collaboration with junior researchers (e.g., co-authorship, co-editorship, joint projects)
Survey: satisfaction	Survey of junior researchers' satisfaction
Started initiatives	Number and weighting of what I have initiated or founded (e.g., periodical, book series, research institution or research cluster, post graduate program, degree program)
Research topics	Number of research topics, approaches, theories, methods, materials, disciplinary areas and languages that I use (e.g., evident in the bibliography of my publications and presentations, information on my research website)
Infrastructure	Number and weighting of infrastructure I have established or I administer (e.g., archive, art collection, specialized library, museum, database)
Current references	Number of current references (e.g., number of sources not older than for instance 5 years that I quote in my publications)
Discussions/debates	Number and weighting of participation, organization or moderation of disputes, debates or discussions about research
Written responses	Number and weighting of written responses (e.g., essay, editorial or newspaper column, open letter)
Opportunities for junior researchers	Career opportunities for junior researchers (e.g., number of positions for junior researchers, number of publications by junior researchers who have been my students, number of co-authorships with junior researchers)
Assessed openness	Assessment of my openness by students and junior researchers
Heterogeneity of junior researchers	Heterogeneity of the junior researchers (e.g., number of exchange students; number of students of other research institutions; number of students from educationally disadvantaged backgrounds, number of minority group students; junior researchers from countries with underdeveloped research structures)
Assistance	Number of collaborations or publications with researchers from institutions with weak reputations or with researchers from countries with underdeveloped research structures

(continued)

Table A.1: Groups of indicators for research quality in the humanities (continued)

Name of group	Specification ^a
Course accessibility	Accessibility to my courses or course of study I offer (e.g., waiving the cost of participation for students, junior researchers and researchers from countries with underdeveloped research structures; public announcement of my courses; access by all to events and courses I offer)
Availability of publications	Availability of my publications (e.g., number of open-access publications in relation to the total number of my publications)
Achievement of own goals	The degree to which I have met the goals I set
Financial independence	The degree of my financial independence (e.g., how much of the research funding is not tied to set targets; the ratio of basic funding to third-party funding)
Absence of requirements	The absence of set targets or output requirements
Sources	Number of sources, materials and original works used in publications or presentations
Research time	Time spent on research (e.g., time spent on research in relation to time spent on teaching and administration; the number of paid hours per year that I can devote to research)
Personal library	Size of my personal library (e.g., size of my personal library, art collection or collection of source material; number of journals subscribed to)
Teaching awards	Number and weighting of awards and prizes for teaching
Survey: enthusiasm – teaching	Survey of students, junior researchers and alumni about whether I arouse passion or enthusiasm for the research or the research topic
Survey: enthusiasm – public	Survey of the public about whether I arouse passion or enthusiasm for the research or the research topic
Strategies	The existence of goals and strategies how to reach these goals
Utilizing sources	Examining and utilizing sources (e.g., number and weighting of historical critical editions; number of appraisals of museum collections; number and weighting of collections or texts in digital database and documenting these collections or texts)
Congruence research - teaching	Degree of congruence between my research and my teaching (e.g., drawing comparisons between my course titles and my publication titles, my course descriptions and my research summaries, the table of contents of my lecture notes and my publications, bibliography of my teaching material and my publications)
Research orientation of teaching	Student satisfaction with the research orientation of the courses
Students' publications	Number and weighting of scholarly publications or presentations by my students
Acknowledging junior researchers	Number of times I refer to my students or junior researchers I have trained in my acknowledgments
Collaboration with students	Collaboration with students (e.g., co-authorship, co-editorship, joint projects)
Publication of course material	Number of publications resulting from revising material I have used in teaching
Invited lectures	Number and weighting of invited lectures

Note: Sometimes the group of indicators can measure more than one aspect (see table A.2). In such a case, the indicators must be slightly adapted to the aspect. For example, the indicator group 'publications' can measure *disciplinary exchange*, *interdisciplinary exchange*, or *international exchange*. In the first case, only the disciplinary publications will be examined, in the second case only interdisciplinary publications will be included, and in the third case only international publications will be counted.

^a The specification of the groups of indicators is displayed in the original wording of the questionnaire for the rating of the indicators. We used the first person singular to establish a close association of the aspects and indicators with the respondent's own research. The scholars had to rate the indicators as to their potential to give good indications of the occurrence of the aspect in their own research. This ensures the link between the ratings and the notions of quality the scholars use in their everyday work.

Table A.2: Groups of indicators and the criteria and aspects they can potentially measure

Name of group	Criterion	Aspect(s)
Publications	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
References	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
	Scholarship, erudition	Knowledge based on own research
Presentations	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
	Connection to other research	Building on current state of research
Editorship	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
	Recognition	Insights are recognized by the research community
	Impact on research community	Stimulating new research Influencing the research community
	Variety of research ^a Relevance ^a	Contributing towards variety and diversity ^a Research is relevant for the research community ^a
Organized events	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
Collaborations	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
	Recognition ^a	Insights are recognized by the research community ^a
Personal contacts	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
Review activities	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
	Recognition	Insights are recognized by the research community
	Impact on research community	Influencing the research community
Academic associations	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange
	Recognition	Insights are recognized by the research community
	Impact on research community	Influencing the research community
Panels	Scholarly exchange	Disciplinary exchange Interdisciplinary exchange International exchange)
	Recognition	Insights are recognized by the research community
	Impact on research community	Influencing the research community
Documentation activities	Fostering cultural memories	Documentation of aspects of the past
	Scholarship, erudition	Rich experience with sources

