

Peer Review under Review

Thomas Widmer
Institut für Politikwissenschaft
Universität Zürich

Referat im Rahmen der Jahrestagung der SEVAL
Eidgenössisches Personalamt, Bern, 10. September 2010

Aufbau meines Vortrags

1. Definitionen
2. Beispiele von Peer Reviews
3. Dimensionen: Vielfalt der Peer Review
4. Diskussionen
5. Fazit

1. Definitionen

Was ist eine Peer Review?

Vorbemerkung zum Geschlecht

siehe Maria Gutknecht-Gmeiner (2008): Exterhne Evaluierung durch Peer Review. Wiesbaden: VS Verlag: 40.

“Evaluation, usually of proposals or college faculty, done by a panel of judges with qualifications approximating those of the author or candidate.”

Michael Scriven (1991): Evaluation Thesaurus. 4th ed. Newbury Park: Sage, p. 255.

“Peer review can be described as the systematic examination and assessment of the performance of a State by other States [...], with the ultimate goal of helping the reviewed State improve its policy making, adopt best practices, and comply with established standards and principles.”

Fabrizio Pagani (2002): Peer Review: A Tool for Co-Operation and Change. An Aanalysis of an OECD Working Method. Paris: OECD, p. 4

Das Peer-Review-Verfahren lässt sich in vielen Kontexten einsetzen, ist aber vor allem aus der Begutachtung von Aufsätzen in Fachzeitschriften bekannt [...]. Zentrales Kennzeichen der Vorgehensweise ist die Begutachtung einer Facharbeit durch andere Experten desselben Fachgebiets.

Wolfgang Meyer (2007): Datenerhebung. In: Stockmann, Reinhard (Hrsg.) (2007): Handbuch zur Evaluation. Münster: Waxmann, pp. 258-9.

“Evaluation carried out by equals, usually prominent colleagues of the same profession.”

Evert Vedung (1997): Public Policy and Program Evaluation. New Brunswick: Transaction, p. 305.

“Peer review is the assessment by an expert of material submitted for publication.”

Carin M. Olson (1990): Peer Review of the Biomedical Literature. American Journal of Emergency Medicine 8(4): 356-358.

„[...] der Peer Review, die Verfahren der Evaluation von wissenschaftlichen Produkten. Er besteht aus einer qualitativen Evaluation von Manuskripten oder Forschungsanträgen durch Kollegen, also andere Mitglieder der Scientific Community.“

Stefan Hirschauer (2006): Wie geht Bewerten? In: Flick, Uwe (Hrsg.): Qualitative Evaluationsforschung. Reinbek: Rowohlt, S. 407.

„Werden die beiden Verfahren der Selbst- und der Fremdevaluation kombiniert, so spricht man von einer Peer Review. Sie besteht aus einer vorangegangenen Selbstevaluation mit nachgelagerter Begutachtung durch externe Fachexperten.“

Margrit Stamm (2003): Evaluation und ihre Folgen für die Bildung. Münster: Waxmann, S. 55

Typische Elemente dieser Definitionen

- Es handelt sich um ein bewertendes Verfahren (~ Evaluation?).
- Bewertung erfolgt durch Gleichgestellte, Gleichartige, die aber nicht in demselben Handlungszusammenhang stehen.
- Bewertende wie zu bewertende Einheiten können auf der Mikro-, Meso- und Makroebene angesiedelt sein.

2. Beispiele von Peer Reviews

Wo gibt es Peer Reviews?

- Ex-ante Projektevaluierungen in der **Forschungs- und Technologieförderung** in Österreich (Rupert Pichler, S. 45)
- Institutionelle Evaluationen bei deutschen **Hochschulen** (Uwe Schmidt, S. 165)
- Universitätsübergreifende Evaluierungen einzelner **Fachbereiche** durch Fachgesellschaften und Ministerien in Österreich (Alexander Kohler, S. 179)
- Peer Reviews zur Überprüfung der Qualität des **Studienangebots** an schweizerischen Fachhochschulen (Heinz Rhyh, S. 190)
- periodische Überprüfung der **Evaluierungssysteme** der Träger in der deutschen **Entwicklungszusammenarbeit** (Michaela Zintl, S. 247)
- Peer Review der **OECD/DAC** der Austrian Development Agency (Anton Mair, S. 261) und der bilateralen schweizerischen Entwicklungszusammenarbeit (Dieter Zürcher, S. 265)
- Peer Review zu individuellen **Forschungsleistungen**, zu **Fachdisziplinen** und Forschungsfeldern in Deutschland (Stefan Kuhlmann, S. 285, 287)
- Internationale Peer Reviews zur **Auswahl förderungswürdiger Projekte** beim Österreichischen Wissenschaftsfonds FWF (Zinöcker/Dinges, S. 299-300)

Beiträge aus Thomas Widmer/Wolfgang Beywl/Carlo Fabian (2009): Evaluation. Wiesbaden: VS Verlag.

3. Dimensionen: Vielfalt der Peer Review

3.1. Funktionen

Peer reviews können summativ (bilanzierend) oder formativ (verbesserungsorientiert) ausgerichtet sein.

Vielfach wird die funktionale Ausrichtung in der Praxis nicht klar festgelegt und auch nicht klar kommuniziert.

3.2. Perspektiven in der Zeit

Peer-Review-Verfahren werden einerseits prospektiv, andererseits aber auch retrospektiv ausgerichtet.

Typische Beispiele

- prospektiv: Bewertung eines Projektantrags (etwa in der Technologieförderung oder in der Entwicklungszusammenarbeit)
- retrospektiv: Bewertung der bisherigen Projektarbeiten (etwa im Hinblick auf die Finanzierung einer zweiten Etappe)

3.3. Bewertungskriterien: Quellen und Form

Grundsätzlich besteht bei Peer Reviews der Grundverdacht, dass die Peers selbst wissen (sollten), welche Bewertungskriterien angemessen, aussagekräftig und bedeutend sind.

