



Exploring and Developing Research Quality Criteria in the Humanities

ECOOM colloquium, December 9th, 2011

Sven Hug, Michael Ochsner & Hans-Dieter Daniel

Social Psychology and Research on Higher Education

ETH Zurich / University of Zurich



Contents

- Background
- Starting point
- Project overview
- Repertory Grid interviews
- Delphi survey
- Preliminary findings



Background

Rectors' Conference of the Swiss Universities (CRUS):

- How can research in the humanities be measured and evaluated?
- Insufficient conventional approaches
- Exploration and development of new approaches
 - which allow an **international comparison**
 - and consider **differences between disciplines**
- Furnishing universities with instruments which offer information for
 - internal evaluation and quality control
 - strategical decisions and profiling



Starting point

- Critique of performance comparison by humanities' scholars:
 - Methods originate from the natural sciences
 - Strong reservations against quantification
 - Fear of dysfunctional effects
 - Lacking consensus over quality criteria
- Bibliometrics: low coverage of books and articles, etc.



Starting point: Critique from the humanities

1. Methods originate from the natural sciences

- Measurement methods do not follow up on **practices of research** in the humanities (e.g. Hicks, 2004; Nederhof, 2006; Vec, 2009)
- Humanities puts its focus on the **expansion of knowledge** rather than a linear understanding of progress (Lack, 2008, p. 14)

*„The humanities and social sciences have always been marked by controversy and competing visions of how things should be.“
(Fisher et al., 2000)*

- Therefore, citations must be interpreted differently (e.g. Hose, 2009)
- Indicators are seen as **tautological**: Citations attract more citations, third-party funding leads to more third-party funds (Charle, 2009, p. 165; Kemp, 2008, p. 14)



Starting point: Critique from the humanities

2. Strong reservations against quantification

- Research has a unquantifiable intrinsic value

„The liberal university has intrinsic value as a freestanding guardian of 'the house of knowledge'." (Fisher et al., 2000)

- „Performance“ as an unquantifiable entity

„Some efforts soar and other sink, but it is not the measurable success that matters, rather the effort. Performance measures are anathema to arts because they narrow whereas the arts expand." (Fisher et al., 2000)



Starting point: Critique from the humanities

3. Fear of dysfunctional effects

- Diversity in research:

“At the moment, we cannot imagine what will happen if it is expanded into an European scale, when in Brussels not Olives but research gets funded. [...] It is going to be a Behemoth of research! [...] This glutton will feast on megatrends, it will permanently transfer quality into measurable items, while in reality, [this] Behemoth likes just one thing: More of the same. ” (Kemp, 2008, p. 148, own translation)

- Conservativity:

„Overall, performance indicators reinforce traditional academic values and practices and in trying to promote accountability, they can be regressive.“ (Fisher et al., 2000)



Starting point: Critique from the humanities

3. Fear of dysfunctional effects

- Participation in civil society (e.g., Fisher et al., 2000):

„Academics play an important social role that is overshadowed by the reductionist focus on publication as the measure of output. They contribute analysis and commentary on issues and the human dimension of all aspects of society.“ (Fisher et al., 2000)



Starting point: Critique from the humanities

4. Lacking consensus over criteria

- A consensus has not been reached yet:

"In a some of the humanities' disciplines a consensus over criteria for good and bad research does not only not exist but there is also no consensus over the objects of research and the meaningful use of the right methods." (Herbert & Kaube, 2008, p. 50, own translation)

- And if, then

[standards of quality] usually refer to the same discipline and are [...] not readily transferable to other subdisciplines. (Herbert & Kaube, 2008, p. 40, own translation)

Starting point: Critique from the humanities

Does this critique free the humanities from accountability?

- Negation by humanities scholars as well as critics of already existing evaluation practices. For example Fisher et al., 2000:

„Nowhere does academic freedom mean freedom from accountability; nor should it. Rather, what is at stake is how accountability will be defined, and what will constitute 'performance' or 'quality'."



