

Effekte hochschulinterner indikatorenbasierter Steuerungs-
und Anreizverfahren auf das Forschungsverhalten von
Professorinnen und Professoren an
sozialwissenschaftlichen Fachbereichen

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von
Dipl.-Soz. Lorenz Tafertshofer
aus Innsbruck

Tübingen

2017

Tag der mündlichen Prüfung:

23.11.2017

Dekan:

Professor Dr. rer. soc. Josef Schmid

1. Gutachter:

Prof. Dr. Bernhard Schmidt-Hertha

2. Gutachter:

Prof. Dr. Josef Schrader

Inhaltsverzeichnis

1. EINFÜHRUNG	6
2. NEW PUBLIC MANAGEMENT UND DIE STEUERUNG DER HOCHSCHULEN	12
2.1 NEW PUBLIC MANAGEMENT ALS NEUES GOVERNANCE-REGIME DER HOCHSCHULEN	12
2.1.1 <i>Die Governance der Hochschulen</i>	12
2.1.2 <i>New Public Management im Hochschulbereich</i>	15
2.2 HOCHSCHULINTERNE STEUERUNGS- UND ANREIZVERFAHREN	21
2.2.1 <i>Indikatorenbasierte Mittelzuweisung</i>	21
2.2.2 <i>Zielvereinbarungen</i>	23
2.2.3 <i>W-Besoldung</i>	25
2.2.4 <i>Weitere Verfahren</i>	26
2.3 MESSUNG DER FORSCHUNGSPERFORMANZ UND FORSCHUNGSINDIKATOREN	28
2.3.1 <i>Quantifizierung der Forschungsleistung</i>	28
2.3.2 <i>Kritische Diskussion der gängigen Forschungsindikatoren</i>	30
2.4 UMSETZUNG UND WIRKUNG INDIKATORENBASIERTER STEUERUNGS- UND ANREIZVERFAHREN	44
3. DIE ORGANISATION UNIVERSITÄT, DIE AKADEMISCHE PROFESSION UND DAS PROBLEM DER STEUERUNG	52
3.1 DIE UNIVERSITÄT ALS BESONDERE FORM DER ORGANISATION.....	52
3.2 DIE AKADEMISCHE PROFESSION	58
3.3 DIE STEUERUNGSRESISTENZ DER AKADEMISCHEN PROFESSION	64
4. THEORETISCHES MODELL UND ABLEITUNG DER FORSCHUNGSHYPOTHESEN	70
4.1 MOTIVATIONSTHEORETISCHER RAHMEN	70
4.1.1 <i>Die Selbstbestimmungstheorie der Motivation</i>	70
4.1.2 <i>Motivation Crowding Theory</i>	74
4.1.3 <i>Das Konzept der Public Service Motivation</i>	77
4.2 VERKNÜPFUNG DER THEORETISCHEN ANSÄTZE UND ABLEITUNG DER FORSCHUNGSHYPOTHESEN	80
5. AUFBAU UND ABLAUF DER UNTERSUCHUNG	89
5.1 FORSCHUNGSDESIGN	89
5.2 ENTWICKLUNG DES ERHEBUNGSINSTRUMENTES	90
5.2.1 <i>Operationalisierung der Variablen</i>	90
5.2.2 <i>Programmierung des Online-Fragebogens und Fragebogenlogik</i>	95
5.2.3 <i>Durchführung und Ergebnisse des Pretests</i>	97
5.3 DATENERHEBUNG	101

5.3.1	<i>Recherche der Stichprobe und Durchführung der Befragung</i>	101
5.3.2	<i>Rücklaufquote und Repräsentativität</i>	103
6.	ERGEBNISDARSTELLUNG	108
6.1	DESKRIPTIVE ANALYSEN IM FÄCHERVERGLEICH	108
6.1.1	<i>Stichprobenbeschreibung</i>	108
6.1.2	<i>Einsatz von Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren</i>	111
6.1.3	<i>Einsatz und Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren</i>	120
6.1.4	<i>Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs in der Scientific Community</i>	129
6.1.5	<i>Beitrag verschiedener Aspekte der Forschungsarbeit für den wissenschaftlichen Erkenntnisgewinn</i>	134
6.1.6	<i>Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit</i>	139
6.1.7	<i>Intrinsische Forschungsmotivation</i>	146
6.1.8	<i>Anmerkungen und Kommentare der Befragten</i>	149
6.2	ÜBERPRÜFUNG DER FORSCHUNGSHYPOTHESEN	154
6.2.1	<i>Bivariate Analysen</i>	154
6.2.2	<i>Multivariate Analysen</i>	182
6.2.2.1	<i>Generelles Vorgehen bei der Berechnung der multiplen Regressionsmodelle</i>	182
6.2.2.2	<i>Ergebnisse der Regressionsanalysen</i>	190
6.3	ZUSAMMENFASSUNG UND INTERPRETATION DER WESENTLICHEN ERGEBNISSE	209
6.3.1	<i>Diskussion der Ergebnisse im Hinblick auf die Fragestellungen der Arbeit und die Forschungshypothesen</i>	209
6.3.2	<i>Zusammenfassende Ergebnisdarstellung</i>	219
7.	ABSCHLIEßENDE DISKUSSION	222
7.1	METHODISCHE REFLEXION	222
7.2	FOLGERUNGEN FÜR DIE FORSCHUNG UND HOCHSCHULPRAXIS	225
	LITERATURVERZEICHNIS	231
	ANHANG	251
	I. FRAGEBOGEN	251
	II. E-MAIL ANSCHREIBEN	258
	III. HISTOGRAMME DER RESIDUEN (MULTIPLE REGRESSION)	259

Abbildungsverzeichnis

Abbildung 1: Motivationstypen nach der Selbstbestimmungstheorie der Motivation.....	70
Abbildung 2: Theoretisches Modell.....	82
Abbildung 3: Histogramm Alter und Beschäftigungsdauer	109
Abbildung 4: Eingesetzte Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren.....	114
Abbildung 5: Eingesetzte indikatorenbasierte Steuerungs- und Anreizverfahren	122
Abbildung 6: Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs im Fächervergleich	132
Abbildung 7: Beitrag verschiedener Aspekte der Forschungsarbeit für den Erkenntnisgewinn im Fächervergleich.....	137
Abbildung 8: Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit im Fächervergleich.....	143
Abbildung 9: Histogramm Index Intrinsische Forschungsmotivation	148
Abbildung 10: Interaktionseffekt Bedeutung Reputationszuwachs x Einsatz Indikator (Publikationen in Zeitschriften mit Peer-Review)	193
Abbildung 11: Theoretisches Modell - Ergebnisdarstellung.....	219

Tabellenverzeichnis

Tabelle 1: Fächerverteilung der Bruttostichprobe	102
Tabelle 2: Rücklauf- und Beendigungsquote je Fach	104
Tabelle 3: Geschlechterverteilung der Stichprobe	106
Tabelle 4: Fächerverteilung der Stichprobe	106
Tabelle 5: Eingesetzte Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren.....	113
Tabelle 6: Eingesetzte Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren im Fächervergleich	116
Tabelle 7: Gewichtung der Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren insgesamt	118
Tabelle 8: Eingesetzte indikatorenbasierte Steuerungs- und Anreizverfahren.....	121
Tabelle 9: Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren (eingesetzte Indikatoren und deren Gewichtung).....	124
Tabelle 10: Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs	129
Tabelle 11: Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs im Fächervergleich	131
Tabelle 12: Beitrag verschiedener Aspekte der Forschungsarbeit für den Erkenntnisgewinn .	135
Tabelle 13: Beitrag verschiedener Aspekte der Forschungsarbeit für den Erkenntnisgewinn im Fächervergleich.....	136
Tabelle 14: Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit	140

Tabelle 15: Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit im Fächervergleich.....	142
Tabelle 16: Intrinsische Forschungsmotivation (Itemwerte)	146
Tabelle 17: Korrelationen zur Variablen AV_1: Publizieren in Zeitschriften mit Peer-Review-Verfahren.....	157
Tabelle 18: Korrelationen zur Variablen AV_2: Publizieren in Zeitschriften ohne Peer-Review-Verfahren.....	159
Tabelle 19: Korrelationen zur Variablen AV_3: Verfassen anderer Publikationen	161
Tabelle 20: Korrelationen zur Variablen AV_4: Publizieren in Zeitschriften mit hohem Impact Factor / in high-ranked journals	163
Tabelle 21: Korrelationen zur Variablen AV_5: Betreuung von Promotionen.....	165
Tabelle 22: Korrelationen zur Variablen AV_6: Einwerben von Drittmitteln der DFG.....	167
Tabelle 23: Korrelationen zur Variablen AV_7: Einwerben von sonstigen Drittmitteln	168
Tabelle 24: Korrelationen zwischen der Variablen „Stellenwert des Publizierens als Koautor / in Mehrautorenschaft“ und dem Stellenwert anderer Publikationstätigkeiten	171
Tabelle 25: Korrelationen zwischen dem Stellenwert von Grundlagenforschung bzw. anwendungsbezogener Forschung und dem Stellenwert anderer Forschungstätigkeiten.....	173
Tabelle 26: Korrelationen zwischen der Variablen „Einsatz Forschungsindikator in Steuerungs- und Anreizverfahren insgesamt“ und der intrinsischen Forschungsmotivation	177
Tabelle 27: Korrelationen zwischen dem Stellenwert verschiedener Forschungstätigkeiten und den Variablen Alter, Beschäftigungsdauer, Besoldungsart und Geschlecht	178
Tabelle 28: Multiple Regression mit der abhängigen Variablen „Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren“	192
Tabelle 29: Bedingte Effekte des Einsatzes des Indikators „Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren“.....	194
Tabelle 30: Multiple Regression mit der abhängigen Variablen „Stellenwert des Publizierens in Zeitschriften ohne Peer-Review-Verfahren“.....	196
Tabelle 31: Multiple Regression mit der abhängigen Variablen „Stellenwert des Verfassens anderer Publikationen“	198
Tabelle 32: Multiple Regression mit der abhängigen Variablen „Stellenwert des Publizierens in Zeitschriften mit hohem Impact Factor / in high ranked journals“	201
Tabelle 33: Multiple Regression mit der abhängigen Variablen „Stellenwert der Betreuung von Promotionen“	203
Tabelle 34: Multiple Regression mit der abhängigen Variablen „Stellenwert des Einwerbens von DFG-Drittmitteln“	205
Tabelle 35: Multiple Regression mit der abhängigen Variablen „Stellenwert des Einwerbens von sonstigen Drittmitteln“	208

1. Einführung

Im Juni 2016 beschlossen Bund und Länder als Nachfolge zur Exzellenzinitiative die Einrichtung eines zeitlich unbegrenzten Programms zur Förderung der Spitzenforschung an deutschen Universitäten. Unter dem Namen *Exzellenzstrategie*¹ werden ab dem Jahr 2019 in einem siebenjährigen Turnus dauerhaft zusätzliche Finanzmittel in der Förderlinie *Exzellenzcluster* vergeben. Gefördert werden sollen hierbei international wettbewerbsfähige Forschungsfelder von einzelnen Universitäten oder Universitätsverbänden. Universitäten, die erfolgreiche Exzellenzcluster etabliert haben, können über die zweite Förderlinie des Programms – in Form von sogenannten *Exzellenzuniversitäten* – dauerhaft zusätzliche Gelder akquirieren, wenn sie sich in einem entsprechenden Auswahlverfahren gegen die anderen Mitbewerber durchsetzen.

Die politische Initiierung eines auf Dauer angelegten Wettbewerbs zwischen den Universitäten um Exzellenz und den damit verbundenen Fördergeldern kann als ein weiteres sichtbares Symptom der zunehmenden Ökonomisierungstendenzen in der deutschen Hochschullandschaft interpretiert werden. Der Bedeutungszuwachs ökonomischer Prinzipien wie Leistungsmessung und der damit verbundenen Vergleichslogik sowie der Schaffung von Konkurrenzdruck spiegelte sich in den vergangenen Jahren bereits in der Bologna-Reform, der Zunahme an Akkreditierungsprozessen im Hochschulbereich oder der wachsenden Popularität und Relevanz von Hochschulrankings wider (vgl. Bogumil et al. 2013; Teichler 2014; Tippelt et al. 2010). In den Universitäten selbst äußert sich der zunehmende Wettbewerb in Form der Etablierung verschiedener output-orientierter Steuerungs- und Anreizverfahren. Solche Verfahren, wie z. B. die indikatorbasierte Mittelzuweisung, versuchen die Forschungsleistung von Fakultäten, Arbeitsgruppen oder Lehrstühlen über verschiedene Forschungsindikatoren wie die Höhe der eingeworbenen Drittmittel oder die Anzahl der Publikationen zu erfassen. So sollen Anreize für ein erhöhtes Forschungsengagement geschaffen werden, indem in Abhängigkeit von der erbrachten Forschungsleistung beispielsweise zusätzliche Finanzmittel an die Fakultät, das Institut oder den Lehrstuhl vergeben werden. Mit dem Einsatz der Verfahren kann zudem der Versuch verbunden sein, das Forschungsverhalten strategisch in eine bestimmte Richtung zu steuern. Ein Verfahren kann z. B. so ausgestaltet sein, dass nur die Erhöhung der Anzahl von Publikationen in Zeitschriften mit Peer-Review-Verfahren mit der Vergabe zusätzlicher Mittel belohnt wird, während Publikationen in Zeitschriften ohne Peer-Review-Verfahren bei der Mittelvergabe nicht berücksichtigt

¹ Informationen zur Exzellenzstrategie von Bund und Ländern unter <https://www.bmbf.de/de/die-exzellenzstrategie-3021.html> (BMBF), <http://www.dfg.de/foerderung/exzellenzstrategie/index.html> (DFG) sowie <https://www.wissenschaftsrat.de/arbeitsbereiche-arbeitsprogramm/exzellenzstrategie.html> (Wissenschaftsrat) (29.05.2017).

werden. Die von dem Einsatz des Verfahrens betroffenen Professorinnen und Professoren sollten dann entsprechend der Zielsetzung des Steuerungsversuches vermehrt bestrebt sein, Publikationen in Peer-Review-Formaten zu veröffentlichen.

Auch wenn die Hochschulforschung die Thematik indikatorenbasierter Steuerungsversuche im Hochschulbereich aus unterschiedlichen fachlichen Perspektiven in den Blick nimmt, ist die Wirkungsweise dieser Verfahren in Bezug auf das Forschungsverhalten von Wissenschaftlerinnen und Wissenschaftlern bislang nur unzureichend empirisch untersucht worden (vgl. unter anderem Kapitel 2.4). In dieser Arbeit sollen daher – mit dem Fokus auf die sozialwissenschaftlichen Fächer Erziehungswissenschaften und Sonderpädagogik, Politikwissenschaften, Psychologie, Soziologie sowie Wirtschaftswissenschaften – mögliche Zusammenhänge zwischen dem Versuch der hochschulinternen indikatorenbasierten Steuerung und dem Forschungsverhalten von Professorinnen und Professoren untersucht werden. Der Bereich der Lehre, der ebenfalls Gegenstand indikatorenbasierter Steuerungsversuche sein kann (vgl. Bogumil et al. 2013), ist dabei nicht Teil der Analyse. Stattdessen liegt das Erkenntnisinteresse dieser Arbeit primär auf dem Bereich der Forschungsarbeit. Der Begriff *indikatorenbasierte Steuerung* bezieht sich hier also auf den Einsatz verschiedener Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren (indikatorenbasierte Mittelzuweisung, Zielvereinbarungen, Benchmarking, Bleibe- und Berufungsverhandlungen und Leistungszulagen im Rahmen der W-Besoldung). Die folgenden acht Forschungsindikatoren werden bei der Analyse berücksichtigt:

- Anzahl Publikationen **mit** Peer-Review-Verfahren
- Anzahl Publikationen **ohne** Peer-Review-Verfahren
- Anzahl anderer Publikationen
- Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals
- Anzahl der Zitationen / Zitationsindizes
- Anzahl betreuter Promotionen
- Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft (DFG)
- Höhe der sonstigen eingeworbenen Drittmittel

Untersucht werden soll nun zum einen, inwieweit der Einsatz von Forschungsindikatoren und der verschiedenen indikatorenbasierten Steuerungs- und Anreizverfahren mittlerweile Verbreitung in den sozialwissenschaftlichen Fachbereichen an deutschen Universitäten gefunden hat. Es wird also erhoben, welcher Anteil der sozialwissenschaftlichen Professorinnen und Professoren im Zuge ihrer Forschungsarbeit mit indikatorenbasier-

ten Steuerungsversuchen konfrontiert ist. Hierbei ist auch die Ausgestaltung der einzelnen Verfahren von Interesse. Überprüft wird daher zudem, welche Forschungsindikatoren mit welcher Gewichtung in den jeweiligen Verfahren zum Einsatz kommen.

Zum anderen liegt das Erkenntnisinteresse der vorliegenden Studie in der Analyse möglicher Zusammenhänge zwischen dem Versuch der indikatorenbasierten Steuerung und dem Forschungsverhalten der Professorinnen und Professoren. Zu berücksichtigen sind hierbei die Spezifika der akademischen Profession als eine ‚besondere‘ Berufsgruppe sowie die Spezifika der Universität als eine spezielle Form der Organisation. Dies betrifft beispielsweise Aspekte wie den Stellenwert der fachspezifischen Scientific Communities als Orientierungspunkt für die Forschungsarbeit oder die mit der Organisationsstruktur von Universitäten verknüpfte Handlungsautonomie der Professorinnen und Professoren, welche die Wirksamkeit universitärer Steuerungsversuche beeinträchtigen können. Von Interesse sind zudem mögliche nichtintendierte Effekte der externen Steuerungsversuche auf die intrinsische Forschungsmotivation der Professorinnen und Professoren.

Zur Analyse der hier skizzierten Themenkomplexe wurden mit Hilfe einer quantitativen Online-Befragung von sozialwissenschaftlichen Professorinnen und Professoren an deutschen Universitäten unter anderem deren Angaben und Einschätzungen zu den folgenden Fragestellungen erfasst:

- Einsatz und Gewichtung von Forschungsindikatoren insgesamt und in den einzelnen hochschulinternen Steuerungs- und Anreizverfahren
- Stellenwert verschiedener Forschungstätigkeiten in der eigenen Forschungsarbeit (Forschungsverhalten)
- Bedeutung verschiedener Forschungstätigkeiten für den Reputationszuwachs in der Scientific Community
- Beitrag verschiedener Forschungstätigkeiten für den wissenschaftlichen Erkenntnisgewinn
- Ausprägung der intrinsischen Forschungsmotivation der Befragten

Unter Verwendung multivariater Analyseverfahren (multiple Regression) sollen mit den erhobenen Daten im Detail die folgenden Forschungsfragen beantwortet werden:

1. Lässt sich ein statistisch signifikanter positiver Zusammenhang zwischen dem Einsatz eines Forschungsindikators (z. B. Höhe der eingeworbenen Drittmittel der DFG) in hochschulinternen Steuerungs- und Anreizverfahren und dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit (z. B. Einwerben von Drittmitteln der DFG) in der Forschungsarbeit der Professorinnen und Professoren nachweisen?

2. Welche Bedeutung hat die Orientierung an der Scientific Community (Reputation) für das Forschungsverhalten? Zeigen sich statistisch signifikante positive Zusammenhänge zwischen der eingeschätzten Bedeutung einer Forschungstätigkeit für den Reputationszuwachs in der Scientific Community und dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit?
3. Welche Relevanz kommt dem wissenschaftlichen Erkenntnisgewinn für die Forschungsarbeit der Professorinnen und Professoren zu? Ergibt sich ein statistisch signifikanter positiver Zusammenhang zwischen dem eingeschätzten Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn und dem Stellenwert der Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren?
4. Kommt es durch den Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren zur Verdrängung autonomer Formen der Motivation?

Den theoretischen Hintergrund für diese Fragestellungen bildet ein Makro-Meso-Mikro-Analyserahmen. Auf der Makroebene werden die Ökonomisierungstendenzen in der deutschen Hochschullandschaft aus einer governancetheoretischen Perspektive in den Blick genommen. Hierbei steht der Wandel des Governance-Regimes der Hochschulen im Mittelpunkt. Von besonderer Relevanz für die Fragestellungen dieser Arbeit ist dabei der zunehmende Wettbewerb im Hochschulbereich, der sich innerhalb der Hochschulen (auch) in der Implementierung von indikatorenbasierten Steuerungs- und Anreizverfahren widerspiegelt. Auf der Mesoebene werden organisations- und professionstheoretische Ansätze herangezogen, die sich sowohl mit den Spezifika der akademischen Profession als auch mit denen der Organisation Universität auseinandersetzen. Dabei wird herausgearbeitet werden, dass sich hochschulinterne Steuerungsversuche mit einer stark ausgeprägten Steuerungsresistenz der akademischen Profession konfrontiert sehen. Die Mikroebene der handelnden Professorinnen und Professoren wird unter Verwendung motivationstheoretischer Ansätze analysiert. Aus dieser motivationstheoretischen Perspektive heraus werden schließlich – unter Berücksichtigung der theoretischen Überlegungen auf der Makro- und Mesoebene – die konkreten zu testenden Forschungshypothesen abgeleitet und überprüft, welche internalen und externalen Faktoren für das Forschungsverhalten als motivational bestimmend angesehen werden können.

Zum Abschluss dieses Einführungskapitels wird im Folgenden nun noch ein kurzer Überblick über den Aufbau der Arbeit und den Inhalt der einzelnen Kapitel gegeben.

In **Kapitel 2** wird zunächst die Makroebene mit einer grundlegenden Beschreibung der Governance von Hochschulen betrachtet. Mit Hilfe des Equalizer-Modells von Schimank (2007) wird dann der im Zuge des *New Public Managements* zu beobachtende Governance-Wandel im Hochschulbereich analysiert. Im Anschluss folgt die Darstellung

und kritische Diskussion hochschulinterner Steuerungs- und Anreizverfahren und der zur Leistungsmessung verwendeten Forschungsindikatoren als Elemente des zunehmenden Wettbewerbs in der deutschen Hochschullandschaft. Das Kapitel schließt mit einer Darstellung aktueller empirischer Erkenntnisse zur Umsetzung und Wirkung indikatorenbasierter Steuerungs- und Anreizverfahren.

Kapitel 3 befasst sich mit der Mesoebene der Organisation Universität sowie der akademischen Profession. Wie bereits angemerkt, soll die Herausarbeitung der organisationalen und professionellen Spezifika dabei als argumentative Grundlage für die der akademischen Profession innewohnende Steuerungsresistenz dienen. Organisationstheoretisch werden dabei drei verschiedene Ansätze herangezogen, die die spezielle Organisationsstruktur von Universitäten als *loosely coupled systems* (Weick 1976), *Profibürokratien* (Mintzberg 1992) und *organisierte Anarchien* (Cohen et al. 1972) charakterisieren. Die professionstheoretischen Überlegungen stützen sich überwiegend auf die Arbeiten von Parsons (1969; Parsons und Platt 1990).

Der motivationstheoretische Rahmen zur Analyse der Mikroebene der handelnden Professorinnen und Professoren ist Gegenstand von **Kapitel 4**. Mit der *Selbstbestimmungstheorie der Motivation* (Deci und Ryan 1993), der *Motivation Crowding Theory* (Frey und Jegen 2001) und dem Konzept der *Public Service Motivation* (Perry und Wise 1990) werden zunächst drei motivationstheoretische Ansätze vorgestellt, die dann zu einem theoretischen Gesamtmodell verknüpft werden. Aus diesem werden schließlich die sechs zu überprüfenden Forschungshypothesen abgeleitet.

In **Kapitel 5** folgen die Darstellung des quantitativ ausgerichteten Forschungsdesigns sowie die Beschreibung der Entwicklung und Umsetzung des Erhebungsinstrumentes in Form eines standardisierten Online-Fragebogens. Im Anschluss wird das Vorgehen bei der Datenerhebung erläutert und Kennzahlen zum Rücklauf und zur Repräsentativität der Stichprobe berichtet.

Kapitel 6 widmet sich der ausführlichen Datenanalyse und der Ergebnisdarstellung. Fächervergleichende deskriptive Analysen geben einen Überblick über die Merkmale der Stichprobe und die Art und Häufigkeit des Einsatzes von Forschungsindikatoren und hochschulinternen Steuerungs- und Anreizverfahren. Weiter werden im Rahmen der fächervergleichenden deskriptiven Analysen verschiedene Verteilungskennzahlen für die zentralen Variablen dieser Studie (Bedeutung verschiedener Aspekte der Forschungstätigkeit für den Reputationszuwachs in der Scientific Community sowie deren Beitrag für den wissenschaftlichen Erkenntnisgewinn, eingeschätzter Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit, intrinsische Forschungsmotivation) präsentiert. Die deskriptiven Analysen schließen mit der zusammenfassenden Darstellung von Anmerkungen und Kommentaren der Befragten.

