



Evaluationen nutzen - Mut zur Qualitätsentwicklung, Selbstreflexion und zum Nein sagen

Dr. Dagmar Simon

16. Jahrestagung des AK Evaluation und Qualitätssicherung der Berliner
und Brandenburger Hochschulen

15./16. März 2018

Universität Potsdam

Agenda

Eine unübersichtliche Evaluationslandschaft

Von Quantitäten und Qualitäten

„Metric Tide“ und Leistungsmessung in einer Universität

Evaluationslandschaft und Organisationsverhalten

Ein vorläufiges Resümee



Hilfe, schon wieder 'ne Evaluation!

Warum immer mehr Evaluationen ?

Exponentielles Wachstum der Wissenschaft

Ausdifferenzierung der Wissenschaft (Forschungsorientierungen, Disziplinen...)

Wettbewerbliche Verfahren, vor allem im Hochschulsystem; Eingang in Forschungsförderungs-, Anreiz- und Bewertungsinstrumente

Legitimationsansprüche für öffentlich geförderte Wissenschaft sind gestiegen (Nachweis von Nützlichkeit)

Wissenschaftsinterne kritische Reflexion der Entscheidungs- und Bewertungsprozesse als Triebkraft (Hornbostel 2016)

Thesen aus der Wissenschaftsforschung

Stärkerer Ökonomisierungs- und Relevanzdruck auf
Wissenschaftssystem (Weingart, Carrier, Krohn 2007)

Governance by Numbers (Heintz 2006)

Deprofessionalisierung der Wissenschaft (Schimank 2009)

The Transition from Quality Control to Quality Monitoring in Science

(nach Hemlin, Rasmussen 2006)

Dimension	Quality Control (Product Orientation)	Quality Monitoring (Process Orientation)
Criteria	Scientific	Scientific and societal
Focus	Individual researchers	Organizations, networks
Goal	Valid, reliable knowledge	Socially robust knowledge, learning
Evaluator	Traditional peers	New peers, users, consultants, lay persons
Evaluation time	After production	Continuously
Science study perspective	First order: philosophy and sociology of knowledge	Second order: knowledge management, organizational learning

Eine unübersichtliche Evaluationslandschaft

Gefahr der „Mehrzweckwaffe“ Evaluation (unterschiedliche Ziele und Konsequenzen)

- Rolle der Gutachter/innen: Critical friends und/oder Sachverwalter professioneller Standards?

Entgegenläufige Tendenzen: Zählen und Messen einerseits, Beitrag zur Organisationsentwicklung andererseits

(implizite) Verständigung auf „informed peer review“

Leistungsindikatoren in der Wissenschaft

Komplexitätsreduktion durch quantitative Indikatoren

Indikatoren zeigen Forschungsleistung an, ohne sie zu erklären

Ausdifferenzierung der Indikatoren ...

... jedoch (implizite) Hierarchisierung: Publikationen und Drittmittel

Herstellung der Vergleichbarkeit von nicht Vergleichbaren
(Heintz 2010, Godin 2006)

Nicht alles an Forschungsleistungen ist in Indikatoren und Zahlen darstellbar

Was nicht gemessen wird, zählt nicht (mehr)

Publikationsindikatoren – ein Problem?

Dominanz der referred journals

Impact-Factor: Änderung der Publikationsstrategien

Fachdisziplinäre Spezifika werden unzureichend berücksichtigt

Unterschiedliche Forschungsorientierungen finden nicht ausreichende Beachtung

Individualisierung von Forschungsleistungen

Bei der Begutachtungspraxis wird unzureichend differenziert



Conclusions and Recommendations of the Metric Tide Report (2015)

Danger in rushing to over-interpret the available data

Metrics should support, not supplant, expert judgement

One size is unlikely to fit all

Indicators can only meet their potential if there are underpinned by an open, transparent and coherent data infrastructure

Inappropriate indicators create perverse incentives

Correlation analysis has shown that individual metrics give significantly different outcomes from the REF peer review process