- Necessity of **definition** and **explication**

„[...] one thing [...] is unavoidable: it must be ensured that a standard is defined or measurement criteria have to be determined respectively.“ (Goebel, 1997, p. 458, own translation)

“Therefore the task is to explicate and to make argumentation accessible across disciplines.“ (Niessen, 2008, p. 264, own translation)



Starting point: Definition of quality

Definition and operationality of quality are not interrelated:

„[The assessments] often still make only a weak connection between theoretical definitions of quality and its measures by asserting a single rank or rating system that obscures the methodological and theoretical assumptions built into it.“ (Brooks, 2005, p. 1-2)

„This leads to the observation, that ultimately ‚quality of research‘ is defined by its measurement.“ (Donovan, 2008, p. 78)



Starting point: Indicators

What do indicators measure?

- e.g. citations:
 - *„It is extremely difficult if not impossible to express what citations measure in one single theoretical concept [...]. Citations measure many aspects of scholarly activity at the same time.“(Moed, 2005, p. 221)*
 - Citations contain nothing which could point to the intrinsic quality of the work (Donovan, 2008, p. 77)
- e.g. third-party funds; prizes, awards
 - *„Quality of research is in fact not measured“ (Donovan, 2008, p. 77)*

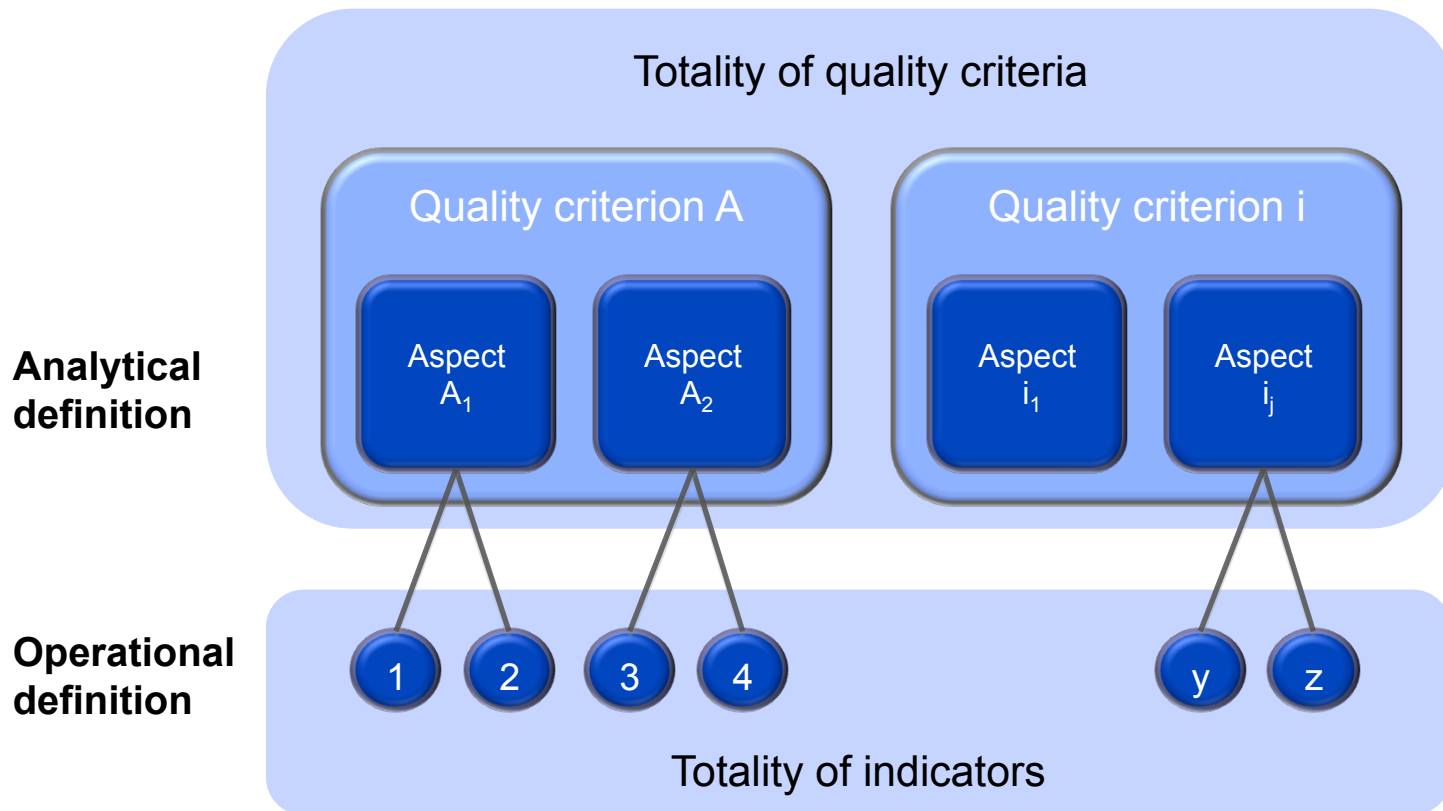


Starting point: Indicators Measurement approach used in the social sciences

„The attempt to create measurability of research and teaching performances presupposes – in view of the necessary operationalisation – an explicit understanding of quality.“ (Schmidt, 2005, own translation)

„[Der] Versuch, Messbarkeit von Forschungs- und Lehrleistungen herzustellen, setzt mit Blick auf die notwendige Operationalisierung ein explizites Verständnis von Qualität voraus.“ (Schmidt, 2005)

Starting point: Indicators Measurement approach





Starting point: Who defines quality?

Most often:

Top-down, e.g.:

- Political stakeholders
- Board of University
- Renowned experts

But:

- **Scholars** want to determine what defines quality in their field
- In this process not individual researchers but rather the **research community** has to determine how research performance is to be measured (see, e.g. Merton, 1973)



Starting point: Tacit knowledge

- Psychology: Judgment and Decision Making (e.g., Thorngate et al. 2009, p. 33):
 - *“I can’t define what merit is, but I know it when I see it.”*
 - *“I trust that my gut will tell me when I have found the winner.”*
- Knowledge of research quality exists as **implicit knowing / tacit knowledge** (Polanyi, 1967)