In den darauf folgenden bivariaten Analysen werden Korrelationen zwischen ausgewählten, insbesondere für die nachfolgenden multivariaten Analysen relevanten Variablen berechnet und damit ein erster Überblick über die Zusammenhangsstruktur und über mögliche Abhängigkeiten der Variablen untereinander gegeben. Nach der Beschreibung des generellen Vorgehens zur Berechnung der multiplen Regressionsmodelle folgt die Ergebnisdarstellung der Regressionsanalysen. Dabei werden die Ergebnisse modell-spezifisch berichtet. Das heißt, für jede untersuchte Forschungstätigkeit (z. B. abhängige Variable: Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren in der Forschungsarbeit) wurde ein eigenes Regressionsmodell gerechnet. Abgeschlossen wird das Kapitel mit der Interpretation und zusammenfassenden Diskussion der Ergebnisse.

In **Kapitel 7** wird das methodische Vorgehen dieser Studie kritisch reflektiert und es werden mögliche Schwächen diskutiert. Soweit möglich werden zudem Handlungsempfehlungen für ähnlich gelagerte Forschungsprobleme gegeben. Das Ende der Arbeit bildet die Darstellung möglicher Forschungsdesiderata sowie die Diskussion der Ergebnisse im Hinblick auf die Hochschulpraxis.

2. New Public Management und die Steuerung der Hochschulen

2.1 New Public Management als neues Governance-Regime der Hochschulen

2.1.1 Die Governance der Hochschulen

Das Konzept der *Governance* findet in verschiedenen wissenschaftlichen Disziplinen und auch in der – beispielsweise politischen oder ökonomischen – Praxis breite Anwendung (vgl. Benz et al. 2007). Mit der Pluralität der Anwendungsbereiche ist auch eine Pluralität von mehr oder weniger klaren Definitionen des Begriffs *Governance* verbunden, so dass eine begriffliche Festlegung hier zwingend notwendig erscheint. Für die Forschungsperspektive dieser Arbeit wird der Begriff *Governance* daher zunächst in Form einer recht allgemein gehaltenen Definition näher bestimmt und im Anschluss durch die Darstellung der für Hochschulsysteme charakteristischen *Governance*-Mechanismen weiter spezifiziert. Darauf aufbauend wird dann im nächsten Teilkapitel mit Hilfe des von Schimank und anderen entwickelten Equalizer-Modells (vgl. Boer et al. 2007; Schimank 2007; Wiesenthal 2000) der Wandel des *Governance*-Regimes im Hochschulsektor skizziert und auf die zentralen Merkmale des *New Public Managements* eingegangen.

Mit Benz et al. (2007) lässt sich der Begriff *Governance* in Abgrenzung zu rein staatlichen bzw. rein hierarchisch strukturierten Steuerungs- und Regelungsmechanismen und unter Berücksichtigung anderer relevanter gesellschaftlicher Akteure wie folgt definieren: *Governance* „(...) steht für alle diese Formen und Mechanismen der Koordination zwischen mehr oder weniger autonomen Akteuren, deren Handlungen interdependent sind, sich also wechselseitig beeinträchtigen oder unterstützen können“ (ebd., S. 9). Die „(...) *Governance*-Perspektive [zielt] darauf, Strukturen, Mechanismen und Wirkungen der Bewältigung von Interdependenz zwischen individuellen, kollektiven oder korporativen Akteuren zu beleuchten“ (ebd., S. 18). Als Beispiel für die in dieser Definition von *Governance* enthaltenen Berücksichtigung der verschiedenen beteiligten Akteure und deren Interdependenzen untereinander kann die recht anschauliche Erläuterung von Alt-richter (2015) zur *Governance* des Schulsystems dienen:

The term “governance” (...) indicates that we assume that school systems – and their reform – are not shaped by a single dominant actor, e.g. by the government and its administrative staff. More actors are involved in the formation of a system. In order to make innovations work, teachers and school leaders must take innovative ideas on board, and they must translate them into actions and organisational arrangements. Students must understand the innovation and reshape at least partially their actions

(and they may need some support and the understanding of their parents). Intermediary institutions, such as the inspectorate, textbook publishers, professional development institutions must act in some accordance. (ebd., S. 26)

Diese Überlegungen lassen sich in ähnlicher Weise leicht auf das Hochschulsystem übertragen: Auch hier sind staatliche (z. B. Ministerien), privatwirtschaftliche (z. B. Verlage zur Veröffentlichung von Forschungsergebnissen, Drittmittelgeber) und eine Reihe weiterer gesellschaftlicher Akteure bzw. Stakeholder (Studierende, Scientific Community, Hochschulleitungen und nachfolgende Hierarchieebenen, Alumni etc.) in ihren Handlungen miteinander verstrickt. Aus der Governance-Perspektive ist nun von zentraler Bedeutung, auf welche Weise die Koordination der Handlungen der beteiligten Akteure erfolgt bzw. erfolgen kann, um die intendierten Effekte (z. B. Lernerfolg von Studierenden, Forschungsleistung von Wissenschaftlerinnen und Wissenschaftlern) zu erzielen. Dabei ist nicht nur die Erreichung bzw. Nicht-Erreichung intendierter Effekte von Interesse, sondern auch die Analyse möglicher nichtintendierter Effekte von Steuerungs- und Regelungsmechanismen (Schrader et al. 2015, S. 9). Zu berücksichtigen ist zudem der Mehrebenencharakter von Governance (vgl. Altrichter und Heinrich 2007, S. 64). In Bezug auf das Hochschulsystem muss die Handlungskoordination beispielsweise zwischen Akteuren auf staatlicher Ebene (Makro), auf der Ebene der Organisation Hochschule (Meso) sowie auf der Mikroebene der einzelnen Wissenschaftlerinnen und Wissenschaftler erfolgen (vgl. Schimank 2008, S. 623). Dabei lässt sich hinsichtlich des Governance-Regimes² des Hochschulsystems die Handlungskoordination zwischen den beteiligten Akteuren anhand von fünf Mechanismen beschreiben (die folgende Beschreibung der fünf Governance-Mechanismen orientiert sich überwiegend an Schimank 2009, S. 125):

- staatliche Regulierung
- externe Steuerung
- akademische Selbstorganisation
- hierarchische Selbststeuerung
- Konkurrenzdruck

Die *staatliche Regulierung* bezieht sich auf die Gesetzgebungskompetenzen des Staates, die mit klaren (Handlungs-)Vorgaben an die Hochschulen verbunden sind. *Externe Steuerung* kann ebenfalls durch den Staat aber auch durch andere vom Staat delegierte gesellschaftliche Stakeholder erfolgen. Im Gegensatz zur staatlichen Regulierung erfolgt

² Graß (2015) beschreibt ein Governance-Regime als „(...) ein analytisches Konstrukt, mit dem aus einer Analyse der Vielzahl von Koordinationsbeziehungen jene wenigen Prinzipien herausgearbeitet werden, die als charakteristisch für die Koordinationsverhältnisse in einem zeitlich und örtlich situierten Feld gelten können“ (ebd., S. 71).

aber keine Detailsteuerung des Handelns der Hochschulen, sondern es werden nur generelle Ziele formuliert. Welche Maßnahmen zur Zielerreichung eingesetzt werden, bleibt dabei den Hochschulen überlassen. Die *akademische Selbstorganisation* umfasst neben der hochschulinternen Selbstverwaltung auch die akademische Selbstverwaltung außerhalb der Hochschulen durch die fachspezifischen Scientific Communities. Schimank nennt in diesem Zusammenhang insbesondere die peer-review-basierten Entscheidungsprozesse in der Forschungsförderung (ebd.). Ergänzend hierzu lassen sich Gutachterprozesse bei wissenschaftlichen Veröffentlichungen sowie die Durchführung von hochschulübergreifenden oder sogar länderübergreifenden Tagungen als Teil der akademischen Selbstorganisation sehen. Die *hierarchische Selbststeuerung* der Hochschulen bezieht sich auf die formale hochschulinterne Organisationsstruktur in Form von Hochschulleitungen auf zentraler Ebene und den Dekanaten auf Fachbereichsebene. Der Governance-Mechanismus *Konkurrenzdruck* schließlich betrifft Wettbewerbssituationen in und zwischen Hochschulen. Hochschulen bzw. Fachbereiche können sich dabei im Wettbewerb um Studierende, Personal, finanzielle Ressourcen (Drittmittel) und nicht zuletzt um Reputation (z. B. Platzierung in Rankings) befinden.

Das auch als Selbstverwaltungsmodell bezeichnete (Bogumil et al. 2013, S. 24) traditionelle Governance-Modell deutscher Hochschulen ist (bzw. war) geprägt durch einen hohen Grad an staatlicher Regulierung und einer stark ausgeprägten akademischen Selbstorganisation, während der externen Steuerung, der hierarchischen Selbststeuerung sowie dem Konkurrenzdruck als Governance-Mechanismen nur eine untergeordnete Bedeutung zukommt (bzw. zukam) (Schimank 2007, S. 239). Der hohe Stellenwert der staatlichen Regulierung resultiert in diesem Modell nicht zuletzt aus der finanziellen Abhängigkeit der Universitäten von staatlichen Mitteln, „welche es dem Staat ermöglicht, detaillierte Vorgaben im Hinblick auf die Nutzung von Mitteln seitens der Universitäten zu machen“ (Bogumil et al. 2013, S. 25). Unberührt von den staatlichen Regulierungen bleiben allerdings die Kernaufgaben Forschung und Lehre, deren Ausgestaltung allein der akademischen Profession obliegt. Dies wird durch die im Grundgesetz festgelegte Wissenschaftsfreiheit garantiert (ebd., S. 26). Trotz der finanziellen Abhängigkeit und der damit verbundenen Entscheidungsbefugnis des Staates in Haushalts- und Personalangelegenheiten (ebd.) verfügt die akademische Profession im Selbstverwaltungsmodell damit über einen hohen Grad an Handlungsautonomie.

Inwieweit dieses traditionelle Governance-Regime aufgrund eines zunehmenden transnationalen und innenpolitischen Veränderungsdrucks (Dobbins und Knill 2016, S. 33) von Reformbemühungen betroffen ist und ob bzw. in welcher Form sich ein Wandel bei den beschriebenen Governance-Mechanismen im deutschen Hochschulsystem beobachten lässt, wird im nun folgenden Kapitel näher untersucht.

2.1.2 New Public Management im Hochschulbereich

Das Selbstverwaltungsmodell als traditionelles Governance-Regime der Hochschulen geriet in den letzten Jahren zunehmend unter Kritik. Diese Kritik bezieht sich unter anderem auf eine mangelnde internationale Wettbewerbsfähigkeit, auf wirtschaftlich ineffizientes Arbeiten aufgrund einer inputorientierten Kameralistik sowie auf eine fehlende Anwendungsorientierung von Forschung und Lehre (Bogumil et al. 2013, S. 26). Hinsichtlich der hochschulinternen akademischen Selbstverwaltung (Kollegialorgane wie Hochschulsenate und Fachbereichsräte) werden die mit hohen Transaktionskosten verbundenen Entscheidungsprozesse kritisiert, die bei gegensätzlichen Interessenlagen durch Entscheidungseffizienz und Entscheidungseffektivität geprägt seien (ebd., S. 27; Trute und Pilniok 2009, S. 27). Nicht zuletzt durch den Bologna-Prozess „als europäische Reaktion auf Entwicklungen wie das Aufkommen der Wissensökonomie, den demografischen Wandel und die Globalisierung“ (Dobbins und Knill 2016, S. 43) entsteht ein erhöhter Reformdruck mit grundlegenden Auswirkungen auf die Hochschulgovernance (vgl. ebd., S. 43–44).

Vor dem Hintergrund dieser Kritik und im Zuge des unter dem Schlagwort *New Public Management* gefassten grundsätzlichen Wandels des Governance-Regimes im öffentlichen Sektor, eröffnete die Novellierung des Hochschulrahmengesetzes im Jahr 1998 auch im Hochschulbereich die Möglichkeiten zu tiefgreifenden Veränderungen in der staatlichen und hochschulinternen Steuerung (vgl. Schimank 2008, S. 624; Trute und Pilniok 2009, S. 22). An welchen Governance-Mechanismen die Änderungen in welcher Weise ansetzen (sollen), lässt sich mit dem bereits erwähnten von Schimank (2007) entwickelten Equalizer-Modell erläutern. Dieses auf den Vorarbeiten von Braun (1999) und Clark (1986; 1998) basierende Modell begreift die einzelnen Governance-Mechanismen als Schieberegler, die je nach Ausprägung graduell verschoben werden können (vgl. Schimank 2007, S. 239–240). Der Equalizer kann damit als Analyserahmen für Governance-Regime unterschiedlichster Konfigurationen genutzt werden.

Das traditionelle Governance-Regime der Hochschulen ist wie erwähnt (vgl. Kapitel 2.1.1) durch die beiden vorherrschenden Governance-Mechanismen der staatlichen Regulierung und der akademischen Selbstorganisation gekennzeichnet. Im Rahmen des Regime-Wandels hin zum New Public Management soll sich die Konfiguration der Governance-Mechanismen nun in der folgenden Weise verschieben: Während die Mechanismen der staatlichen Regulierung und der akademischen Selbstorganisation nach ‚unten‘ gefahren werden, werden die Schieberegler für die externe Steuerung, die hierarchische Selbststeuerung sowie für Konkurrenzdruck bzw. Wettbewerb nach ‚oben‘ gefahren (Schimank 2009, S. 131–132). Dabei sind entsprechend der Logik des Equalizer-Modells graduelle Unterschiede bei der Positionierung der Schieberegler denkbar. Das heißt, die Governance-Regime verschiedener Hochschulen können z. B. je nach

Bundesland durchaus anders akzentuiert sein. Im Detail sind mit der Neuausrichtung des Governance-Regimes die folgenden mehr oder weniger weit vorangeschrittenen Veränderungen im deutschen Hochschulsystem empirisch beobachtbar (vgl. für eine detaillierte empirische Analyse zum Umsetzungsstand auch die Studie von Bogumil et al. 2013; Heinze et al. 2011):

Der Mechanismus der *staatlichen Regulierung* wird durch Deregulierungsmaßnahmen heruntergefahren. Die kameralistisch geprägte Detailsteuerung der Länder hinsichtlich von Personalentscheidungen und der Mittelverwendung wird ersetzt durch eine outputorientierte Steuerung mittels Globalbudgets³ und Zielvereinbarungen zwischen den Ländern und den Hochschulen (Bogumil und Heinze 2009a, S. 8). Während die Änderung des Haushaltsrechts in Form der Etablierung von Globalbudgets relativ weit vorangeschritten ist, sind die Deregulierungsmaßnahmen hinsichtlich des Personalrechts (Stellenpläne, Dienstrecht) allerdings weniger weit fortgeschritten (vgl. Bogumil et al. 2013, S. 98–102; Schimank 2009, S. 128–129).

Insbesondere durch die Implementierung von Zielvereinbarungen als Steuerungsinstrument gewinnt der Mechanismus der *externen Steuerung* im Governance-Regime des Hochschulsystems an Bedeutung. Zielvereinbarungen können sich dabei auf verschiedenste Aspekte der universitären Forschung oder Lehre beziehen (z. B. Internationalisierung, Studiendauer etc.; vgl. zu Zielvereinbarungen auch Kapitel 2.2.2). Ebenfalls zur externen Steuerung kann die Einrichtung von Hochschulräten gezählt werden. Welche Konstellation von Akteuren als ‚Hochschulrat‘ klassifiziert werden kann, ist dabei nicht immer ganz eindeutig (vgl. Bogumil et al. 2013, S. 47). Als wesentliche Merkmale können zumindest die Verortung des Gremiums unterhalb der ministeriellen Ebene sowie die Beteiligung von universitätsexternen Mitgliedern an hochschulrelevanten Entscheidungsprozessen bestimmt werden (vgl. ebd.). Die Hochschulräte können sich aus Vertretern der Politik und Wirtschaft oder anderer gesellschaftlicher Teilbereiche sowie aus Mitgliedern der Hochschulen selbst zusammensetzen. Die Ausgestaltung der Hochschulräte differiert dabei erheblich zwischen den Ländern (Schimank 2009, S. 129–130) und selbst die Benennung des Gremiums unterscheidet sich je nach Bundesland (z. B. Aufsichtsrat, Landeshochschulrat, Kuratorium) (Bogumil et al. 2013, S. 47). Auch wenn die dem Hochschulrat zugesprochenen Kompetenzen je nach Bundesland teils sehr unterschiedlich ausfallen (vgl. ebd., S. 87–97), kann festgehalten werden, „dass der Hochschulrat über signifikante formale Einflussmöglichkeiten auf die strategische Ausrichtung und teilweise sogar das operative Geschäft von Universitäten verfügt“ (ebd., S. 91). Dies

³ „Bei einem Globalbudget erhalten die dezentralen Einheiten ein festes Budget, dessen Verausgabung nicht von der Zentrale vorgegeben wird. Es ist also nicht nach einzelnen Titeln gegliedert, wie ein kameralistisches System, in dem die Zentrale den dezentralen Einheiten detailliert vorschreiben kann, wofür Finanzmittel im Einzelnen zu verausgaben sind“ (Bogumil et al. 2013, S. 31).

kann als deutliche Stärkung des Mechanismus der externen Steuerung interpretiert werden.

Eine Steigerung des Mechanismus *Konkurrenzdruck* bzw. des hochschulinternen sowie hochschulübergreifenden Wettbewerbs kann als das zentrale Element im Governance-Regime des New Public Managements gesehen werden. Um den Schieberegler dieses Governance-Mechanismus nach oben zu fahren, wurde in den letzten Jahren eine Reihe von Maßnahmen initiiert. Zu nennen ist hier insbesondere die 2005 ins Leben gerufene Exzellenzinitiative⁴. Diese kann als deutschlandweiter Wettbewerb der Universitäten um Forschungsfördergelder verstanden werden, um innovative Forschungskonzepte und Spitzenforschung zu ermöglichen und die deutsche universitäre Forschung auch international wettbewerbsfähig zu machen (vgl. Jansen 2009, S. 37). Der steigende Konkurrenzdruck wird weiter durch den Bedeutungszuwachs internationaler und nationaler Hochschulrankings deutlich. Vor allem in den internationalen Rankings werden die Hochschulen dabei meist als ganzheitliche Organisation behandelt und miteinander verglichen. Aufgrund großer fachspezifischer Unterschiede hinsichtlich der Ausprägungen der in den Rankings verwendeten Indikatoren (vgl. Kapitel 2.2.4 und 2.3) ist die fehlende Berücksichtigung der fachlichen Profile der Hochschulen als methodisch höchst zweifelhaft zu werten und wird auch vielfach kritisiert (vgl. Federkeil 2013, S. 43; Lenzen 2012, S. 17–18). Hochschulrankings können einerseits als Repräsentant für die gestiegene Bedeutung des Wettbewerbs gewertet werden und ihn somit lediglich abbilden, andererseits können sie durch eine vermehrte Fokussierung der Hochschulen auf Rankingergebnisse selbst Motor des Wettbewerbs sein. Diese den Hochschulrankings innewohnende Dualität beschreibt Federkeil (2013) wie folgt:

Auf der einen Seite ist das Entstehen globaler Rankings und die hohe Aufmerksamkeit, die sie finden, *Ausdruck* eines zunehmenden internationalen Wettbewerbs zwischen Hochschulen (um Reputation, Talente und Ressourcen). Gleichzeitig sind die Rankings aber auch *Medium* dieses Wettbewerbs; sie reproduzieren und verstärken ihn durch ihre eigenen Ergebnisse. Dies wird z. B. daran deutlich, dass sich Maßnahmen zur Förderung von Spitzenforschungsuniversitäten in vielen Ländern explizit auf das Abschneiden in den weltweiten Rankings beziehen. (ebd., S. 35–36)

In Anbetracht der Entwicklungen der letzten Jahre kann angenommen werden, dass Rankings zunehmend als Medium des Wettbewerbs fungieren und Hochschulen aktiv

⁴ Detailliertere Informationen zu den Förderlinien Graduiertenschulen, Exzellenzcluster und Zukunftskonzepte sowie zu den Fördervolumina unter <http://www.dfg.de/foerderung/programme/exzellenzinitiative/> (02.11.2016).

versuchen, den Rankingkriterien (besser) zu entsprechen.⁵ Betrachtet man beispielsweise die Außendarstellung der Universitäten auf ihren Webseiten, zeigt sich anhand der regelmäßigen Berichte über das (gute) Abschneiden in verschiedenen Hochschulrankings, welche Rolle Rankingplatzierungen mittlerweile im Reputationsmanagement der Hochschulen einnehmen.

Weitere Instrumente, die den Governance-Mechanismus Konkurrenzdruck stärken, stellen die indikatorenbasierte Mittelzuweisung und die Leistungszulagen im Rahmen der W-Besoldung dar. Beide Verfahren werden im Detail in Kapitel 2.2.1 und 2.2.3 dargestellt. An dieser Stelle soll nur darauf hingewiesen werden, dass die zugrundeliegende Logik der Verfahren darin besteht, zusätzliche Finanzmittel in Abhängigkeit der erbrachten Leistung – gemessen anhand des Outputs hinsichtlich verschiedener Indikatoren wie Drittmittelhöhe oder Publikationsanzahl – zu vergeben. Die indikatorenbasierte Mittelzuweisung kann dabei als outputorientiertes Steuerungsinstrument neben dem Governance-Mechanismus Konkurrenzdruck auch dem Governance-Mechanismus der externen Steuerung zugerechnet werden. Für die W-Besoldung trifft dies nicht zu, da die Aushandlung der Leistungszulagen hochschulintern erfolgt.

In Bezug auf die Mechanismen der *hierarchischen Selbststeuerung* und der *akademischen Selbstorganisation* analysierten Bogumil et al. (2013) die bestehenden Kompetenzen von Leitungs- (z.B. Rektorat) gegenüber Kollegialorganen (z.B. Senat) auf Ebene der verschiedenen Bundesländer. Hierbei konnte eine Stärkung der hierarchischen Selbststeuerung verbunden mit einer Schwächung der akademischen Selbstorganisation – insbesondere bei Finanzfragen – beobachtet werden (ebd., S. 103–110). Allerdings sind diese ermittelten Kompetenzverschiebungen zunächst rein formaler Natur und sagen noch nichts über die tatsächlich ablaufenden Entscheidungsprozesse aus (vgl. ebd., S. 112). Schimank attestiert den Reformbemühungen entsprechend bislang nur geringe Fortschritte. Die Kompetenzen der Leitungsorgane seien zwar innerhalb der Landeshochschulgesetze gestärkt und die Rechte der Fachbereichsräte und Senate beschränkt worden, de facto aber würden die Entscheidungsprozesse nach einem ähnlichen Muster wie bisher ablaufen (vgl. Schimank 2009, S. 131). Dies führt Schimank auf die Dominanz der bisherigen Konsens-Kultur der akademischen Selbstorganisation zurück, in welcher sowohl die Leitungsfiguren als auch die Professorenschaft sozialisiert wurden (ebd.). Eine rein formale Kompetenzverschiebung durch entsprechende Anpassung der Landeshochschulgesetze wäre damit nicht ausreichend, um die Schieberegler

⁵ Vgl. hierzu auch das vom Auswärtigen Amt geförderte Pilotprojekt "Verbesserung internationaler Rankingergebnisse deutscher Universitäten – Die Technische Universität Dresden und die Universität Tübingen als Vorreiter für das internationale Bildungsmarketing des Standorts Deutschland" (<http://www.uni-tuebingen.de/aktuelles/newsletter-uni-tuebingen-aktuell/2014/4/forschung/1.html>) (02.11.2016).

des Governance-Equalizers auf die im Sinne des New Public Managements gewünschte Position zu bringen.

Zusammenfassend lässt sich festhalten, dass insbesondere der Governance-Mechanismus Konkurrenzdruck bzw. Wettbewerb durch die dargestellten Maßnahmen und Instrumente deutlich an Bedeutung gewonnen hat. Auch bei den Mechanismen staatliche Regulierung, externe Steuerung sowie hierarchische Selbststeuerung sind Resultate der Reformprozesse erkennbar. Die akademische Selbstorganisation scheint von den Reformbemühungen dagegen noch weniger stark betroffen zu sein, auch wenn formal Kompetenzeinbußen zu beobachten sind. Eine moderate Entwicklung der Hochschul-governance hin zu einem mehr marktorientierten Paradigma machen auch Dobbins und Knill (2016) in ihrer vergleichenden Studie von 14 westeuropäischen Ländern aus. Sie sprechen in diesem Zusammenhang von einer „gezügelter“ Einführung wettbewerbsorientierter Instrumente“ (ebd., S. 73), bei der allen hochschulpolitischen Akteuren nach wie vor eine bedeutende Rolle in der Governance der Hochschulen zukommt (vgl. ebd.).

