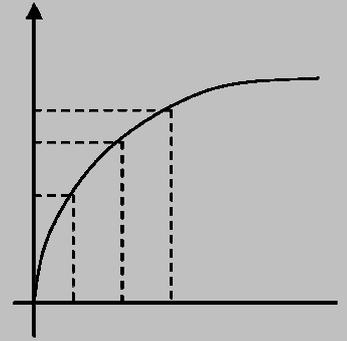


Diskussionspapiere des  
Europäischen Instituts für Sozioökonomie e. V.

Working Papers of the  
European Institute for Socioeconomics



# **Evaluation zwischen Angebot und Nachfrage – Vom Ethos der Forschung und dessen Wirkung auf die Wissensmärkte**

Eike Emrich

Nr. 9

Jahr 2014

**Prof. Dr. Eike Emrich**  
Sportökonomie und Sportsoziologie  
Universität des Saarlandes  
Sportwissenschaftliches Institut  
Campus  
66123 Saarbrücken  
Tel: +49(0)681 302 4170  
E-Mail: e.emrich@mx.uni-saarland.de

© 2014

Europäisches Institut für Sozioökonomie e. V. / European Institute for Socioeconomics

c/o Universität des Saarlandes

Sportwissenschaftliches Institut  
Arbeitsbereich Sportökonomie und Sportsoziologie  
Campus, Gebäude 8.1

66123 Saarbrücken

<http://www.soziooekonomie.org>

Eike Emrich

## **Evaluation zwischen Angebot und Nachfrage - Vom Ethos der Forschung und dessen Wirkung auf die Wissensmärkte**

### **Zusammenfassung**

Politisch vorangetriebene Mechanismen der Außensteuerung treiben Wissenschaftler in ein auch persönlich einkommensrelevantes Rattenrennen um Drittmittel und Impact-Faktoren, das die Bindung an das reine Ethos der Forschung im Sinne Mertons (1938) schwächt und die Wahrscheinlichkeit affirmativer Forschung erhöht, und zwar um so intensiver, je geringer die Bindung an das reine Ethos der Forschung wird. Bezogen auf Evaluationsforschung als Form angewandter Wissenschaft, die sowohl universitär wie außeruniversitär angeboten wird, tauchen dabei folgende Fragen auf: Welchen konkreten institutionellen Einflüssen ist Evaluationsforschung innerhalb und außerhalb der Universität jeweils ausgesetzt? Zu welchen spezifischen Konkurrenzverhältnissen kann es zwischen den Anbietertypen von Evaluationswissen kommen und welche Auswirkungen hat dies auf die jeweilige Wissensproduktion? Welche strukturellen Aspekte und Interaktionsdynamiken prägen die soziale Beziehung zwischen Auftraggeber und Auftragnehmer, und zwar sowohl für Anbieter innerhalb als auch außerhalb der universitären Wissenschaft? Zwecks Beantwortung wird wie folgt vorgegangen: Zunächst wird in einem kurzen Überblick die Frage der Unabhängigkeit der Evaluationsforschung kontextabhängig (universitäre versus nicht-universitäre Evaluationsforschung) diskutiert. Dann wird mit Hilfe empirischer Befunde das Problem der Berufs- bzw. Professionsethik von in der Evaluation tätigen Forschern aufgezeigt und am reinen Ethos der Forschung (Merton 1938) gespiegelt. Nachfolgend wird die soziale Beziehungen zwischen Auftraggeber und – nehmer mit Hilfe spieltheoretischer Überlegungen illustriert und mit Hilfe des in der Institutionenökonomik diskutierten Konzepts unvollständiger Verträge (vgl. Richter & Furubotn, 1999) in ihrer Interaktionsdynamik und in ihren wechselnden strukturellen Ungleichgewichten diskutiert, und zwar sowohl für die außeruniversitären Leistungsanbieter als auch für die universitären.

# Evaluation zwischen Angebot und Nachfrage - Vom Ethos der Forschung und dessen Wirkung auf die Wissensmärkte<sup>1</sup>

## 1. Einleitung

Bruno S. Frey (2007) spricht von *Evaluitis* und bezeichnet mit diesem nicht ganz wertfreien Begriff den Sachverhalt einer sehr häufig geforderten und auch durchgeführten, keineswegs immer problemadäquaten und Ziel führenden Anwendung von Evaluationsverfahren auf verschiedene Gegenstände.<sup>2</sup> Offensichtlich trifft Evaluationswissen im Zuge dieser „*Evaluitis*“ auf eine rege Nachfrage. Gleichzeitig ist aber auch nicht auszuschließen, dass im Markt vorhandene Evaluationsanbieter im Sinne des Saischen Theorems auch endogen Nachfrage nach ihren Evaluationsleistungen vor allem im öffentlichen Bereich erzeugen und so die „*Evaluitis*“ befeuern. Zu diesem Zweck propagieren Evaluatoren im politischen Raum oder in medialen Kontexten Nachfrage steigernd die Vorzüge ihrer Leistungen und damit implizit deren strategische Verwendbarkeit für Legitimationszwecke<sup>3</sup>. Zum Beispiel legen die vielfältigen Evaluierungen im Bildungssystem unter dem Stichwort IGLU- und TIMSS-Studie, PISA oder die Rankings der deutschen Universitäten durch das CHE-Ranking (vgl. Berghoff et al., 2009) den Verdacht nahe, dass hier die OECD bzw. die Bertelsmann Stiftung<sup>4</sup> die Nachfrage nach ihren Wissensangeboten im Wissensmarkt aktiv erhöhen. Dazu variieren sie

---

<sup>1</sup> Diskussionen des Themas mit Christian Pierdzioch, Werner Pitsch, Freya Gassmann, Michael Fröhlich, Werner Pitsch und Lutz Thieme (allesamt Mitglieder im Europäischen Institut für Sozioökonomie) verdanke ich wertvolle Einsichten und Erkenntnisse.

<sup>2</sup> „*Evaluitis*“: -itis deutet immer auf eine Entzündung hin, mit den Symptomen *rubor, dolor, vapor* und *functio laesa*. Zu vermuten ist, dass der von Frey verwendete Begriff einen gewissen Zwangscharakter von Evaluationen erfassen soll, die öffentlich gefordert werden, in die man Betroffene treibt, und die an pseudosakrale Passionswege erinnern: Man wird evaluiert, leidet, verspricht Buße und handelt danach weiter so wie zuvor. Damit ist letztlich ein Ritualismus begrifflich eingefangen, in dem Evaluation zu einer Art zwangsweise stattfindenden, leerlaufenden Handlung wird.

<sup>3</sup> Im Einzelnen muss man Legitimation unterscheiden nach der Frage: Für welchen Zweck wird Legitimation benötigt? Im Zusammenhang mit Evaluation könnte man zwischen Legitimierung von Ressourcenverwendung zur Bearbeitung eines Problems und Legitimierung von Veränderungen institutioneller Rahmenbedingungen infolge politischer Interventionen unterscheiden. Im zweiten Fall wäre noch zu unterscheiden zwischen der Legitimation vor der (beabsichtigten) politischen Intervention und nach der erfolgten politischen Intervention (zur nachträglichen Begründung/Stabilisierung bereits erfolgter Veränderungen).

<sup>4</sup> Das Centrum für Hochschulentwicklung (CHE) wurde 1994 in Gütersloh von der Bertelsmann Stiftung und der Hochschulrektorenkonferenz als gemeinnützige GmbH gegründet (vgl. <http://www.che.de/cms/?getObject=5&getLang=de>, Zugriff am 19.03.2014). Die oben erwähnten IGLU und TIMSS-Studien werden vom Institut für Schulentwicklungsforschung an der Technischen Universität Dortmund (IFS) durchgeführt (s. <http://ifs-dortmund.de/>; Zugriff am 22.3.2014). TIMSS wurde dabei vom Bundesministerium für Bildung und Forschung (BMBF) sowie der Ständigen Konferenz der Kultusminister (KMK) der Länder in der Bundesrepublik Deutschland (zu gleichen Anteilen) gefördert (<http://ifs-dortmund.de/1269.html>; Zugriff am 22.3.2014).

in einem komplexen Prozess wiederholt die Fragen, Ziele, Kriterien und vor allem Gegenstände von Evaluierungen und liefern dazu mehr oder minder komplexe Sachstandsmessungen.<sup>5</sup> Die auf der politischen Ebene sowohl für eine aufmerksamkeitssteigernde Krisenrhetorik als auch als implizite Bestätigung für getroffene bildungspolitische Entscheidungen der Vergangenheit (oder in der Zukunft notwendige) verwertbaren Ergebnisse von Sachstandsmessungen ohne theoretische Begründung der ausgewählten Indikatoren führen dann zu immer neuen Evaluationsaufträgen bzw. Sachstandsmessungen, als deren Resultat schon aus mathematischen Gründen infolge der Variation der ausgewählten Indikatoren immer wieder Auf- und Absteiger sichtbar werden (zur Modellierung des damit verbundenen rasenden Stillstands vgl. Emrich & Pierdzioch, 2012). Durchgeführt werden sie auf Kosten Dritter, gewöhnlich des Steuerzahlers, und produzieren so einen stetigen Einkommensstrom für Evaluatoren bzw. Evaluationagenturen auf Kosten des Steuerzahlers.

Wie in jedem Markt kommt es abgesehen vom geräumten Markt mit seinem Gleichgewicht von Angebot und Nachfrage zu über die Zeit variierenden unterschiedlichen Angebots- und Nachfragemengen und damit zu unterschiedlichen Preisen. Das Angebot von Evaluationsleistungen und die Nachfrage danach bestimmen auch die Einkommensmöglichkeiten von Evaluationsanbietern, wobei grundsätzlich zwischen verschiedenen Typen von Anbietern zu unterscheiden ist:

1. Erstens der außeruniversitäre Anbieter von Evaluationsleistungen, der sich im klassischen Sinn unternehmerisch im Sinne angewandter Wissenschaft im Markt betätigt und dann Einkommen erzielt, wenn er Evaluationsaufträge erhält und ausführt.
2. Zweitens außeruniversitäre Anbieter, die
  - a) als Einrichtungen der Öffentlichen Hand fungieren (wie z. B. die Agentur für Qualitätssicherung, Evaluation und Selbstständigkeit von Schulen (AQS) in Rheinland Pfalz<sup>6</sup> oder das Deutsche Evaluierungsinstitut der Entwicklungszusammenarbeit (DEVal), das derzeit über 38 feste Mitarbeiterinnen und Mitarbeiter verfügt und eine gemeinnützige Gesellschaft mit beschränkter Haftung (gGmbH) ist),<sup>7</sup>

---

<sup>5</sup> So wird aktuell die Problemlösekompetenz von Schülerinnen und Schülern außerhalb der Schule untersucht, ein neues Geschäftsfeld in Erweiterung von PISA (vgl. Janker, 2014).