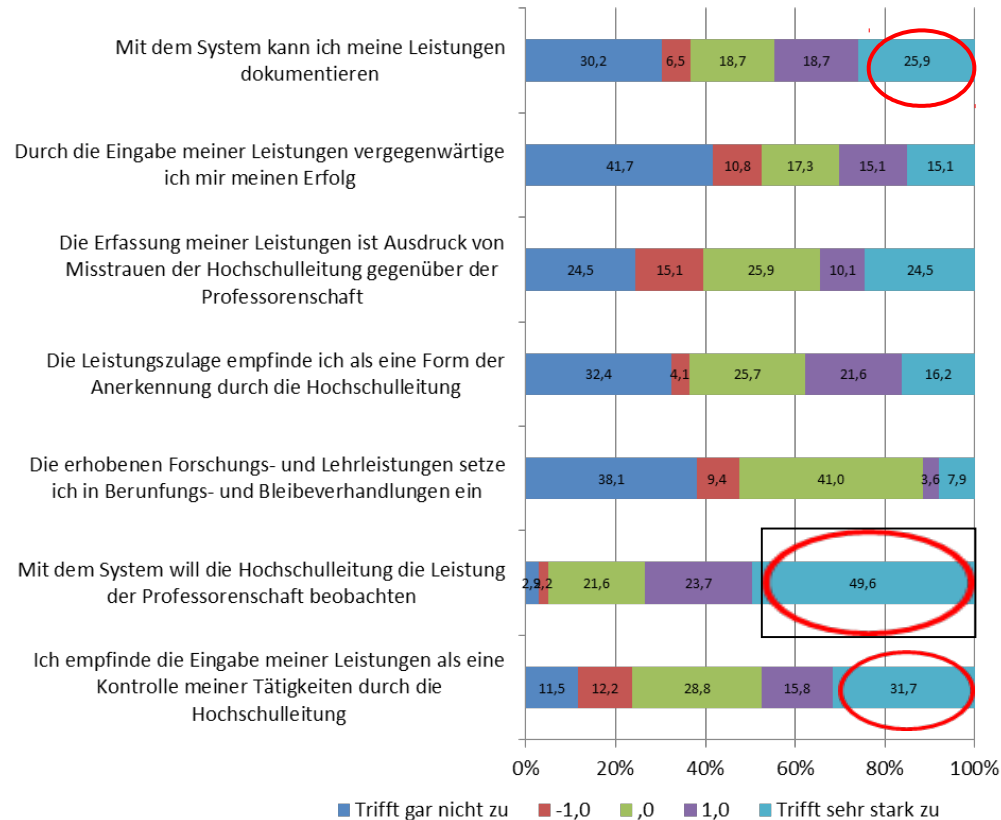
Need of „science of science policy“

Messung wissenschaftlicher Leistungen

Wissenschaftliche Leistung lässt sich nicht messen. Ob die Leistung hoch oder niedrig, gut oder schlecht ist, kann nur annäherungsweise durch Indikatoren ausgedrückt werden.