Der Governance-Mechanismus mit dem stärksten Bedeutungszuwachs – der Konkurrenzdruck – wird in der Diskussion um die Absichten der Reformbemühungen in Richtung New Public Management meist auch als deren primäre Zielsetzung ausgemacht: Es gehe um die Schaffung bzw. Intensivierung eines Wettbewerbs innerhalb und zwischen den Hochschulen, um den Akteuren des Hochschulsystems Anreize zu Leistungssteigerungen sowie effizientem und effektivem Handeln zu geben. Nicht zuletzt soll im Zuge des (internationalen) Konkurrenzdrucks auch die Qualität und Innovativität wissenschaftlicher Arbeit, insbesondere der Forschung, steigen (vgl. Schimank und Volkmann 2008, S. 389–390). Schimank (2009) betont allerdings, dass eine Interpretation der Ziele des New Public Managements, in der nur die Entfesselung der Marktkräfte in den Blick genommen wird, zu kurz greift:

Es geht in der Tat darum, in den Hochschulen durch gesteigerten Konkurrenzdruck Effizienz- und Leistungssteigerungspotentiale zu entfesseln. Das ist aber eben nicht alles. Es [New Public Management; Anm. d. Verf.] steht vielmehr im Kontext mit abgestimmten Verschiebungen auch der anderen vier Regler. Die Konkurrenzintensivierung soll durch politische Zielvorgaben gerahmt, teilweise auch angesichts von Marktversagen korrigiert werden – daher tritt eine starke Außensteuerung der Hochschulen neben den Konkurrenzdruck. (ebd., S. 126)

Zudem soll das Zurückfahren der staatlichen Regulierung die nötigen Freiräume für den Wettbewerb bereitstellen, die Schwächung der akademischen Selbstorganisation die notwendigen Entscheidungsprozesse optimieren (keine Blockadehaltung) und die damit

verbundene Kompetenzverschiebung zugunsten der hierarchischen Selbststeuerung die Rolle der Hochschulen als handlungsfähigen Akteur stärken (ebd., S. 126–127).

Vor dem Hintergrund dieser Mehrdimensionalität der Hochschulgovernance wird im Folgenden dennoch der Fokus primär auf den Wettbewerbsmechanismus bzw. die externen Steuerungsversuche und den mit diesen Mechanismen verknüpften Steuerungs- und Anreizverfahren gerichtet. Entsprechend dem Forschungsinteresse dieser Arbeit wird der Blick dabei weg von der Makroebene der außeruniversitären Governance-Mechanismen auf die Mesoebene der Universität und der hochschulinternen Steuerungsinstrumente gelegt, die unmittelbar die Mikroebene der Forschungsarbeit⁶ der Professorinnen und Professoren betreffen. Diesen hochschulinternen Steuerungsinstrumenten und Anreizverfahren (indikatorenbasierte Mittelzuweisung, Zielvereinbarungen, Leistungszulagen W-Besoldung, Berufungs- und Bleibeverhandlungen, Benchmarking; vgl. das folgende Kapitel) ist allen gemeinsam, dass sie eine outputorientierte Steuerungs- bzw. Anreizwirkung über die Verwendung von Forschungsindikatoren (z. B. Drittmittel, Anzahl Publikationen, Anzahl Promotionen) erzielen wollen. Das Ziel dieser indikatorenbasierten Steuerungsversuche besteht wie bereits ausgeführt primär darin, Effektivitäts- und Effizienzsteigerungen in der Forschungsarbeit des wissenschaftlichen Personals zu erzielen. Durch das Setzen von gezielten Anreizen kann – so die Hoffnung – das Verhalten zudem in eine bestimmte Richtung gesteuert werden (z. B. vermehrtes Einwerben von Drittmitteln aufgrund daran geknüpfter Leistungszulagen). Daneben können auch andere Aspekte wie Leistungsgerechtigkeit und Transparenz der Mittelvergabe einen Grund für den Einsatz dieser Verfahren darstellen (vgl. Schröder 2004, S. 40).

Die Implementierung von Governance-Mechanismen im Sinne des New Public Managements und die damit verbundenen Steuerungsversuche innerhalb der Hochschulen wurden und werden vielfach kritisch diskutiert. Mehr grundsätzliche Kritik am neuen Governance-Regime betrifft den hohen Aufwand der zur Leistungsmessung erbracht werden muss und der eventuell erzielte Effizienzgewinne wieder zunichtemacht (Schimank und Volkmann 2008, S. 390). Daneben werden eine Reihe nichtintendierter Effekte wie z. B. eine Fokussierung auf leichter zu publizierende Mainstreamforschung, eine übertriebene Drittmittelfokussierung, die Vernachlässigung der Lehre, die Bildung von Zitationskartellen, eine ‚Salamitaktik‘ beim Publizieren von Forschungsergebnissen oder die Senkung qualitativer Standards zur Erhöhung der Promotionsquote befürchtet (vgl. Jansen 2009; Kieser 2010; Kühl 2015a; Minssen und Wilkesmann 2003). Diese und weitere mögliche nichtintendierte Effekte werden in Kapitel 2.3 im Rahmen der Diskussion

⁶ Hochschulinterne Steuerungs- und Anreizverfahren kommen natürlich auch im Bereich Lehre zum Einsatz. In dieser Arbeit beschränkt sich die Analyse aber wie bereits angemerkt nur auf den Bereich Forschung an Universitäten.

von gängigen Indikatoren zur Messung der Forschungsperformanz noch näher beleuchtet. Davor folgt zunächst die detaillierte Darstellung der hochschulintern verwendeten Steuerungs- und Anreizverfahren, welche im Zuge der dargestellten Governance-Reformen zunehmend Verbreitung in den Hochschulen gefunden haben.

2.2 Hochschulinterne Steuerungs- und Anreizverfahren

2.2.1 Indikatorenbasierte Mittelzuweisung

Für die indikatorenbasierte Mittelzuweisung finden in der Praxis und im wissenschaftlichen Diskurs eine Reihe von unterschiedlichen Begriffen Verwendung. Am gebräuchlichsten sind dabei, neben dem Begriff der indikatorenbasierten Mittelzuweisung selbst, die Bezeichnungen *leistungsorientierte Mittelvergabe (LOM)* (z. B. Krempkow und Schulz 2012), *formelgebundene Mittelvergabe* (z. B. Jaeger 2009) oder *parameterorientierte Mittelvergabe* (z. B. Minssen et al. 2003).⁷ Gemeint ist dabei stets „eine Finanzierungsformel, welche aus einer Reihe von Indikatoren besteht“ (Bogumil et al. 2013, S. 32). In Bezug auf die universitäre Forschung können solche Indikatoren beispielsweise die Höhe der eingeworbenen Drittmittel, die Anzahl an Publikationen oder die Anzahl betreuter Promotionen sein (vgl. zu den Indikatoren auch das nachfolgende Kapitel 2.3.2). Die in einem festgelegten Referenzzeitraum erbrachte Leistung hinsichtlich der verwendeten Indikatoren (Output) bestimmt, inwieweit die geleistete Forschungsarbeit durch die Vergabe zusätzlicher Finanzmittel belohnt oder durch den Wegfall dieser zusätzlichen Mittel entsprechend sanktioniert wird. Die eingesetzten Mittelzuweisungsverfahren können sich dabei nicht nur hinsichtlich der verwendeten Indikatoren, sondern auch in Bezug auf deren Gewichtung innerhalb des Verteilungsmodells unterscheiden. Sowohl die bewusste Auswahl der zu verwendenden Indikatoren als auch deren höhere oder niedrigere Gewichtung bietet für den Mittelgeber die Möglichkeit, gezielte Anreize zu setzen. Gewichtungen werden aber auch eingesetzt, um fachspezifische Unterschiede (z. B. Drittmittelhöhe bei naturwissenschaftlichen Fächern gegenüber sozialwissenschaftlichen Fächern) auszugleichen (ebd., S. 33–34). Aufgrund eben solcher Fachunterschiede, aber auch Unterschieden im Forschungsverhalten von Wissenschaftlerinnen und Wissenschaftlern derselben Disziplin, kann ein formuliertes Mittelverteilungsmodell nicht allen betroffenen Personen(gruppen) gleichermaßen gerecht werden. Die verwendeten Indikatoren und deren Gewichtung entsprechen mal mehr und mal weniger stark dem jeweiligen spezifischen ‚Forschungstyp‘ der Professorinnen und Professoren,

⁷ Im Sinne der Einheitlichkeit wird in dieser Arbeit primär der Begriff „indikatorenbasierte Mittelzuweisung“ verwendet.

so dass es nicht verwundert, dass die Ausgestaltung der Mittelvergabemodelle oft Gegenstand von vielfältiger Kritik ist. Hierzu lässt sich allerdings festhalten, dass gegebene Mitwirkungsmöglichkeiten bei der Ausarbeitung eines Mittelverteilungsmodells dazu beitragen, dass dieses von den Professorinnen und Professoren positiver wahrgenommen (Minssen et al. 2003, S. 56–58) und entsprechend weniger kritisch gesehen wird.

Indikatorenbasierte Mittelzuweisungsverfahren werden auf unterschiedlichen Ebenen eingesetzt. Zum einen kann auf Ebene der Bundesländer die staatliche Hochschulfinanzierung der Universitäten über Indikatorenmodelle erfolgen. Dies geschieht je nach Bundesland in unterschiedlich starkem Maße (vgl. hierzu im Detail Wespel und Jaeger 2015). Zum anderen findet die indikatorenbasierte Mittelzuweisung auch hochschulinternen Verwendung, wobei hier nochmals zwischen der universitätsinternen (Ebene Hochschulleitung – Fakultät bzw. andere untergeordnete Organisationseinheiten)⁸ und der fakultätsinternen (Fakultät – Institut/Lehrstuhl) Mittelzuweisung unterschieden werden kann (Bogumil et al. 2013, S. 79–85).⁹

Auf der universitätsinternen Ebene wurden im Jahr 2010 (bereits) an 96 % der Universitäten Mittel über indikatorenbasierte Verfahren verteilt (ebd., S. 81). Der Anteil am Globalbudget, der über solche Mittelvergabeverfahren verteilt wird, reicht dabei von weniger als fünf Prozent bis hin zur nahezu vollständigen Verteilung der Mittel über indikatorenbasierte Verfahren (ebd.). Hinsichtlich des in der Studie von Bogumil et al. ermittelten hohen Anteils an Universitäten, die auf universitätsinterner Ebene indikatorenbasierte Mittelzuweisungsverfahren anwenden, ist also erstens zu beachten, dass die Höhe der auf diese Weise vergebenen Gelder auch vergleichsweise gering ausfallen kann (bei 60 % der Universitäten unter 25 % des Globalbudgets (ebd.)). Zweitens bedeutet der generelle Einsatz des Verfahrens auf der universitätsinternen Ebene nicht, dass auch alle Fachbereiche (gleichermaßen) von der Vergabepaxis betroffen sind. Denkbar ist z. B., dass die Mittelvergabe in der Medizin, bei der Modelle der indikatorenbasierten Mittelzuweisung auf Fakultätsebene bereits nahezu flächendeckend implementiert sind (Karmann 2014, S. 462), auch auf der Ebene Hochschulleitung-Fakultät mehr als bei anderen Fächer nach Formelmodellen erfolgt. Inwieweit der Einsatz der indikatorenbasierten Mittelzuweisung auf universitätsinterner Ebene nach Einschätzung

⁸ Im Folgenden bezeichnen die Begriffe der „*universitätsinternen* indikatorenbasierten Mittelzuweisung“ und der „*universitätsinternen* Zielvereinbarungen“ in Abgrenzung zum *fakultätsinternen* Einsatz der Verfahren, den Einsatz von diesen auf der Ebene der Hochschulleitung. Dagegen umfasst der Begriff „*hochschulintern*“ als Oberbegriff alle Steuerungs- und Anreizverfahren, die innerhalb der Hochschulen (ohne Differenzierung nach Hierarchieebene) eingesetzt werden, also sowohl universitätsintern als auch fakultätsintern eingesetzte Verfahren.

⁹ Denkbar und vereinzelt auch implementiert sind indikatorenbasierte Mittelzuweisungen auf weiteren Ebenen wie z. B. zwischen der Hochschulleitung und dem Institut oder einzelnen Arbeitsgruppen. In dieser Arbeit wird sich aber auf die am häufigsten vorkommenden Varianten beschränkt. Sofern die Ergebnisse der hier durchgeführten Befragung darauf hindeuten, dass die betrachteten Steuerungsverfahren in hohem Maße auch auf anderen Ebenen eingesetzt werden (Anmerkungen der Befragten), wird dies im Ergebnisteil dieser Arbeit thematisiert.

der betroffenen Professorinnen und Professoren in den sozialwissenschaftlichen Fächern verbreitet ist, wird daher – auch für alle weiteren hier untersuchten Steuerungs- und Anreizverfahren – im Zuge dieser Arbeit noch diskutiert werden.

Fakultätsintern, das heißt auf der Ebene Fakultät-Institut/Lehrstuhl, finden indikatorenbasierte Mittelzuweisungsverfahren an Universitäten ebenfalls sehr oft Verwendung. Die deutschlandweite Befragung von Dekanen in der Studie von Bogumil et al. (2013) ergab einen Implementierungsstand von 91 %, wobei sich die Ausgestaltung des Verfahrens in 43 % der Fälle an der Mittelzuweisung auf der Ebene Hochschulleitung-Fakultät orientiert (ebd., S. 85). Zu vermuten ist aber auch hier, dass sich hinsichtlich des prozentualen Implementierungsstandes fachbereichsspezifische Unterschiede ergeben (vgl. Karmann 2014, S. 462).

2.2.2 Zielvereinbarungen

Steuerungsverfahren in Form von Zielvereinbarungen lassen sich mit Jaeger et al. (2006) wie folgt beschreiben:

Im Unterschied zu Formelmodellen werden bei Zielvereinbarungen bzw. Kontrakten die Ziele in der Regel nicht einseitig durch die nächsthöhere Instanz festgelegt, sondern sind das Ergebnis eines Aushandlungsprozesses zwischen Zuweisungsgeber und Zuweisungsnehmer (z. B. Ministerium und Hochschulen bzw. Hochschulleitung und Fakultäten/Fachbereiche). Die finanziellen Zuweisungen können dabei an die Durchführung von Maßnahmen oder an das Ausmaß der Zielerreichung gekoppelt werden. Anders als bei der formelgebundenen Mittelvergabe verfügt der Zuweisungsgeber bei Zielvereinbarungen über einen diskretionären Entscheidungsspielraum. Zudem ermöglichen Zielvereinbarungen die Berücksichtigung von innovativen und zukunftsorientierten Vorhaben, was bei der vergangenheitsorientierten Formelvergabe nur bedingt möglich ist. (ebd. S. 6–7)

Charakteristisch für Zielvereinbarungen ist also der Vertragscharakter, der sich aus einem *beidseitigen* Aushandlungsprozess über die Inhalte der zu erreichenden Ziele ergibt. Wie Jaeger et al. in obiger Definition anmerken, ist diese Form von Zielvereinbarungen die Regel. Möglich ist allerdings auch die *einseitige* Vorgabe der zu erreichenden Ziele durch den Mittelgeber mit wenig oder keinem Mitspracherecht des Zuweisungsempfängers. In beiden Fällen beziehen sich die Vereinbarungen und die gegebenenfalls daran gekoppelten Mittelzuweisungen aber auf noch zukünftig zu erbringende Leistungen. Darin besteht ein zentraler Unterschied zum Verfahren der indikatorenbasierten Mittelzuweisung, das bereits erbrachte Leistungen belohnen oder sanktionieren soll.

Wie auch bei der indikatorenbasierten Mittelzuweisung können sich Zielvereinbarungen auf verschiedene Indikatoren (z. B. Drittmittelhöhe, Publikationsoutput) beziehen, denen innerhalb der Vereinbarung ein unterschiedlich starkes Gewicht zugesprochen werden kann. Die vereinbarten Ziele können hierbei auch quantitativ nicht direkt messbare Inhalte betreffen (Bogumil et al. 2013, S. 36). Da der Formulierung der zu erreichenden Ziele in der Regel ein beidseitiger Abstimmungsprozess zwischen den Verhandlungspartnern vorausgeht, ist aufgrund dieses Aushandlungscharakters zu erwarten, dass Zielvereinbarungen als Steuerungsverfahren weniger kritisch wahrgenommen werden als beispielsweise die indikatorenbasierte Mittelzuweisung.

Zielvereinbarungen werden, wie auch das Verfahren der indikatorenbasierten Mittelzuweisung, auf verschiedenen Ebenen eingesetzt. Nahezu alle Universitäten schließen Zielvereinbarungen mit den Ländern bzw. den zugehörigen Wissenschaftsministerien ab (Bogumil et al. 2013, S. 77). Universitätsinterne Zielvereinbarungen zwischen der Hochschulleitung und den untergeordneten Organisationseinheiten (Fakultäten oder Instituten) wurden im Jahr 2010 an 75 % der Universitäten abgeschlossen, wobei von diesen wiederum drei Viertel auch die Allokation von Finanzmitteln an die Erreichung der vereinbarten Ziele knüpften (ebd., S. 79–80). Fakultätsinterne Zielvereinbarungen (Ebene Fakultät – Institut/Lehrstuhl) wurden dagegen nur von 26 % der in der Studie von Bogumil et al. befragten Dekane abgeschlossen. Weniger als die Hälfte dieser Zielvereinbarungen waren dabei mit der Vergabe von zusätzlichen Finanzmitteln an die Institute bzw. Lehrstühle verknüpft (ebd., S. 83). Zielvereinbarungen sind also im Gegensatz zur indikatorenbasierten Mittelzuweisung nicht zwingend mit der Vergabe (oder Kürzung) von Finanzmitteln verbunden. Je nachdem, wie spezifisch die zu erreichenden Ziele bzw. die hierfür notwendigen Maßnahmen innerhalb der Zielvereinbarungen definiert werden und inwieweit überhaupt eine Überprüfung der Zielerreichung erfolgt, haben Zielvereinbarungen damit einen mehr oder weniger stark verbindlichen Charakter (vgl. ebd., S. 84).

Abschließend sei darauf hingewiesen, dass es beim parallelen Einsatz von Zielvereinbarungen und indikatorenbasierten Mittelzuweisungsverfahren zu Zielkonflikten kommen kann. Dies ist dann der Fall, wenn die Aushandlung von Zielvereinbarungen und die Ausarbeitung indikatorenbasierter Mittelzuweisungsmodelle nicht aufeinander abgestimmt sind (vgl. Trute und Pilniok 2009, S. 25–26) und in beiden Verfahren sich widersprechende oder in Bezug auf zeitliche oder finanzielle Ressourcen konfligierende Ziele festgesetzt werden. Solche, aus fehlender Abstimmung resultierende Zielkonflikte können natürlich auch alle anderen Steuerungs- und Anreizverfahren wie z. B. die in dem nächsten Teilkapitel dargestellte W-Besoldung betreffen.

2.2.3 W-Besoldung

Die W-Besoldung nimmt in der Diskussion der universitären Steuerungs- und Anreizverfahren eine Sonderstellung ein. Im Gegensatz zu den bisher beschriebenen Verfahren erfolgte die Einführung der neuen Besoldungsart an den Universitäten nicht auf freiwilliger Basis, sondern nach gesetzlichen Vorgaben. Durch den Erlass des Professorenbesoldungsreformgesetzes im Jahr 2002 wurden die Bundesländer verpflichtet, bis Anfang 2005 das neue Besoldungssystem umzusetzen (Bogumil et al. 2013, S. 44). Während die ursprüngliche C-Besoldung der Professorinnen und Professoren dem Senioritätsprinzip folgt und die Besoldung an Dienstaltersstufen gekoppelt ist, wird im Rahmen der W-Besoldung ein festes Grundgehalt gezahlt, das variable Zulagen für bestimmte Leistungen vorsieht (ebd.). Diese Leistungszulagen können „aus Anlass von Berufungs- und Bleibeverhandlungen, für besondere Leistungen in Forschung, Lehre, Kunst, Weiterbildung und Nachwuchsförderung sowie für die Wahrnehmung von Funktionen oder besonderen Aufgaben im Rahmen der Hochschulsebstverwaltung oder der Hochschulleitung vergeben werden“ (ebd., S. 44–45). Kriterien, nach denen Leistungszulagen gewährt werden, sind wie bei den anderen Verfahren auch meist quantifizierbare Indikatoren wie die eingeworbenen Drittmittel, die Anzahl von Publikationen oder die Anzahl von Promotionen (vgl. ebd., S. 86–87). Jochheim (2015) kommt in einer Fallstudie mit sieben Universitäten allerdings zu dem Schluss, dass „die untersuchten Universitäten die W-Besoldung nur eingeschränkt als Steuerungsinstrument [nutzen]. Das Budget wird weitestgehend für Berufungs- und Bleibeleistungsbezüge verwendet, besondere Leistungen werden nur in Ausnahmefällen belohnt“ (ebd., S. 288). Als Grund werden hier die begrenzten finanziellen Mittel genannt. Diese würden lieber in Neuberufungen bzw. Bleibeverhandlungen investiert, als bereits erbrachte Leistungen durch Leistungsbezüge zu honorieren (ebd.).

Während alle ab Januar 2005 eingestellten Professorinnen und Professoren automatisch nach der W-Besoldung bezahlt werden (W2, W3 sowie W1 für Juniorprofessuren), besteht für Professorinnen und Professoren, die bereits vor diesem Stichtag im Amt waren, die Möglichkeit, in der C-Besoldung (C2, C3, C4) zu verbleiben. Sofern die C-besoldeten Professorinnen und Professoren in erneute Berufungs- oder Bleibeverhandlungen einsteigen, ist allerdings ein Wechsel in die W-Besoldung erforderlich. Auf Antrag ist auch ein freiwilliger Wechsel in die W-Besoldung möglich (Bogumil et al. 2013, S. 46), wobei dies nur für einen kleinen Teil der Professorinnen und Professoren eine attraktive Option darstellt (ebd., S. 87). Durch das kontinuierliche Ausscheiden von C-besoldeten Professuren aus dem Dienst und den W-besoldeten Neueinstellungen steigt der Anteil an W-besoldeten Professorinnen und Professoren in den letzten Jahren stetig an. Im Jahr 2015 betrug dieser in der Fächergruppe der Rechts-, Wirtschafts- und Sozialwissenschaften (ohne Juniorprofessuren) bereits 58.6 % (Statistisches Bundesamt 2016).

2.2.4 Weitere Verfahren

Im Folgenden werden knapp noch einige weitere universitäre Steuerungsinstrumente vorgestellt, die teils auch im Rahmen der Fragebogenerhebung dieser Arbeit thematisiert wurden. Dabei werden auch Verfahren beschrieben, die nicht zwingend als Steuerungsinstrument zu verstehen sind, die aber als solches genutzt werden können bzw. von welchen eine handlungsleitende Anreizwirkung ausgehen kann (z. B. Hochschulrankings).

Als ein Steuerungsverfahren im weiteren Sinne können *Bleibe- und Berufungsverhandlungen* aufgefasst werden. Auch hier können Indikatoren zur Messung der Forschungsleistung wie die Höhe der eingeworbenen Drittmittel, die Anzahl der Publikationen oder verschiedene Zitationsindizes eine Rolle bei den Verhandlungen zur Ausgestaltung der alten oder neuen Stelle spielen. Ähnlich wie bei Zielvereinbarungen stehen sich in diesen Verhandlungen zwei gleichberechtigte Parteien gegenüber, wobei je nach Ausgangslage (z. B. mehrere vorliegende Rufe eines Kandidaten) die eine oder die andere Seite mehr ‚Druckmittel‘ in der Hand hält. Anzunehmen ist nun, dass Professorinnen und Professoren bestrebt sein werden, den für Bleibe- und Berufungsverhandlungen relevanten Anforderungen bzw. Kriterien in ausreichendem Umfang gerecht zu werden, um sich so bei (zukünftigen) Verhandlungen in einer aussichtsreichen Position zu befinden. Gerade weil Leistungen in der Lehre bei Bleibe- und Berufungsverhandlungen oft von nur untergeordneter Bedeutung sind (vgl. Wissenschaftsrat 2008a, S. 43), kann von der Definition der für die Verhandlungen relevanten Forschungsindikatoren eine Anreizwirkung für ein entsprechendes Forschungsverhalten ausgehen.

Ein weiteres Verfahren, das nicht zwingend mit finanziellen Anreizen verknüpft sein muss, stellt das *Benchmarking verschiedener Lehrstühle innerhalb einer Fakultät* im Rahmen einer fakultätsinternen Forschungsevaluation dar (vgl. Sieweke et al. 2014). Hierbei wird die Forschungsleistung der Lehrstühle anhand verschiedener Indikatoren miteinander bzw. hinsichtlich eines festgelegten Referenzwertes verglichen. Der indikatorenbasierte Vergleich soll dabei primär organisationsintern zur Positionsbestimmung der einzelnen Organisationseinheiten genutzt werden. Im Gegensatz zu anderen Steuerungsverfahren dient der interne Vergleich mehr dem Aufdecken von Schwachstellen und der Ermittlung von besonders ‚erfolgreichen‘ Organisationseinheiten und soll dadurch Anstoß von Lernprozessen zur Verbesserung von Arbeitsabläufen sein. In diesem Sinne ist das Benchmarking als Teil des Qualitätsmanagements von Universitäten zu sehen (vgl. Oberschelp und Jaeger 2015; Schedler 2003). Die inhaltliche Abgrenzung zu Verfahren wie der indikatorenbasierten Mittelzuweisung ist dennoch nicht immer ganz eindeutig, da auch für diese die entsprechenden Ausprägungen der herangezogenen (Forschungs-)Indikatoren ermittelt werden müssen und für (interne) Vergleiche herange-

zogen werden können. Der Unterschied lässt sich wohl am besten mit der bereits erwähnten fehlenden Verknüpfung mit finanziellen Anreizen und der im Benchmarking explizit enthaltenen Vergleichslogik – verbunden mit gegenseitigen Lernprozessen zwischen den Organisationseinheiten – kennzeichnen. Schedler (2003) spricht in diesem Zusammenhang Benchmarkingverfahren eine stärker ausgeprägte qualitative Komponente zu als reinen nicht weiter interpretierten Kennzahlenvergleichen (vgl. ebd., S. 30–31). Zur Verbreitung des Einsatzes von Benchmarkingverfahren im universitären Bereiche merken Oberschelp und Jaeger (2015) an, dass diese vorrangig in der Hochschulverwaltung und in Serviceeinrichtungen der Hochschulen und nur selten in den Bereichen Forschung und Lehre Anwendung finden (ebd., S. 489).