<sup>6</sup> In ihrer Beschreibung heißt es (<http://aqs.rlp.de/externe-evaluation/>, Zugriff am 20.März 2014): „Sie ist eine Landeseinrichtung, die auf der Grundlage des rheinland-pfälzischen Schulgesetzes regelmäßig an allen Schulen in öffentlicher Trägerschaft eine externe Evaluation durchführt. Die externe Evaluation ist Bestandteil einer umfassenden Strategie schulischer Qualitätsentwicklung mit dem Ziel, Schulen Impulse zur nachhaltigen Verbesserung des Unterrichts und der schulischen Prozesse zu geben“

<sup>7</sup> Alleiniger Gesellschafterin der DeVal ist die Bundesrepublik Deutschland, die durch das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) vertreten wird (<http://www.deval.org/de/struktur.html>, Zugriff am 20.März 2014).

b) außeruniversitäre Anbieter sind, wie z.B. das CHE, das als gemeinsame Einrichtung der Bertelsmann Stiftung und der Hochschulrektorenkonferenz eine Art Public Private Partnership darstellt oder die OECD.<sup>8</sup>

3. Drittens universitäre Anbieter, die sich im Evaluationsmarkt quasi unternehmerisch betätigen, also eine Art verbeamteter Kleinunternehmer.<sup>9</sup>

Neben inhaltlichen und methodologischen Fragen der Evaluation (vgl. z. B. Kromrey, 2003, Dessauer et al., 2014, Emrich & Pierdzioch, 2012) tauchen damit folgende Fragen auf:

- 1) Welchen konkreten institutionellen Einflüssen ist Evaluationsforschung innerhalb und außerhalb der Universität jeweils ausgesetzt?
- 2) Zu welchen spezifischen Konkurrenzverhältnissen kann es zwischen den Anbietertypen von Evaluationswissen kommen und welche Auswirkungen hat dies auf die jeweilige Wissensproduktion?
- 3) Welche strukturellen Aspekte und Interaktionsdynamiken prägen die soziale Beziehung zwischen Auftraggeber und Auftragnehmer, und zwar sowohl für Anbieter innerhalb als auch außerhalb der universitären Wissenschaft?

Zur Beantwortung dieser Fragen wird im nachfolgenden Beitrag aus wissenschaftssoziologischer und -ökonomischer Perspektive wie folgt vorgegangen:

- 1) Zunächst wird in einem kurzen Überblick die Frage der Unabhängigkeit der Evaluationsforschung kontextabhängig (universitäre versus nicht-universitäre Evaluationsforschung) diskutiert.
- 2) Dann wird mit Hilfe empirischer Befunde das Problem der Berufs- bzw. Professionsethik von in der Evaluation tätigen Forschern aufgezeigt und am reinen Ethos der Forschung (Merton 1938) gespiegelt.
- 3) Nachfolgend wird die soziale Beziehungen zwischen Auftraggeber und –nehmer mit Hilfe spieltheoretischer Überlegungen illustriert und
- 4) mit Hilfe des in der Institutionenökonomik diskutierten Konzepts unvollständiger Verträge (vgl. Richter & Furubotn, 1999) in ihrer Interaktionsdynamik und in ihren wechselnden strukturellen Ungleichgewichten diskutiert, und zwar sowohl für die außeruniversitären Leistungsanbieter als auch für die universitären

---

<sup>8</sup> Ihnen ist gemeinsam, dass sie unter zumindest partieller Aushebelung des Wettbewerbs mehr oder minder exklusiv bestimmte, gesellschaftlich positiv oder bedeutsam bewertete Gegenstände (Bildung, Entwicklungshilfe) im öffentlichen Auftrag und mit öffentlichen Mitteln evaluieren.

<sup>9</sup> Die Tätigkeit der universitären Evaluationsanbieter wird solange mit Steuermitteln, die in die Personal- und Sachkosten der Universität fließen, subventioniert, wie die neuen Overhead-Regeln der EU nicht konsequent angewendet werden. Durch Ausgründung von Forschungsinstituten als eigenständige Einheit aus dem universitären Raum und mit Mitarbeitern, die hauptberuflich an der Universität arbeiten, kann dieser Sachverhalt umgangen werden.

Evaluationsforscher.

## **2. Das Grundverständnis universitärer Forschung – das Ethos der Forschung**

In der universitären Institution Wissenschaft haben sich zahlreiche Handlungsimperative für die Tätigkeit des Wissenschaftlers herausgebildet, insbesondere ein spezifisches Wissenschaftsethos, das als Form normativer Selbstverpflichtung die Tätigkeit des Forschers steuert und das aufgrund von im System verankerten Überwachungs- und Selbstreinigungskräften im Fall seiner Verletzung durch abweichendes Verhalten wiederhergestellt wird. Das Ethos der reinen Wissenschaft im Sinne Mertons (1938) (pure ethos of science) verpflichtet universitäre Wissenschaftler zur Einhaltung bestimmter Regeln und trägt damit zur Erfüllung normativer Erwartungen bei, wodurch die Vertrauenskosten der Kollegen aus der Scientific Community, aber auch anderer Wissensnachfrager außerhalb der Universität erheblich sinken. Konkret sind hier die Verpflichtungen zu nennen, werturteilsfrei, ohne Ansehen der Person die jeweiligen Forschungsergebnisse unter Offenlegung aller methodischen Schwächen der Scientific Community nach erstinstanzlicher Prüfung durch Fachkollegen (Begutachtung) zur weiteren Beurteilung mittels Publikation zu übergeben und damit wahrheitsgetreu ohne Ansehen der Person Befunde zu berichten, nichts zu verschweigen und nichts zu manipulieren. Wird gegen diese normativen Verpflichtungen verstoßen, ist die Institution der Wissenschaft mit wirksamen Selbstheilungskräften ausgestattet, in deren Rahmen abweichendes Verhalten innerhalb des Wissenschaftssystems sanktioniert werden, indem abweichendes Verhalten mit extremen Reputationsverlusten des Wissenschaftlers bestraft wird und dadurch die Normen letztendlich stabilisiert werden. Reputation als immaterieller Faktor ist dabei die „Währung“, in der Wissenschaftler vornehmlich entlohnt werden bzw. die für sie den höchsten Anreiz darstellt.

Im Zuge außerwissenschaftlicher Impulse auf die universitäre Forschung befindet sich die Universität derzeit in einem exogen aus dem politischen System heraus befeuerten institutionellen Wandel, in dessen Kern die Etikettierung der Institution Universität als reformbedürftig steht (vgl. Emrich & Fröhlich, 2010). Von der Idee her war die Universität eine Institution, die einen funktionalen Rahmen für die Organisierbarkeit des Nicht-Organisierbaren lieferte. Unter Lehrstuhlinhabern wurde Zurückhaltung auf Gegenseitigkeit praktiziert (Plessner, 1985) und Werte und normative Selbstverpflichtungen stellten die vertrauensbasierte Grundlage für das Wirken derer dar, die sich dem Ethos der Wissenschaft verpflichtet fühlen. Ihrer Idee nach war die in der Universität organisierte Wissenschaft „Kirche der Vernunft“ und „Werthaltung und Lebensform“ (Mittelstrass, 1982) zugleich.“ Aktuell wird sie immer mehr zu einer Massenausbildungsstätte, in der der Gedanke von Bildung als Entwicklung eigener Urteilskraft am kulturellen Objekt zusehends verdrängt wird und die spezifische berufliche

Ausbildung in den Vordergrund rückt. Die Massenuniversität heutiger Tage wird im Zuge dieses Prozesses in ihrem Betrieb, also auf der Ebene der Organisation, verstärkt durch außen gesteuerte Erwartungen an ihre „betriebliche“ Leistungsfähigkeit im Sinne hoher Absolvenzzahlen gesteuert und soll vorrangig Absolventen spezialistisch auf die aktuellen Anforderungen des Arbeitsmarktes vorbereiten. Letztendlich sollen die Universitäten nicht mehr dem reinen Bildungsideal verpflichtet sein, sondern zunehmend die employability ihrer „Kunden“ herstellen.

Zahlreiche normative Spannungslinien zwischen der Idee der Universität und ihrer starken Verbetrieblichung sind die Folge des geschilderten Prozesses (Paris, 2001; dazu schon früh Weber, 1988). Der einzelne Wissenschaftler wird im Zuge dieses institutionellen Wandels intensiv auf Basis der Summe eingeworbener Drittmittel bewertet - Evaluationsprojekte sind sehr häufig drittmittelfinanziert - und so in eine scharfe drittmittel- und impactgesteuerte Konkurrenz getrieben, die aufgrund der politisch gesteuerten universitären Ressourcenknappheit hohen Konkurrenzdruck erzeugt und den Charakter eines Rattenrennens (vgl. zum Begriff Akerlof, 1976) annimmt. Gleichzeitig wird auf organisatorischer Ebene die Universität zunehmend bürokratisch und betriebswirtschaftlich nach den durch häufige Evaluationen befeuerten Ergebnissen dieses Konkurrenzkampfes gesteuert, und zwar mit Hilfe jener Ressourcen, die man nicht mehr direkt ins System gibt, sondern programm- und antragsgesteuert durch staatliche (z. B. Ministerien), intermediäre (wie die staatlich finanzierte und durch die Wissenschaft verwaltete DFG) und nicht staatliche Forschungsfördereinrichtungen (wie Stiftungen) (vgl. zu den Wirkungen des CHE-Rankings auf eine Wissenschaftsdisziplin Dessauer et al., 2014; s. dazu das PARK-Modell von Emrich & Pierdzioch, 2012) vergibt.

Der dem Wissenschaftsethos und seinen Werten verpflichtete und sich an den daraus abgeleiteten Normen orientierende Wissenschaftler ist durch diese Außensteuerung bei zunehmender Ressourcenknappheit im System in seiner Unabhängigkeit gefährdet. Er wird durch die universitäre Außensteuerung gelockt, letztlich auch gedrängt, an diesem Ressourcen verschwendenden Rattenrennen (Akerlof, 1976) um Drittmittel teilzunehmen. Jede Verweigerung der Teilnahme würde als mangelnde Leistungsfähigkeit interpretiert und die Einhaltung der Regeln des organisierten Skeptizismus würde ihn im Konkurrenzkampf mit weniger stark normativ gebundenen Wissenschaftlern innerhalb und außerhalb der Wissenschaft behindern.