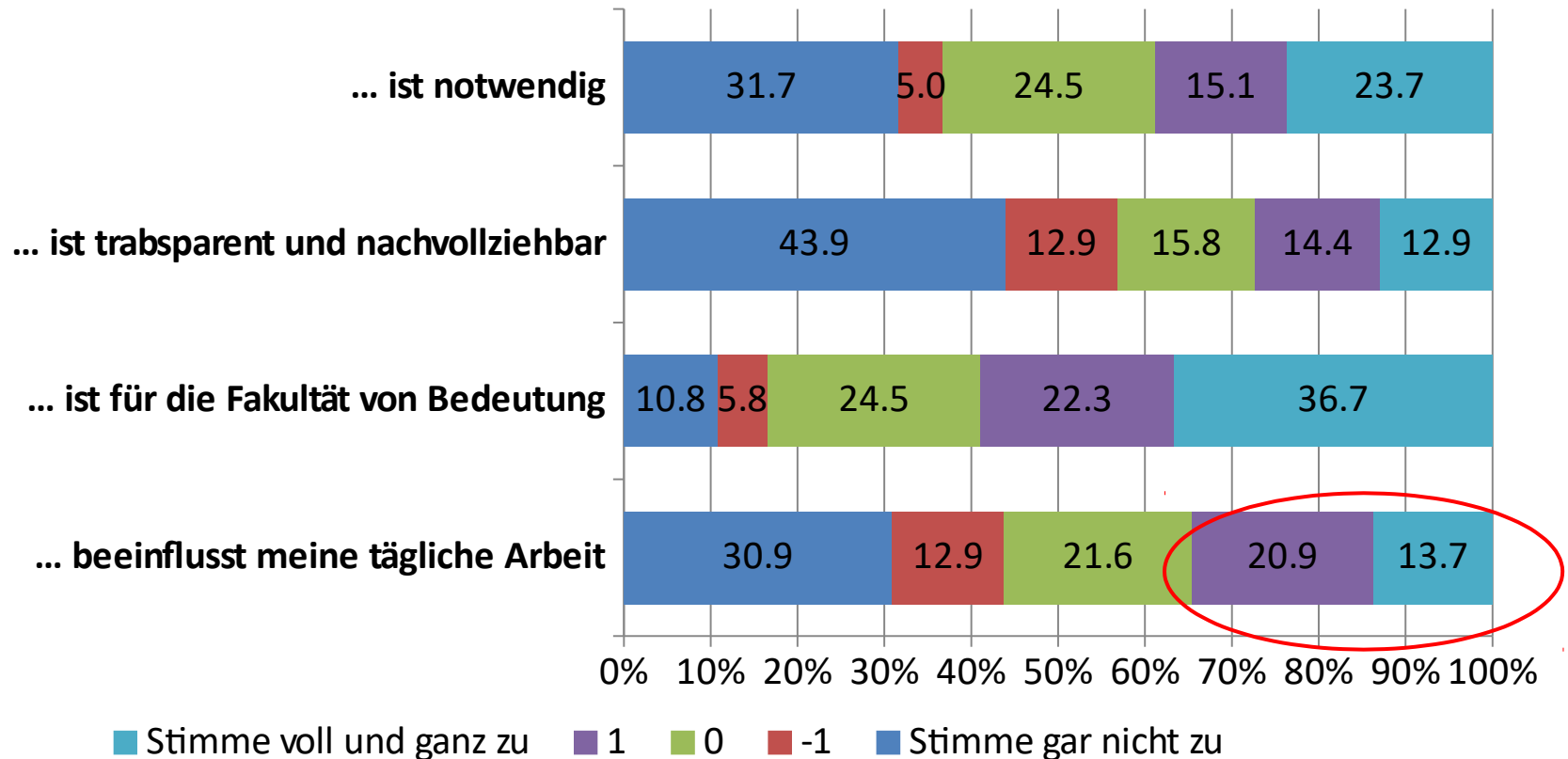
Die Wissenschaftsphilosophie hat es seit langem aufgegeben, Kriterien für die Qualität wissenschaftlicher Leistungen zu entwickeln, stattdessen sich auf ein Verfahren geeinigt: „Vernünftige Wetten der Fachkollegen“ (Toulmin).