(continued)

Table A.2: Groups of indicators and the criteria and aspects they can potentially measure
(continued)

Name of group	Criterion	Aspect(s)
Output of documentation activities	Fostering cultural memories Scholarship, erudition	Documentation of aspects of the past Rich experience with sources
Activities for the public	Fostering cultural memories	Documentation of aspects of the past
Outputs for the public	Fostering cultural memories	Documentation of aspects of the past
Survey: renewal of interpretations	Fostering cultural memories	Renewal of interpretations of aspects of the past
Reviews of the researcher's work	Recognition	Insights are recognized by the research community
	Impact on research community Relevance	Influencing the research community Research is relevant for the research community
Citations	Recognition	Insights are recognized by the research community
	Impact on research community	Stimulating new research Influencing the research community
	Relevance	Research is relevant for the research community
Acknowledgments	Recognition	Insights are recognized by the research community
	Impact on research community	Stimulating new research Influencing the research community
	Passion, enthusiasm Relevance	Arouse passion for research Research is relevant for the research community
Peer-reviewed channels	Recognition	Insights are recognized by the research community
	Connection to other research Relevance	Building on current state of research Research is relevant for the research community
Usage statistics	Recognition	Insights are recognized by the research community
	Impact on research community Relevance	Influencing the research community Research is relevant for the research community
Third party funding	Recognition	Insights are recognized by the research community
	Relevance	Research is relevant for the research community
Prizes	Recognition	Insights are recognized by the research community
	Impact on research community Relevance	Influencing the research community Research is relevant for the research community
Appointments to professorship	Recognition	Insights are recognized by the research community
	Impact on research community Relevance	Influencing the research community Research is relevant for the research community

(continued)

Table A.2: Groups of indicators and the criteria and aspects they can potentially measure (continued)

Name of group	Criterion	Aspect(s)
Attractivity to junior researchers	Recognition	Insights are recognized by the research community
	Impact on research community Passion, enthusiasm ^a	Influencing the research community Arouse passion for research ^a
Monographs	Reflection, criticism Scholarship, erudition	Visualizing complexity Knowledge based on own research
Monographs relative to articles	Reflection, criticism	Visualizing complexity
Qualification of junior researchers	Continuity, continuation Passion, enthusiasm ^a	Promotion of young academics Arouse passion for research ^a
Success of junior researchers	Continuity, continuation Impact on research community	Promotion of young academics Stimulating new research Influencing the research community
	Passion, enthusiasm	Arouse passion for research
Teaching	Continuity, continuation	Promotion of young academics
External education	Continuity, continuation	Promotion of young academics
Collaboration with junior researchers	Continuity, continuation	Promotion of young academics
Survey: satisfaction	Continuity, continuation	Promotion of young academics
Started initiatives	Impact on research community	Stimulating new research Influencing the research community
	Variety of research	Contributing towards variety and diversity
	Vision of future research	Pointing out important research for the future
Research topics	Variety of research	Contributing towards variety and diversity
Infrastructure	Variety of research	Contributing towards variety and diversity
Current references	Connection to other research	Building on current state of research
Discussions/debates	Connection to other research	Engaging in ongoing research debates
Written responses	Connection to other research	Engaging in ongoing research debates
Opportunities for junior researchers	Openness to ideas and persons	Openness to other persons
Assessed openness	Openness to ideas and persons	Openness to other persons
Heterogeneity of junior researchers	Openness to ideas and persons	Openness to other persons
Assistance	Openness to ideas and persons	Openness to other persons
Course accessibility	Openness to ideas and persons	Openness to other persons
Availability of publications	Openness to ideas and persons	Openness to other persons
Achievement of own goals	Selfmanagement, independence	Realization of own research goals
Financial independence	Selfmanagement, independence	Research outcomes are unpredictable
Absence of requirements	Selfmanagement, independence	Research outcomes are unpredictable
Sources	Scholarship, erudition	Knowledge based on own research
Research time	Scholarship, erudition	Knowledge based on own research

(continued)

Table A.2: Groups of indicators and the criteria and aspects they can potentially measure (continued)

Name of group	Criterion	Aspect(s)
Personal library	Scholarship, erudition	Knowledge based on own research
Teaching awards	Passion, enthusiasm	Arouse passion for research
Survey: enthusiasm – teaching	Passion, enthusiasm	Arouse passion for research
Survey: enthusiasm – public	Passion, enthusiasm	Arouse passion for research
Strategies	Vision of future research	Pointing out important research for the future
Utilizing sources	Vision of future research	Pointing out important research for the future
Congruence research - teaching	Connection between research and teaching, scholarship of teaching	Research-based teaching
Research orientation of teaching	Connection between research and teaching, scholarship of teaching	Research-based teaching
Students' publications	Connection between research and teaching, scholarship of teaching	Research-based teaching
Acknowledging junior researchers	Connection between research and teaching, scholarship of teaching	Teaching-based research
Collaboration with students	Connection between research and teaching, scholarship of teaching	Teaching-based research
Publication of course material	Connection between research and teaching, scholarship of teaching	Teaching-based research
Invited lectures	Relevance	Research is relevant for the research community

Note: We assigned the groups of indicators only to the aspects that are consensual in at least one of the three disciplines covered in our study.

^a While the group of indicators points to the occurrence of this aspect or criterion, not all of the indicators in the group are a sufficient condition for the occurrence of the aspect or criterion.