Starting point: The Zurich approach

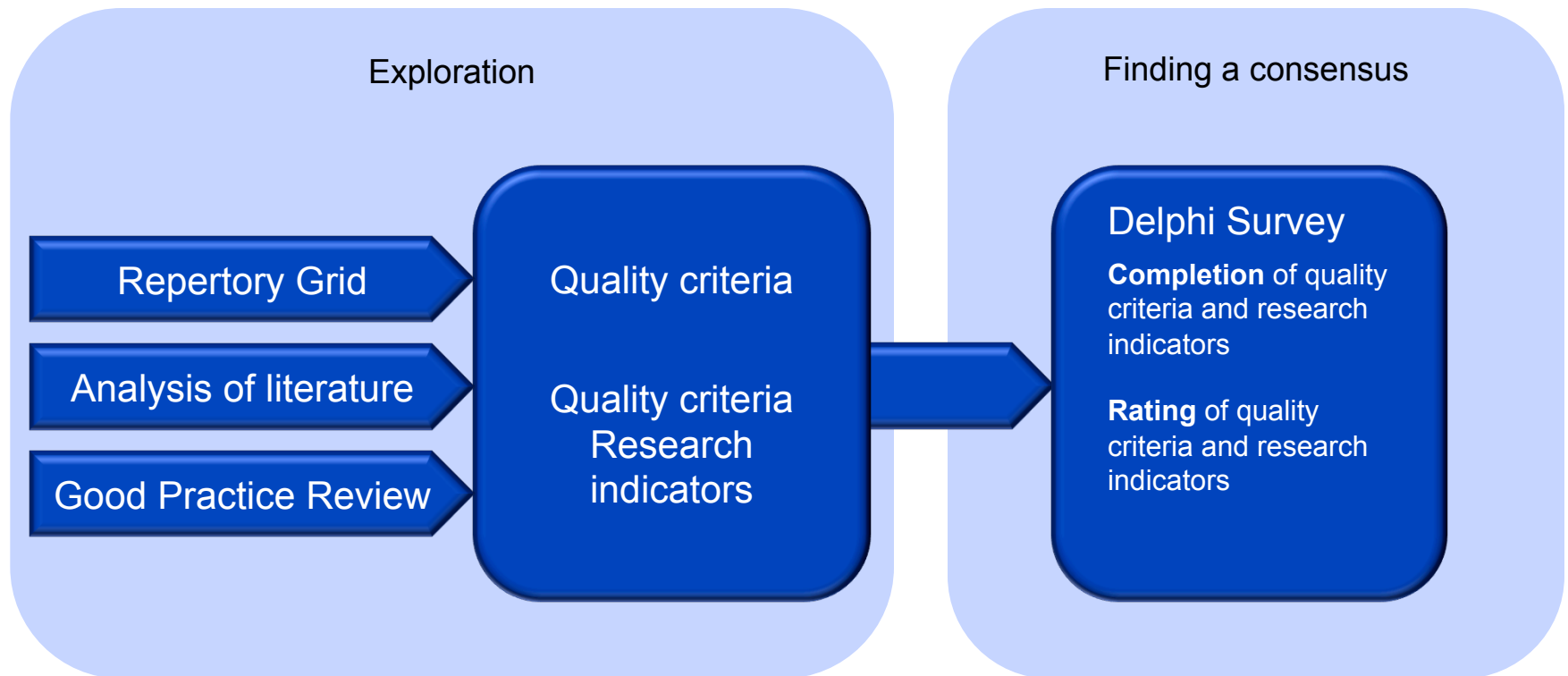
- **Bottom-up approach:**
 - Scholars' concepts of quality
 - Open-outcome
- **Making explicit:**
 - Making implicit knowledge explicit (Repertory Grid technique)
 - Capture the quality criteria in the scholars' words and intention as good as possible
- **Measurement approach:**
 - Measurement approach from the social sciences
 - Determination of quantifiable and non-quantifiable quality criteria
- **Acceptance:**
 - Reaching a consensus on a broad basis (Delphi survey)



Starting point: Summary

Issues	Approach
Methods from the natural sciences	<ul style="list-style-type: none">• Scholars' concept of quality• Open-outcome approach
Fear of dysfunctional effects	<ul style="list-style-type: none">• Capture the quality criteria in the scholars' words and intentions as good as possible
Reservations against quantification	<ul style="list-style-type: none">• Measurement approach from social sciences• Determination of quantifiable and non-quantifiable quality criteria
Lacking consensus	<ul style="list-style-type: none">• Inclusion of all scholars• Reach consensus