Der Wissenschaftler im Sinne der Idee der Universität verpflichtete sich selbst zum reinen Ethos der Wissenschaft und seine Forschungsmittel wurden in der Universität als eine Art „Staatsbetrieb“ bereitgestellt. Die politisch vorangetriebenen Veränderungen der Rahmenbedingungen begünstigen und legitimieren nun eine Entwicklung hin zu einem Akteur, der sich

in einem zunehmend nach Angebot und Nachfrage gesteuerten Wissensmarkt bewähren muss, wenn er Reputationsverluste vermeiden, seine Arbeitsmittel erhalten und teilweise sein infolge der Besoldungsreform in Teilen leistungsabhängiges, an Drittmitteln und Publikation ausgerichtetes Einkommen behalten oder ausweiten will. Zwar wird auf institutioneller Ebene im politischen Talk (Brunsson, 2003) nach wie vor eine Rhetorik werturteilsfreier Wissenschaft gepflegt, faktisch aber werden durch die Außensteuerung der universitären Wissenschaft Handlungsanreize wirksam, welche die Einhaltung normativer Standards in der Forschung erschweren und den reinen Wissenschaftler in einen Wettbewerb mit außengeleiteten Individuen treiben. Der Wissenschaftlertypus, der sich außengeleitet insbesondere jener Themen annimmt, die gerade im Wissensmarkt nachgefragt sind, hat dabei Konkurrenzvorteile, weil er auf solche äußeren Handlungsanreize rational reagiert und seinen Nutzen maximiert. Dies muss nicht zwingend heißen, dass dieser Typus die Standards des reinen Ethos der Forschung verletzt. Die Chance steigt jedoch, dass er sich im Wissensmarkt nach Angebot und Nachfrage richtet, kurzum Mainstreaming betreibt, und nicht exklusiv oder vorwiegend für ihn wissenschaftlich relevante Fragen stellt, sofern bzw. weil diese im Rennen um Drittmittel Nachteile brächten. Wenn in diesem Zusammenhang ein Thema „Konjunktur“ hat, kommt es zwangsläufig zu einem Missverhältnis von wissenschaftlichen Fragen und Antworten. Dies wird wiederum sichtbar in einer Homogenisierung von wissenschaftlich bearbeiteten Themen, da vorrangig solche beforscht werden, die in höherem Maße drittmittelfähig sind und für die beim Forscher bereits im Markt bewährte Instrumente und Vorwissen vorliegen. Daraus resultiert auch das Phänomen, dass vielfach Antworten gegeben werden, für die gar keine Fragen gestellt worden waren oder dass Evaluationsforschung zu Phänomenen betrieben wird, deren Existenz gar nicht nachgewiesen ist (vgl. zu Evaluationen ohne nachgewiesenes Explanandum Klein & Emrich, 2013).

Aufgrund der Außensteuerung der Wissenschaft und der zunehmend ressourcen- und impactabhängigen Bewertung der individuellen Leistung wird somit der an der Idee der Universität und ihren institutionell verankerten Verpflichtungen orientierte Forscher nicht nur in seiner Existenz bedroht. Vielmehr kann er in seinem Handeln aufgrund der geänderten äußeren Erwartungen, sofern er ihnen genügen kann und will, gar nicht mehr von jenen unterschieden werden, die sich von jeher stärker an äußeren Erwartungen ausrichteten bis hin zu jenen, die sich von Anfang der Universität nur als schützender Hülle bedienten, um Eigeninteressen zu verfolgen (vgl. Paris, 2001; Emrich & Thieme, 2012).

Die Überlebenszwänge im universitären System führen also zu einer Homogenisierung von Verhalten und Forschungsfragen, bringen jedoch für einen Teil der Beteiligten normative

Spannungslinien, die um so stärker werden dürften je intensiver das reine Ethos der Forschung und das Bild selbstbestimmter Forschung ausgeprägt sind.<sup>10</sup> Forschungsergebnisse werden im Rahmen dieses Prozesses externer Steuerung zunehmend nicht mehr wissenschaftsintern diskutiert, sondern verlagern sich in Richtung Kommunikations- und Publikumsforschung. Sie werden zu medial angeheizten Debatten genutzt, deren einfache Talkshow geprägte Pro- und Con-Diskussionen anfangs Aufmerksamkeitssteigerungen produzieren, insbesondere wenn sie von allgemein als bedrohlich empfundenen Krisenbehauptungen begleitet werden. Gleichzeitig nutzen auch Wissenschaftler die Medien, um auf sich und die jeweiligen Forschungsleistungen aufmerksam zu machen und so die Chancen im Kampf um Drittmittel zu erhöhen bzw. Nachteile gegenüber den in die Öffentlichkeit drängenden Kollegen im Kampf um knappe Ressourcen zu vermeiden. Damit wechselt die Diskussion wissenschaftlicher Ergebnisse zumindest teilweise von der Wissenschaft in den öffentlichen Raum. Zudem findet die Bewertung von Qualität und Wichtigkeit der Forschung nun neben der Scientific Community auch in den Medien statt, was sich dann wiederum auf das System der Wissenschaft auswirkt, indem Wissenschaftler mit geringer medialer Präsenz und geringer Prominenz gleichzeitig Nachteile im Konkurrenzkampf um Drittmittel insbesondere aus dem außeruniversitären Bereich haben. Die medialen Debatten werden zusätzlich dadurch angeheizt, dass sich auch die Intention von Forschung wandelt. Sie soll jetzt Entscheidungen öffentlich legitimieren, statt im aufklärenden Sinn für eine politische Entscheidung verlässliche Informationen zu liefern. Im Zuge der zunehmenden Außensteuerung der Wissenschaft sind also gleichzeitig eine zunehmende Instrumentalisierung der Medien durch die Wissenschaft zur Austragung von wissenschaftlichen Prioritätskonflikten und eine Instrumentalisierung der Wissenschaft durch die Medien zur Aufmerksamkeitssteigerung zu beobachten.<sup>11</sup> Das eigentlich zu diskutierende wissenschaftliche Problem kann in den angeheizten medialen Debatten von Laien nur durch Glauben entscheiden werden, wodurch die Wissenschaft allmählich

---

<sup>10</sup> Beide Typen zeigen Überschneidungen mit dem Konzept des vorrangig an Werten und Normen orientierten Homo sociologicus und dem einkommensmaximierenden, stärker äußeren Anreizen folgenden Homo oeconomicus. Allerdings kann auch ein Homo oeconomicus, der mit aller Intensität redlich forscht, um einen herausragenden Preis zu gewinnen und besondere Reputation durch herausragende Leistungen zu erlangen, der Verletzung des reinen Ethos der Forschung die kalte Schulter zeigen, solange er begründet hoffen kann, dieses Ziel zu erreichen. Insofern haben wir es schlicht mit normkonformen und abweichendem Verhalten zu tun, mit einer wahrscheinlich höheren Wahrscheinlichkeit im Mittel für abweichendes Verhalten beim Homo Oeconomicus (den Hinweis verdanke ich Christian Pierdziach, Helmut Schmidt Universität Hamburg). Gleichzeitig ist zu vermuten, dass Junior-Professoren durch ihre Sozialisation eher dem Pol des außengesteuerten Forschers zuneigen, da sie ja häufig gar keine Universität mehr kennen, die eine Organisationsform bietet, die freie Forschung nach eigenem Antrieb als Normalfall akzeptiert.

<sup>11</sup> Mittel zur Aufmerksamkeitssteigerung sind in das öffentliche Bewusstsein gerückte Krisen mit bedrohlichem Charakter (Weingart 2001; schon früh Kurucz 1986; zu moralischem Unternehmertum Klein & Emrich 2013).

entwertet wird (s. die Diskussionen um das Waldsterben, das Ozon-Loch, den Rinderwahnsinn usw.). Letztlich könnte es sogar zu Spezialisierungseffekten im Markt der Evaluation bzw. Politikberatung auf Basis von Evaluationen kommen (siehe am Beispiel der Politikberatung den Fall des Ökonomenstreits im Zusammenhang mit der Euro-Krise Haucap & Mödl, 2013).

### **3. Normative Selbstbindungskräfte für Evaluationsforscher**

Ein Blick auf die Internetseiten der Deutschen Forschungsgemeinschaft fördert die normativen Forderungen gegenüber einer angemessenen Evaluation zutage ([Grundhttp://www.dfg.de/dfg\\_profil/foerderatlas\\_evaluation\\_statistik/programm\\_evaluation/evaluationsstandards/](http://www.dfg.de/dfg_profil/foerderatlas_evaluation_statistik/programm_evaluation/evaluationsstandards/), Zugriff am 16. März 2014). Dort heißt es: „Evaluationsstudien werden durch qualifizierte Dritte, unabhängig und objektiv, durchgeführt. Die Auftragnehmer nehmen die Expertise der DFG in Bezug auf ihre Programme auf, um die Qualität und Richtigkeit der Berichte sicherzustellen, sind aber bei der Formulierung ihrer Ergebnisse und Schlussfolgerungen vollständig frei.“ Gleichzeitig stellen Evaluationsstudien „für die DFG ein wichtiges Instrument dar, um die Dynamiken und wechselnden Prioritäten in der Forschungslandschaft zu erkennen.“ Forschungsnehmer evaluieren demnach als qualifizierte Dritte unabhängig und objektiv die Wirkung der DFG-Programme, was angesichts der Veränderungen in der universitären Forschungslandschaft fast zwangsläufig zu Interessenkonflikten führen muss. Die DFG lässt also jene Programmwirkungen evaluieren, die sie selbst initiiert hat, und zwar von Forschern, die prinzipiell an der Konkurrenz um die über die Programme bereitgestellten Mittel teilnehmen können bzw. schon an anderen Programmen der DFG teilgenommen haben, was die Frage der Unabhängigkeit der Auftragnehmer aufkommen lässt (vgl. Stockmann, 2000, S. 19, zu „Standards for Evaluation“ vom Joint Committee on Standards for Educational Evaluation, vgl. auch Joint Committee, 2000; vgl. Deutsche Gesellschaft für Evaluation [DeGEval], 2004, S. 5).

Stockmann, Meyer und Schenke (2011) haben sich generell mit den normativen Erwartungen von Evaluatoren an ihre eigene Tätigkeit beschäftigt. Die abgefragten Wertvorstellungen waren auf einer Likert-Skala entsprechend der Ausprägungen 1 (stimme überhaupt nicht zu) bis 6 (stimme voll zu) zu beantworten und repräsentieren im Sinne von Durkheim (1973) Elemente einer beruflichen Sondermoral, die aufgrund ihrer normativen Selbstverpflichtungsstruktur als Substitut für die Bindung an das reine Ethos der Forschung dienen könnte. Die Ergebnisse (s. Tabelle 1) sind insofern interessant, als (sicherlich verstärkt durch Effekte sozialer Erwünschtheit) die „hauptsächliche Verpflichtung gegenüber den Evaluationsstan-

dards“ mit dem Mittelwert von 5,1 zwar hohe, aber eben keine völlige Zustimmung erreicht. Mit absteigender Zustimmung folgen die „Darstellung der tatsächlich ermittelten Ergebnisse (4,7), „keine Beeinflussung in der methodischen Vorgehensweise“ (4,3) und „schonungslose Offenlegung von Schwachstellen“ (4,2) sowie „Verantwortung gegenüber den Stakeholdern ist prioritär“ (3,0). Die geringste Zustimmung, aber keine völlige Ablehnung, erzielt das Item „Ausrichtung der Evaluationsergebnisse nach den Interessen der Auftraggeber“ (2,0; s. dazu Tabelle 1). Dies zeigt, dass die von der DeGeval formulierten Standards einer guten Evaluation, nämlich Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit keineswegs vollständig im Sinne einer spezifischen Berufsmoral und damit als Richtschnur für das berufliche Handeln von Evaluatoren verinnerlicht wurden (zu den Standards der DeGeval s. <http://www.degeval.de/degeval-standards/standards/>, Zugriff am 16. März 2014).