Anerkennung und Kontrollerfahrung

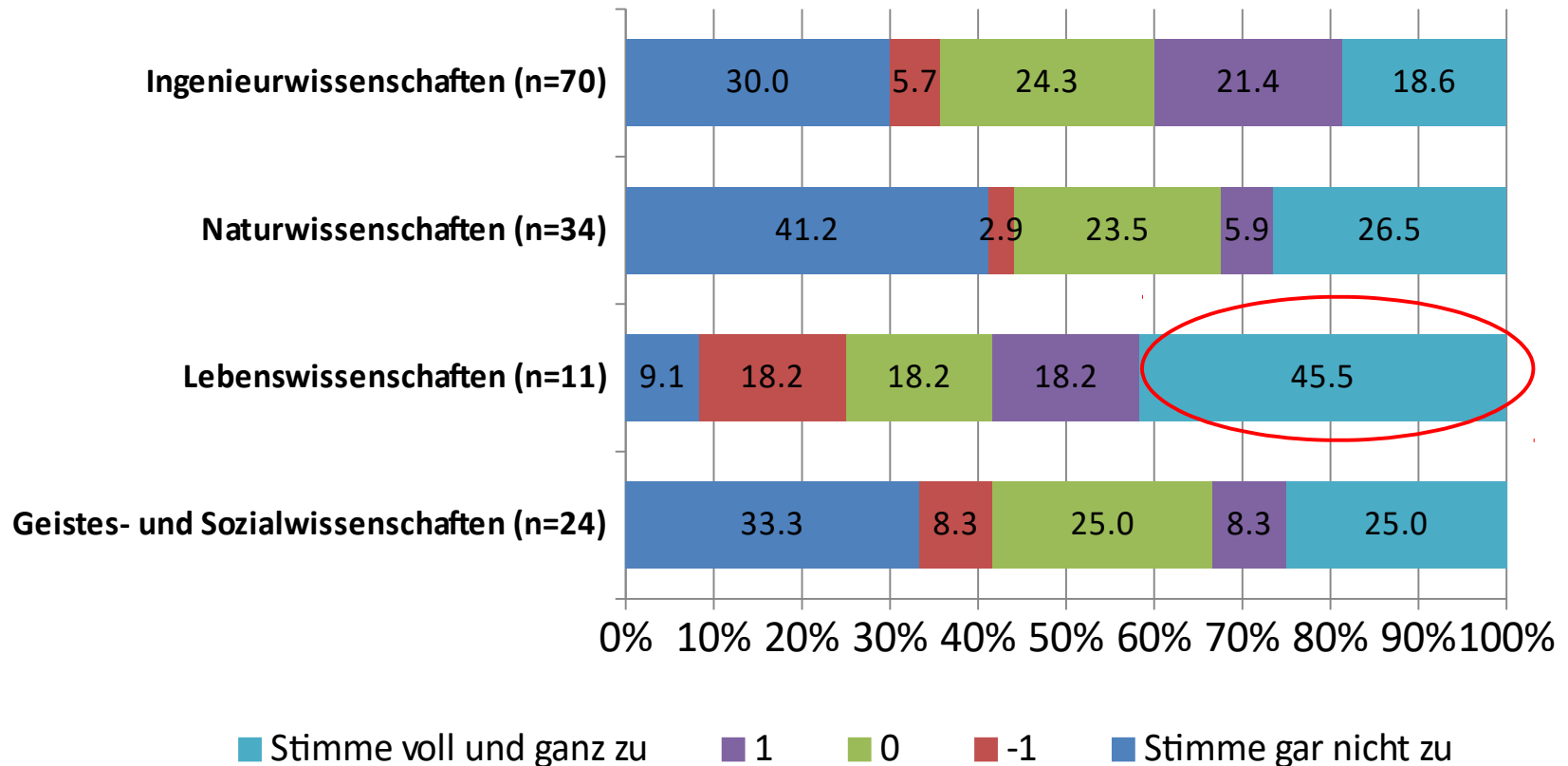


Ergebnisse Wahrnehmung von Leistungsmessung

Steuerung mit [Leistungsmessungssystem]...