Spätestens mit dem erstmaligen Erscheinen des *CHE-Rankings* 1998 verzeichnen *Hochschulrankings* in den letzten Jahren auch in Deutschland einen stetigen Bedeutungszuwachs (Bogumil et al. 2013, S. 37–38). Neben nationalen Rankings, wie dem bereits genannten CHE-Ranking, spielen dabei auch internationale Rankings wie das *Times Higher Education World University Ranking*, das *Academic Ranking of World Universities* oder das von der EU-Kommission geförderte und seit 2014 veröffentlichte *U-Multirank* (vgl. zu U-Multirank Ziegele und van Vught 2013) eine wichtige Rolle im vergleichenden Wettbewerb zwischen den Universitäten (vgl. für einen Überblick Hazekorn 2013; Rauhvargers 2014; Roessler 2013). Mit zunehmendem Bedeutungszuwachs der Rankings – auch in der Öffentlichkeit, der Wirtschaft oder bei den (zukünftigen) Studierenden – sind Hochschulen zunehmend darauf bedacht, in diesen besonders gut abzuschneiden. Aufgrund dessen sehen Kritiker die Gefahr einer zunehmenden Orientierung der Universitäten an den Rankingkriterien, also an den in den Rankings verwendeten Indikatoren (vgl. Federkeil 2012, 2013; Lenzen 2012; Michael 2005; Tippelt et al. 2010). Beispielsweise kann der Hochschulleitung daran gelegen sein, zur Verbesserung der Rankingplatzierung den Publikationsoutput zu erhöhen. Über die bereits beschriebenen hochschulinternen Steuerungsverfahren kann dann versucht werden, Anreize zu vermehrten Publikationsaktivitäten auf der Mikroebene der Forschungsarbeit der Professorinnen und Professoren zu schaffen, und damit den für das Ranking relevanten Pro-Kopf-Output zu steigern. Als Folge der verstärkten Fokussierung auf Rankings wird unter anderem befürchtet, dass die Heterogenität von Universitätsprofilen einer zunehmenden Vereinheitlichung der Hochschullandschaft weicht, bestimmte für Rankings weniger relevante Bereiche der universitären Arbeit wie z. B. die Lehre oder eher drittmittelschwache Fachgebiete vernachlässigt werden und es zu sogenanntem *gaming*, also zu Manipulationen und Trickereien zur Erfüllung der Rankingkriterien kommt (vgl. Kieser 2010; Kroth und Daniel 2008; Marginson und van der Wende 2007). Demgegenüber werden als positive Aspekte von Rankings unter anderem die Schaffung von Transparenz für Studierende und andere Stakeholder, die Orientierungsfunktion für Studieninteressierte

und akademisches Personal und nicht zuletzt der den Rankings inhärente Wettbewerbsmechanismus als Anreiz für die (Weiter-)Entwicklung der Hochschulen hervorgehoben (vgl. Burckhart 2012; Lenzen 2012; Tippelt et al. 2010; Ziegele und van Vught 2013).

Unabhängig von den erwarteten positiven oder negativen Folgen von Rankings für die Hochschullandschaft bezieht sich ein weiterer Diskurs auf die den Rankings zugrundeliegende Methodik. Insbesondere die verwendeten Indikatoren, die größtenteils auch in den anderen bereits beschriebenen Steuerungs- und Anreizverfahren herangezogen werden, sind Gegenstand von vielfältiger Kritik. Welche Indikatoren für den Bereich Forschung in Rankings und den universitären Steuerungsverfahren generell genutzt werden, woran sich die Kritik festmacht und welche möglichen nichtintendierten Effekte mit dem Einsatz der jeweiligen Forschungsindikatoren verbunden sein können, ist Gegenstand des nächsten Kapitels.

2.3 Messung der Forschungsperformanz und Forschungsindikatoren

2.3.1 Quantifizierung der Forschungsleistung

Für den Einsatz der im vorangegangenen Kapitel beschriebenen hochschulinternen Steuerungs- und Anreizverfahren ist die Erfassung der Forschungsleistung der betroffenen Arbeitseinheiten notwendig. Dabei kann versucht werden, die Forschungsleistung unter qualitativen oder quantitativen Gesichtspunkten zu erfassen. In den meisten forschungsevaluativen Verfahren wird sich allerdings überwiegend auf quantifizierbare Aspekte der Forschungsarbeit beschränkt. Dies ist vermutlich darauf zurückzuführen, dass rein qualitative Kriterien, wie z. B. der inhaltliche Gehalt eines Zeitschriftenartikels, nur mit hohem Aufwand zu ermitteln sind, notwendigerweise auf subjektiven Urteilen beruhen und eine vergleichende Perspektive auf höherem Aggregationsniveau nur schwer umzusetzen ist. Wie Hornbostel (1997) allerdings anmerkt, können Forschungs- bzw. Wissenschaftsindikatoren nicht auf ihren rein quantitativen Charakter reduziert werden:

Anders als Kennzahlen, die in der Regel nur eine Beschreibung von schon vorfindlichen [sic] Quantitäten liefern (z. B. Anzahl des Forschungspersonals) und Operationalisierungsprobleme nur im Sinne angemessener Klassifikationen haben, beziehen sich Indikatoren auf ein Konstrukt, auf etwas das nicht als Quantität vorfindlich ist, sondern über ein Konzept definiert wird (Qualität, Effizienz, Leistung etc.). (ebd. S. 180)

Ist hier von Forschungsindikatoren die Rede, geht es damit immer auch um das dahinterstehende zu messende Konstrukt, also um die Forschungsleistung oder die For-

schungsqualität. Wenn beispielsweise in Hochschulrankings die Höhe der eingeworbenen Drittmittel aufgeführt wird, dient das meist nicht der neutralen Beschreibung der finanziellen Ausstattung der Hochschule, sondern soll als *Indikator* die Forschungsleistung oder Forschungsqualität *anzeigen*. Inwieweit die Indikatoren dieser Aufgabe gerecht werden, ist Gegenstand zahlreicher wissenschaftstheoretischer Diskussionen und wird auch bei der im nächsten Kapitel folgenden näheren Beschreibung verschiedener Forschungsindikatoren thematisiert werden. Festzuhalten bleibt, dass die meisten der in der Praxis verwendeten Forschungsindikatoren quantitativer Natur sind, dabei aber auch den Anspruch erheben, zumindest Hinweise auf qualitative Aspekte der Forschungsleistung bzw. zur Forschungsqualität zu liefern.

Forschungsindikatoren lassen sich weiter danach klassifizieren, auf welche Aspekte der universitären Forschung sich diese beziehen. Es kann hierbei zwischen den Dimensionen *Kontext*, *Input*, *Prozess*, *Output* und *Outcome* unterschieden werden (vgl. zu den einzelnen Dimensionen Tippelt et al. 2010, S. 100), wobei der überwiegende Teil der zur Forschungsevaluation verwendeten Indikatoren den Dimensionen Input, Output und Outcome zugeordnet werden kann. Als Inputindikatoren können beispielsweise die Anzahl der forschenden Mitarbeiter oder der Professuren verstanden werden, welche als (personelle) Ressource in die Forschungsarbeit einfließen. Outputindikatoren, wie die Anzahl der Publikationen oder der Promotionen, beziehen sich auf kurzfristige Ergebnisse (ebd.) der Forschungsarbeit. Ergebnisse der Forschung, die erst nach einem längeren oder über einen längeren Zeitraum wirksam werden, werden dagegen als Outcome definiert (ebd.). In diesem Sinne können z. B. zitationsbasierte Kennzahlen als Outcomeindikatoren verstanden werden. Ob ein Indikator den Input, Output oder den Outcome misst, ist nicht immer eindeutig und teils definitionsabhängig. So wird das Drittmittelvolumen je nach Sichtweise sowohl als Input- als auch als Outputgröße interpretiert. Als Inputgröße können Drittmittel insofern definiert werden, als dass sie als zusätzliche externe finanzielle Ressource in die Forschungsarbeit einfließen (vgl. Hornbostel 1997, S. 211). Eine Outputgröße stellen Drittmittel in dem Sinne dar, dass das erfolgreiche Einwerben eines Drittmittelprojektes bereits als Ergebnis der geleisteten Vorarbeiten (Verfassen des Antrags, im Vorfeld erbrachte Forschungsarbeiten) angesehen werden kann und somit ein Zwischenprodukt der Forschungsarbeit darstellt (ebd., S. 234).¹⁰ Hinsichtlich der Validität von Forschungsindikatoren wird mit Bezug zu den Bewertungsdimensionen *Input* und *Output/Outcome* argumentiert, dass Indikatoren, die Outputs oder Outcomes messen, zur Erfassung der Forschungsleistung besser geeignet seien als Indikatoren, welche sich auf Inputs beziehen (Kroth und Daniel 2008, S. 554). Die intuitiv

¹⁰ In dieser Arbeit wird sich an der erstgenannten Definition orientiert und der Drittmittelindikator primär als Inputindikator verstanden.

nachvollziehbare Feststellung von Kroth und Daniel kann sich dabei auf eine Untersuchung dreier internationaler Rankings stützen, bei denen die verwendeten Indikatoren unter dem Gesichtspunkt der Konstruktvalidität und der Reliabilität analysiert wurden (vgl. Kroth und Daniel 2008).

Vor dem Hintergrund dieser generellen Ausführungen zur indikatorenbasierten Messung der Forschungsperformanz werden nun nachfolgend die gängigsten Forschungsindikatoren näher beschrieben und kritisch im Hinblick auf ihre Validität diskutiert. Auch mögliche nichtintendierte Effekte, die mit dem Einsatz der jeweiligen Indikatoren verbunden sein können, werden thematisiert. Die Auswahl der zu diskutierenden Indikatoren beschränkt sich dabei auf diejenigen Forschungsindikatoren, die in hochschulinternen Steuerungs- und Anreizverfahren sowie in verschiedenen Hochschulrankings (*Times Higher Education World University Ranking, Academic Ranking of World Universities, QS World University Rankings, CHE Ranking* etc.) am häufigsten eingesetzt werden. Auf die Betrachtung reiner Inputindikatoren in Form von personellen Ressourcen wie der Anzahl der Professorinnen und Professoren oder der wissenschaftlichen Mitarbeiterinnen und Mitarbeiter wird dabei verzichtet. Diese stellen primär ein Maß für die Größe eines Fachbereichs oder einer ganzen Hochschule dar und stehen somit nur in indirektem Zusammenhang mit der Forschungsleistung, sofern man nicht an dem absoluten Forschungsoutput (z. B. alle Publikationen eines Fachbereiches in einem bestimmten Zeitraum), sondern an einer größenunabhängigen Betrachtung (z. B. Anzahl Publikationen pro Professur eines Fachbereiches in einem bestimmten Zeitraum) interessiert ist.

2.3.2 Kritische Diskussion der gängigen Forschungsindikatoren

Drittmittel

Die Höhe der eingeworbenen oder auch verausgabten Drittmittel zählt wohl zu den am weitesten verbreiteten Indikatoren, wenn es um den Versuch geht, Forschungsleistung zu quantifizieren. In den hochschulinternen Steuerungs- und Anreizverfahren, wie z. B. der indikatorenbasierten Mittelzuweisung oder den Leistungszulagen im Rahmen der W-Besoldung, wird das Drittmittelvolumen mit am häufigsten als Forschungsindikator eingesetzt (Bogumil et al. 2013, S. 81–87). Das Drittmittelvolumen, teils auch die Anzahl drittmittelfinanzierter Mitarbeiterinnen und Mitarbeiter, findet zudem fast durchgehend Verwendung in namhaften Hochschulrankings wie dem *Times Higher Education Ranking* oder dem *U-Multirank* (vgl. Federkeil 2013; Rauhvargers 2011, 2013; Ziegele und van Vught 2013).

Die Verwendung der Drittmittelhöhe zur Beurteilung der Forschungsleistung ist insofern problematisch, als dass es sich hierbei um eine Inputgröße handelt (Hornbostel 1997, S. 211–212; Rassenhövel und Dyckhoff 2006, S. 95–96; vgl. zur Diskussion, ob Drittmittel als Input- oder Outputgröße zu verstehen sind, auch das vorangegangene

Kapitel), die zunächst nichts über den eigentlichen Forschungsoutput aussagt. Zudem ist davon auszugehen, dass die Einwerbung von Drittmitteln in vielen Fällen gar nicht zwangsläufig als (überdurchschnittliches) Forschungsengagement in Form der Akquirierung prestigeträchtiger Forschungsprojekte zu werten ist. Vielmehr kann der Versuch der Drittmiteleinwerbung auch schlicht als ein notwendiges ‚Übel‘ angesehen werden, um Defizite in der oft eher dürftigen Grundfinanzierung der Hochschulen zu kompensieren (Forschergruppe "Governance der Forschung" 2009, S. 132). Das Problem ergibt sich hier also daraus, dass etwas als Maßstab für Forschungsleistung gewertet wird, was oft schlicht getan werden muss, bzw. von dem zumindest erwartet wird, dass es getan wird (Jacob und Teichler 2011, S. 68), um so die universitäre Forschungsarbeit eines Lehrstuhls oder einer Arbeitsgruppe grundsätzlich erst zu ermöglichen.

Wenn zu der ohnehin vorhandenen Notwendigkeit Drittmittel einzuwerben zusätzliche Steuerungs- und Anreizmechanismen ins Spiel kommen, die das Einwerben von externen Forschungsgeldern noch weiter fördern und belohnen, besteht zudem die Gefahr, dass Professorinnen und Professoren zu „Antragsprofis“ (Schimank und Volkmann 2008, S. 390) werden. Aufgrund der zunehmenden zeitlichen Belastung durch die Bearbeitung ständig neuer Projektanträge können diese dann nur noch wenig oder keine Zeit mehr für die eigentliche Forschungsarbeit aufwenden (ebd.). Hiermit verbunden werden auch mögliche negative Auswirkungen auf die Qualität (Quantität vor Qualität) und der Innovativität (Mainstreamforschung) der Forschung befürchtet (Weingart 2005, S. 125–128). Zudem konnte in einer Studie von Jansen et al. (2007) für Arbeitsgruppen der Astrophysik, Nanotechnologie und Mikroökonomik nachgewiesen werden, dass die Einwerbung von Drittmitteln nur bis zu einer bestimmten kritischen Höhe zu einer Produktivitätssteigerung in Form eines erhöhten Publikationsoutputs führt. Ab diesem kritischen Schwellenwert ist nicht nur mit einer stagnierenden, sondern sogar mit einer abnehmenden Produktivität (Publikationen absolut und Publikationen pro Kopf) zu rechnen. Dies wird auf überoptimale Arbeitsgruppengrößen und den damit gestiegenen Transaktionskosten innerhalb der Arbeitseinheit zurückgeführt (Jansen 2009, S. 54). Die Gleichsetzung von hohem Drittmittelvolumen mit hoher Forschungsleistung erscheint also aufgrund der beschriebenen möglichen negativen Effekte auf die Forschungsarbeit nicht immer gerechtfertigt. Das betrifft in erster Linie die Qualität der Forschung, kann aber, wie die Studie von Jansen zeigt, auch quantitative Aspekte der Forschungsleistung betreffen.

In der Diskussion um die Eignung von Drittmittelvolumina zur Messung der Forschungsleistung wird die Relevanz von Drittmitteln als Indikator oft damit begründet, dass der Vergabe von Drittmitteln Entscheidungsprozesse vorausgehen, bei denen Qualitätsaspekte der Forschungsvorhaben ausschlaggebend sind (Burckhart 2012, S. 823). Das ist sicherlich richtig, und Forschungsvorhaben, die nicht bestimmten qualitativen

Standards entsprechen, werden in den meisten Fällen auch keinen Erfolg beim Akquirieren der Finanzmittel haben. In diesem Sinne können eingeworbene Drittmittelprojekte, insbesondere wenn sie von hoch reputierlichen Mittelgebern wie der Deutschen Forschungsgemeinschaft (DFG) oder dem Bundesministerium für Bildung und Forschung (BMBF) nach Durchlaufen eines anerkannten Begutachtungsverfahrens vergeben wurden, durchaus als Maß für die Forschungsleistung angesehen werden. Allerdings kann dieser Argumentationslinie entgegengesetzt werden, dass sich die (qualitative) Begutachtung meist auf das *Forschungsvorhaben* in Form des Projektantrages beschränkt und das eigentliche *Forschungsergebnis* bei der Vergabe der Drittmittel keine Rolle spielt (Gillet 1989, S. 28). Das Drittmittelvolumen wäre damit mehr als Indikator für die Qualität (und Quantität) der Projektanträge zu verstehen denn als ein Maß für die Qualität der Forschungsergebnisse. Daneben ist zu beachten, dass auch die Gutachterurteile selbst im Hinblick auf ihre Reliabilität und Validität nicht unumstritten sind (vgl. Bornmann und Daniel 2003).

Der hier dargestellten Kritik können wiederum zum einen empirische Befunde entgegengehalten werden, die einen positiven Zusammenhang zwischen drittmittelgeförderten Projekten und der Qualität der Forschung – gemessen anhand der Anzahl der Zitationen der im Rahmen des Forschungsprojektes entstandenen Publikationen – sowie der Publikationsaktivität generell nahelegen (Hornbostel 2001, S. 150–154; Schmidt und Weishaupt 2008, S. 131–132).¹¹ Zum anderen ist zu berücksichtigen, dass das Drittmittelvolumen im Zusammenspiel mit anderen Forschungsindikatoren durchaus geeignet erscheint, Informationen zur Forschungsaktivität eines Fachbereiches oder eines Lehrstuhls zu liefern. So zeigte sich in einer Analyse von Berghoff und Hornbostel (2003) ein starker Zusammenhang zwischen der eingeschätzten Forschungsreputation einer Fakultät durch Fachkollegen und der indikatorengestützten Messung der Forschungsaktivität anhand der Indikatoren Drittmittel, Promotionen und Publikationen (ebd., S. 191). Entsprechend finden sich in der aktuellen Diskussion zahlreiche Plädoyers für eine mehrdimensionale Sichtweise in (Forschungs-)Rankings und bei der Evaluation von Forschungsleistung, die der eindimensionalen Betrachtung einzelner Indikatoren vorzuziehen sei (vgl. Grözinger und Leusing 2006, S. 12; Ziegele und van Vught 2013, S. 55).

Es lässt sich festhalten, dass sich die Verwendung von Drittmittelvolumina bzw. die Verwendung verwandter Indikatoren (z. B. Anzahl drittmittelfinanzierter Forschungsprojekte oder drittmittelfinanzierter Mitarbeiterinnen und Mitarbeiter) zur Messung der Forschungsleistung mit verschiedenen Problemen konfrontiert sieht, welche im Hinblick auf die Validität der Indikatoren durchaus kritisch gesehen werden müssen. Unbestreitbar

¹¹ Wobei hier wiederum eingeworfen werden könnte, dass die höhere Anzahl an Zitationen bei drittmittelgeförderten Publikationen gerade als ein Beleg für die kritisierte Förderung der Mainstreamforschung durch die zunehmende Drittmittelorientierung dient, da aktuell populäre Forschungsthemen auch öfter zitiert werden.

ist aber auch, dass die Einwerbung von Drittmitteln und die Durchführung drittmittelfinanzierter Forschungsprojekte mittlerweile mehrheitlich ein fester Bestandteil der (sozialwissenschaftlichen) Forschung an deutschen Universitäten ist (vgl. zur Entwicklung der Drittmiteleinnahmen in den Sozialwissenschaften Schmidt-Hertha, Tafertshofer, Tippelt 2016, S. 160–167). Die Drittmittelforschung ist damit in vielen Fällen eng mit der Forschungsarbeit und auch -leistung eines Fachbereiches verknüpft. Auch andere für die Messung der Forschungsleistung wichtige Aspekte der Forschungsarbeit, wie z. B. die Anzahl von durchgeführten Promotionsprojekten, stehen mit den eingeworbenen Drittmitteln in direktem Zusammenhang (Schmidt-Hertha 2012, S. 160). Will man die Forschungsleistung quantitativ erfassen, müssen Drittmittelindikatoren daher entsprechend Berücksichtigung finden. Dies sollte allerdings stets mit Hilfe eines mehrdimensionalen Ansatzes in Verbindung mit weiteren relevanten Indikatoren geschehen. Einen solchen Indikator stellt die bereits angesprochene Anzahl der durchgeführten Promotionsprojekte dar, welcher – neben den Habilitationen – Gegenstand des nun folgenden Abschnitts ist.

Promotionen und Habilitationen

Werden Promotionen oder Habilitationen als Kriterium zur Beurteilung der Forschungsleistung herangezogen, wird entweder die Anzahl der in einem bestimmten Zeitraum erfolgreich abgeschlossenen Promotions- bzw. Habilitationsverfahren oder die Anzahl der zum Messzeitpunkt laufenden Promotions- bzw. Habilitationsverfahren, also die Anzahl der Doktorandinnen und Doktoranden oder Habilitandinnen und Habilitanden, berücksichtigt (Schenker-Wicki 1996, S. 115). Die Diskussion um die Eignung dieser Indikatoren beginnt dabei bereits mit der Überlegung, ob die Betreuung von Doktorandinnen und Doktoranden nicht eher eine Form von Lehrleistung sei (Müller 2012, S. 83). Natürlich kann die Betreuungsbeziehung während der Promotionsphase mehr oder weniger stark durch lehrbezogene Aspekte geprägt sein (z. B. Durchführung von Kolloquien und anderen promotionsrelevanten Lehrveranstaltungen, Unterstützung bei auftretenden Problemen oder offenen Fragen etc.). Primär steht bei Promotions- oder Habilitationsprojekten aber die Durchführung einer meist umfangreichen Forschungsarbeit und damit klar der forschende Aspekt im Vordergrund. Daher erscheint es durchaus legitim, Indikatoren auf Basis von Promotionen oder Habilitationen als Forschungsindikatoren zu verstehen. Promotionen und Habilitationen werden zudem oft auch als Indikatoren zur quantitativen (und qualitativen) Bewertung der Nachwuchsförderung interpretiert, wobei Nachwuchsförderung hierbei meist als eine Dimension des Konstrukts Forschung verstanden wird (Jansen 2009, S. 47; Schmidt und Weishaupt 2008, S. 113).

Wie bei fast allen der hier dargestellten Indikatoren ergibt sich auch bei der Nutzung der Indikatoren Promotionen und Habilitationen das Problem, dass qualitative Kriterien

außer Acht gelassen werden und somit keine Rückschlüsse auf die Qualität der Forschungsarbeit einer Arbeitseinheit gezogen werden können (Müller 2012, S. 83). Die Anzahl der (abgeschlossenen) Promotions- und Habilitationsverfahren sagt zunächst nur etwas über den Umfang bzw. die Intensität der Nachwuchsförderung in der Forschung aus und erlaubt keine Aussage über die Qualität der einzelnen Arbeiten oder den forschenden Nachwuchs generell.

Hinsichtlich der Verbreitung des Einsatzes von Promotionen und Habilitationen als Indikatoren zur Quantifizierung der Forschungsleistung zeigt sich ein ähnliches Bild wie bereits bei den Drittmitteln. Innerhalb der hochschulinternen Steuerungs- und Anreizverfahren haben Promotionen und Habilitationen als genutzter Indikator insgesamt einen hohen Stellenwert, wobei die Nutzungshäufigkeit zwischen den einzelnen Verfahren teils erheblich differiert (Bogumil et al. 2013, S. 81–87). Insbesondere die Anzahl der Promotionen, seltener die Anzahl der Habilitationen, wird auch in Hochschulrankings als Indikator herangezogen. Die Promotionsquote soll dabei entweder als Indikator für die Lehrleistung, wie im *Times Higher Education Ranking* (Rauhvargers 2013, S. 32), oder als Indikator für die Forschungsleistung, wie im *CHE Ranking Vielfältige Exzellenz* (vgl. Berghoff et al. 2011), fungieren. Dies spiegelt die bereits angesprochene Problematik wider, dass die Anzahl der Promotionen als Leistungsindikator nicht trennscharf einem der Bereiche Forschung oder Lehre zuzuordnen ist. Es herrscht damit ein gewisser Interpretationsraum vor, der den jeweils Verantwortlichen (z. B. den Herausgebern von Rankings oder den Entscheidern in der Ausgestaltung von Mittelverteilungsverfahren) die Definitionsmacht überlässt.