Zu vermuten ist sogar, dass die Evaluatoren einfach mit der Zeit ihre Standards im Sinne des Grenzmoralprinzips (Briefs, 1957) an den Markt angepasst haben. Bedenkt man dabei, dass die Wirkungen sozialer Erwünschtheit eigentlich eine spezifische Darstellung der Befragten in Richtung ausgeprägter Wertvorstellungen erwarten lässt, wird die in den Ergebnissen sichtbar werdende Anpassung an Marktstandards durchaus plausibel.<sup>12</sup> Eine berufliche Sondermoral für Evaluatoren, die adäquat zum reinen Ethos der Wissenschaft im Sinne Mertons (1938) in normativer Selbstverpflichtung ein vergleichbares Maß an Eigenkontrolle fordert und das auch überwacht würde, zeigt sich also (noch) nicht.

Analog zur oben geschilderten Wirkung der zunehmenden Außensteuerung der Universität können wir auch eine Homogenisierung von Verhalten vermuten. Je mehr sich in der Universität beheimatete Evaluatoren an den oben geschilderten außengesteuerten Veränderungen der Bewertung ihrer universitären Tätigkeit orientieren, um so mehr nähern sich die beiden Typen von Evaluatoren an, indem für am freien Markt tätige und aus der Universität heraus operierende die äußeren Anreize gleichermaßen wichtig sind bzw. werden.

---

<sup>12</sup> Die höheren Werte der Standardabweichung bei den Items „schonungslose Offenlegung von Schwachstellen“ und „Moralische Verantwortung gegenüber den Stakeholdern ist prioritär“ zeigt dabei eine gewisse Spreizung bei den Antworten und damit Uneinheitlichkeit im Antwortverhalten.

<b>Tabelle 1: Wertvorstellungen zum Berufsbild des Evaluators</b> (Skala von 1 =stimme überhaupt nicht zu bis 6 =stimme voll zu; SD=Standardabweichung; n=Gesamtzahl valider Antworten, Stockmann, Meyer & Schenke, 2011)				
<b>Item</b>	<b>MW</b>	<b>SD</b>	<b>N</b>	<b>Weiß nicht</b>
Hauptsächliche Verpflichtung gegenüber den Evaluationsstandards	5,1	1,0	177	1
Darstellung der tatsächlich ermittelten Ergebnisse	4,7	1,3	175	1
Keine Beeinflussung in der methodischen Vorgehensweise	4,3	1,4	175	3
Schonungslose Offenlegung von Schwachstellen	4,2	1,5	176	2
Moralische Verantwortung gegenüber den Stakeholdern ist prioritär	3,3	1,6	158	15
Methodische Orientierung an den Interessen der Stakeholder	3,0	1,4	168	6
Abschwächung der negativen Evaluationsergebnisse gegenüber dem Auftraggeber	2,2	1,2	173	3
Ausrichtung der Evaluationsergebnisse nach den Interessen der Auftraggeber	2,0	1,3	176	3

#### **4. Die soziale Beziehung zwischen Auftraggeber und Auftragnehmer in der Evaluationsforschung**

Soziales Handeln wird von Weber (1980) als subjektiv sinnhaft auf das Handeln anderer bezogen definiert. Insofern gehören auch Dulden und Unterlassen zu zentralen Formen sozialen Handelns. Insbesondere Dulden und Unterlassen können Merkmale höchster Sozialkompetenz sein. Sie sind z.B. wesentlicher Bestandteil des sozialen Taktes, wenn man mit einer gewissen Zartheit im Umgang miteinander sein Gegenüber sich so darstellen lässt, wie es sich wohl ursprünglich darstellen wollte und deshalb von bestimmten Handlungen gerade absieht. Aus dem sinnhaft subjektiv aufeinander bezogenen sozialen Handeln von mindestens zwei Personen kann sich abhängig von der Interaktionsdichte aufeinander bezogener sozialer Handlungen eine soziale Beziehung entwickeln. Diese soll nach Webers (1980, § 3) Nominaldefinition „ein seinem Sinngehalt nach aufeinander gegenseitig eingestelltes und dadurch orientiertes Sichverhalten mehrerer heißen. Die soziale Beziehung besteht also durchaus und ganz ausschließlich: in der Chance, dass in einer (sinnhaft) angebbaren Art sozial gehandelt wird, einerlei zunächst: worauf diese Chance beruht“. (Weber 1980, § 3). Die soziale Beziehung kann soziologisch dabei durch Vergemeinschaftung und/oder Vergesellschaftung als grundlegende Handlungsorientierung der Interaktionspartner gekennzeichnet sein, wobei das für die Vergesellschaftung typische zweckrationale Handeln

durch seine soziale Einbettung in Vergemeinschaftungsprozesse erleichtert werden kann (vgl. Emrich, 2006).

Um das soziale Handeln von Auftraggeber und –nehmer in seiner zweckrationalen Interessengeleitetheit bzw. Interaktionsdynamik zu analysieren, nutzen wir neben den analytischen Begriffen der Soziologie spieltheoretische Konzepte und die neuere Institutionenökonomik (vgl. Richter & Furubotn, 1999). Die Soziologie betont den normativen Kontext der Interaktion, die Spieltheorie stellt den Akteur und die Rationalität seiner Entscheidungen in den Vordergrund. Die Institutionenökonomik erweitert den neoklassischen Zugang eines unter vollständiger Information und Knappheitsbedingungen rational entscheidenden Individuums und geht davon aus, dass Menschen auch bei unvollständiger Information und unter Zeitknappheit rational handeln müssen.<sup>13</sup> Im Sinne rationaler Nutzenmaximierer streben nun zwei potentielle Vertragspartner danach, ihren jeweiligen Nutzen auf Kosten des anderen zu maximieren. Sie nutzen hierzu Informationsasymmetrien, die dadurch entstehen, dass jeder Vertragspartner über sich selbst und seine konkreten Absichten mehr weiß als über den jeweils anderen. Je nachdem, ob man sich zeitlich in der Phase vor oder nach einem Vertragsabschluss befindet, ergeben sich zwangsläufig wechselnd asymmetrische Informationsverteilungen zwischen den Vertragspartnern, also Auftraggeber (Prinzipal) und -nehmer (Agent). Der Auftraggeber kann nach Abschluss eines Vertrages über eine Evaluation während der Vertragslaufzeit das Handeln des Auftragnehmers nicht vollständig beobachten und er hat darüber hinaus gewöhnlich auch weniger Zugang zu relevanten Informationen als der Auftragnehmer (Evaluator), der diese zudem selektiv nutzen kann. Im Fall des versteckten Handelns ist es für den Auftraggeber unmöglich oder zu kostspielig, den Agenten dauerhaft zu überwachen, bei versteckter Information ist es für den Auftraggeber unmöglich oder zu kostspielig, Informationen über den Agenten zu erhalten. Insofern werden infolge erprobter sozialer Beziehungen gewachsene Reputation und Vertrauen zu wichtigen Faktoren, die Überwachungskosten senken können. Auch im Fall öffentlicher Ausschreibungen werden diese Mechanismen nicht grundsätzlich außer Kraft gesetzt, es kommt lediglich zu erweiterten Principal Agent-Beziehungen, die im

---

<sup>13</sup> Institutionen wie die Universität stehen ebenso wie etwa Evaluationsagenturen im freien Markt formelle (objektives Recht) oder informelle Normen (subjektives Recht) zur Steuerung individuellen Verhaltens zur Verfügung. Da Gesetze nicht alle Fälle der Zusammenarbeit zwischen mindestens zwei Personen antizipieren können, werden subjektive Rechtsnormen benötigt, die in Form freiwillig geschlossener Verträge Verfügungsrechte sichern. Zentral für die Institutionenökonomik ist in diesem Zusammenhang die Theorie unvollständiger Verträge, wobei unter Vertrag „im ökonomischen Sinne [...] jede bindende explizite oder implizite Vereinbarung über den Austausch von Gütern oder Leistungen zwischen Menschen, die dieser Vereinbarung zustimmen, weil sie sich davon eine Besserstellung versprechen [...]“, verstanden wird (Richter & Furubotn 1999, S. 156 f.). Verträge sowohl zwischen Personen als auch korporativen Akteuren fixieren, welche Rechte und Pflichten sich die einzelnen Mitglieder einer durch den jeweiligen Vertrag bestimmten sozialen Einheit wechselseitig zugestehen bzw. zumuten.

Sinne informeller Kartelle Entscheidungsträger im Bereich der Öffentlichen Hand mit einbeziehen.

<b>Tabelle 2: Erfahrungen mit verschiedenen Beeinflussungsformen</b> (Skala von 1 =habe ich noch nie erlebt bis 6 =habe ich schon sehr oft erlebt; n=Gesamtzahl valider Fälle; AG=Auftraggeber, Stockmann, Meyer & Schenke, 2011)							
<b>Item</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>N</b>
AG verlangte Umformulierung einzelner Sätze	12	13	16	21	20	17	168
AG hat verlangt, dass ich die Ergebnisse positiver darstelle	25	21	12	21	15	7	169
AG verlangte anderes methodisches Vorgehen	28	24	20	13	13	3	168
AG verlangte andere Schlussfolgerungen aus den Ergebnissen	37	23	14	11	13	3	168
AG verlangte inhaltliche Verzerrungen	45	22	12	12	7	2	170
AG verlangte andere Datenquellen	46	22	13	12	5	2	165
AG verlangte negativere Darstellung der Ergebnisse	74	17	5	3	1	1	168

Tabelle 2 illustriert mit empirischen Befunden die in diesem Kontext grundsätzlich vorkommenden Beeinflussungsformen, in denen sowohl Asymmetrien vor Vertragsabschluss (etwa Auftraggeber verlangt anderes methodisches Vorgehen) als auch nach Vertragsabschluss (Auftraggeber verlangt andere Schlussfolgerungen) deutlich werden.<sup>14</sup>

## 5. Der Evaluator als Unternehmer im Wissensmarkt

Weber (1980) hat nachdrücklich darauf hingewiesen, wie innerhalb eines von ihm als Vergesellschaftung bezeichneten Prozesses, für den der geldgebundene Austausch von Leistungen dominant wird, sittlich-normative Selbstverpflichtungen durch formales Recht ersetzt werden. Wenn jetzt Evaluation als wissenschaftliche Leistung gegen Geld gehandelt wird, gleichzeitig aber für die damit verbundene Forschungstätigkeit die normativen Erwartungen des reinen Ethos der Forschung bindend sein sollen, dann stellt sich um so intensiver die Frage, wie der universitäre Wissenschaftler, der Evaluationsforschung betreibt, richtig handelt, nämlich ob er Agent seines Auftraggebers ist oder ob er nur der Wissenschaft und der Wahrheit verantwortlich ist, wobei sich beides nicht grundsätzlich ausschließen muss. Entscheidend dürften in diesem Zusammenhang die Erwartungen des Auftraggebers sein, der spezifische Erwartungen an die Erkenntnis- und Legitimationsfunktionen (vgl. Stockmann, 2000, vgl. Stockmann &

<sup>14</sup> Hier wäre ein Feld für Befragungen mit indirekten Fragetechniken, mit deren Hilfe wohl verlässlicher der Anteil der Merkmalsträger im Kollektiv ermittelt werden könnte, die schon einmal Beeinflussungsversuchen nachgegeben haben.