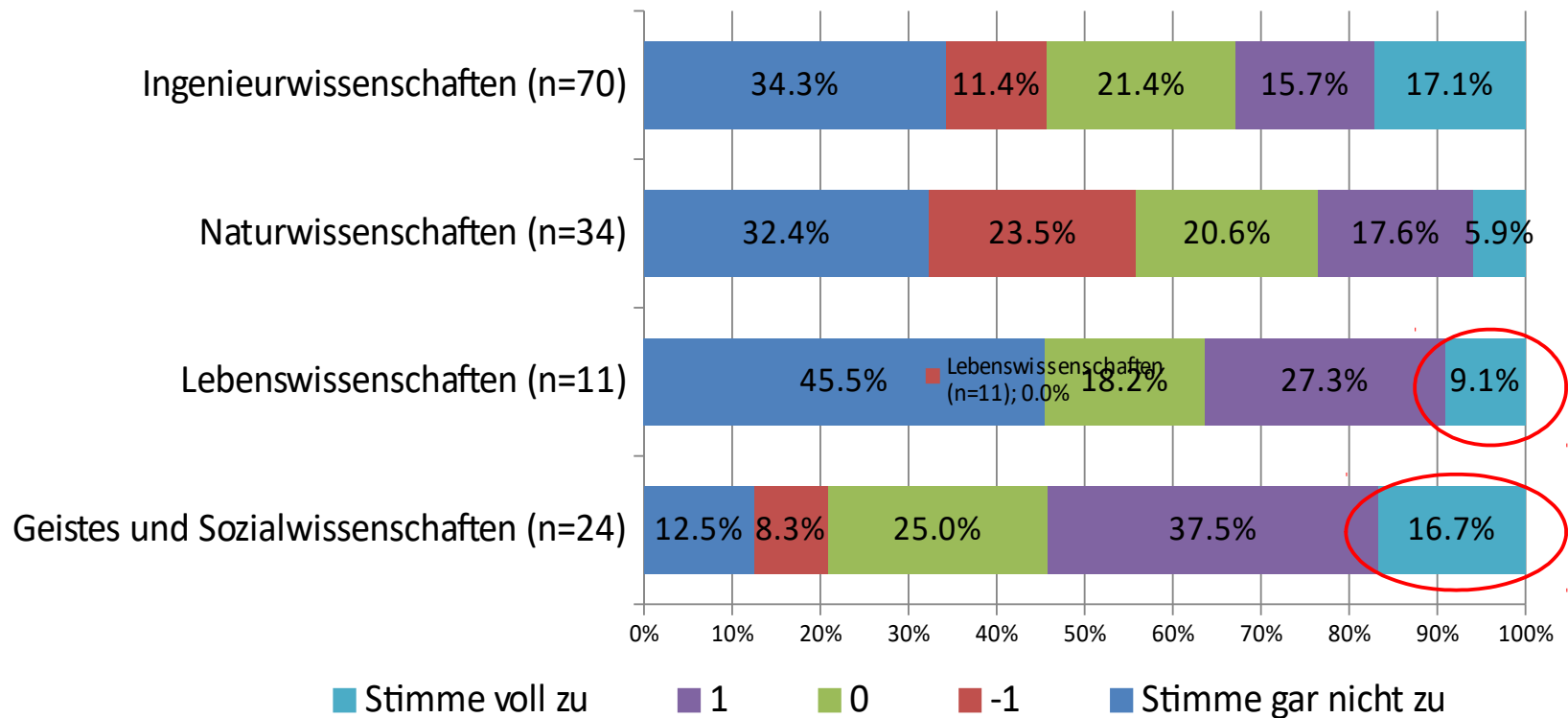


Steuerung ist notwendig

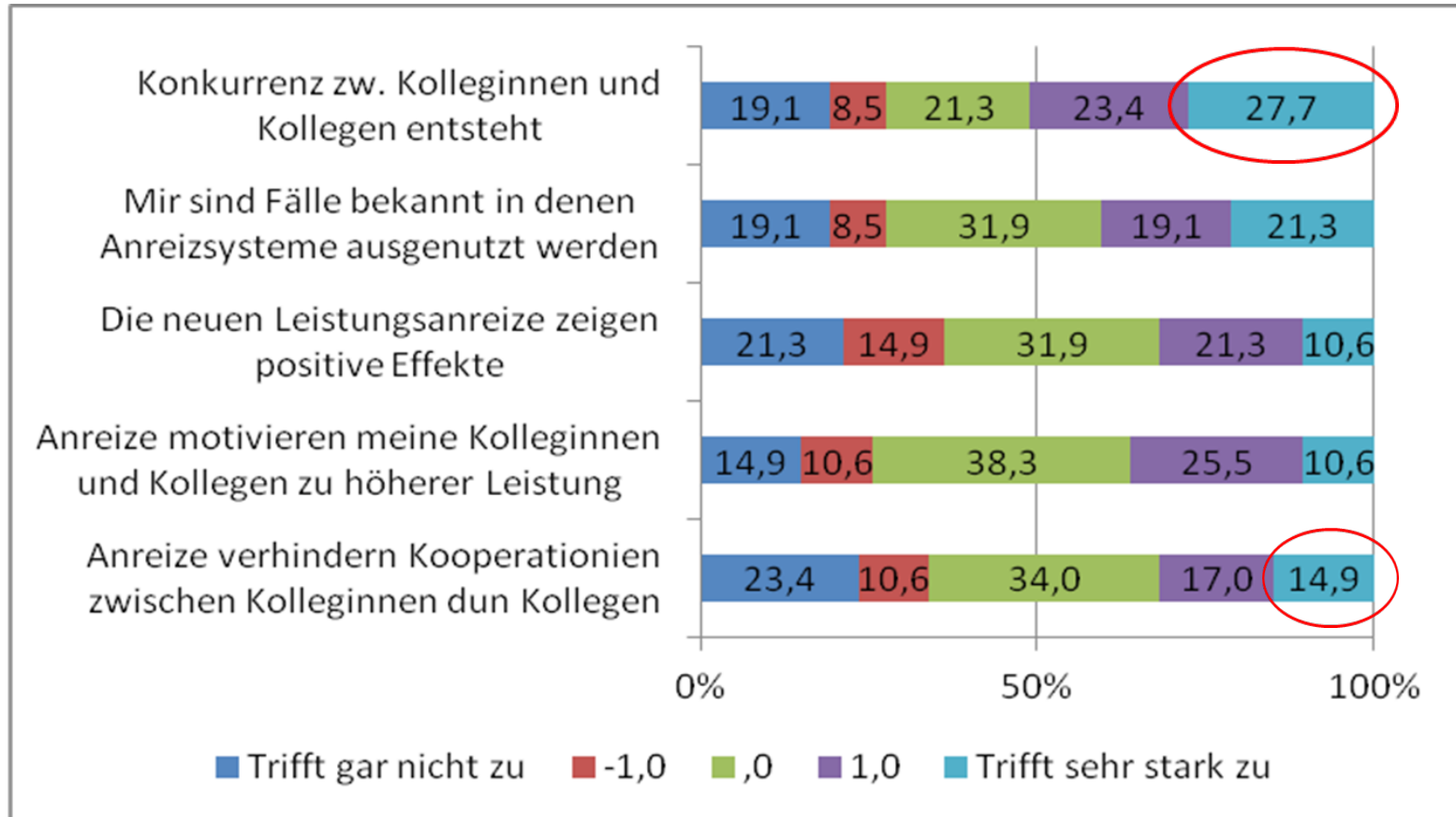


Die Steuerung mit dem Leistungsmessungssystem

beeinflusst meine tägliche Arbeit



Folgen Kennzahlen basierter Steuerung



Wirkungen von quantitativen Bewertungen aus einer Organisationsperspektive

Isomorphismus (z.B. U.S. News & Word Report Ranking auf die Law Schools; Espeland und Sauder 2007)

Unterschiedliche (strategische) Antworten von Organisationen auf quantitative Bewertungen

Bedeutung der „Evaluationslandschaften“ („Evaluative landscape, in contrast to a specific ranking, rating or award, describe the universe of relevant practices in the environment as a set of organizations that aim at evaluating and comparing these organizations' performance in a standardized way“ (Brandtner 2017)

Die „Evaluationslandschaft“ und Organisationsverhalten

Je pluraler die Evaluationslandschaft und je fragmentierter die Kriterien, desto weniger konvergent das Organisationsverhalten

Je mehr externe Gruppen die institutionelle Umwelt von Organisationen konstituieren, desto pluraler ist die Evaluationslandschaft der Organisation(en)

Je weniger institutionalisiert und mehr umstritten Werte in der institutionellen Umwelt einer Gruppe von Organisationen sind, desto pluraler ist die Evaluationslandschaft

Peer Review – die beste aller Welten?

Der Mertonsche Universalismus und Begutachtungen als sozialer Prozess

Bedeutung gemeinsamer Denkschulen, Orientierung am „mainstream“, Innovationsfeindlichkeit (Frey 2003, Gillies 2014)

Unterschiede in den Disziplinen, Schwierigkeiten inter- und transdisziplinäre Forschung zu beurteilen, jedoch ID und TD findet in der Evaluationspraxis zunehmend mehr Anerkennung

Peer Review – die beste aller Welten?