Wie auch beim Drittmittelindikator werden hinsichtlich der Nutzung von Promotions- und Habilitationsquoten als Forschungsindikatoren mögliche negative Effekte befürchtet. So wird diskutiert, ob gesetzte Anreize (z. B. zusätzliche Gelder im Rahmen der indikatorenbasierten Mittelverteilung) zur Erhöhung der Promotions- und Habilitationsquoten nicht dazu führen könnten, dass es zu einer Absenkung der qualitativen Standards kommt (Kühl 2015a, S. 3; Minssen und Wilkesmann 2003, S. 124). Durch das Herunterschrauben der Anforderungen könnten dann auch qualitativ weniger gute Arbeiten für eine erfolgreiche Promotion ausreichen oder auch die Bearbeitungsdauer verkürzt und damit die Promotionsquote erhöht werden.

Es zeigt sich, dass auch die Nutzung von Promotionen und Habilitationen als Indikatoren zur Messung der Forschungs- oder – je nach Definition – Lehrleistung nicht unumstritten ist und teils negative Effekte durch das Setzen von Fehlanreizen erwartet werden. Nichtsdestotrotz ist insbesondere der Einsatz von Promotionsquoten in indikatorenbasierten Steuerungs- und Anreizverfahren und Hochschulrankings gängige Praxis. Neben Drittmitteln finden Promotionen als Indikatoren im Bereich der Forschungsevaluation mit am häufigsten Verwendung. Einen weiteren Indikator, der oft in Kombination mit den

beiden erstgenannten eingesetzt wird, stellen der Publikationsoutput bzw. verschiedene Varianten publikationsbasierter Indikatoren dar. Diese bilden zusammen mit drittmittelbasierten Indikatoren und den Promotionen die „vielbenutzte Trinität Drittmittel – Publikationen – Promotionen“ (Grözinger und Leusing 2006, S. 13) bei der Messung von Forschungsleistung und werden nun im folgenden Kapitel näher erläutert.

Publikationen

Publikationsbasierte Indikatoren kommen in unterschiedlichen Variationen zum Einsatz, wobei allen Varianten gemeinsam ist, dass stets der Publikationsoutput, ermittelt anhand der Anzahl relevanter Veröffentlichungen innerhalb eines bestimmten Zeitraumes, die interessierende Größe ist. Qualitative Aspekte der Publikationen bleiben also auch hier zunächst unberücksichtigt. Allerdings wird über eine Wertung der Qualität der Publikationsorte, z. B. der Reputation einer veröffentlichenden Zeitschrift, indirekt versucht, auch die Qualität von Publikationen mit einzubeziehen. Hierauf wird im Folgenden noch genauer eingegangen werden.

Die zur Messung der Forschungsleistung verwendeten Publikationsindikatoren unterscheiden sich zunächst hinsichtlich des Publikationsformats (Monografie, Zeitschriftenartikel, Sammelband, Tagungsband etc.), welches als für den Indikator relevant definiert wurde. So können alle Veröffentlichungen eines Autors unabhängig vom Publikationsformat oder beispielsweise nur Zeitschriftenartikel zur Ermittlung des Publikationsoutputs berücksichtigt werden. Publikationsbasierte Indikatoren können sich weiter darin unterscheiden, ob eine Gewichtung nach der Anzahl der an einer Veröffentlichung beteiligten Autorinnen und Autoren oder auch nach der Anzahl der Seiten der Publikation vorgenommen wird. Eine dritte Differenzierung betrifft den Erscheinungsort von Zeitschriftenartikeln. Hier spielt zum einen die Unterscheidung zwischen Zeitschriften, bei denen der Veröffentlichung eines Artikels ein Peer-Review-Verfahren vorausgeht, und solchen ohne Begutachtungsverfahren eine Rolle. Zum anderen wird auch zwischen Veröffentlichungen in nationalen und internationalen Zeitschriften unterschieden (vgl. zur Ausgestaltung verschiedener Publikationsindikatoren Fidler und Parsons 2008; Müller 2012; Rauhvargers 2011; Sieweke et al. 2014; Wissenschaftsrat 2008b). Auch Zeitschriftenrankings, die vor allem in den Wirtschaftswissenschaften populär sind, werden teils als Gewichtungskriterium zur Bestimmung der Qualität oder des ‚Impacts‘ eines Artikels herangezogen. Die Rankings basieren dabei entweder auf Häufigkeitsangaben, wie oft aus einer Zeitschrift zitiert wurde (z. B. *Journal Citation Reports* von Thomson Reuters), oder auf Befragungen von Wissenschaftlerinnen und Wissenschaftlern, die die Qualität einer Zeitschrift beurteilen und diese in vorgegebene Kategorien eingruppierten sollen (*VHB-JOURQUAL* des *Verbands der Hochschullehrer für Betriebswirtschaft*) (vgl.

zu Zeitschriftenrankings Baum 2011; Kieser 2010, 2012). Zur Berechnung publikationsbasierter Indikatoren werden die zu berücksichtigenden Artikel dann anhand der Rankingposition der veröffentlichenden Zeitschrift gewichtet (Sieweke et al. 2014, S. 290–291).

Welche Variante der Indikatorengestaltung zum Einsatz kommt, hängt insbesondere auch von der Publikationskultur des zu evaluierenden Faches ab. Je nach Fach haben bestimmte Publikationsformate und Publikationsstrategien einen unterschiedlich hohen Stellenwert (vgl. zur Beschreibung der Publikationskulturen verschiedener Fächer Alexander von Humboldt Stiftung 2009). So ist z. B. die Gewichtung nach der Anzahl der Autorinnen und Autoren nicht zielführend, wenn in einer Fachdisziplin Artikel meist ohnehin von der gesamten Forschergruppe veröffentlicht werden. Ebenso ist die Verwendung der Anzahl der veröffentlichten Monografien als Indikator wenig sinnvoll, wenn in der betrachteten Disziplin das Verfassen von Monografien in der Praxis keinen Stellenwert (mehr) hat. Aufgrund dieser fachlichen Unterschiede in den Publikationskulturen sehen sich Versuche, den Publikationsoutput nicht differenziert nach Fach, sondern mit einem einheitlichen Standard zu erfassen (z. B. in Hochschulrankings), mit vielfältiger Kritik konfrontiert. Meist bilden die Datenbanken *Web of Science* (vormals *ISI Web of Knowledge*) oder *Scopus* die Basis solcher Publikationsanalysen. In diesen Datenbanken sind aber englischsprachige Zeitschriftenartikel überrepräsentiert, während Monografien und Sammelbandbeiträge deutlich seltener erfasst sind (Federkeil 2013, S. 43). Hieraus ergeben sich zwei wesentliche Probleme bei vergleichenden Publikationsanalysen: Zum einen wird der Publikationsoutput von Fächern, deren Publikationskultur sich nicht überwiegend auf Veröffentlichungen in Zeitschriften beschränkt (z. B. bei geisteswissenschaftlichen Fächern) systematisch unterschätzt (Kroth und Daniel 2008, S. 554). Es ergibt sich ein falsches Bild des Publikationsverhaltens, das in Richtung Zeitschriftenpublikationen hin verzerrt ist, und andere Formate nur in Teilen in die Analyse mit einbezieht. Zum anderen werden durch den englischsprachigen Fokus der Datenbanken Fächer, deren Veröffentlichungssprache nicht primär Englisch ist, bei der Erfassung des Publikationsoutputs benachteiligt. Durch den Sprachbias können sich z. B. bei ländervergleichenden Hochschulrankings zudem Nachteile für nicht-englischsprachige Hochschulen ergeben (Amsler und Bolsmann 2012, S. 287; Lenzen 2012, S. 18). Daneben bleibt festzuhalten, dass sowohl die genannten, aber auch andere Publikationsdatenbanken eine Reihe weiterer Mängel aufweisen, die zu methodischen Problemen bei Publikationsanalysen führen und die Validität der darauf aufbauenden Indikatoren deutlich einschränken können. So folgt die Erfassung von Institutionen und Namen der Autorinnen und Autoren oft keinem einheitlichen Schema, sondern diese sind in unterschiedlichen Schreibvarianten erfasst, was erschöpfende Analysen des Publikationsoutputs erschwert. Hinsichtlich einer erschöpfenden Erfassung aller relevanten Publikationen ist

zudem, neben dem Sprachbias und der Vernachlässigung anderer Publikationsformate, generell mit Erfassungslücken bei den Datenbanken zu rechnen (vgl. zur Diskussion der Datenbanken Fidler und Parsons 2008; Müller 2012; Robinson-García und Calero-Medina 2014).

Es wurde deutlich, dass bereits der Versuch der rein quantitativen Erfassung des Publikationsoutputs mit zahlreichen Problemen behaftet ist. Sollen nun zur Bewertung der Publikationsleistung auch qualitative Aspekte der Publikationen in den verwendeten Indikatoren Berücksichtigung finden, sind hierbei weitere mögliche Einschränkungen zu diskutieren. Wie bereits angemerkt, wird zur Beurteilung der Qualität des Publikationsoutputs oft auf die Reputation der veröffentlichenden Zeitschrift zurückgegriffen (z. B. ermittelt anhand von Zeitschriftenrankings) bzw. darauf, ob der veröffentlichte Artikel einen Peer-Review-Prozess durchlaufen hat. Inwieweit ein erfolgreich durchlaufenes Peer-Review-Verfahren als Qualitätsindikator angesehen werden kann, ist allerdings umstritten. So konnte in Studien empirisch eine nur geringe Gutachterreliabilität, also eine nur geringe Übereinstimmung zwischen zwei oder mehreren Gutachterurteilen nachgewiesen werden (vgl. Fischer 2011; Miller 2006). Will man die Anzahl von Zitationen eines Artikels als Qualitätsmerkmal interpretieren, so zeigt sich in anderen Studien auch im Hinblick auf die spätere Zitationshäufigkeit nur eine schwache Korrelation mit den Gutachterurteilen (vgl. Singh et al. 2007; Starbuck 2005). Es lassen sich aber auch empirische Befunde finden, die hinsichtlich der Reliabilität und Validität von Gutachterurteilen zu einem deutlich positiveren Ergebnis kommen und Peer-Review-Prozessen durchaus prognostische Validität zusprechen (Daniel et al. 1998, S. 35–47). Die teils widersprüchlichen Ergebnisse der verschiedenen Studien sind vermutlich auch auf die unterschiedliche Datenbasis (untersuchte Fächer) und auf (qualitative) Unterschiede der verschiedenen Gutachterverfahren zurückzuführen. Festzuhalten bleibt, dass ein erfolgreiches Peer-Review-Verfahren noch kein Garant für die Qualität eines Artikels ist, auch wenn es empirische Hinweise auf eine höhere Resonanz eines begutachteten Artikels innerhalb der Scientific Community in Form von Zitationen gibt. Auf die Frage, inwieweit die Zitationshäufigkeit als Indikator für die Qualität einer Veröffentlichung oder Zeitschrift geeignet ist, wird noch im folgenden Kapitel näher eingegangen.

Da das Publizieren wissenschaftlicher Erkenntnisse eines der zentralen Elemente der Forschungsarbeit darstellt, welches zudem quantitativ vermeintlich leicht zu erfassen ist, verwundert es nicht, dass publikationsbasierte Indikatoren zu den am häufigsten eingesetzten Forschungsindikatoren zählen. So arbeiten die meisten der bekannten Hochschulrankings mit publikationsbasierten Indikatoren und einzelne Rankings wie das *CWTS Leiden Ranking* (vgl. zur kritischen Diskussion des Leiden Rankings Bornmann et al. 2013) beschränken sich sogar gänzlich auf Analysen des Publikationsoutputs. Im Zeitverlauf werden die von den Rankings verwendeten Publikationsindikatoren und die

zugrundeliegende Methodik oft überarbeitet und angepasst (vgl. Rauhvargers 2013, 2014), was sicherlich auch auf die anhaltende kritische Diskussion bezüglich der Aussagekraft publikationsbasierter Analysen zurückzuführen ist. Das *CHE Ranking* verzichtete beispielsweise im Jahr 2011 aufgrund der „Unzulänglichkeit der [...] für das *CHE Ranking* verwendeten Publikationsdatenbank“ (CHE Centrum für Hochschulentwicklung 2012, S. 1) im Fach Soziologie gänzlich auf bibliometrische Analysen. Auch in anderen Fächern wie der Politikwissenschaft werden im *CHE Ranking* keine Publikationsanalysen (mehr) durchgeführt.¹²

Innerhalb der hochschulinternen Steuerungs- und Anreizverfahren zeigen sich bei der Verwendung publikationsbasierter Indikatoren, wie bereits bei den Promotionen, wieder deutliche Unterschiede zwischen den einzelnen Verfahren. Beispielsweise werden nach der Erhebung von Bogumil et al. (2013) Publikationen innerhalb der universitätsinternen indikatorenbasierten Mittelzuweisung von allen abgefragten Indikatoren am seltensten genutzt (13 % der befragten Kanzler), während sie in der fakultätsinternen indikatorenbasierten Mittelzuweisung am vierthäufigsten (37 % der befragten Dekane) zum Einsatz kommen (ebd., S. 82–85).

Auch mit dem Einsatz von publikationsbasierten Forschungsindikatoren können nicht-intendierte Effekte verbunden sein. Vielfach werden negative Auswirkungen auf die Qualität der Publikationen diskutiert, die aus dem Druck resultieren, möglichst viel zu publizieren, um möglichst gut in der hochschulinternen Forschungsevaluation abzuschneiden. Damit verbunden wird auch die zunehmende Tendenz zur sukzessiven Veröffentlichung von Forschungsergebnissen in mehreren Publikationen („Salomitaktik“) kritisiert (vgl. Alberts 2013; Kieser 2010; Oancea 2014; Weingart 2005). Die Ausrichtung an Mainstreamforschung und die Wahl von Forschungsthemen, die sich leichter publizieren lassen (vgl. Kieser 2012; Kieser und Osterloh 2012; Lenzen 2012), sind weitere mögliche Folgen einer zu starken Orientierung an den Kriterien der Publikationsindikatoren. Zudem kann es in der Forschungsarbeit der Professorinnen und Professoren zur Vernachlässigung von Buchpublikationen und anderssprachigen Publikationen zugunsten von Artikeln in englischsprachigen Zeitschriften kommen (vgl. Fejes und Nylander 2014; Kieser 2010).

Wenn es um die Quantifizierung von wissenschaftlichem Publikationsverhalten geht, spielen in vielen Fällen auch Zitationen eine Rolle. Die Funktion von Zitationen als Indikator für die Qualität einer wissenschaftlichen Publikation wurde bereits angesprochen. Im nächsten Kapitel wird nun ausführlich auf die Ausgestaltung und Verbreitung von Forschungsindikatoren auf der Basis von Zitationen und Zitationskennzahlen eingegan-

¹² Vgl. zur Methodik und den verwendeten Indikatoren im *CHE Ranking* das Methoden-Wiki des CHE unter <http://www.che-ranking.de/methodenwiki/index.php/Hauptseite> (03.02.2016).

gen und diskutiert, wie sich die bereits dargestellten nichtintendierten Effekte publikationsbasierter Forschungsevaluation durch die Verwendung zitationsbasierter Indikatoren noch verschärfen können.

Zitationen und Zitationskennzahlen

Das Zitieren anderer Veröffentlichungen ist ein wesentlicher Bestandteil der wissenschaftlichen Forschungspraxis und ermöglicht durch die strukturierte Bezugnahme auf die (Vor-)Arbeiten anderer Autorinnen und Autoren das systematische Fortschreiten des wissenschaftlichen Erkenntnisgewinns. Besonders einflussreiche Arbeiten, auf die sich weitere Forschung stützen kann, werden auch besonders häufig zitiert. Daher ist es wenig verwunderlich, dass beim Versuch Forschungsleistung zu messen, auch auf zitationsbasierte Indikatoren zurückgegriffen wird. Mit diesen soll die Qualität bzw. der Einfluss der Arbeiten auf das wissenschaftliche Feld erfasst werden. Allerdings werden zitationsbasierte Indikatoren weniger oft zur Evaluation der Forschungsleistung eingesetzt, als man aufgrund des Stellenwerts von Zitationen in der wissenschaftlichen Forschungsarbeit vermuten könnte. So finden sie in den hochschulinternen Steuerungs- und Anreizverfahren bislang nur selten Verwendung, was unter anderem auf eine mangelnde Akzeptanz hinsichtlich der Eignung zitationsbasierter Forschungsindikatoren zurückgeführt werden kann (Sieweke et al. 2014, S. 295). In internationalen Rankings und auch zur landesweiten Forschungsevaluation in Ländern wie Schweden und England kommen zitationsbasierte Indikatoren dagegen häufiger zum Einsatz (vgl. Federkeil 2013; Fejes und Nylander 2014; Rauhvargers 2013).

In der Diskussion um die Eignung von Zitationen als Forschungsindikator stellt sich zunächst die Frage, ob die Anzahl der Zitationen einer Veröffentlichung überhaupt ein Qualitätskriterium ist. Schließlich werden Arbeiten auch dann zitiert, wenn die publizierten Ergebnisse falsch sind, angezweifelt werden oder ihnen widersprochen wird (Kieser 2010, S. 352). Auch erhöht eine hohe Sichtbarkeit der eigenen Veröffentlichung innerhalb der relevanten Scientific Community die Chance von anderen Autorinnen und Autoren zitiert zu werden. Einen Einfluss auf die Sichtbarkeit kann z. B. die Bekanntheit und die Auflage der veröffentlichenden Zeitschrift haben oder ob ein Artikel (zusätzlich) im Internet veröffentlicht und als Online-Ressource verfügbar gemacht wird. Wie oft eine Veröffentlichung zitiert wurde, ist also nicht zwangsläufig als Qualitätsmerkmal zu verstehen.

Meist wird zur Forschungsevaluation aber nicht einfach die Anzahl aller Zitationen einer Autorin oder eines Autors herangezogen oder diese auf Ebene einer Arbeitsgruppe aggregiert, sondern es werden Zitationskennzahlen wie der *Journal Impact Factor* oder der *h-Index* als Indikatoren genutzt. Diese sind konstruktionsbedingt in Bezug auf die zugrundeliegende Datenbasis in hohem Maße selektiv. Der *Journal Impact Factor*, der

eine zitationsbasierte Kennzahl für den Einfluss einer wissenschaftlichen Zeitschrift darstellt (vgl. zur Berechnung Baum 2011), und der *h-Index*, als eine personenbezogene Zitationskennzahl (vgl. zur ausführlichen Beschreibung und Berechnung Ferrara und Romero 2013), basieren z. B. beide auf der *Web of Science*-Datenbank. Wie bereits bei den publikationsbasierten Indikatoren ergibt sich damit das Problem, dass Zitationen in Quellen, die nicht in der Datenbank erfasst sind, für die Berechnung der Kennzahlen unberücksichtigt bleiben. Aufgrund dieses Datenbank-Bias wird – je nach Publikationskultur – für Teile der Sozialwissenschaften der Einsatz zitationsbasierter Kennzahlen als ungeeignet angesehen (Rauhvargers 2011, S. 42). Baum (2011) zweifelt die Aussagekraft des *Journal Impact Factors* generell an, da aufgrund der stark schiefen Verteilung der Zitationshäufigkeiten der verschiedenen Artikel einer Zeitschrift, Rückschlüsse vom Wert des *Journal Impact Factors* auf die Qualität eines einzelnen Artikels nicht zulässig seien:

If the distribution of citations to a journal's articles is highly skewed, with few articles near the mean, it is inappropriate to ascribe its IF [Impact Factor; Anm. d. Verf.] to individual articles published within it. High skewness in the citedness of a journal's articles also undermines the IF as an index of journal citedness by making a small number of highly-cited articles disproportionately influential in determining the journal's IF and its resultant ranking. If being cited is taken as an indicator of quality, then the practice of associating a journal's IF with the quality of its articles has the potential to be seriously misleading: The majority of articles published in high-IF journals may be basking, along with the journals themselves, in the reflected glory of a small minority of citation-worthy articles. (ebd., S. 451)

Neben der methodischen Kritik werden im Hinblick auf die Verwendung zitationsbasierter Forschungsindikatoren negative Auswirkungen auf das Publikationsverhalten der Wissenschaftlerinnen und Wissenschaftler befürchtet. Analog zu den nichtintendierten Effekten publikationsbasierter Forschungsindikatoren besteht die Gefahr, dass überwiegend nur noch in den für die Zitationsindikatoren relevanten Formaten und Zeitschriften publiziert wird. Dies kann dazu führen, dass bestimmte Forschungsfelder vernachlässigt werden, wenn thematisch keine passenden Publikationsformate zur Verfügung stehen. Lau, Cisco und Delgado-Romero (2008) nennen als Beispiel Artikel aus dem Bereich der *Multicultural Psychology*, welche für die Veröffentlichung in hoch gerankten Zeitschriften traditionell nicht berücksichtigt werden (ebd., S. 203–204). Weiter werden Trickserien und manipulatives Verhalten beim Publizieren, sogenanntes *Gaming*, als mögliche Konsequenzen einer zu starken Orientierung an zitationsbasierten Indikatoren diskutiert. Zu nennen sind hier unter anderem vermehrte Selbstzitationen (Kieser 2010, S. 355), das

Veröffentlichen von Artikeln, die eine hohe Zitationsrate versprechen, wie z. B. Metaanalysen (Kieser 2012, S. 105), und die Kooperation „mit berühmten Kollegen [...], um in den Genuss besserer Zitationsraten zu gelangen“ (Lenzen 2012, S. 17). Zudem kommen Wilhite und Fong (2012) in einer großangelegten Befragungsstudie zu dem Schluss, dass auch externe Einflussnahmen auf das Zitationsverhalten der Wissenschaftlerinnen und Wissenschaftler durch die Zeitschriften selbst vorkommen:

Overall, the empirical results from the author survey and the journal-based data tell a consistent story. Coercive self-citation exists and is more common in the business disciplines than in economics, sociology, and psychology. Additionally, some editors seem to target specific articles and authors. (ebd., S. 543)

Insbesondere zeitschriftenbasierte Zitationskennzahlen, welche andere Formate unberücksichtigt lassen, sind also mit einigen methodischen Problemen konfrontiert und bergen zudem die Gefahr, dass verstärkt nichtintendierte Effekte in Bezug auf das Publikations- und Zitationsverhalten auftreten können. Durch die Verwendung zitationsbasierter Kennzahlen, bei denen die Gesamtheit der Zitationen einer Autorin oder eines Autors generell, unabhängig von der veröffentlichenden Zeitschrift oder dem Publikationsformat erfasst werden, können einige Nachteile zeitschriftenbasierter Zitationskennzahlen kompensiert werden (Müller 2012, S. 78). Probleme wie das Fehlen einer geeigneten Datenbank für einzelne Fächer (ebd.) oder Fehlanreize zur manipulativen Selbstzitation oder sogar dem Bilden von Zitationskartellen (ebd., S. 82) bestehen aber auch hier.

Lässt man die genannten methodischen Schwierigkeiten außer Acht, ergibt sich bei der Verwendung von Zitationen als Forschungsindikator weiterhin die bereits am Anfang des Kapitels erwähnte Problematik, dass Zitate nicht zwangsläufig als Wertschätzung der Qualität oder der Erkenntnisse anderer Arbeiten zu verstehen sind (vgl. Kieser 2010, S. 352). Dieser Argumentation, welche die generelle Eignung von Zitationen bzw. Zitationskennzahlen als Forschungsindikator anzweifelt, wird von Kroth und Daniel (2008) entgegengehalten, dass zwar nicht jede Zitation die Qualität der zitierten Arbeit widerspiegelt, aber „(...) in der Regel [zeigt], dass die Publikation einen anerkannten Beitrag zum wissenschaftlichen Fortschritt oder zur wissenschaftlichen Debatte darstellt“ (ebd., S. 554). Ähnlich argumentiert Müller (2012):

Fraglich ist allerdings, ob dieser Umstand [das Vorhandensein negativer Zitationen; Anm. d. Verf.] die Validität von Zitationsindikatoren zur Forschungsleistungsmessung tatsächlich einschränkt. Thesen, die sich als falsch erwiesen haben, werden schließlich nur dann zitiert, wenn sie trotzdem als hinreichend bedeutsam anerkannt wurden und den neuen Erkenntnissen im Weg standen. (ebd., S. 73)

Folgt man dieser Argumentationslinie und akzeptiert die grundsätzliche Eignung von Zitationen als Qualitätsmerkmal einer wissenschaftlichen Arbeit, lässt sich zusammenfassend festhalten, dass die Validität zitationsbasierter Forschungsindikatoren mit dem Vorhandensein geeigneter Datenbanken, die alle relevanten Publikationsformate einer (Sub-)Disziplin erfassen und damit erschöpfende Zitationsanalysen ermöglichen, steht und fällt.

Mit den Drittmitteln, Promotionen und Habilitationen, Publikationen sowie den Zitationen und Zitationskennzahlen wurden die wichtigsten und am weitesten verbreiteten Forschungsindikatoren dargestellt. Abschließend sollen im Sinne der Vollständigkeit noch einige weitere Indikatoren, die vereinzelt zur Forschungsevaluation oder in Rankings Verwendung finden, beschrieben und kurz diskutiert werden.