Meyer, 2014) der zu beauftragenden Evaluation hat. Je nach Erwartung des Auftraggebers sollen Ergebnisse entweder zur Stabilisierung oder auch zur Aufdeckung unangenehmer Fakten (Aufklärungsfunktion) dienen, insofern können Erkenntnis- und Legitimationsfunktion in Widerspruch treten, wenn falsche oder gelenkte Erkenntnisse für die Legitimation eingesetzt werden bzw. werden sollen.<sup>15</sup> Denkbar ist, dass es darüber hinaus auch spezifische Erwartungen an die von Stockmann diskutierte Kontroll- und Lernfunktion der Evaluationen gibt. Wenn in jeder Evaluation entweder vorrangig die Erkenntnisfunktion, die Kontrollfunktion, die Lernfunktion oder die Legitimationsfunktion bzw. bestimmte Kombinationen dieser Funktionen dominieren, ergeben sich auch spezifische Erwartungen seitens des Auftraggebers. Wenn es diesem tatsächlich dominant um Erkenntnis oder organisationales Lernen geht, dann hat der Auftraggeber ein Interesse an den tatsächlichen Ergebnissen, also an ehrlicher Forschung. Wenn es dem Auftraggeber um Legitimation beabsichtigten oder bereits erfolgten Handelns geht, dann ist er an affirmativen Erkenntnissen interessiert. Geht es um die Kontrollfunktion (muss also kein Handeln legitimiert werden), dann sind die Ergebnisse dem Auftraggeber wahrscheinlich egal, weil er gegenüber dem zu erfassenden Sachstand neutral ist. Allerdings haben sowohl Auftraggeber als auch Auftragnehmer ein Interesse an der Wiederholung dieser Sachstandsmessung. Beiden fehlen dabei die Anreize zur Verwendung einer Messmethode, die sich theoretisch eingebettet an wissenschaftlichen Kriterien orientiert und die Indikatorenauswahl theoretisch begründet, weil die Adressaten der Ergebnisse nur in wenigen Fällen in der Wissenschaft sind, sondern gewöhnlich die Öffentlichkeit/Medien, die Politik o. ä.<sup>16</sup>

Nachfolgend wird vorrangig die Legitimationsfunktion als zentrale Funktion betrachtet. In einer einfachen Vierfeldertafel für Interaktionen zwischen Auftraggeber und –nehmer (Spiele mit zwei Spielern, jeweils zwei Strategien, wobei das Handeln des Anderen nicht direkt beobachtet werden kann) gehen wir davon aus, dass man auf beiden Seiten entweder die Verhaltensstrategie ehrliche Forschung oder gelenkte Ergebnisse wählen kann (s. Tabelle 3).

Wir nehmen dabei für alle Spiele mit 2 Strategien  $|S1|=|S2|=2$  und 2 Spielern die in Tabelle 3 dargestellte Matrix an. Dabei dürfte der marktorientierte Forscher hoch flexibel sein, während der innerlich dem Ethos der Forschung verpflichtete nur unter extremer Not (Hochkostensitu-

---

<sup>15</sup> Stockmann (2000) listet vier Funktionen von Evaluation auf: die Gewinnung von Erkenntnissen (Erkenntnisfunktion), die Ausübung von Kontrolle (Kontrollfunktion), die Schaffung von Transparenz (Lernfunktion), um einen Dialog zu ermöglichen und die Dokumentation des Erfolgs (Legitimationsfunktion).

<sup>16</sup> .Dass Qualität an sich nicht messbar ist, sei hier nur am Rande erwähnt.

ation) bestimmte Varianten der Unehrlichkeit präferieren bzw. akzeptieren dürfte, wenn überhaupt.<sup>17</sup>

Für den Wissenschaftler, der sich am reinen Ethos der Forschung in normativer Selbstverpflichtung orientiert und/oder Reputationsverlust bei entdeckter Zuwiderhandlung fürchtet, gilt:

- 1) Er präferiert stets ehrliche Ergebnisse, für ihn gilt  $a > b$  und  $a' > b'$ .
- 2) Unehrliche Forschung betreibt er grundsätzlich nicht, also  $c < b$  und  $c' < b'$
- 3) Spielt er ehrlich und trifft auf einen unehrlichen Auftraggeber, den er nicht überzeugen kann, bricht er den Auftrag ab, also  $a > d$  und  $a' > d'$ .

Für den ehrlichen Auftraggeber, der reine Erkenntnis will, gilt:

- 1) Er präferiert stets ehrliche Ergebnisse ( $a > c$  und  $a > c'$ ).
- 2) Liefert ein Auftragnehmer affirmative Forschung, bricht er ab ( $c < a$  und  $c' < a'$ ).

Treffen ehrliche Wissenschaftler und ehrliche Auftraggeber aufeinander, haben beide einen Anreiz, E/E zu wählen. Der eine erhält ehrlich Ergebnisse, die er konstruktiv im Sinne der Erkenntnisfunktion von Evaluation einsetzen kann, der andere hält sich an die Regeln der Wissenschaft und wird für diese Leistung bezahlt, was sich als Drittmittelwerbung reputations- und einkommenssteigernd für ihn auswirkt. Für den völlig affirmativen Wissenschaftler in der Universität und außerhalb, der das reine Ethos der Forschung vielleicht noch kennt, es aber für völlig irrelevant hält und der sich am „Verkauf“ seiner Dienstleistung im Wissensmarkt ausrichtet, für den der Kunde immer König ist, und der deshalb situativ entscheidet, gilt:

- 1) Er ist völlig indifferent. Für ihn gilt  $a = b$  und  $b = b'$ , er liefert, was gekauft wird.
- 2) Er zieht  $a = b$  und  $a' = b'$  gegenüber  $(c, c')$  und  $(d, d')$  vor.
- 3) Spielt er ehrlich und trifft auf einen unehrlichen Spieler, bricht er das Spiel nicht ab, sondern liefert affirmative Ergebnisse, da dann gilt  $a < d$  und  $a' < d'$  bzw.  $(b, b') > (a, a')$ .
- 4) Spielt er unehrlich und trifft auf einen Spieler, der ehrliche Forschung will, liefert er situativ angepasste affirmative Ergebnisse, also in diesem Fall ehrliche Befunde, da dann  
gilt  
 $d < a$  und  $d' < a'$  bzw.  $(a, a') > (b, b')$ .

Für den unehrlichen Auftraggeber, der die gelenkte Erkenntnisfunktion zwecks Legitimation braucht, gilt:

- 1) Er präferiert gelenkte Ergebnisse gegenüber un gelenkten ( $d > a$  und  $d' > a'$ ).

---

<sup>17</sup> Der Markt für Gutachten ist hier nicht unser Thema, obgleich ähnliche Mechanismen wie in der Evaluationsforschung zu vermuten sind (vgl. <http://www.aerzteblatt.de/archiv/154014/Gerichtsgutachten-Oft-wird-die-Tendenz-vorgegeben>, Zugriff am 31.3.2014).

- 2) Trifft er auf einen Auftragnehmer, der konsequent ehrliche Forschung präferiert, bricht er ab, da  $d > a$  und  $d' > a$ .

Treffen unehrlicher Wissenschaftler und unehrlicher Auftraggeber aufeinander, haben beide einen Anreiz, V/V zu wählen, der eine erhält unehrliche Ergebnisse, die er konstruktiv im Sinne der Legitimationsfunktion von gelenkter Evaluation einsetzen kann, der andere wird für diese Leistung bezahlt. Ein Gleichgewicht in reinen Strategien liegt übrigens dann vor, wenn die Initiative beim Auftraggeber liegt (Fall V/E) und für diesen die Präferenz  $d > b$  und  $d' > a'$  gilt. Der Auftraggeber, der gelenkte Ergebnisse will, präferiert also eher eine Beauftragung für gelenkte Forschung bei einem bisher ehrlichen Forscher, da dann die gelenkten Ergebnisse zwangsläufig für ihn höheren Wert haben. Voraussetzung hierfür wäre jedoch auf Seite des Auftraggebers das strategische Zurückhalten und Lenken von Informationen bei einem ehrlichen Auftragnehmer, um diesem die Fiktion der Ehrlichkeit zu belassen, aber gleichzeitig das Ergebnis zu beeinflussen. Wird im Verlauf der Evaluation deutlich, dass die Gefahr unerwünschter Ergebnisse besteht, kann man die Evaluation immer noch abbrechen und den Auftrag zurückziehen. Liegt die Initiative beim Auftragnehmer, liegt ein solches Gleichgewicht vor, wenn für diesen  $c' > b'$  und  $b > a$  gilt. Der Auftragnehmer ist also offensichtlich eher bereit, gelenkte Forschung bei einem unehrlichen Auftraggeber zu liefern. Dies lässt ihm für die Zukunft noch die Chance zur Fortsetzung ehrlicher Forschung bei ehrlichem Auftraggeber.

Tabelle 3: Präferenzen von Auftraggeber und Auftragnehmer in der Evaluationsforschung

Spieler B (Auftraggeber)	Spieler A (Auftragnehmer)/ E / ehrliche Forschung	V / gelenkte Ergebnisse
E/ ehrlicher Forschung	a a'	c c'
V/ gelenkte Ergebnisse	d d'	b b'

Faktisch ist hier zu bedenken, dass sich in der Forschungspraxis wohl keine Extrema, sondern meist graduelle Näherungen ergeben, die sich eben im vereinzelt Verschweigen einzelner Fakten (Verdecken kritischer neuralgischer Punkte durch Offenlegen vereinzelter Schwachpunkte) und vereinzelt Bestätigen gewünschter Ergebnisse zeigen dürften, also eine

schwierige Grenznutzenabwägung zwischen Inszenieren als kritischer Wissenschaftler und gezieltem Bestätigen einzelner kritischer Punkte.

Die in Tabelle 4 exemplarisch gezeigten Präferenzen mit kardinalen Auszahlungswerten wurden für einen Forscher gewählt, für den der „Kunde König ist“ und der „auf Bestellung“ affirmative Befunde liefert. Das Risiko beim „Liefen auf Bestellung“ wird hier für den Auftragnehmer mit fast gegen null gehend angenommen, denn die Befunde werden keiner Scientific Community vorgestellt, durchlaufen keinen Begutachtungsprozess und werden gewöhnlich nicht publiziert, zudem sind sie Eigentum des Auftraggebers, der sie in seinem Sinn strategisch nutzt und häufig gestützt durch den abgeschlossenen Vertrag Anrecht auf mehrere Besprechungsrunden hat, in denen das Ergebnis diskutiert wird. Deutlich wird, dass unehrliche Auftraggeber und unehrliche Auftragnehmer weiterhin V/V und ehrliche Auftraggeber und –nehmer weiterhin E/E präferieren. Zudem ist davon auszugehen, dass beim „Liefen auf Bestellung“ höhere Preise erzielt werden als beim ehrlichen Evaluieren, deshalb wird der Wert von V/V mit 3, der von E/E mit 2 angesetzt, da ehrliche Auftraggeber und –nehmer niedrigere Preise vereinbaren dürften, da sie beide kein Entdeckungs- oder Reputationsrisiko einpreisen müssen.

Tabelle 4: Auszahlungsmatrix für einen affirmativen Evaluationsforscher ohne hohes Entdeckungsrisiko bei einem Auftraggeber, der gelenkte Ergebnisse will.