- Moreover, the customary rules of methodological pluralism and cognitive contextualism (evaluating proposals according to the discipline of the applicant) are more salient in the humanities and social science panels than they are in the pure and applied science panels, where disciplinary identities may be unified around the notion of scientific consensus, including the definition of shared indicators of quality (*Lamont et al. 2011*).

Peer Review – die beste aller Welten?

Überlastung des Gutachtersystems (Wissenschaftsrat 2017)

Gutachter/innen-Kreise erweitern (Postdocs...)

Losverfahren ?!?

Förderformat „Experiment“ der Volkswagenstiftung

Resümee (1)

Wenn schon Evaluationen, dann sollten sie die Weiterentwicklung von wissenschaftlichen Einrichtungen unterstützen und nicht ausschließlich Kontrollfunktionen ausüben.

Vertrauen in die Peers schafft die Grundlagen für Veränderungen der Organisationen

Empfehlungen der Peers werden durch die Organisation „gefiltert“

Verfahren hat Auswirkungen auf die Entwicklung einer „Evaluationskultur“

Zunehmende Verknüpfung von externen und internen Prozessen der Organisationsentwicklung

Wissenschaftliche Einrichtungen nutzen Evaluationen für ihre interne Organisationsentwicklung, insbesondere zur

Legitimation und „Validierung“ interner Entwicklungen

Resümee (2)

Bedeutung der Auswahl und Zusammensetzung der Gutachter/innen in externen Evaluationen

Begrenztheit der Steuerung über Kriterien und Indikatoren

Kontinuität, jedoch Gefahr der Routinisierung

Evaluationsverfahren evaluieren und weiterentwickeln

Kurze Evaluationszyklen vermeiden

Vielen Dank für Ihre Aufmerksamkeit!



Kontakt:

Dr. Dagmar Simon

dagmar.simon@wzb.eu

Wissenschaftszentrum Berlin für Sozialforschung

Forschungsgruppe Wissenschaftspolitik

Reichpietschufer 50, 10785 Berlin