Weitere Forschungsindikatoren

Ein weiteres mögliches Bewertungskriterium der Forschungsleistung betrifft den Grad der internationalen Ausrichtung bzw. der internationalen Vernetzung einer Hochschule, eines Fachbereiches oder einer Arbeitsgruppe. Je nach zugrundeliegender Definition und Operationalisierung von Internationalität wird der entsprechende Indikator zur Bewertung der Dimension Lehre oder Forschung herangezogen. Wird der Grad der Internationalität einer Hochschule beispielsweise anhand der Anzahl ausländischer Studierender ermittelt, erscheint die Zuordnung zur Dimension Lehre relativ eindeutig. Weniger klar wird die Zuordnung, wenn – wie z. B. im *Times Higher Education Ranking* – der Anteil ausländischer Mitarbeiterinnen und Mitarbeiter als Indikator herangezogen wird (vgl. zu den verwendeten Indikatoren im Ranking Hazelkorn 2013, S. 18). Da die Mitarbeiterinnen und Mitarbeiter in den meisten Fällen sowohl in der Lehre als auch in der Forschung aktiv sind und beide Bereiche ohnehin nicht trennscharf voneinander gedacht werden können, ist eine eindeutige inhaltliche Zuordnung schwierig. Oft wird die internationale Ausrichtung, Internationalität, internationale Vernetzung etc. daher auch als eigenständige Bewertungsdimension verstanden bzw. dargestellt.¹³ Der Aspekt der *Internationalisierung* spielt dabei nicht nur in einzelnen Hochschulrankings eine Rolle, sondern ist auch Gegenstand von hochschulinternen Zielvereinbarungen und – in deutlich geringerem Maße – der indikatorenbasierten Mittelzuweisung (Bogumil et al. 2013, S. 80–85).

Im Rahmen universitärer Steuerungsinstrumente werden vereinzelt auch forschungsbezogene Tätigkeiten innerhalb der Scientific Community wie z. B. gehaltene Vorträge,

¹³ Vgl. z. B. das Entscheidungsmodell des *CHE Ranking* mit dem eigenen ‚Baustein‘ *Internationale Ausrichtung* unter <http://www.che-ranking.de/methodenwiki/index.php/Indikatoren> (09.02.2016).

organisierte Konferenzen oder die Anzahl erstellter Gutachten als Indikatoren herangezogen (ebd., S. 85). Ebenso werden gelegentlich erhaltene Wissenschafts- und Forschungspreise zur hochschulinternen Forschungsevaluation eingesetzt (Sieweke et al. 2014, S. 289). Hinsichtlich der Wissenschafts- und Forschungspreise ist zu ergänzen, dass diese nicht nur als Indikatoren für gute Forschungsleistung angesehen werden können, sondern auch selbst als Anreiz für gute Forschungsleistung wirken und als entsprechendes anreizbasiertes Steuerungsinstrument verstanden werden können. Die Anreizwirkung basiert dabei zum einen auf dem materiellen Gewinn und zum anderen, in vermutlich höherem Maße, auf dem mit dem Erhalt der Preise verknüpften Reputationszuwachs in der Scientific Community (Krempkow et al. 2013, S. 47).

Die Reputation selbst, genauer die Reputation von ganzen Hochschulen bzw. Fachbereichen, wird in einzelnen Hochschulrankings wie dem *QS World University Ranking*, dem *Times Higher Education Ranking* und dem *CHE Ranking* ebenfalls zur Bewertung der Forschungsleistung eines Standortes genutzt. Dies geschieht meist auf der Basis von Befragungen von Wissenschaftlerinnen und Wissenschaftlern oder Arbeitgebern, welche, teils anhand vorgegebener Listen von Hochschulen, diejenigen Standorte mit besonders hoher Forschungsreputation benennen sollen (Federkeil 2013, S. 40–41; Rauhvargers 2011, S. 65). Diese Formen der Reputationsbefragungen sind vielfacher Kritik ausgesetzt. So sind die erreichten Rücklaufquoten für Aussagen mit repräsentativem Anspruch vielfach als zu gering anzusehen (Glaser 2012, S. 211; Rauhvargers 2011, S. 65; Williams 2008, S. 55–56). Zudem werden in den großen internationalen Rankings zur Stichprobe selbst nur wenig Informationen veröffentlicht (Federkeil 2013, S. 41), so dass die Darstellung der Methodik unvollständig und nicht nachvollziehbar bleibt. Gerade im Hinblick auf regionale Verzerrungen bei Reputationsmessungen (vgl. zu regionalen Effekten bei Reputationsmessungen Honecker 2011, S. 21–22; Kroth und Daniel 2008, S. 551) wären Informationen zum Stichprobenaufbau und eventuell durchgeführten Gewichtungungsverfahren wünschenswert. Grundsätzliche Zweifel gibt es an der generellen Eignung von Reputation als Qualitätsindikator. So weist Ridder (2010) darauf hin, dass in „(...) der öffentlichen Wahrnehmung (...) die Reputation einer Hochschule zunächst eher ihrem zumeist über Jahre traditionell gewachsenen Nimbus zugeordnet [wird] und weniger ihrer Kompetenz in Forschung oder gar Lehre“ (ebd., S. 53). Auch Kroth und Daniel (2008) halten fest, dass „Reputationswerte (...) keine adäquaten Indikatoren für die Qualität von Forschung und Lehre [sind], weil der Ruf einer Universität noch lange nach der tatsächlichen Leistung anhält und traditionsreiche Universitäten bevorteilt sind“ (ebd., S. 550). Die Reputation einer Hochschule oder eines Fachbereiches kann also nicht zwangsläufig als Indikator für die aktuelle Qualität der Forschungsarbeit aufgefasst werden, sondern ist eher ein Repräsentant für vergangene gute Leistungen.

Abschließend sei noch darauf hingewiesen, dass die hier vorgenommene Darstellung von Forschungsindikatoren nicht als erschöpfend verstanden werden kann. Die zur Forschungsevaluation herangezogenen Indikatoren und die Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren kann zwischen einzelnen Hochschulen erheblich differieren und in der Praxis können noch weitere Indikatoren, insbesondere auch solche mit einem mehr qualitativen Fokus, zum Einsatz kommen. In diesem Kapitel lag der Schwerpunkt auf der Beschreibung der gängigsten Forschungsindikatoren. Auf die Darstellung teils exotischer Indikatoren, wie die Anzahl der (ehemaligen) Mitarbeiter, die einen Nobelpreis erhalten haben (vgl. zur Verwendung im *ARWU Ranking* Rauhvargers 2011, S. 25), wurde daher verzichtet.

Im folgenden Kapitel werden nun die wichtigsten empirischen Erkenntnisse verschiedener Studien zur Thematik der indikatorenbasierten Leistungsmessung im Bereich Forschung und zum Einsatz, der Ausgestaltung und der Wirkung universitärer Steuerungsinstrumente näher in den Blick genommen.

2.4 Umsetzung und Wirkung indikatorenbasierter Steuerungs- und Anreizverfahren

Mit dem Bedeutungszuwachs des New Public Managements und den ersten Implementierungen von staatlichen und hochschulinternen Steuerungsverfahren in Deutschland rückten diese alsbald in den Fokus der wissenschaftlichen Forschung. Wissenschaftlerinnen und Wissenschaftler verschiedener Disziplinen (unter anderem Erziehungswissenschaft, Geschichtswissenschaft, Jura, Politikwissenschaft, Psychologie, Medizin, Soziologie, Wirtschaftswissenschaften) befassen sich z. B. mit Fragestellungen zum Umsetzungsstand von indikatorenbasierten Steuerungs- und Anreizverfahren in Ländern und Universitäten, zur Wahrnehmung und Bewertung von diesen durch die betroffenen Akteure und zu den intendierten und nichtintendierten Wirkungen auf die Hochschullandschaft generell und die Bereiche Lehre und Forschung im Speziellen. In diesem Zusammenhang sind hier insbesondere die vom BMBF finanzierten Förderlinien „Leistungsbeurteilung in der Wissenschaft“¹⁴, in deren Rahmen auch diese Dissertation entstanden ist, sowie „Neue Governance der Wissenschaft - Forschung zum Verhältnis von Wissenschaft, Politik und Gesellschaft“¹⁵ zu nennen.

Im Rahmen der letzteren untersuchten Bogumil et al. (2013) in einem von der Hans-Böckler-Stiftung mitfinanzierten Forschungsprojekt mit Hilfe von Fallstudien und einer

¹⁴ <http://www.hochschulforschung-bmbf.de/de/1333.php> (25.10.2016)

¹⁵ <https://www.bmbf.de/foerderungen/bekanntmachung.php?B=317> (25.10.2016)

quantitativen Befragung von Hochschulakteuren (unter anderem Rektorate, Kanzlerinnen und Kanzler, Dekanate, Professorinnen und Professoren) inwieweit neue Steuerungsinstrumente wie die indikatorenbasierte Mittelzuweisung oder Zielvereinbarungen in Bundesländern und Hochschulen umgesetzt und ausgestaltet sind. Ergebnisse der Studie hierzu wurden in dieser Arbeit bereits teilweise berichtet. Daneben wurde in der Untersuchung auch die Einschätzung der verschiedenen Hochschulakteure zu den Wirkungen der eingesetzten Verfahren ermittelt. Im Rahmen eines Folgeprojektes¹⁶ wurden die Ergebnisse der im Jahr 2010 durchgeführten Befragung in einer längsschnittlichen Trendstudie mit den Ergebnissen einer Erhebung aus dem Jahr 2014 abgeglichen. Hierbei zeigte sich, dass die Performanzwirkung neuer Steuerungsinstrumente (z. B. Anreize zur Erhöhung der Drittmittelwerbungen und der Promotionszahlen) von den Hochschulakteuren zwar tendenziell positiv eingeschätzt wird, dieser allerdings hohe Transaktionskosten in Form von gestiegenem Verwaltungsaufwand und gestiegenem Konfliktpotential (z. B. zwischen Hochschulleitung und untergeordneten Organisationseinheiten) gegenüberstehen. Die Anreizwirkung der Steuerungsinstrumente wird zudem im Jahr 2014 teils deutlich schwächer beurteilt als im Jahr 2010. Als Gründe werden eine veränderte Ausgestaltung der Verfahren bzw. eine Diskrepanz zwischen der Formalstruktur und der Aktivitätsstruktur von Universitäten, das heißt eine in der Realität anders praktizierte Anwendung der Steuerungsverfahren als formal vorgegeben, vermutet. Die Steuerungsverfahren würden demnach teils weniger strikt angewendet, um hochschulinterne Konflikte zu vermeiden (vgl. zu den Ergebnissen der Studie Bogumil et al. 2015).

Innerhalb der Förderlinie „Leistungsbewertung in der Wissenschaft“ nehmen sich verschiedene, teils noch laufende Projekte mit zum jetzigen Zeitpunkt noch nicht veröffentlichten Ergebnissen¹⁷ der Thematik der indikatorenbasierten Steuerung oder verwandten Fragestellung an: Wollersheim et al. (2014) befragten ca. 9000 Wissenschaftlerinnen und Wissenschaftler aus den Wirtschaftswissenschaften zu den an ihren Hochschulen zur Messung der Forschungsleistung verwendeten Indikatoren und zu deren Einschätzungen, anhand welcher Indikatoren eine adäquate Messung von Forschungsleistung zu erfolgen hätte. Dabei zeigte sich, dass die Bewertung von Forschungsleistung an wirtschaftswissenschaftlichen Fakultäten weit verbreitet ist und hierfür vorrangig quantitative Indikatoren wie die Anzahl der Publikationen oder die Höhe der eingeworbenen

¹⁶ Vom BMBF gefördertes Verbundprojekt „Wirkungen neuer Governance auf Aktivitätsstrukturen und Leistungsprofile deutscher Hochschulen“.

¹⁷ Zu nennen sind hier mit Bezug zum Thema der vorliegenden Arbeit vor allem das Projekt „Einfluss von Leistungsbewertungen auf die Integration von Lehre und Forschung im Fächervergleich“ (LILF), welches die Wirkungen von Leistungsbewertungen auf die professionelle Identität in den Blick nimmt, sowie das Projekt „Nicht-intendierte Effekte der Leistungsbewertung in der Wissenschaft“ (NEL), welches sich der Thematik mit Hilfe zweier internationaler Fallstudien annähert.

Drittmittel verwendet werden. Demgegenüber wird bei der Frage danach, welche Indikatoren eingesetzt werden sollten, qualitativen Indikatoren wie dem „Erkenntnisgewinn für die Wissenschaft“ der Vorrang gegeben. Auf die Schwierigkeit des Erfassens solcher qualitativer Indikatoren weisen die Autorinnen hin. Ähnliche Ergebnisse zeigten sich bei einer deutschlandweiten Befragung sozialwissenschaftlicher Professorinnen und Professoren im Projekt *QualRep*. Auch hier wurde deutlich, dass insbesondere die Höhe der Drittmittel und die Anzahl von (peer-reviewed) Publikationen eine wichtige Rolle bei der Bewertung der Forschungsleistung spielen, während andere Indikatoren eher von untergeordneter Bedeutung sind (vgl. Schmidt-Hertha, Tafertshofer, Werner 2016).

Ein weiteres Projekt der Förderlinie befasst sich unter anderem mit der Rolle der Leistungsbewertung in Berufungsverfahren vor dem Hintergrund des Wandels des Governance-Regimes im Hochschulbereich. Mit Hilfe einer quantitativen Inhaltsanalyse von Ausschreibungstexten für Professuren in Ausgaben der ZEIT der Jahrgänge 1995, 2009 und 2012 analysierten Kleimann et al. (2015) die sich verändernden Anforderungen, die an die (geleistete) Forschungsarbeit von Bewerberinnen und Bewerbern gestellt werden. Dabei können die Autoren zeigen, dass im Zeitverlauf statt einer generellen Forschungserfahrung mit zunehmender Häufigkeit konkretisierte Anforderungen hinsichtlich Aspekten wie Drittmittelinwerbung, Publikationen, Nachwuchsförderung und Forschungsoperationen genannt werden. Dies ist im Hinblick auf eine vermutete Steuerungswirkung von Berufungs- und Bleibeverhandlungen auf das Forschungsverhalten als interessanter Befund zu werten.

Am *Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW)*¹⁸ wurden und werden ebenfalls Forschungsprojekte – teils auch im Rahmen der oben genannten Förderlinie – mit thematischem Bezug zu der in dieser Arbeit verfolgten Forschungsfrage durchgeführt: Jaeger et al. (2006) konnten in einer bundesweiten Befragung bereits für das Jahr 2003 eine starke Verbreitung von universitätsinternen indikatorenbasierten Zuweisungsverfahren (85 % der befragten Universitäten setzten diese ein, 12 % planten dies in naher Zukunft zu tun) feststellen. In einer weiteren Fallstudie (Jaeger 2006) konnte gezeigt werden, dass sich die Fakultäten bei der fakultätsinternen Mittelverteilung überwiegend an den universitätsintern verwendeten Modellen zur Verteilung der Gelder an die Fakultäten orientieren und diese um weitere Kriterien wie Gutachtertätigkeiten oder anderes wissenschaftliches Engagement erweitern. Die in der Fallstudie befragten Hochschulleitungen selbst sahen das primäre Ziel einer indikatorenbasierten Mittelzuweisung dabei meist mehr in der Schaffung von Transparenz bei der Mittelvergabe denn in einer Anreizfunktion für gesteigerten Wettbewerb.

¹⁸ Vormalig *HIS-Institut für Hochschulforschung*.

Neben den genannten Publikationen sind beim DZHW eine Reihe von weiteren (Fall-)Studien veröffentlicht, die insbesondere die Ausgestaltung und Wirkung (Evaluation) von indikatorenbasierten Finanzierungsmodellen und Zielvereinbarungen auf der Ebene Staat-Hochschule in den Blick nehmen. Da in der vorliegenden Arbeit vor allem die hochschulinternen Steuerungsversuche auf die Mikroebene der Professorinnen und Professoren von Interesse sind, sei an dieser Stelle lediglich auf die entsprechenden Publikationen von Jaeger und In der Smitten (2009; 2010; In der Smitten und Jaeger 2012) verwiesen.

Eine umfangreiche Studie zur fakultätsinternen indikatorenbasierten Mittelzuweisung im Fach Medizin wurde von Krempkow et al. (2013) vorgelegt. Mit Hilfe von Dokumentenanalysen, Experteninterviews, bibliometrischen Analysen sowie einer standardisierten Online-Befragung von Professorinnen und Professoren und wissenschaftlichem Personal wird die Ausgestaltung, Bewertung und Wirkung der indikatorenbasierten Mittelzuweisung untersucht. Hinsichtlich der Wirkungen wird im Rahmen der Experteninterviews mit den betroffenen Akteuren von einer Zunahme der Publikationstätigkeit, insbesondere in den für die Mittelzuweisung relevanten Zeitschriften mit hohem Impact-Factor, berichtet. Dagegen würden medizinische Fortbildungsarbeiten (so genannte *case reports*) trotz eines großen Leserkreises weniger oft verfasst, da diese bei der indikatorenbasierten Mittelzuweisung keine Berücksichtigung finden. Die Ergebnisse der Befragung offenbaren im Gegensatz zu anderen Fächern (vgl. Minssen et al. 2003, S. 52) eine hohe Kenntnis des wissenschaftlichen Personals über die Ausgestaltung des Mittelzuweisungsverfahrens und ein hohes Interesse an den zugrundeliegenden Indikatorenwerten. In Bezug auf die eingeschätzten Wirkungen zeigt sich, dass jeweils mehr als die Hälfte der Befragten eine Tendenz hin zu vermehrter Mainstreamforschung, dem Verfassen von kürzeren Artikeln sowie zu zunehmenden Konflikten um Koautorenschaften sehen. Allerdings wird auch eine positive Wirkung auf die Leistungsmotivation und die Leistungsgerechtigkeit wahrgenommen. Weiter stimmten die Befragten mehrheitlich zu, dass die mit der Einführung der indikatorenbasierten Mittelzuweisung verbundenen Ziele eines Wettbewerbs zwischen den fakultätsinternen Einrichtungen und einer erhöhten Transparenz der Forschungsleistung erreicht wurden. Eine knappe Mehrheit sieht auch eine Effizienz- sowie Qualitätssteigerung in der Forschung.¹⁹ In der Medizin scheint die indikatorenbasierte Mittelzuweisung in der Wahrnehmung der betroffenen Akteure also durchaus Effekte auf das Forschungsverhalten zu haben. Dabei zeigen sich sowohl intendierte (Wettbewerb, Leistungsmotivation, erhöhte Publikationstätigkeit, Transparenz) als auch nichtintendierte (Mainstreamforschung, Konflikt um Koautorenschaften, Vernachlässigung anderer Publikationsformate) Folgen. Anzumerken ist hierbei, dass

¹⁹ Für den Bereich Lehre wird mehrheitlich *keine* Effizienz- oder Qualitätssteigerung der Lehrleistung gesehen. Dasselbe gilt für die Schaffung von Transparenz bei der Lehrleistung.

die Ergebnisse aus dem Fach Medizin aufgrund der unterschiedlichen Forschungs- und Publikationskultur nicht ohne weiteres auf die in dieser Arbeit untersuchten sozialwissenschaftlichen Fächer übertragbar sind.

Ein bereits etwas älteres Forschungsprojekt von Minssen et al. (2003) nimmt ebenfalls die Folgen der indikatorenbasierten Mittelzuweisung in den Blick. In einer auf das Land Nordrhein-Westfalen begrenzten schriftlichen Vollerhebung wurden Rektorate, Dekanate und Professorinnen und Professoren aller Hochschulen befragt und im Nachgang qualitative Wirkungsanalysen in Form von zwölf Expertengesprächen durchgeführt. 49 % der befragten Professorinnen und Professoren beurteilten die indikatorenbasierten Mittelzuweisungsverfahren dabei positiv (ebd., S. 59). Mit Hilfe einer Faktorenanalyse auf Grundlage der Items zur Beurteilung des Verfahrens konnten Minssen et al. zwei Gruppen von Professorinnen und Professoren herausarbeiten, die sie entsprechend ihrer Einstellungen als „Ökonomen“ und „Traditionalisten“ bezeichnen:

Der erste Faktor lädt hoch auf Items, die Leistung in Forschung, Lehre und Selbstverwaltung sowie Transparenz als Folge der Indikatorisierung betonen. Hier werden also ökonomische Steuerungskriterien hervorgehoben, die nicht im Gegensatz zur Freiheit von Lehre und Forschung gesehen werden. Wir bezeichnen diese Gruppe deswegen als „Ökonomen“. Sie umfasst 33 % der befragten Professoren. Der zweite Faktor lädt hoch auf den Items, die Kontrolle der Professoren, Einschränkung der Freiheit von Forschung und Lehre sowie die materielle Steuerung des Verhaltens durch die parametergestützte Mittelvergabe in den Vordergrund stellen. In diesem Fall werden also die traditionellen Werte der Selbstregulation der Wissenschaft betont, die im Widerspruch zu einer Indikatorisierung der Mittelvergabe stehen; wir bezeichnen diese Gruppe deswegen als „Traditionalisten“. In dieser Gruppe versammeln sich 67 % der Professoren. (ebd., S. 60)

In der Stichprobe unterscheiden sich beide Gruppen hinsichtlich soziodemografischer Variablen oder der Fachzugehörigkeit nicht, so dass die Autoren über den wissenschaftlichen Sozialisationsprozess vermittelte generelle Einstellungsmuster als Erklärung für die Gruppenunterschiede heranziehen (vgl. ebd., S. 63). Wie Minssen et al. anmerken, ist die Differenzierung in diese beiden Gruppen bzw. die zugrundeliegenden Einstellungsmuster insbesondere im Hinblick auf motivationstheoretische Aspekte interessant. In Übereinstimmung mit der *Motivation Crowding Theory* wäre nämlich davon auszugehen, dass bei den Traditionalisten durch die indikatorenbasierte Mittelzuweisung negative Effekte auf die intrinsische Motivation zu befürchten wären, während bei den Ökonomen kein bzw. ein positiver Effekt zu erwarten ist (vgl. ebd., S. 61–62; Frey und Jegen 2001

und das Kapitel 4.1.2 in dieser Arbeit). In Bezug auf die konkrete Wirkung der indikatorenbasierten Mittelzuweisung gaben 36 % der Befragten – überwiegend Ökonomen – an, dass die eigene Aufgabenwahrnehmung durch das Verfahren beeinflusst wurde. Dies betreffe insbesondere eine verstärkte Bemühung um Drittmittel (Minssen et al. 2003, S. 65–67). Dennoch bewerten die Autoren der Studie die Auswirkungen des Verfahrenseinsatzes auf das tatsächliche Verhalten als eher überschaubar. Vermutet wird stattdessen, dass die Selbsteinschätzungen der Ökonomen zur veränderten Aufgabenwahrnehmung mehr „(...) ihre generelle Befürwortung einer Parameterorientierung zum Ausdruck bringen“ (ebd., S. 65), als dass es sich um eine tatsächliche Verhaltensänderung handelt.

Eine weitere Studie von Schröder (2004) beschäftigt sich allgemein mit der Ausgestaltung und Wirkungsweise indikatorenbasierter Steuerungsverfahren. Anhand einer Fallstudie von fünf Universitäten werden unter anderem die indikatorenbasierte Mittelzuweisung und Zielvereinbarungen analysiert. Die Ergebnisse von insgesamt 67 teilstandardisierten Interviews mit Hochschulleitungen, Verwaltungsmitarbeitern und wissenschaftlichem Personal auf Fakultätssebene zeigen dabei, dass indikatorenbasierte Steuerungsverfahren an den untersuchten Universitäten nicht primär zur Schaffung eines hochschulinternen Wettbewerbs bzw. zur Schaffung von Leistungsanreizen eingesetzt werden, sondern mehr der effizienten und transparenten Allokation der Mittel dienen sollen. Entsprechend positiv bewerten die befragten Wissenschaftlerinnen und Wissenschaftler die Verfahren im Hinblick auf eine effiziente Mittelverteilung, die Schaffung von Transparenz und auch im Hinblick auf Leistungsgerechtigkeit, während die Möglichkeiten der Verfahren zur Steigerung der Leistungsbereitschaft sowie deren Steuerungswirkung eher gering eingeschätzt werden. Im Gegensatz zur Studie von Minssen et al. (2003) sieht Schröder dabei disziplinspezifische Unterschiede hinsichtlich der Bewertungen der Verfahren. Zu beachten ist hier allerdings die sehr kleine Stichprobe von $n = 44$, so dass sich in der Gruppe der Geistes- und Sozialwissenschaften beispielsweise gerade einmal $n = 11$ Befragte befinden. Hinsichtlich der konkreten Wirkungen des Einsatzes indikatorenbasierter Steuerungsverfahren wird vor allem eine Erhöhung der Drittmittelaktivität erwartet. Negative Effekte im Sinne einer Abnahme von bestimmten Aktivitäten (z. B. verminderte Förderung von Promotionen und Habilitation) werden überwiegend nicht vermutet. Schröder schließt daraus, dass die Verfahren zu einer moderaten Leistungssteigerung führen, die Anreizwirkung aber noch erhöht werden könnte, wenn die notwendigen Rahmenbedingungen (z. B. Erhöhung der zugewiesenen Mittel) geschaffen werden würden (vgl. Schröder 2004, S. 47).