Spieler B (Auftraggeber)	Spieler A (Auftragnehmer)/ E / ehrliche Forschung	V/ in Ergebnissen teilweise oder ganz gelenkte For- schung
E/ ehrliche Forschung	2	3
V/ in Ergebnissen teilweise oder ganz gelenkte Forschung	0	3
	3	3

Hier kann man klar ein Spiel mit gemischten Strategien erkennen. Es handelt sich also um ein Koordinationsspiel, da die Interessen zwischen dem affirmativen Wissenschaftler und dem unehrlichen Auftraggeber übereinstimmen und nicht gegenläufig sind wie etwa in einem Gefangenendilemma. Für den affirmativen Auftragnehmer gilt  $(c, c') > (a, a')$ , der affirmative Auftragnehmer wird sich grundsätzlich nach dem unehrlichen Auftraggeber richten und wird

das tun, was der Auftraggeber möchte. Also kann  $(c,c')=(b,b')$  nicht sein und es muss gelten  $(b,b') >$  alle anderen Alternativen.

In der Realität kommt es nun wie bereits angesprochen zu wiederholten Verhandlungssituationen zwischen Auftraggeber und -nehmer (z. B. mehrfach jährlich), wie wir sie oben beschrieben haben. Dabei gehen wir davon aus, das gemäß dem Tit for Tat-Prinzip Spieler 1 in Periode  $(t + 1)$  kooperiert, wenn sein Gegenüber vorher in  $t$  kooperiert hat. Weicht sein Gegenüber in  $t$  ab, weicht Spieler 1 in  $(t + 1)$  ab. Der Spieler, der die Tit for-Tat Strategie anwendet, macht also immer dasselbe wie sein Gegenspieler, allerdings eine Runde verzögert. Das heißt bezogen auf unseren affirmativen Forscher folgt auf V des Auftraggebers in der nächsten Runde V des Auftragnehmers und auf E des Auftraggebers folgt in der nächsten Runde E des Auftragnehmers. Uns interessiert nun, unter welchen Bedingungen der Mechanismus des Tit for Tat des affirmativen Forschers unterbrochen wird bzw. die Strategie verlassen wird und auf V des Auftraggebers in der nächsten Runde mit E geantwortet wird.

Theoretisch können solche Verhandlungen sogar unendlich oft stattfinden, bestehen doch Institutionen aus über die Zeit jeweils unterschiedlichen Menschen. Um dieses Problem einzufangen, muss man unterscheiden zwischen Erhalt eines Auftrages jetzt und dem stets nicht ganz sicheren Erhalt eines Folgeauftrages in der jeweils nächsten Periode. Wenn beide etwa einmal pro Jahr verhandeln, kann man das Problem des wegen der Konkurrenz anderer affirmativer Anbieter unsicheren Erhalts von Nachfolgeaufträgen „einpreisen“, indem man z.B. den Zinssatz  $p$  als zukunftsbezogenen Diskontfaktor nutzt. Dabei gehen wir davon aus, dass der Erhalt eines Auftrages nebst damit verknüpftem Einkommen und Konsummöglichkeiten in Höhe eines vereinbarten Betrages von  $e$  besser ist als ein Vertrag im kommenden Jahr mit dem Betrag  $E$ , wenn  $(1+p)e > E$  gilt.

Alternativ kann man auch annehmen, dass die Bewertung der antizipierten Konsummöglichkeit, die aus dem im nächsten Jahr erwarteten Vertrag folgt, auf das  $q$ -fache geschmälert ist mit:  $q=1/(1+p)<1$ . Will man nun wissen, wann sich bei wiederholten Spielen für den affirmativen Forscher, für den der Kunde König ist, ein Wechsel von der Tit for Tat-Strategie bei einem unehrlichen Auftraggeber lohnt, gilt anknüpfend an unser Beispiel folgende Ungleichung (zur Spieltheorie allgemein vgl. Holler & Illing, 2005).

Es muss gelten:  $2 + 2q + 2q^2 + 2q^3 + \dots < 3 + 1q + 1q^2 + \dots$  (geometr. Reihe)

Das gilt genau dann, wenn

$$2 \cdot \frac{1}{1-q} < 3 + \frac{1}{1-q} - 1 \Leftrightarrow \frac{1}{1-q} < 2 \Leftrightarrow \frac{1}{2} < 1-q \Leftrightarrow q < 0,5$$

Damit muss also

$$q = \frac{1}{1+p} < 0,5 \Leftrightarrow 1+p > 2 \Leftrightarrow p > 1^{18}$$

gelten, damit sich eine Abweichung von der Tit for Tat-Strategie für den affirmativen Forscher lohnen würde.

Tabelle 5 zeigt eine Ungleichung, ab der sich ein Wechsel von der Tit for Tat-Strategie für den affirmativen Forscher angesichts künftiger Sanktionen bzw. weniger wahrscheinlich gewordener künftiger Konsummöglichkeiten lohnt. Wenn sich also der Zukunftswert der Strategie V/V für den affirmativen Forscher und unehrlichen Auftraggeber in der Ungleichung in Tabelle 5 auf den Wert 1 verringert hat, im Vergleich zu 2 für E/E, lohnt sich die Aufgabe der Strategie „Wie Du mir so ich Dir“ und der Auftragnehmer enttäuscht dann den Auftraggeber, um in der nächste Periode E/E zu präferieren bzw. E/V zu riskieren. Gleichzeitig wird klar, dass ein bloßes Reden über Sanktionen, also Talk, ohne erkennbares Vollziehen, also hohe Sanktion bei gegen null gehender Entdeckungswahrscheinlichkeit, den Wechsel der Strategie nicht verändern dürfte. Dies bedeutet gleichzeitig, dass, wenn man durch Einführung verstärkter Sanktionsmechanismen die Auszahlungswerte wie in Tabelle 5 gezeigt verringert, das Ethos der Forschung als normative Regulierung substituiert werden kann.<sup>19</sup>

Tabelle 5: Wechselbedingungen für einen affirmativen Forscher hin zur Ehrlichkeit

Spieler B (Auftraggeber)	Spieler A (Auftragnehmer)/ E / ehrliche Forschung	V/ in Ergebnissen teilweise oder ganz gelenkte For- schung
E/ ehrliche Forschung	2	3
V/ in Ergebnissen teilweise oder ganz gelenkte Forschung	0	1
	3	1

<sup>18</sup> Für allgemeine Auszahlungsgrößen  $x, y > 0$  mit  $y > x$  und  $y$  als Auszahlung im Fall gelenkter Forschung gilt:

$$k := \frac{x-1}{y-x} < p$$

Für den Fall, dass  $y = x + 1$  gilt  $x \in (0,1) \Rightarrow k < 0; x \in (1,n) \Rightarrow k \in (0,1); x \geq n + 1 \Rightarrow k \geq 1$ .

Ist  $y = n \cdot x$ , so gilt für  $x \in (0,1] \Rightarrow k \leq 0; x > 1 \Rightarrow k < 1$ .

<sup>19</sup> Die Ergebnisse zeigen auch, warum bei wiederholten Verträgen bestehende kartellähnliche Zustände zwischen Auftraggeber und -nehmer so stabil sind bzw. so lange aufrechterhalten werden.

Geeignete Maßnahmen für die Substituierung könnten etwa ein verstärkter Zwang zur Publikation und der Ausschluss von Evaluationsnehmern aus der einschlägigen Fachorganisation wie z.B. der Deutschen Gesellschaft für Evaluation (DeGEval) sein, wenn etwa Forscher Evaluationsberichte nicht publizieren<sup>20</sup> bzw. Einblick in dieselben gewähren, sowie Schulungen von Evaluatoren im Bereich der forschungsethischen Grundsätze. Damit wächst einerseits die Chance, dass Abweichler, die affirmative Forschung liefern, entdeckt und öffentlich bekannt werden und vermehrt Wissensbisse bekommen. Auftraggeber kämen andererseits dadurch nicht mehr in den Nutzen einer falschen Legitimation durch gelenkte Befunde. Denkbar wäre auch eine unabhängige Ethikkommission, die dem Auftragnehmer nach vollzogenem Auftrag und vorliegendem Bericht eine Art Gütesiegel erteilt.

Insbesondere durch die externe Überwachung von Forschungsstandards kann mittels der damit verbundenen Kontrolle affirmativer Forscher ein Substitut für das Ethos der Forschung geschaffen werden, wenn man ehrliche Forschung will (Homann & Pies, 1994). Dieses Instrument greift dabei um so wirksamer, je schwächer sich die Forscher am reinen Ethos der Forschung orientieren und je weniger problematisch Erwartungsenttäuschungen desselben sind.<sup>21</sup> Gleichzeitig wäre es ein Beitrag zur Wettbewerbsgerechtigkeit und hätte damit ähnliche Funktion wie die eingangs erwähnten Overhead-Regeln (s. Fußnote 9), nur diesmal mit Wirkung in anderer Richtung. Die Entwicklung der Nachfrage nach ehrlicher Evaluationsforschung hängt dann jedoch davon ab, in welchem Umfang tatsächlich seitens der Auftraggeber Erkenntnis gewünscht ist oder falsche Legitimationsfunktion. Wird also Evaluation bisher eher zur Legitimation mit Hilfe affirmativer Forschung benötigt, wird die Nachfrage stark schrumpfen. Nehmen wir nun weiter an, dass Auftraggeber mittlerweile die Legitimationsfunktion mittels gelenkter Befunde wichtiger nehmen als die reine Erkenntnisfunktion und dass sie deshalb affirmative Forschung präferieren<sup>22</sup>, ist gleichzeitig zu erwarten, dass bei Nichteinführung der Sanktionsmechanismen universitäre und außeruniversitäre Anbieter mit variablen Präferenzen zusehends mehr Aufträge erhalten, dadurch erhöhen sie den Druck auf universitäre Anbieter, die sich normativ eng am Ethos der Forschung ausrichten und exklusiv ehrliche Forschung anbieten.

Verweigern die am Ethos der Forschung ausgerichteten universitären Forscher affirmative Evaluationen, wird sich das drittmittelbezogene Nahrungsangebot für die ehrlichen Forscher

---

<sup>20</sup> Datensensible Evaluationsbefunde könnte man entsprechend anonymisieren.

<sup>21</sup> Unbedingt zu vermeiden ist, dass zwecks Legitimation eine Überwachung kommuniziert wird, durch Entkoppelung von Reden und Handeln aber eine Beeinträchtigung der Produktionsfunktion vermieden wird (s. zur organisierten Heuchelei von Organisationen Brunsson, 2003)

<sup>22</sup> Unter diesen Bedingungen ähnelt die selbst beauftragte Evaluation einem echten Bűßerritual. Man simuliert nicht Buße, um hinter weiter zu machen wie zuvor, sondern will echte Läuterung.

in der Universität immer mehr verknappen, ebenso auch für die im außeruniversitären Markt tätigen ehrlichen Forscher, die grundsätzlich keine affirmative Forschung liefern (zu ähnlichen Mechanismen Emrich & Thieme. 2012). Wird sich die Zahl der Auftraggeber stark erhöhen, die ehrliche Forschung wollen, wird sich auch die Konkurrenz für die ehrlichen Auftragnehmer weiter verschärfen, weil auch die variablen Forscher ehrliche Forschung liefern können. Je nach Ausgestaltung der Regeln ist aber Pareto-Effizienz in diesem Markt zumindest grundsätzlich denkbar.