Fachlichkeit, Disziplin und/oder Profession als Quelle legitimer, aber selten explizierter Bewertungskriterien

Fragen etwa zur Effizienz oder zum Gemeinwohl sind eher nicht von Interesse.

Bei Peer-Review-Verfahren werden Bewertungskriterien teils gar nicht, teils recht offen definiert.

Durchaus üblich sind – etwa – bei der Review eines Zeitschriftenartikels – Fragen wie:

- “Is the methodology sound?”
- “Are the results credible and do they address the research question(s)?”
- “Does the paper make a contribution to the field?”
- “Does the paper add to what is already known or current thinking?”

3.4. Bewertungsgegenstände

Evaluanda aller Art auf allen Ebenen:

- Staaten / Organisationen (Strukturen)
- Programme / Projekte (Prozesse)
- Produkte / Leistungen (Resultate)
- Personen
- ...

3.5. Organisatorische Form der Peer Review

„on site“		distanziert
konferenziell		aktenbasiert
kollektiv		individuell
offen		blind (ein- / beidseitig)

3.6. Reziprozität

Häufig ist das einseitige Modell:
A bewertet B

Seltener geht es um eine gegenseitige Review:

A bewertet B
B bewertet A.

Beispiele:

gegenseitige Visitationen der Lehrpersonen („critical friends“)
Verbund Technischer Universitäten in Deutschland

3.7. Bewertungsmethodik

Gewicht der

- Expertise der Peers („connaissanceurship“)

und der

- Evaluationsexpertise (etwa mit der Erstellung von Selbstevaluationsberichten)

im Peer-Review-Verfahren

4. Diskussionen

4.1. Validität und Reliabilität

4.2. Retrospektiv – prospektiv

4.3. Prozesse, Produkte und Wirkungen

4.5. Innovation und Konservatismus

4.5. Vertrauen und Akzeptanz

4.6. Betrug und andere Verbrechen

4.7. Peer review ~ Laienbewertung

4.1. Validität und Reliabilität

“Editorial peer review, although widely used, is largely untested and its effects are uncertain.”

Tom Jefferson et al. (2002): Effects of Editorial Peer Review. A Systematic Review. Journal of the American Medical Association 287(21): 2784.

„Thus, although recommendations made by reviewer have considerable influence on the fate of both papers submitted to journals and abstracts submitted to conferences, agreement between reviewers in clinical neuroscience was little greater than would be expected by chance alone.“

Peter M. Rothwell/Christopher N. Martin (2000): Reproducibility of peer review in clinical neuroscience. Brain 123: 1964.

“Peer reviews frequently produce shaky results. Matched panels use widely different merit criteria and performance standards and reach miscellaneous conclusions.”

Evert Vedung (1997): Public Policy and Program Evaluation. New Brunswick: Transaction, pp. 90

“The studies confirm that peer review represents a quality filter and works as an instrument for the self-regulation of science.”

Lutz Bornmann (2010): Does the Journal Peer Review Select the “Best” from the Work Submitted? The State of Empirical Research. IETE Technical Review 27(2): 93.

4.2. Retrospektiv – prospektiv

Peer Reviews haben nicht zuletzt pragmatische Vorteile bei der prospektiven Sicht. Andere Evaluationsansätze zeigen Schwächen bei der Einschätzung zukünftiger Entwicklungen, sie haben ihre Stärken eher dort, wo Evidenz erhoben werden kann.

4.3. Prozesse, Produkte und Wirkungen

Kausale Attribution ist nicht die Stärke der Peer Review, hier sind andere Ansätze bei weitem leistungsfähiger

4.4. Innovation und Konservatismus

lineare Akkumulation in der ‚normal science‘ (Popper)

wissenschaftlicher Fortschritt durch Revolution (Kuhn)

4.5. Vertrauen und Akzeptanz

Peer Reviews haben primär im entsprechenden Feld Glaubwürdigkeit und Akzeptanz, nicht aber gegenüber Aussenstehenden.

Peer Reviews eignen sich daher primär für die Selbstregulierung und –steuerung.

4.6. Betrug und andere Verbrechen

Robert K. Merton: Normen der Wissenschaft bekannt auch unter dem Akronym Cudos für

- ⇒ Communalism
- ⇒ Universalism
- ⇒ Disinterestedness
- ⇒ Organised Scepticism

Zuckerman, Harriet/Merton, Robert K. (1971): Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. *Minerva* 9(1): 66-100.
Reprinted in Merton, Robert K. (ed.) (1973): *The Sociology of Science*. Chicago: University of Chicago Press, p. 460-496

4.7. Peer review ~ Laienbewertung

Peers qualifizieren sich als Peers nicht wegen ihrer Evaluationsexpertise; es handelt sich also in diesem Sinne bei der Peer Review um eine Laienbewertung

Ausnahme: Peer Review under Review

5. Fazit

“Peer review is slow, expensive, profligate of academic time, highly subjective, prone to bias, easily abused, poor at detecting gross defects, and almost useless in detecting fraud.”

Richard Smith (1999): Opening Up BMJ Peer Review. *British Medical Journal* 318 (7175): 4-5.

"Peer review is – under specific conditions – the worst form of evaluation, except for all those other forms that have been tried from time to time."

Sir Winston Churchill from a House of Commons speech on Nov. 11, 1947, slightly adapted

Original quote: "Democracy is the worst form of government, except for all those other forms that have been tried from time to time." The Official Report, House of Commons (5th Series), 11 November 1947, vol. 444, cc. 206–07.

Besten Dank!