Sieweke et al. (2014) befragten in einer explorativen Studie Fakultätsleitungen von rechts- und wirtschaftswissenschaftlichen Fakultäten zum Einsatz, zur Ausgestaltung und zu den Auswirkungen von forschungsbezogenen Leistungsmessungen im Rahmen

von fakultätsinternen Steuerungsverfahren. Hierbei zeigte sich, dass die Bewertung der Forschungsleistung vor allem im Rahmen der indikatorenbasierten Mittelzuweisung erfolgt und hierfür am häufigsten die Indikatoren Drittmittel, (nach Zeitschriftenrankings gewichtete) Publikationen und Promotionen herangezogen werden. Die Auswirkungen nach Einführung der Verfahren sehen die befragten Fakultätsleitungen entsprechend der verwendeten Indikatoren in einer erhöhten Publikations- und Drittmittelorientierung sowie in einer generell erhöhten Forschungsorientierung.

Nichtintendierte Effekte der Leistungsmessung werden in einer qualitativen Studie von Frost und Brockmann (2014) in den Fokus gerückt. Anhand von 30 narrativen Interviews mit wirtschaftswissenschaftlichen Universitätsprofessorinnen und –professoren zeigen die Autorinnen, dass indikatorenbasierte Leistungsmessungen durchaus mit Produktivitätssteigerungen einhergehen können, dass aber eine starre Orientierung an den für die Leistungsmessung relevanten Kriterien ebenso zu „tunnel vision, measure fixation, sub-optimisation and pigeonholing“ (ebd. S. 41) mit negativen Auswirkungen auf die Innovativität der Forschungsarbeit führen kann.

Daneben finden sich weitere Studien, die andere Aspekte der indikatorenbasierten Steuerung wie z. B. die Auswirkungen auf den Bereich Lehre (Wilkesmann und Schmid 2012) oder die Wirkungen der indikatorenbasierten Steuerung auf staatlicher Ebene (Henke und Dohmen 2012) in den Blick nehmen. Diese stehen sicherlich in einem Zusammenhang mit der Thematik dieser Arbeit, sind aber für die Bearbeitung der konkreten Forschungsfrage von untergeordnetem Interesse. Auf eine detaillierte Darstellung dieser und ähnlicher Studien wird daher verzichtet.

Die bisherigen Arbeiten bieten bereits einen breiten Überblick, insbesondere zum Einsatz und zur Ausgestaltung hochschulinterner indikatorenbasierter Steuerungsverfahren. Hinsichtlich der Wirkung des Einsatzes von Leistungsbewertungs- bzw. Steuerungsverfahren ergeben sich in den Studien zwar Hinweise auf einen moderaten Steuerungseffekt bzw. Leistungssteigerungen (Bogumil et al. 2013; Bogumil et al. 2015; Krempkow et al. 2013; Schröder 2004; Sieweke et al. 2014), allerdings scheinen Funktionen wie die Schaffung von Transparenz und Leistungsgerechtigkeit bei der Anwendung der Verfahren eine bedeutendere Rolle zu spielen und im Hinblick auf ihre Wirkung deutlich positiver eingeschätzt zu werden als eine intendierte Produktivitätssteigerung (Krempkow et al. 2013; Schröder 2004). Zudem deutet die bisherige Forschung darauf hin, dass mit dem Einsatz der Verfahren nichtintendierte Effekte auf das Forschungsverhalten und die Motivation von Wissenschaftlerinnen und Wissenschaftlern verbunden sein können (Frost und Brockmann 2014; Krempkow et al. 2013; Minssen et al. 2003).

Die Aussagekraft der vorgestellten Studien ist allerdings mit einigen Einschränkungen verbunden. Zunächst ist zu berücksichtigen, dass der Einsatz und die Ausgestaltung von

hochschulinternen Steuerungs- und Anreizverfahren im Zeitverlauf einem stetigen Wandel unterworfen ist. Gerade nach der ersten Implementierungsphase der Verfahren ist damit zu rechnen, dass diese z. B. im Hinblick auf die verwendeten Indikatoren oder des zu verteilenden Mittelvolumens angepasst wurden. Aufgrund der teils länger zurückliegenden Erhebungszeitpunkte einiger Studien können die Ergebnisse damit nicht direkt auf die jetzige Situation in der Hochschullandschaft übertragen werden. Einschränkungen ergeben sich weiter daraus, dass fachspezifische (z. B. medizinische Fakultäten) und regionale Besonderheiten (z. B. nur Hochschulen in Nordrhein-Westfalen in der Stichprobe) berücksichtigt werden müssen. Zudem beruht die Analyse der Wirkungen überwiegend auf reinen Selbst- und Fremdeinschätzungen der befragten Personen (Hochschulleitungen, Dekaninnen und Dekane, Professorinnen und Professoren) zu den (erwarteten) Auswirkungen des Einsatzes der Steuerungsverfahren auf die (eigene) Forschungsarbeit. Das konkrete Forschungsverhalten kann dabei nicht erfasst werden.

Insgesamt haben die Studien überwiegend deskriptiven oder explorativen Charakter. Auf hypothesentestende Verfahren wird weitgehend verzichtet. Um diese Lücke zu schließen, wurde aufbauend auf den bisherigen Erkenntnissen und mit Hilfe eines professions- und motivationstheoretischen Interpretationsrahmens für diese Arbeit eine quantitativ-deduktive Vorgehensweise gewählt, in welcher mittels multivariater statistischer Analyseverfahren auch Drittvariableneffekte in den Blick genommen werden sollen. Die interessierende abhängige Variable stellt dabei das konkrete Forschungsverhalten dar, welches über eine Proxy-Variable erfasst werden soll (vgl. Kapitel 5.2.1). Auch bei der vorliegenden Studie ist die Generalisierbarkeit der Ergebnisse eingeschränkt und auf die Fächergruppe der Sozialwissenschaften begrenzt. Mit Ausnahme der Wirtschaftswissenschaften liegen zu den anderen hier betrachteten Fächern (Erziehungswissenschaften, Politikwissenschaften, Psychologie und Soziologie) nach Kenntnis des Autors aber noch keine umfangreichen Analysen zum Einsatz und zur Wirkung von hochschulinternen Steuerungs- und Anreizverfahren vor, so dass die explizite Betrachtung dieser Fächergruppe ein weiteres Forschungsdesiderat darstellt.

Im nächsten Kapitel werden nun die Besonderheiten des Untersuchungsgegenstandes der Organisation Universität und der akademischen Profession herausgearbeitet. Diese sollen dann als Kontrastfolie für die darauffolgende Erarbeitung des motivationstheoretisch begründeten Modells und der aus diesem abgeleiteten Forschungshypothesen dienen.

3. Die Organisation Universität, die akademische Profession und das Problem der Steuerung

3.1 Die Universität als besondere Form der Organisation

Wie sich bei der Diskussion um die Governance von Hochschulen bereits angedeutet hat, erfolgt die Koordinierung der Tätigkeiten der verschiedenen Akteure innerhalb von Universitäten nicht einfach in Form eines hierarchisch gegliederten Steuerungsmechanismus im Sinne einer Top-Down-Steuerung. Dies ist der Tatsache geschuldet, dass es sich bei Universitäten um einen speziellen Typus von Organisationen handelt, der im Gegensatz zu beispielsweise privatwirtschaftlichen Organisationen durch eine Reihe von besonderen Merkmalen gekennzeichnet ist. Diese speziellen Merkmale sollen nun im Folgenden herausgearbeitet werden.

Aus der Vielfalt der theoretischen Zugänge zum Begriff *Organisation* (vgl. Kieser und Kubicek 1978, S. 11–12) wird hier für eine erste allgemeine Definition die systemtheoretische Perspektive von Luhmann herausgegriffen. Nach Luhmann (1975a) grenzen sich Organisationen von anderen Sozialsystemen wie Interaktionssystemen durch das Merkmal der an formalen Bedingungen geknüpften *Mitgliedschaft* ab, wobei die individuellen Motivlagen der Organisationsmitglieder und die Ziele des Organisationssystems nicht übereinstimmen müssen:

Als organisiert können wir Sozialsysteme bezeichnen, die die Mitgliedschaft an bestimmte Bedingungen knüpfen, also Eintritt und Austritt von Bedingungen abhängig machen. Man geht davon aus, daß die Verhaltensanforderungen des Systems und die Verhaltensmotive der Mitglieder unabhängig voneinander variieren können, sich aber unter Umständen zu relativ dauerhaften Konstellationen verknüpfen lassen. Mit Hilfe solcher Mitgliedschaftsregeln – etwa Autoritätsunterwerfungen gegen Gehalt – wird es möglich, trotz frei gewählter variabler Mitgliedschaft hochgradig künstliche Verhaltensweisen relativ dauerhaft zu reproduzieren. (ebd., S. 12)

Die Mitgliedschaftsregeln als *formalisierte Verhaltenserwartungen* (vgl. Luhmann 1964) erlauben also eine Steuerung der Handlungen der Organisationsmitglieder im Sinne des Organisationssystems. Sie tragen so zur Reduktion von Komplexität im Luhmannschen Sinne und der damit verbundenen gleichzeitigen Steigerung des Komplexitätsverarbeitungspotentials des Organisationssystems insgesamt bei (vgl. Luhmann 1975b).

Nach obiger Definition lassen sich Universitäten zweifelsohne als organisierte Sozialsysteme verstehen. Professorinnen und Professoren, wissenschaftliches Personal und Verwaltungspersonal oder die Studierenden sind alle durch mehr oder weniger formalisierte Erwartungen an die Organisation gebunden und richten ihr Handeln entsprechend

am Systemziel aus: Die Professorinnen und Professoren lehren und forschen, die Studierenden lernen und schreiben Prüfungen, das Verwaltungspersonal regelt die bürokratischen Angelegenheiten etc. Aufgrund der Spezifika des Wissenschaftssystems kann der Formalisierungsgrad der Verhaltenserwartungen bei Universitäten im Vergleich zu anderen Organisationssystemen allerdings als eher gering angesehen werden – man denke beispielsweise an die Wahlfreiheit hinsichtlich der Forschungsthemen oder der inhaltlichen Ausgestaltung von Seminaren. Vor dem Hintergrund dieser und anderer Besonderheiten der Organisation Universität werden daher im Folgenden drei enger gefasste Theorieansätze dargestellt, die eine für die Zwecke dieser Arbeit nähere Charakterisierung der Organisation Universität erlauben.

Universitäten als lose gekoppelte Systeme

Der erste Ansatz geht auf Weick (1976) zurück und beschreibt Universitäten als *loosely coupled systems*. Organisationen können sich zunächst aus Elementen (z. B. verschiedene Abteilungen oder Personen) zusammensetzen, die z. B. im Hinblick auf ähnliche Aufgabengebiete (*technical coupling*) oder aufgrund von Abhängigkeiten innerhalb des Positionsgefüges (*authority of office*) relativ eng aneinander gebunden sind:

At the outset the two most commonly discussed coupling mechanisms are the technical core of the organization and the authority of office. The relevance of those two mechanisms for the issue of identifying elements is that in the case of technical couplings, each element is some kind of technology, task, subtask, role, territory and person, and the couplings are task-induced. In the case of authority as the coupling mechanism, the elements include positions, offices, responsibilities, opportunities, rewards, and sanctions and it is the couplings among these elements that presumably hold the organization together. (ebd., S. 4)

Wie Weick anmerkt, sind aber weder enge Kopplungen in Form des *technical coupling* noch des *authority of office* die dominierenden strukturerhaltenden Mechanismen in (amerikanischen) Bildungsorganisationen (ebd.). Stattdessen seien diese Organisationen durch überwiegend lose Kopplungen zwischen den Elementen gekennzeichnet. Das Prinzip der losen Kopplung lässt sich mit Weick wie folgt definieren:

By loose coupling, the author intends to convey the image that coupled events are responsive, *but* [Hervorhebung im Original] that each event also preserves its own identity and some evidence of its physical or logical separateness. (ebd., S. 3)

Diese Beschreibung verweist bereits auf den Aspekt der Autonomie einzelner Organisationselemente als ein Spezifikum der Organisation Universität (vgl. Kapitel 3.2). Das *technical coupling* ist aufgrund der unterschiedlichen Untersuchungsgegenstände der einzelnen Fachgebiete und auch innerfachlicher Diversität, insbesondere in der Forschung aber auch in der Lehre, nur schwach ausgeprägt. Zwischen einzelnen Fächern oder Lehrstühlen kann und wird auch oft kooperiert. Diese Kooperationen sind aber funktional meist nicht zwingend erforderlich (vgl. Meier und Schimank 2010, S. 110). Der Mechanismus *authority of office* ist aufgrund der vergleichsweise schwachen Position von Leitungsorganen ebenfalls von eher untergeordneter Bedeutung (vgl. ebd., S. 110–111). Meier und Schimank merken dabei an, dass eben diese lose Kopplung die notwendigen Freiräume für innovative Forschung garantiert (ebd., S. 111).

Universitäten als Profibükratien

Auch der Ansatz von Mintzberg (1992), der Universitäten als *Profibükratien* bzw. Expertenorganisationen charakterisiert, betont den Autonomieaspekt der akademischen Profession. Nach Mintzberg (1980) lassen sich Organisationen anhand der fünf Organisationsbausteine *strategic apex* (strategische Spitze), *operative core* (operativer Kern), *middle line* (Mittellinie), *technostructure* (Technostruktur) und *support staff* (Hilfsstab) sowie dem jeweils dominanten Koordinierungsmechanismus (*direct supervision*, *standardization of work process*, *standardization of skills*, *standardization of outputs*, *mutual adjustment*) und Dezentralisierungstyp (*vertical*, *horizontal*, *selective*) typisieren (vgl. ebd.).

Der Typus der Profibükratie (*professional bureaucracy*) ist nun geprägt durch einen großen operativen Kern, der sich aus hochqualifizierten *professionals* zusammensetzt (ebd., S. 333). Dieser besteht in Universitäten aus den Professorinnen und Professoren sowie dem wissenschaftlichen Personal, „(...) who themselves produce the basic products and services of the organization“ (ebd., S. 323). Der zentrale Koordinierungsmechanismus besteht hierbei in der *standardization of skills*, indem Personal in der Regel aufgrund der (hohen) fachlichen Qualifikationen rekrutiert wird und diesem im Anschluss ein hoher Grad an Autonomie hinsichtlich der Aufgabenerfüllung gewährt wird (ebd., S. 333–334). Die Dezentralisierung, definiert als „the extent to which power over decision making in the organization is dispersed among its members“ (ebd., S. 326), erfolgt sowohl auf horizontaler (z. B. Entscheidungsbefugnisse einzelner Lehrstühle) als auch vertikaler (Hochschulleitung, Dekane, Lehrstuhl) Ebene. Der operative Kern von Profibükratien hat dabei einen wesentlichen Teil der formellen und informellen Macht inne (ebd., S. 334) und ist eng mit der organisationalen Mittellinie verknüpft:

Not only the professionals control their own work, but they also tend to maintain collective control of the administrative apparatus of the organization. Managers of the middle line, in order to have power in the Professional Bureaucracy, must be professionals themselves, and must maintain the support of the professional operators. Moreover, they typically share the administrative tasks with the operating level, tasks require a good deal of mutual adjustment, achieved in large part through standing committees, task forces, and other liaison devices. (ebd.)

Die Organisationsbausteine Mittellinie sowie die strategische Spitze sind in Profibürokratien eher von geringer Größe (vgl. ebd.). Sie bestehen in Universitäten aus den Dekanaten und Selbstverwaltungsgremien bzw. aus der Hochschulleitung (Würmseer 2010, S. 81). Die Technostruktur, bestehend aus der Hochschulleitung untergeordneten Stabspositionen für Planung, Forschung und Studienangelegenheiten, ist ebenfalls nur von geringem Umfang, während die Hilfsstäbe in Form der (Zentral-)Verwaltung, Bibliotheken oder dem Studentenwerk stark ausgeprägt sind (ebd.; Mintzberg 1980, S. 334).

Die Beschreibung des strukturellen Aufbaus der Universität als Profibürokratie macht die tragende Rolle des operativen Kerns – also der Professorinnen und Professoren bzw. des wissenschaftlichen Personals – im Organisationsgefüge der Universität deutlich. Dieser Aspekt wird in Kapitel 3.3 nochmals aufgegriffen und die sich hieraus ergebenden Besonderheiten in Bezug auf die Frage der organisationsinternen Steuerung diskutiert.

Universitäten als organisierte Anarchie

Eine dritte Charakterisierung von Universitäten als *organisierte Anarchien* nimmt vor allem die organisationsinternen Entscheidungsprozesse (*garbage can model*) in den Blick. Nach Cohen et al. (1972) sind organisierte Anarchien dadurch gekennzeichnet, dass sie erstens keine kohärenten Präferenzen bzw. Zielsetzungen besitzen (*problematic preferences*), sondern auf der Basis einer Reihe von unscharfen und inkonsistenten Zielen operieren (ebd., S. 1). Zweitens sind die organisationsinternen Prozesse zur Zielerreichung nicht eindeutig geregelt bzw. für die Organisationsmitglieder unklar (*unclear technology*). Operiert wird daher „on the basis of simple trial-and-error procedures, the residue of learning from the accidents of past experience, and pragmatic inventions of necessity“ (ebd., S. 1–2). Das dritte Merkmal organisierter Anarchien bezieht sich auf die fluide Teilnehmerschaft (*fluid participants*) in verschiedenen organisationsinternen Gremien und Entscheidungsinstanzen: „Participants vary in the amount of time and effort they devote to different domains; involvement varies from one time to another“ (ebd., S. 2).

Allen drei Merkmalen begegnet man typischerweise in Universitäten: Mögliche Zielsetzungen von Universitäten können vielfältiger und nicht immer eindeutiger Natur sein

(Grundlagenforschung, Transfer von Forschungsergebnissen in die Wirtschaft, Reputationsaufbau, Lehre, Wirtschaftlichkeit der Organisation etc.) und dabei auch miteinander in Konflikt geraten (z. B. hohes Anforderungsniveau in der Lehre zur Sicherstellung qualitativer Standards vs. Erzielung eines hohen Outputs an erfolgreichen Absolventen). Unklare Technologien ergeben sich vor allem hinsichtlich der Aufgabenbereiche Forschung und Lehre, für die es keine klaren Handlungsanweisungen gibt, die zu einem garantierten Erfolg, z. B. im Sinne eines Erkenntnisgewinns oder der ‚Produktion‘ gut qualifizierter Studierender, führen (vgl. Meier und Schimank 2010, S. 109). Der Aspekt der fluiden Teilnehmerschaft findet sich in allen möglichen Entscheidungsgremien an Universitäten wieder (z. B. Promotionsausschuss, Berufungskommission, Fakultätsrat etc.). Zum einen sind die Mitglieder der Universität, allen voran die Professorinnen und Professoren, durch die Pluralität der Aufgaben (Forschung, Lehre, Selbstverwaltung) und der Mitwirkung in verschiedenen Gremien meist nicht in der Lage, in allen Entscheidungsgremien mit derselben zeitlichen Intensität und demselben Engagement mitzuwirken. Zum anderen ist die wechselnde Mitgliedschaft aufgrund des Wahlprinzips auf Zeit, welches in den meisten Fällen die personelle Zusammensetzung der Gremien bestimmt, begründet.

Entscheidungen in solchen organisierten Anarchien laufen nun in vielen Fällen nach einem mehr oder weniger zufälligen Muster ab, indem Probleme, mögliche Lösungen und am Entscheidungsprozess beteiligte Personen bei einer Entscheidungsgelegenheit (z. B. Berufungsverhandlung) aufeinandertreffen (vgl. Cohen et al. 1972, S. 16). Welche Lösung in einer solchen Entscheidungssituation dann welchem Problem zugeordnet wird, hängt von der Art und Anzahl der vorhandenen Probleme, der verfügbaren Lösungen sowie den Entscheidungsträgern ab (vgl. ebd., S. 16–17). Dabei sind diese Bedingungen für Entscheidungsfindungen relativ unabhängig voneinander und auch nicht zeitlich geordnet. Beispielsweise kann die Lösung für ein Problem schon vor Auftreten des Problems existieren (vgl. Kieser und Kubicek 1978, S. 74–75). Ein kurzes Beispiel soll einen solchen möglichen Entscheidungsmoment verdeutlichen: So kann es vorkommen, dass ein Fach bereits institutsinterne Lehrevaluationen durchführt (Lösung), bevor seitens der Fakultät im Rahmen einer Fachbereichssitzung (Entscheidungsgelegenheit) der Versuch unternommen wird, regelmäßige Evaluationen verpflichtend vorzuschreiben (Problem). Für dieses Fach (Teilnehmer 1) stellt sich die Frage nach der Einführung eines Evaluationssystems also gar nicht, während ein anderes Fach (Teilnehmer 2) angesichts des befürchteten Mehraufwands vehement die Zustimmung verweigert und ein dritter Fachvertreter (Teilnehmer 3) aufgrund anderer Verpflichtungen der Fachbereichssitzung gar nicht erst beiwohnt. Die Entscheidung wird daraufhin möglicherweise vertagt, bis die vier Komponenten der Entscheidungsfindung in anderer Konstellation erneut aufeinandertreffen.

Ohne die Entscheidungsprozesse nach dem *garbage can model* noch weiter in der Tiefe zu beleuchten, macht der Ansatz der organisierten Anarchie doch deutlich, dass sich Entscheidungen in Universitäten nicht (immer) nach einer rein rationalen Logik vollziehen, sondern oft von Unsicherheit und Variabilität geprägt sind. Meier und Schimank (2010) merken allerdings an, dass Entscheidungsprozesse in Form des Mülleimermodells dennoch nicht als „Defekt“ (ebd., S. 110) angesehen werden sollten, sondern eher als sich bietende Möglichkeit „unter den gegebenen Bedingungen (...), überhaupt zu Entscheidungen zu kommen“ (ebd.).

Vor dem Hintergrund der in Kapitel 2 dargestellten Governancereformen des Hochschulsystems, die ja auch intraorganisationale Aspekte der Universität betreffen (Stärkung der hierarchischen Selbststeuerung und Schwächung der akademischen Selbstorganisation), stellt sich allerdings die Frage, ob die Charakterisierungen von Universitäten als organisierte Anarchie, Profibürokratie oder lose gekoppeltes System noch zutreffend oder bereits ins Wanken geraten sind (vgl. Kloke und Krücken 2012). Krücken und Meier (2006) weisen in diesem Zusammenhang auf einen Wandel der Organisation Universität zu einem *organisationalen Akteur* hin. Dieser organisationale Akteur ist gekennzeichnet durch eine zunehmende Rechenschaftspflicht (Evaluationen, Akkreditierungen), der Definition von (ganzheitlichen Organisations-)Zielen (Leitbilder), dem Ausbau formaler Strukturen (Abteilungen für Controlling, Beratungsstellen, Technologietransfer etc.) sowie zunehmenden Managementaufgaben, die entweder vom wissenschaftlichen Personal oder einem professionellen Management übernommen werden (vgl. Krücken und Meier 2006, S. 247–256; Maasen und Weingart 2006, S. 22–23).

Die beschriebenen Wandlungsprozesse sind sicher nicht von der Hand zu weisen und bleiben auch nicht ohne Einfluss auf universitäre Entscheidungs- und Steuerungsprozesse. Allerdings kann angenommen werden, dass den ursprünglichen Organisationsstrukturen und –logiken nach wie vor ein sehr hoher Stellenwert zukommt. Schon in der Diskussion der Governancereformen in Kapitel 2 zeigte sich, dass sich zwar ein Bedeutungsgewinn bei den Governance-Mechanismen der hierarchischen Selbststeuerung und des Konkurrenzdrucks abzeichnet, die akademische Selbstorganisation aber noch relativ unberührt von den Reformbemühungen bleibt. Kloke und Krücken (2012) kommen in ihrer Analyse von universitären Zielkonflikten zu dem Schluss, „(...) dass die im hochschulpolitischen Diskurs vielfach kritisierte interne lose Kopplung der Organisation ‚Universität‘ (...) nach wie vor von hoher Bedeutung ist“ (ebd., S. 27). Krücken (2006) merkt zudem an, dass sich der Wandel individueller und organisationaler Praktiken langsamer vollzieht, „(...) als es der gegenwärtige Wettbewerbs- und Organisationsdiskurs vermuten lässt“ (ebd., S. 13). Die hier beschriebenen Theorieansätze zur Charakterisie-

Die Universität als besondere Form der Organisation haben daher nichtsdestotrotz nach wie vor analytische Relevanz und werden in den folgenden Argumentationslinien mitberücksichtigt.

In diesem Teilkapitel war die Universität als Organisation Gegenstand der Analysen, wobei vor allem Aspekte der Organisationsstruktur in den Blick genommen wurden. Dabei deutete sich bereits an, dass die besondere Stellung der Universität als Organisation vor allem in den Spezifika der akademischen Profession begründet ist. Was unter der akademischen Profession generell zu verstehen ist und wie sich diese Spezifika im Detail darstellen, wird nun im nächsten Kapitel näher beleuchtet.