Die Aushebelung der Marktmechanismen durch Privat Public Partner-Anbieter wie das CHE-Ranking ist hier ohne größeren Einfluss. Quasi-Monopolisten wie die OECD oder das CHE-Ranking drücken zwar mit erheblicher politischer und medialer Unterstützung eine erhöhte Menge des Guts „Evaluation“ in den Markt, fördern jedoch gleichzeitig die steuerfinanzierte Nachfrage durch politischen Lobbyismus. Diese Evaluationsanbieter befriedigen mittels der in hohem Maße verfügbaren Steuermittel den endogen erzeugten Bedarf von Politik und Öffentlichkeit nach Evaluationswissen. Mit dem Ausmaß der steuerlichen Alimentierung nimmt auch die Wettbewerbsverzerrung zu, kommen doch kaum andere Anbieter in dieses Marktsegment. Je mehr Druck diese PPP-Anbieter medial entfalten, um so mehr Bedarf erzeugen sie und um so mehr Steuermittel benötigen sie. Institutionell verbundene Auftraggeber/Auftragnehmer schaffen sich somit einen Markt und sorgen durch ausgeprägtes rent seeking für dessen Fortbestand und Ausweitung. Dabei sind die Ergebnisse egal, weil im Mittelpunkt die durch bloße Sachstandsmessungen (und damit meist naiven Empirismus) dokumentierten relativen Veränderungen stehen.

Damit ist im Sinne des oben skizzierten institutionenökonomischen Zugangs auch klar, welche relevanten Überlegungen der evaluierende Forscher anstellen muss, um entscheiden zu können, welche Intentionen und versteckten Aktionen vor und nach Vertragsabschluss zwischen Auftraggeber und –nehmer relevant sind. Der Auftragnehmer muss Informationen darüber sammeln, was der Auftraggeber wohl wirklich will und umgekehrt gilt dies auch für den Auftraggeber, der einen Auftragnehmer möchte, der seine realen Erwartungen unabhängig von den kommunizierten nicht enttäuscht. Abhängig davon, ob es sich hier um einen Erstvertrag handelt oder ob es sich um eine stabile Geschäftsbeziehung handelt, sind hier der soziale Ruf in seiner Signaling-Funktion sowie Screening- und Monitoring-Strategien jeweils relevant. Je mehr man weiß, was man voneinander zu erwarten hat und je häufiger diese Erwartungen schon im Sinne erprobter Vertrauensbeziehungen erfüllt worden sind, um so stabiler entwickelt sich über differenzierte Signaling-Strategien ein stummer Handel (zum stummen

Handel s. Herodot), der zu stabilen Geschäftsbeziehungen führt (s. zu den Mechanismen vor Beauftragung Benedikt & Gresser, 2014, am Beispiel von Gerichtsgutachten).

Hat der Auftraggeber also einen Vertrag abgeschlossen und sich beide vorher vorsichtig „abgetastet“, bleibt zur Vermeidung des Moral Hazard Problems für den Auftraggeber nur die vertragliche Zusicherung, dass die Ergebnisse ihm exklusiv zur Verfügung stehen und die Möglichkeit zur Einführung von mehrfachen Besprechungen der Ergebnisse z.B. in Beiräten, die wiederum das Einkommen des Auftragnehmers infolge erhöhten Zeitaufwands reduzieren. Der Forscher wird seine Antworten auf die Fragen, was können wir wirklich gesichert sagen, je nachdem, ob er eher redliche oder affirmative Forschung liefert, abwägen zwischen der Ehrlichkeit und Unbeeinflusstheit wissenschaftlicher Befunde, unabhängig welche Erwartungen der Auftraggeber signalisiert hat, und der Erfüllung der Erwartungen des Auftraggebers. Damit wird der anerkanntermaßen ehrliche und völlig unbestechliche Forscher wiederum zum Spezialanbieter in all jenen Fällen, in denen die Ergebnisse kritikresistent und möglichst unangreifbar sein müssen.

<b>Tabelle 6: Asymmetrien zwischen Auftraggeber und Evaluator</b>		
Vor- und nach Vertragsabschluss	<b>Typ der Asymmetrie zwischen Evaluator (Auftragnehmer) und Auftraggeber</b>	
	Hidden Characteristics Vor Vertragsabschluss	Hidden Action Nach Vertragsabschluss
	Negative Eigenschaften werden beiderseits verborgen. (Bevorzugt der Auftraggeber bestimmte Ergebnisse? Ist der Wissenschaftler wirklich in der Lage, zugesagte Leistungen zu erbringen usw.?)	Forschung ist nicht direkt beobachtbar Welche Interessen vertritt der Wissenschaftler wirklich? Verwertungsklauseln, Veröffentlichungsmöglichkeit usw. Wer hat überhaupt die Verfügungsrechte an Ergebnissen, wie lange sind sie eventuell eingeschränkt?
Zu bearbeitende Probleme	Informationsbeschaffung (Informationsasymmetrien vermeiden) Über welche Informationen verfügen die Partner vor Vertragsabschluss?)	Nachteilsvermeidung (Moral Hazard) Wie können unliebsame Ergebnisse nach Vertragsabschluss verhindert werden?
Lösung	Wirksames Signaling auf beiden Seiten Verlässlichkeit erfragen (Reputationsrecherche)	Monitoring Beständige Sitzungen / Beirat Nachbesserungsrecht beim Schlussbericht

Tabelle 6 skizziert die vor bzw. nach Vertragsabschluss relevanten Informationsasymmetrien. Abschließend wird damit deutlich, warum der einschlägige Ruf von Evaluationsunternehmen, Empfehlungen und die Pflege von Kontakten zwischen potentiellen Auftraggebern und –nehmern so hohe Bedeutung haben, sie verringern die gegenseitigen Transaktionskosten und belegen die Plausibilität stabiler kartellähnlicher Zustände, wie sie implizit aus der Berechnung im Kontext von Tabelle 5 deutlich wurden.

## 6. Zusammenfassung und Diskussion

Deutlich konnte gezeigt werden, dass bei wiederholten Spielen für affirmative Forscher und unehrliche Auftraggeber ein Verlassen der Tit for Tat-Strategie von V/V hin zu E/E nur lohnt, wenn erhöhte Risiken bestehen. Bestehen solche Risiken also nicht, was wir für die Gegenwart annehmen können, insbesondere für außeruniversitäre Evaluationsnehmer, ist die V/V Variante stabil die lohnendste. Angesichts vielfältiger Principal Agent-Beziehungen zwischen verschiedenen Auftraggebern und einem Auftragnehmer bedeutet dies, dass über die Zeit eher diejenigen im Markt bleiben, die als Auftragnehmer sowohl E/E als auch V/V beherrschen.<sup>23</sup> Wenn also Auftragnehmer flexibel genug sind, sowohl Auftraggeber zu „bedienen“, die V erwarten als auch diejenigen, die E erwarten (hier gilt: der Kunde ist König und geliefert wird, was bestellt wurde), werden sie wesentlich mehr Aufträge auf Dauer haben als etwa der am reinen Ethos der Forschung orientierte Typus, der unabhängig nur E bieten will und damit die Erwartung V/V nicht „bedienen“ will. Dies bedeutet aber auch, dass angesichts der geschilderten zunehmenden Außensteuerung von Wissenschaft die Wahrscheinlichkeit für V/V solange hoch bleibt, wie nicht solche entsprechende Schutzmechanismen konstruiert, durchgesetzt und sanktionsbewehrt überwacht werden, die moralisch wünschenswertes Verhalten begünstigen.

Evaluationsforschung als sowohl universitäre Wissenschaft als auch außerhalb der Universität betriebene angewandte Wissenschaft kann sich also wirkungsvoll gegenüber einer Außensteuerung durch marktliche Mechanismen schützen. Die Aushebelung der Marktmechanismen durch z. B. Privat Public Partner-Anbieter ist ungleich problematischer. Während man durch Einführung der EU-Overhead-Regeln die Wettbewerbssituation für universitäre Evaluationsanbieter verschlechtert, indem Angebote dadurch verteuert werden, setzen auf der Seite außeruniversitärer Anbieter Evaluationsagenturen wie etwa das CHE-Ranking oder die OECD als mit Steuergeld subventionierte Einrichtungen die Wettbewerbsregeln ebenfalls außer Kraft, indem sie die Politik teilweise „vor sich her treiben“ und den Bedarf nach Evaluation endogen solange erhöhen, bis die Politik nicht mehr anders kann und sie beauftragt, während anderen Evaluationsagenturen der Zugang in diesen Markt faktisch kaum gelingt. So wird der Wettbewerb faktisch auf die wenigen im politischen und medialen Bewusstsein präsenten Anbieter reduziert, die schon frühere Sachstandsmessungen produziert haben. Vor dem Hintergrund der Tournament-Theorie könnte man auch argumentieren, dass die Politik, nicht selten getrie-

---

<sup>23</sup> Hier wurde vorrangig die Legitimationsfunktion betrachtet. Denkbar ist auch, dass sich entsprechende der von Stockmann angesprochenen Funktionen von Evaluationen (s. Fußnote 11) vier Marktsegmente ausprägen mit Spezialisierungen für die vier Funktionen. Je nach deren Bedeutung haben dann die „richtigen“ Wissenschaftler mehr oder weniger Marktmöglichkeiten.

ben durch seitens der Evaluationsagenturen medial erzeugten öffentlichen Druck, einen Wettbewerb organisiert, der die Teilnehmer in einen speziellen Typ eines Rattenrennens verwickelt. In diesem Rattenrennen finden permanente Rennen statt. Der Ressourcenzustrom wird allerdings immer wieder erhöht, so dass bei Simulation von harter Konkurrenz die ins scheinbare Rattenrennen getriebenen Anbieter mit höherer Wahrscheinlichkeit und ausreichend bis opulent mit Ressourcen ausgestattet überleben als Konkurrenten in nicht staatlich subventionierten Märkten (vgl. Lazear & Rosen, 1981). Mit der staatlichen Subventionierung werden somit die Ergebnisse der Evaluationen so gelenkt, dass die Nachfrage nach mehr Evaluation im Sinne von bloßen Sachstandsmessungen verstärkt wird (man ist ja häufig nicht an wirklicher Erkenntnis interessiert, sondern an der nachträglichen Legitimation politischer Entscheidungen), was wiederum den steten Fluß von Steuergeldern auf die Mühlen dieser Quasi-Monopolisten verstärkt.

Gleichzeitig erhöht dies den Druck auf sonstige Evaluationsanbieter zu einer affirmativen Haltung gegenüber der Politik, da ja die staatlich subventionierten, sich im Scheinwettbewerb befindlichen Quasi-Monopolisten den Mainstream vorgeben. Die eingangs zitierte Formulierung von Frey (2007), der von Evaluitis sprach, hat also einschließlich ihres wertenden Untertons durchaus ihre Berechtigung. Ob es sich bei der offensichtlichen Präferenzierung und Belohnung affirmativer Evaluations“forschung“ im Zuge der Universitätsreformen um eine erwünschte oder unerwünschte Wirkung der aktuellen Förderungsstrukturen handelt, kann allerdings hier nicht beantwortet werden.