Project overview



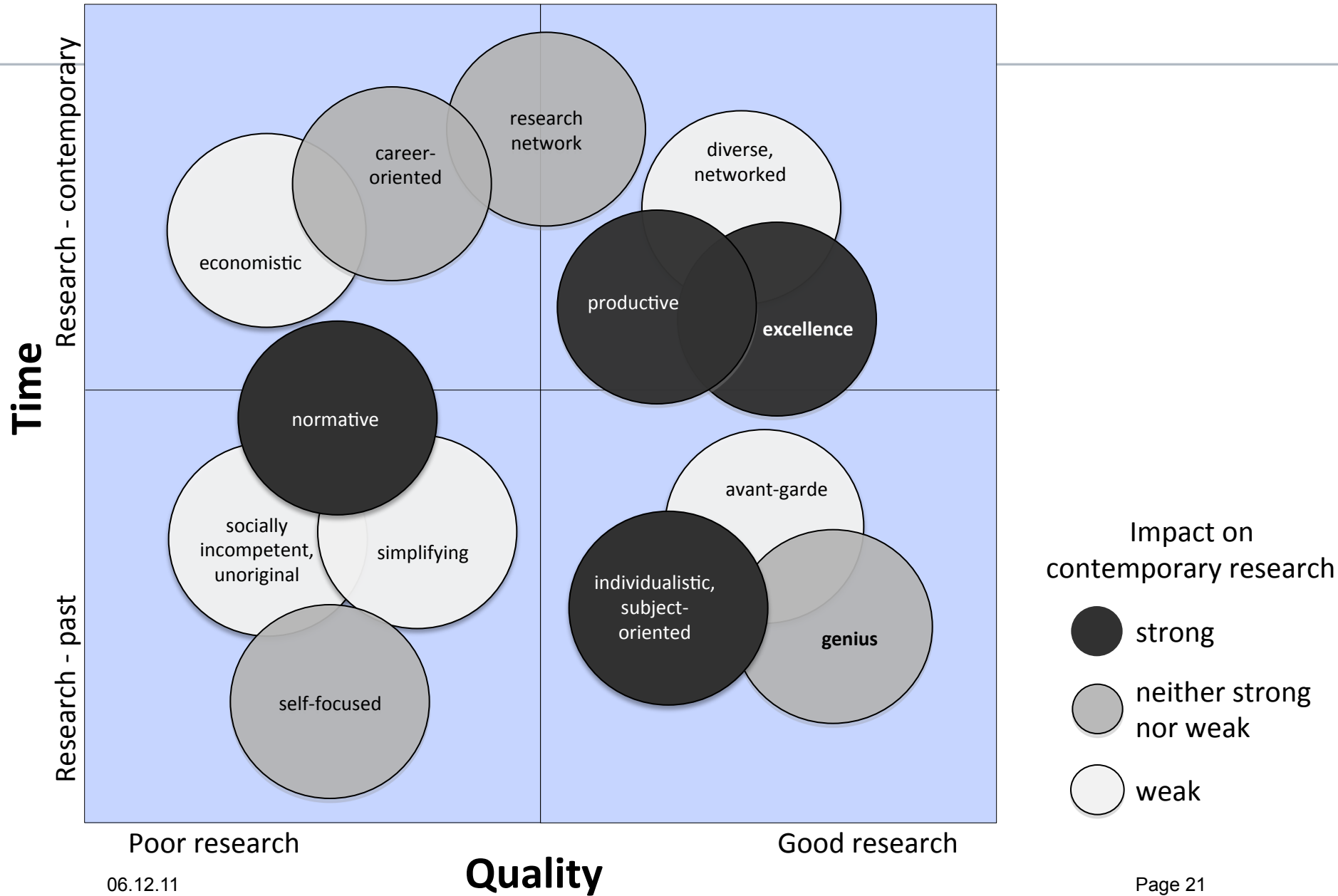
Repertory Grid interviews

- **Exploration** of **implicit** concepts of quality („tacit knowledge“)
- Repertory Grid technique was founded by George A. Kelly within the scope of his Psychology of Personal Constructs (1955)



- Sample: 21 researchers from Art History, English Literature Studies and German Literature Studies
- Face-to-face interviews, 2 hours length

German Literature Studies: (implicit) quality concepts





Repertory Grid interviews: Example for derived quality criteria

Criterion: **Erudition, Scholarship, Being Well-read**

- Rich experience with sources, materials, original works (knowledge of the material)
- Rich knowledge and insights that have been obtained through own research and reflection

Criterion: **Passion, Enthusiasm, Intrinsic Motivation**

- Inducement / elicitation of passion or enthusiasm for the research or the research topic



Delphi: Overview

Basic concept of Delphi:

- Repeated, written survey of experts with anonymous feedback after each survey round

Pool of experts (German Literature, English Literature, Art History):

- Switzerland: 130 professors at 8 universities
- LERU (League of European Research Universities): 520 professors at 17 universities

Three survey rounds:

- Round 1: **Completion** of quality criteria and research indicators
- Round 2: **Rating** of quality criteria
- Round 3: **Rating** of research indicators



Delphi: Round 1

- Purpose: **Completion** of **quality criteria** and **research indicators** from the explorational phase
- Sample:
 - Switzerland: 30 professors per field (total: 90)
 - LERU: 30 professors per field (total: 90)
- Paper-and-pencil survey in German and English
- Field work: May 20 until July 18, 2010
- Response rate: 28% (50 of 180 professors)



Name
of criterion

Selection of aspects

Supplemental aspects

Naming of indicators

Name of the criterion: **Recognition, Acknowledgement**

Please mark those aspects you personally find fit or match the criterion :

- Good research is acknowledged by the research community.
- Good research is valued by society.

In your view, are aspects you consider important missing? Please describe them below:

-
-
-

Please specify how each aspect marked or noted above can be identified or determined.

.....

.....

.....

.....

.....

.....



Delphi: Round 2

- Purpose: **Rating** of **quality criteria**
- Sample:
 - Switzerland: 130 professors
 - LERU: 520 professors
- Online-survey in German and English
- Field work: March 29 until April 24, 2011
- Response rate: 30% (196 of 650)



Delphi: Round 2 – Findings

Consensual quality criteria

Consensus in all three fields:

- Erudition, Scholarship, Being Well-read
- Fostering Cultural Memory
- Vision of the Future
- Inspiration for and Impact on Research Community
- Innovation, Originality
- Rigour
- Openness to Ideas and Persons
- Passion, Enthusiasm, Intrinsic Motivation
- Scholarly Exchange



Delphi: Round 2 – Findings

Consensual quality criteria

Consensus in two fields:

- Ability to Link or Connect to other Research, Topicality
- Diversity, Variety
- Connection between Research and Teaching, Scholarship of Teaching
- Reflection, Criticism
- Self-Management, Independence

Consensus in one field:

- Continuity, Continuation
- Relevance, Significance
- Recognition, Reputation, Acknowledgement



Delphi: Round 3

- Purpose: **Rating** of **research indicators**
- Sample:
 - Switzerland: 130 professors
 - LERU: 520 professors
- Online-survey in German and English
- Field work: October 17 until December 31, 2011
- Response rate: currently 17%



Delphi: Round 3 - Example

Criterion: Fostering Cultural Memory

- Aspect: Documentation, conservation, preservation or keeping alive of aspects of the past:
 - Number, weighting and duration of **documentation** or **preservation activities**
 - Number and weighting of **outputs** reflecting documentation and preservation activities
 - Number and weighting of **activities for the public** (e.g.: guided tours, public lectures, readings, media appearances, performances)
 - Number and weighting of **outputs for the public** (e.g.: popular books or articles, exhibitions, documentary films)
 - **Other/additional quantitative statements**, such as ...
- Aspect: Renew the understanding and the interpretation of aspects of the past
 - No research indicators available → assessment by peers only



Delphi: Round 3 - Operationalizability

Field	No. of criteria	No. of aspects	No. of aspects with indicators
German Literature Studies	16	36	19 (= 53%)
English Literature Studies	13	29	15 (= 52%)
Art History	13	31	15 (= 48%)



Preliminary findings

- **Interest** in development of **quality criteria** is high, if the specific fields are part of the process (response rate: 30%)
 - But interest seems to be low when **research indicators** come into play (response rate: currently 17%)
- **Consensus** of a set of quality criteria is possible **within a given field**
- **Common denominators** between fields are possible, but field differences are vast
 - 9 criteria, which all fields have in common, versus 8 criteria which are not common to all
- In order to adequately evaluate research performances in the humanities, **a broad spectrum** of quality criteria has to be considered.
- Only **50%** of the quality criteria's aspects can be measured with indicators



Contact details

Sven Hug

ETH Zurich

Social Psychology and Research on Higher Education

Muehlegasse 21

8001 Zurich

Phone: +41 44 632 46 85

E-Mail: sven.hug@gess.ethz.ch