## **Literatur**

- Akerlof, G. (1976). The economics of caste and the rat race and other wonderful tales. In: *Quarterly Journal of Economics*, 90, 599-617.
- Benedikt, J. & Gresser, U. (2014) Gerichtsgutachten: Oft wird die Tendenz vorgegeben. *Deutsches Ärzteblatt* 2014; 111(6): A-210 / B-180 / C-176. Online verfügbar unter: [www.aerzteblatt.de/archiv/154014/Gerichtsgutachten-Oft-wird-die-Tendenz-vorgegeben](http://www.aerzteblatt.de/archiv/154014/Gerichtsgutachten-Oft-wird-die-Tendenz-vorgegeben)
- Berghoff, S.; Federkeil, G.; Giebisch, P.; Hachmeister, C.-D.; Hennings, M.; Roessler, I. & Ziegele, F. (2009). *Das CHE-Forschungsranking deutscher Universitäten 2009*, Gütersloh, Centrum für Hochschulentwicklung.
- Briefs, G. (1957). Grenz-moral in der pluralistischen Gesellschaft. In: Beckerath, E.; Meyer, F. W. & Müller-Armack, A. (Hrsg.): *Wirtschaftsfragen der freien Welt*. Festschrift für Ludwig Erhard. Frankfurt am Main, 97-108.
- Brunsson, N. (2003). *The organization of hypocrisy: talk, decisions, and actions in organizations*, Copenhagen, Business School Press.
- DeGEVal (Deutsche Gesellschaft für Evaluation): *Empfehlungen zur Anwendung der Standards für Evaluation im Handlungsfeld der Selbstevaluation*. Alter, Dezember 2004.
- Dessauer, B., Emrich, E. Klein, M. & Pierdzioch, C. (2014). Zur Evaluation wissenschaftlicher Publikationsleistungen in der Sportwissenschaft. Angenommen zur Publikation in der Zeitschrift für Evaluation 1/2014.

- Durkheim, E. (1973): *Erziehung, Moral und Gesellschaft*. Vorlesung an der Sorbonne 1902/1903, übers. von L. Schmitz, Neuwied: Luchterhand (orig. 1925).
- Lazear E. P. & Rosen, S. (1981). "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89(5), pp. 841-864.
- Holler, M. & Illing G. (2005). *Einführung in die Spieltheorie*. 6. Aufl. Springer. Berlin.
- Emrich, E.; Papathanassiou, V. & Pitsch, W. (1996). Klettertechnik für Aufsteiger - Seilschaften als soziales Phänomen, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 141-155.
- Emrich, E. (2006). Sportwissenschaft zwischen Autonomie und außerwissenschaftlichen Impulsen. In: *Sportwissenschaft*, 36, 151-170.
- Emrich, E.: „Ars Corruptendi“ (2006). Zur Interaktions- und Beziehungsdynamik bei Bestechungen. In: *sozialersinn* 7, 2, S.327-343.
- Emrich, E. & Fröhlich, M. (2010). Universität in Deutschland zwischen Institution und Organisation. Reflexionen zur Idee der Universität und ihrer betrieblichen Ausgestaltung. In: *sozialersinn* 11, 1, S.125-144.
- Emrich, E. & Pierdzioch, C. (2012). *Das Vademecum der Evaluologie*. Neue Arten im Biotop der Wissenschaft. Saarbrücken. Universaar. Online verfügbar unter: <http://universaar.uni-saarland.de/monographien/volltexte/2012/98/pdf/Emrich.pdf>.
- Emrich, E. & Pierdzioch, C. (2011) *Im Biotop der Wissenschaft*. Das PARK-Modell der Makroökonomie. Universaar: Saarbrücken. Online verfügbar unter: [http://universaar.uni-saarland.de/monographien/volltexte/2011/86/pdf/emrich\\_pierdzioch.pdf](http://universaar.uni-saarland.de/monographien/volltexte/2011/86/pdf/emrich_pierdzioch.pdf).
- Emrich, E. & Thieme, L. (2012). Überlegungen zur Schließung von Wissensmärkten am Beispiel angewandter Sportwissenschaft. In: Körner, S.; Frei, P. (Hrsg.): *Die Möglichkeit des Sports*. Kontingenz im Brennpunkt sportwissenschaftlicher Analysen. Bielefeld 2012, S. 219-253.
- Eßbach, W. (2009). *Jenseits der Fassade*. Die deutsche Bachelor-/Master-Reform. In: Kaube, J. (Ed.): *Die Illusion der Exzellenz*. Lebenslügen der Wissenschaft, Berlin, Verlag Klaus Wagenbach, 329-372.
- Goffman, E. (2003 [1959]). *Wir alle spielen Theater*. Die Selbstdarstellung im Alltag, München, Zürich, Piper Verlag.
- Harvey, L. & Green, D. (2000). Qualität definieren. Fünf unterschiedliche Ansätze, in: *Zeitschrift für Pädagogik*, 41, 17-39.
- Haucap, J. & Mödl, M. (2013). Zum Verhältnis von Spitzenforschung und Politikberatung: Eine empirische Analyse vor dem Hintergrund des Ökonomenstreits. *Perspektiven der Wirtschaftspolitik* 14 (3/4), S. 346-378.
- Herodot, Buch 4 Dareios: Skythenfeldzug. Feldzug gegen Libyen, Beschreibung Libyens, Kapitel 196, 490-424 v. Chr.
- Homann, K. & Pies, I. (1994). *Wirtschaftsethik in der Moderne: Zur ökonomischen Theorie der Moral*. In: *Ethik und Sozialwissenschaften* 5, 3-12.
- Hornbostel, S. (2001). Die Hochschule auf dem Weg in die Audit Society, In: Stölting, E.; Schimank, U. (Ed.): *Die Krise der Universitäten*, Wiesbaden, Westdeutscher Verlag, 139-158.
- Huber, M. (2005). Reform in Deutschland. Organisationssoziologische Anmerkungen zur Universitätsreform. In: *Soziologie*, 34, 391-403
- Janker, K. (2014). Jeder fünfte Schüler ist im Alltag abgehängt. *Sueddeutsche.de* vom 1. April 2014. Online verfügbar unter: <http://www.sueddeutsche.de/bildung/pisa-studie-jeder-fuenfte-schueler-ist-im-alltag-abgehaengt-1.1926567>, Zugriff am 2. April 2014.
- Joint Committee on Standard for Educational Evaluation (Hrsg.) (2000). *Handbuch der Evaluationsstandards*. 2. durchgesehene Auflage, Leske + Budrich, Opladen 2000.

- Kaube, J. (2009). *Excellenz per Beschluss*, in: Kaube, J. (Ed.): *Die Illusion der Excellenz. Lebenslügen der Wissenschaft*, Berlin: Verlag Klaus Wagenbach, 82-89.
- Klein, M. & Emrich, E. (2013). *Konstruktion sozialer Probleme. Zur Rolle von moralischen Kreuzfahrern im wissenschaftlichen Legitimationsprozess*. In: *sozialersinn* 14, 1, S. 103-121.
- Kromrey, H. (2003). *Evaluation in Wissenschaft und Gesellschaft*. Vortrag bei der Feierstunde zur Eröffnung des Centrums für Evaluation am 21.2.2003 in Saarbrücken / zur Veröff. in: *ZfEv, Zeitschrift für Evaluation*, H. 1/2003, online verfügbar unter: [http://hkromrey.de/Kromrey\\_EvalWissG.pdf](http://hkromrey.de/Kromrey_EvalWissG.pdf), Zugriff am 31.3.2014.
- Küppers, B.-O. (2008). *Nur Wissen kann Wissen beherrschen. Macht und Verantwortung der Wissenschaft*, Köln, Fackelträger Verlag.
- Kurucz, J. (1986). *Ideologie, Betrug und naturwissenschaftliche Erkenntnis. Eine wissenssoziologische Untersuchung*, Saarbrücken, Verlag d. Reihe. *Saarländische Beiträge zur Soziologie* 6.
- Lange, S. (2008). *New Public Management und die Governance der Universität*. In: *dms - der moderne Staat - Zeitschrift für Public Policy, Recht und Management*, 1, 235-248.
- Merton, R. K. (1938). *Science and the social order*. In: *Philosophy of Science*, 5, 321-337.
- Mittelstraß, J. (1982). *Wissenschaft als Lebensform. Reden über philosophische Orientierung in Wissenschaft und Universität*. Frankfurt am Main: Suhrkamp.
- Münch, R. (2009). *Globale Eliten, lokale Autoritäten. Bildung und Wissenschaft unter dem Regime von PISA, McKinsey & Co*, Frankfurt am Main, Suhrkamp.
- Paris, R. (2001). *Machtfreiheit als negative Utopie. Die Hochschule als Idee und Betrieb*. In: Stölting, E. & Schimank, U. (Ed.): *Die Krise der Universitäten*, Wiesbaden, Westdeutscher Verlag, 194-222.
- Parsons, T. & Platt, G. M. (1990). *Die amerikanische Universität*, Frankfurt am Main, Suhrkamp.
- Plessner, H. (1985). *Zur Soziologie der modernen Forschung und ihrer Organisation in der deutschen Universität (1924)*, In: Plessner, H. (Ed.): *Gesammelte Schriften in zehn Bänden - X: Schriften zu Soziologie und Sozialphilosophie*, Frankfurt am Main, Suhrkamp, 7-30.
- Stichweh, R. (2009). *Autonomie der Universitäten in Europa und Nordamerika: Historische und systematische Überlegungen*. In: Kaube, J. (Ed.): *Die Illusion der Excellenz. Lebenslügen der Wissenschaft*, Berlin, Verlag Klaus Wagenbach, 38-49.
- Stockmann, R., Meyer, W. & Schenke, H. (2011). *Unabhängigkeit von Evaluationen*. In: *Zeitschrift für Evaluation*, 10 (1), S. 39-67.
- Stockmann, R. (2000). *Evaluation in Deutschland*. In: R. Stockmann (Hrsg.): *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. Leske & Budrich, Opladen 2000.
- Stockmann, R. & Meyer, W. (2014). *Evaluation. Eine Einführung*. Opladen & Toronto: Budrich.
- Weber, M. (1980 [1921]). *Wirtschaft und Gesellschaft*, Tübingen, Mohr Siebeck.
- Weber, M. (1988). *Politik als Beruf*. In: Ders., *Gesammelte politische Schriften*. Hrsg. von Johannes Winckelmann (5. Aufl.: S. 505-560). Tübingen: Mohr Siebeck (orig. 1919).
- Weber, M. (2002). *Wissenschaft als Beruf*. In: Kaesler, D. (Ed.): *Max Weber Schriften 1894-1922*, Stuttgart, Kröner, 474-511 (orig. 1919).
- Weingart, P. (2001). *Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft*, Göttingen, Velbrück Wissenschaft.