3.2 Die akademische Profession

Wie auch beim Begriff der Organisation treffen im wissenschaftlichen Diskurs über den Begriff *Profession* bzw. *Professionalisierung* eine Vielzahl theoretischer Perspektiven aufeinander. Lediglich auf die Beschreibung der Merkmale einer Profession in Abgrenzung zu nicht-professionellen Berufen beschränken sich die sogenannten Professionskriterienansätze oder Attribute-Modelle (vgl. Pfadenhauer 2003, S. 32). Eine metaanalytische Betrachtung bestehender Professionskriterienansätze erlaubt die Herausarbeitung der zentralen Definitionskriterien. Kurtz (2002) benennt mit Rückgriff auf Cogan (1953) dabei die folgenden Merkmale einer Profession:

- Organisation der Berufsangehörigen in einem selbstverwalteten Berufsverband, welcher spezifische Verhaltensregeln in Form einer Berufsethik aufstellt (*code of ethics*)
- Vorhandensein einer besonderen Wissensbasis (fundiertes Fach-/Expertenwissen)
- professionelle Arbeit als Dienst an der Allgemeinheit und mit Bezug auf zentrale gesellschaftliche Werte, verbunden mit einer mehr altruistischen als egoistischen Dienstmotivation
- Handlungs- und Entscheidungsautonomie aufgrund des Expertenstatus
- ausgeprägtes Verantwortungsbewusstsein gegenüber den Klienten (z. B. Patienten, Studierende) gepaart mit einem Vertrauen der Klienten in die fachliche Kompetenz und moralische Integrität des Professionellen
- exklusives Handlungskompetenzmonopol (z. B. *nur* Ärzte behandeln Patienten, *nur* Professoren betreuen Doktorarbeiten) mit einem hohen Maß an gesellschaftlicher Wertschätzung
- in der Regel Verbot öffentlicher Werbung für professionelle Berufe (ebd., S. 49)

Legt man diese Auflistung als Maßstab zur Definition einer Profession an, erscheint es gerechtfertigt, Berufsgruppen wie Ärzte, Juristen oder eben Wissenschaftler als Profession zu begreifen, auch wenn einzelne Merkmale nicht (mehr) immer gleichermaßen zutreffend erscheinen.²⁰

Neben diesen rein deskriptiven Professionskriterienansätzen, die für sich allein noch keine Professionstheorie darstellen (vgl. ebd.), lassen sich strukturfunktionalistische (Parsons, Goode), interaktionistische (Hughes, Abbott), machttheoretische (Larson), struktur- (Oevermann) sowie systemtheoretische (Luhmann, Stichweh, Kurtz) Professionskonzepte unterscheiden (vgl. ebd., S. 50–57). Daneben finden sich auch eine Reihe neuerer professionstheoretischer Ansätze, „die eine Verlagerung von Konzepten der Profession und der Professionalisierung hin zu einer Analyse des Konzeptes von Professionalität und professionellen Handelns vornehmen“ (Kloke 2014, S. 139). In diesen neueren Ansätzen werden auch Aspekte einer zunehmenden Ökonomisierung (vgl. ebd., S. 136) sowie das sich wandelnde Verhältnis zwischen Organisation und Profession behandelt. Dieses wird als weniger konfliktär angesehen und stattdessen die Rolle von Organisationen bei Prozessen der Professionalisierung in den Blick genommen (ebd., S. 158).

Im Folgenden wird hier vor allem auf den Ansatz von Parsons (Parsons 1969; Parsons und Platt 1990) Bezug genommen, da dieser unter anderem explizit die akademische Profession betrachtet und dabei auch deren Rolle in den Governance-Strukturen von Universitäten bzw. des Universitätssystems beleuchtet (vgl. Schimank 2005, S. 143). Für die Thematik der vorliegenden Arbeit erscheint der Ansatz von Parsons somit am geeignetsten. Zudem werden vor allem im darauffolgenden Kapitel einzelne Aspekte neuerer professionstheoretischer Ansätze herausgegriffen.

Viele der zu Beginn dieses Kapitels dargestellten Merkmale einer Profession finden sich auch bei Parsons wieder. So geht dieser von einer Kollektiv- bzw. Dienstleistungsorientierung von Professionen aus und spricht diesen ein Expertenwissen zu (vgl. Wenzel 2005, S. 50–51). Diesem Expertenwissen steht das Nicht-Wissen der Klienten (z. B. Patienten, Studierende) gegenüber. Das durch diese Wissensasymmetrie geprägte Verhältnis zwischen Experte und Klient wird nun durch Vertrauen stabilisiert. Der Klient vertraut darauf, dass der Professionelle nicht nach egoistischen Interessen, sondern stets kompetent und dienstleistungsorientiert handelt. Dieses Vertrauen gewährt dem Experten die für die Ausübung seiner Tätigkeit notwendige Autonomie – er ist in seinem Handeln nur seinem Wissen und seiner fachlichen Kompetenz verpflichtet. Der Professionelle wiederum muss auf die Kooperationsbereitschaft des Klienten vertrauen können

²⁰ Beispielsweise wurde in den letzten Jahren das Werbeverbot für Ärzte gelockert (vgl. hierzu im Detail Kock 2013). Werbemaßnahmen sind demnach teilweise zulässig, allerdings nicht in dem Maße wie in anderen privatwirtschaftlichen Branchen üblich.

(der Patient muss seine Medikamente einnehmen, der Studierende die Literatur für das nächste Seminar durcharbeiten etc.). Das Vertrauensverhältnis hat daher einen wechselseitigen Charakter (ebd., S. 51). Ein weiteres Element von Professionen sieht Parsons in dem Prozess der Wissenserzeugung, das heißt „Professionen wenden Wissen nicht einfach nur an, sie bleiben in den Prozess der Wissenserzeugung eingebunden. Vermittlung, Anwendung, Erweiterung und Verbesserung von Wissen bilden einen Nexus, den Parsons als ‚professionellen Komplex‘ bezeichnet“ (ebd., S. 52). Professionen tragen also selbst dafür Sorge, dass das notwendige Fachwissen zur Lösung von professionsspezifischen Problemen weiterentwickelt und an (zukünftige) Angehörige der Profession weitergegeben wird. Dieses Wissen und dessen Anwendung tragen dabei stets den Charakter kognitiver Rationalität. Parsons sieht die Universität aufgrund ihrer wissenserzeugenden Funktion nun als Zentrum des professionellen Komplexes (ebd.).

Zusammenfassend lässt sich festhalten, dass sich die Profession bei Parsons durch eine Kollektivorientierung an einem gesellschaftlichen Zentralwert wie z. B. Gesundheit, Erziehung oder Wahrheit (vgl. Kurtz 2002, S. 50) auszeichnet, für welchen die Professionsangehörigen als Experten fungieren und Problemlösungen für Klienten anbieten. Die Professionellen arbeiten dabei mit rationalem professionsspezifischem Fachwissen, woraus sich eine Wissensasymmetrie zu den nicht-wissenden Klienten ergibt. Diese begegnen dieser Asymmetrie ihrerseits mit Vertrauen in die fachliche Kompetenz und das rationale, auf den gesellschaftlichen Zentralwert ausgerichtete Handeln der Professionellen. Dieses Vertrauen im Zusammenspiel mit dem Expertenstatus der Professionsangehörigen bildet schließlich die Basis für die gewährte Autonomie der Professionen.

Durch welche Spezifika zeichnet sich nun die akademische Profession konkret aus? Zunächst einmal gilt auch für diese, dass die Autonomie der Berufsausübung ein wesentliches Element darstellt und in Deutschland sogar im Rahmen der Wissenschaftsfreiheit grundgesetzlich geschützt ist. Die methodische und inhaltliche Ausgestaltung von Forschung und Lehre soll damit allein den Professorinnen und Professoren bzw. dem wissenschaftlichen Personal obliegen und nicht von außen gesteuert werden. Die akademische Profession ist zudem dadurch gekennzeichnet, dass die jeweilige fachspezifische Scientific Community und nicht die Organisation Universität den zentralen Bezugspunkt für die Professionsangehörigen darstellt (vgl. Enders 1998, S. 58; Jacob und Teichler 2011, S. 82; Minssen et al. 2003, S. 36). Die Scientific Communities bilden in diesem Sinne die jeweils fachspezifischen kollektiven Orientierungspunkte und bestimmen den – um im Vokabular der Professionsforschung zu bleiben – *code of ethics*. Diese Community-Orientierung (vgl. zur Community-Orientierung von Professionen auch Stichweh 2005, S. 38–39) kann sich schließlich auch in der formalen Mitgliedschaft in institutionalisierten Fachverbänden (z. B. Deutsche Gesellschaft für Erziehungswissenschaft) und Unterverbänden (z. B. Sektion Erwachsenenbildung) niederschlagen.

Den Scientific Communities kommt dabei noch eine weitere bedeutende Rolle als 'Produktionsstätte' wissenschaftlicher (Forschungs-)Reputation zu (vgl. Minssen et al. 2003, S. 36). Durch Publikationen, (eingeworbene) Forschungsprojekte, Vorträge auf Tagungen, Zitationen und Zitationsindizes, Gremienarbeit etc. werden Wissenschaftlerinnen und Wissenschaftler in der Scientific Community sichtbar und können sich Reputation aufbauen. Diese Auflistung macht bereits deutlich, dass insbesondere der Forschungsleistung eine tragende Rolle beim Reputationsaufbau zukommt, da die Ergebnisse der Forschungsarbeit leichter als die Leistungen in der akademischen Lehre nach außen sichtbar gemacht werden können. Zwar kann auch die Lehrleistung einzelner Wissenschaftlerinnen und Wissenschaftler auch außerhalb der eigenen Hochschule sichtbar gemacht werden, indem etwa öffentlichkeitswirksam Lehrpreise für herausragende Leistungen vergeben werden. Dies geschieht allerdings noch in relativ geringem Ausmaß, auch wenn in den letzten Jahren Bemühungen zur Aufwertung der Lehre zu erkennen sind (vgl. Wissenschaftsrat 2008a). Studentische Lehrevaluationen stellen eine weitere Möglichkeit dar, die Lehrleistung nach außen sichtbar zu machen. Außer im Rahmen von Berufungsverfahren oder zum hochschulinternen Benchmarking werden personenbezogene Lehrevaluationen aber meist nicht offengelegt.²¹ Die Aussagekraft von studentischen Lehrevaluationen als Indikator für die Lehrqualität ist zudem nicht unumstritten und die Ergebnisse können teils durch äußere Einflussfaktoren verzerrt sein (vgl. Schmidt et al. 2005; Wolbring 2013).

Im Reputationsmanagement der Professorinnen und Professoren können neben der Scientific Community noch andere gesellschaftliche Stakeholdergruppen von Bedeutung sein (Medien, privatwirtschaftliche Kooperationspartner, politische Entscheidungsträger). Honecker (2011) weist allerdings darauf hin, dass für Wissenschaftlerinnen und Wissenschaftler die Reputation innerhalb des eigenen akademischen Feldes einen deutlich höheren Stellenwert besitzt als die Reputation im außerwissenschaftlichen Bereich (ebd., S. 21).²² Erbrachte wissenschaftliche Leistungen werden also von den *peers* mit Anerkennung belohnt und tragen so – mehr oder weniger langfristig – zu einem Aufbau von Reputation bei. Eine hohe Reputation verbessert wiederum die Stellung einer Wissenschaftlerin oder eines Wissenschaftlers in der Scientific Community und eröffnet (neue) Karrieremöglichkeiten. Reputation wird in diesem Sinne oft auch als *Währung* der

²¹ In Hochschulrankings, wie z. B. dem CHE-Ranking, werden Studierendenbewertungen allerdings durchaus nach außen sichtbar gemacht. Diese Bewertungen beziehen sich aber nicht auf einzelne Lehrpersonen, sondern auf verschiedene Aspekte (z. B. Studierbarkeit insgesamt, Betreuung durch Lehrende, etc.) des gesamten Lehrangebotes des betroffenen Faches eines Standortes.

²² Vermutlich gibt es aber hier durchaus Unterschiede hinsichtlich der Fachzugehörigkeit. Kühl (2015b) merkt beispielsweise an, dass in Fächern wie der Philosophie, Germanistik oder Soziologie medial stark präsente Kolleginnen und Kollegen eher kritisch beäugt werden, während in Teilen der Politikwissenschaft oder der Betriebswirtschaftslehre mediale Präsenz zusätzliche Anerkennung verspricht (ebd., S. 806).

Wissenschaft bezeichnet. Der Besitz eines Währungsmittels ist mit der Möglichkeit der Einflussnahme assoziiert und entsprechend kann wissenschaftliche Reputation mit Parsons auch als eine Form des Kommunikationsmediums *Einfluss* verstanden werden (vgl. Schluchter 1974, S. 378; Schneider 2009, S. 211; Wenzel 2005, S. 59). Auch Luhmann betont die hohe Bedeutung der Reputation für die Karrierechancen in der Wissenschaft (Luhmann 1992, S. 680). Das daraus resultierende Reputationsstreben von Wissenschaftlerinnen und Wissenschaftlern sollte dabei allerdings nicht als reine persönliche Befriedigung narzisstischer Bedürfnisse missverstanden werden. Vielmehr kommt der Reputation im Wissenschaftssystem eine wichtige Funktion zu, indem diese die Aufmerksamkeit steuert (Kühl 2015b, S. 806). Gerade im Hinblick auf eine zunehmend undurchschaubare Flut von wissenschaftlichen Veröffentlichungen (vgl. hierzu die Studie von Parolo et al. 2015) kann die Reputation einer Autorin oder eines Autors als erster Anhaltspunkt dienen, dass die Publikation gewissen qualitativen Standards genügt – eine hundertprozentige Garantie für einen hohen wissenschaftlichen Gehalt einer Publikation kann die Reputation der Autorin oder des Autors natürlich dennoch nicht sein (vgl. Kühl 2015b, S. 806).

Die vorangegangenen Anmerkungen zur Bedeutung der Reputation und dem Reputationsstreben in der Wissenschaft machen bereits deutlich, dass man sich von dem Bild einer rein altruistisch, allein im Sinne der Wahrheitsfindung handelnden akademischen Profession verabschieden muss. Diesem, auch teils im Rahmen der anfänglich diskutierten Professionskriterienansätze weit verbreiteten Bestimmungsmerkmal von Professionen erteilte bereits Parsons eine Absage. Hierzu Kurtz (2002):

Die Unterscheidung zwischen der Kollektivitätsorientierung der professionellen Berufe und der Selbstorientierung der Wirtschaft ist für Parsons keine Differenz auf der Ebene der Motivation der beruflich Handelnden, sondern eine auf der institutionellen Ebene. Weder ist die ›typische Motivierung‹ [Hervorhebung im Original] in den Professionen ausschließlich altruistisch noch die der Geschäftsleute ausschließlich egoistisch. (ebd., S. 50)

Die gewinnorientierte Ausrichtung von Unternehmen und die Verpflichtung der Wissenschaft auf die mehr altruistischen Ziele des Erkenntnisgewinns und der Wahrheitsfindung dürfen also nicht generell mit den individuellen Motivationslagen der einzelnen Akteure verwechselt werden (vgl. auch Schmeiser 2006, S. 300).²³

²³ Nichtsdestotrotz gibt es Hinweise darauf, dass in bestimmten Berufsgruppen bestimmte Motivationslagen häufiger vorkommen als in anderen - hierauf wird in dieser Arbeit im Kapitel 4.1.3 zur *Public Service Motivation* noch einmal zurückgekommen.

Dies führt zu einem weiteren vieldiskutierten Merkmal der akademischen Profession, nämlich der den Wissenschaftlerinnen und Wissenschaftlern zugesprochenen intrinsischen Motivation. Diese ist nun nicht gleichzusetzen mit einem altruistisch motivierten Verhalten, also einem „(...) Verhalten [, das] dadurch definiert [ist], dass der Helfende weder einen materiellen noch einen psychologischen Nutzen aus seiner Hilfeleistung gegenüber dem Hilfeempfänger erfährt“ (Fetchenhauer und Bierhoff 2004, S. 132).²⁴ Nach Deci (1972) lässt sich intrinsische Motivation vielmehr wie folgt definieren: „We say a person is intrinsically motivated to perform an activity if there is no apparent reward except the activity itself or the feelings which result from the activity“ (ebd., S. 217). Im Gegensatz zum rein altruistischen Verhalten wird man bei intrinsisch motivierten Handlungen also durch die Freude an der Ausübung der Tätigkeit an sich belohnt.²⁵ Nun kann argumentiert werden, dass gerade die im vorangegangenen Abschnitt besprochene Reputation die Belohnung bzw. den Anreiz für die wissenschaftliche Arbeit darstellt und das Verhalten damit mehr extrinsisch als intrinsisch motiviert wäre. Das ist sicherlich teilweise zutreffend, allerdings ist die ‚Währung‘ Reputation äußerst knapp und viele Forschungsarbeiten führen nicht zum Erfolg oder bleiben unbeachtet (vgl. Gustin 1973, S. 1121–1122; Kieser 2010, S. 349). Allein die unsichere Hoffnung auf eine zukünftige Anerkennung der eigenen Arbeit in der Scientific Community erscheint als alternative Erklärung, warum Menschen – zudem oft in Form mehr oder weniger prekärer Beschäftigungsverhältnisse – Wissenschaft betreiben, nicht ausreichend. Die Annahme, dass der Grund hierfür (auch) in der intrinsischen Motivation, also in der Freude mit der wissenschaftlichen Tätigkeit an sich, zu finden ist, liegt daher nahe (vgl. Minssen et al. 2003, S. 37). Dies gilt insbesondere auch für den Bereich Lehre, dem nur wenig Anerkennung durch die Scientific Community zuteilwird (vgl. zur intrinsischen Lehrmotivation auch die Studie von Wilkesmann 2012).²⁶

Die der akademischen Profession zugesprochene intrinsische Motivation spielt auch hinsichtlich des Vertrauensaspektes eine zentrale Rolle. Das Vertrauen in die Profession wird, wie bereits angemerkt, aufgrund der Wissensasymmetrie zwischen dem Professionellen (Expertenwissen) und dem Klienten notwendig und bildet so die Basis für die Autonomie der Professionen. Für die akademische Profession stellen die Klienten nun gesellschaftliche Stakeholder (z. B. politische Akteure) bzw. im Bereich Lehre die Stu-

²⁴ Auf die Schwierigkeit der Definition von Altruismus wird an dieser Stelle nicht weiter eingegangen. Für eine Diskussion und weiterführende Literatur siehe z. B. Fetchenhauer und Bierhoff (2004) oder Batson (1991).

²⁵ Auf die theoretischen Hintergründe zur intrinsischen Motivation wird in Kapitel 4.1.1 und 4.1.2 noch näher eingegangen.

²⁶ Hier können allerdings – neben der intrinsischen Motivation – auch extrinsische Anreize wirksam werden wie z. B. Wertschätzung durch die Studierenden bzw. die Hochschule oder Zufriedenheit über den Lernerfolg der Studierenden.

dierenden dar. Aufgrund des spezifischen Fachwissens der Mitglieder der akademischen Profession ist es für Außenstehende zum einen schwierig, die Arbeit zu kontrollieren. Zum anderen lassen sich für die zwei Kernaufgaben der akademischen Profession – die Forschung und die Lehre – von außen nur schwer klare Vorgaben und Handlungsanweisungen formulieren, die den Erfolg der wissenschaftlichen Arbeit garantieren. Die gesellschaftlichen Akteure vertrauen also darauf, dass sich die Professorinnen und Professoren dem Wert der kognitiven Rationalität verpflichtet sehen (Schimank 2005, S. 152) und „(...) durch ein intrinsisches Erkenntnisstreben geleitet werden (...)“ (ebd.) und gewähren entsprechende Handlungsautonomie. Es wird darauf vertraut, dass Professorinnen und Professoren aufgrund ihrer intrinsischen Motivation und ihrer Bindung an den gesellschaftlichen Zentralwert der Wahrheitsfindung keiner externen Vorgaben oder Anreize bedürfen.

Bevor der Autonomie- und Vertrauensaspekt im nächsten Teilkapitel nochmals hinsichtlich der Thematik der Steuerungsversuche im Wissenschaftssystem aufgegriffen wird, soll hier nicht unerwähnt bleiben, dass durchaus umstritten ist, ob die wissenschaftliche Tätigkeit überhaupt als eine Profession angesehen werden kann. Es wird beispielsweise argumentiert, dass wesentliche Merkmale einer Profession wie ein eindeutiger Klientenbezug (zumindest im Bereich Forschung) oder ein praktisches Anwendungsfeld fehlen (vgl. zur Diskussion Enders 1998, S. 56–58). Wie anhand der vorangegangenen Ausführungen zu den Spezifika der akademischen Profession deutlich wurde, wird in der vorliegenden Arbeit die Berufsgruppe der Wissenschaftlerinnen und Wissenschaftler in Übereinstimmung mit der Sichtweise von Parsons und Platt (1990) – ohne nun tiefer in die Diskussion einzusteigen – dennoch als Profession angesehen.

Nachdem im vorangegangenen Kapitel die Universität als Organisation und in diesem Kapitel die Merkmale der akademischen Profession näher betrachtet wurden, wird im Folgenden das „durchaus spannungsreiche[] Verhältnis“ (Meier und Schimank 2010, S. 108) zwischen diesen analysiert. Dabei werden insbesondere der (zunehmende) Versuch der Steuerung wissenschaftlichen Verhaltens sowie die mit den Besonderheiten der akademischen Profession verbundene Steuerungsproblematik in den Blick genommen.

3.3 Die Steuerungsresistenz der akademischen Profession

Wie sich im vorangegangenen Teilkapitel gezeigt hat, gründet die Autonomie der akademischen Profession primär auf dem Vertrauen in das intrinsisch motivierte und allein der wissenschaftlichen Wahrheitsfindung verpflichtete Handeln der Professionsmitglieder. Die hervorgehobene Stellung der akademischen Profession macht dabei zu einem großen Teil die Besonderheit der Organisation Universität aus (vgl. Kapitel 3.1 und 3.2).

Nun mehren sich in den letzten Jahren Positionen, die einen zunehmenden Vertrauensverlust in die akademische Profession und ein sich wandelndes Verhältnis zwischen Organisation und Profession wahrnehmen. In einer Umfrage des *Instituts für Demoskopie Allensbach* unter deutschen Führungskräften aus dem Jahr 1993 ergaben sich beispielsweise mehrheitlich negative Bewertungen hinsichtlich der Arbeit deutscher Universitäten (Daniel et al. 1998, S. 15–16). Im Rahmen einer international vergleichenden Hochschul-lehrerstudie von Enders und Teichler zeigte sich, dass auch die Hochschullehrer selbst ein schwindendes Ansehen ihres Berufes in der Öffentlichkeit wahrnehmen (Enders 1998, S. 67). Neuere professionstheoretische Ansätze thematisieren einen steigenden Ökonomisierungsdruck auf die Professionen, der Aspekte der wirtschaftlichen Effizienz und Effektivität im professionellen Handeln statt Gemeinwohlorientierung und Klientenbezug betont (vgl. Kloke 2014, S. 136). Power (1999) prägte in diesem Zusammenhang den Begriff der *Audit Society*, in welcher das Vertrauen in die sich selbst regulierenden Professionen einem zunehmenden Zwang zur quantitativen Vermessung und Bewertung der erbrachten Leistung weicht (Lange 2008, S. 240). Den Professionen wird nicht mehr ‚blind vertraut‘, sondern sie müssen sich durch Leistung legitimieren. Die Erfassung der Leistung erfolgt nun über vor allem quantitativ messbare Indikatoren. Das grundlegende Vertrauen in die Profession wird damit ersetzt durch das Vertrauen in die „scheinbar sichere (...) Welt der Leistungskennziffern und Indikatoren“ (ebd.). Dass diese nicht immer als valide Repräsentanten für das zu messende Konstrukt (Forschungsleistung oder Forschungsqualität) angesehen werden können und eine einseitige Fokussierung auf Leistungsindikatoren mit verschiedenen nichtintendierten Effekten einhergehen kann, wurde bereits in Kapitel 2.3 aufgezeigt.

Auf was lässt sich der beschriebene Vertrauensverlust in die akademische Profession aber nun zurückführen? Nach Schimank (2005) resultieren dieser und das sich wandelnde Verhältnis zwischen Staat und akademischer Profession bzw. zwischen der Organisation Universität und der akademischen Profession vor allem aus den folgenden beiden Umständen (vgl. im Folgenden ebd., S. 153–157):

Erstens kam es im Zuge der Bildungsexpansion in den 60er und 70er Jahren aufgrund des gestiegenen Bedarfs an Lehrpersonal zu einer eiligen Rekrutierung einer Vielzahl von neuen Professorinnen und Professoren, die allerdings nicht alle das notwendige Qualitätsniveau mitbrachten. Das Problem wurde weiter durch eine Überlastung der akademischen Profession aufgrund der hohen Studierendenzahlen verschärft, worunter teilweise die Qualität der Arbeit in Forschung und Lehre zu leiden hatte. In der Wahrnehmung der Bevölkerung führte dies nun nach Schimank zu einer zunehmenden öffentlichen Desillusionierung der akademischen Profession (ebd., S. 154). In diesem Zusammenhang sind sicherlich auch zwei von Stichweh (2005) thematisierte Aspekte von Bedeutung. Zum einen:

Die *Massenmedien* (Hervorhebung im Original), einer der gesellschaftlichen Funktionszusammenhänge, zu deren gesellschaftlichen Funktionen es zu gehören scheint, Kritik und Misstrauen auf alle traditionellen gesellschaftlichen Institutionen anzuwenden, applizieren diese Strategien auch und gerade auf die Professionen und deren traditionell gesicherte Privilegien und dies wirkt vermutlich in die Richtung einer graduellen Erosion der Bereitschaft, an Kompetenz zu glauben und Vertrauen zu investieren. (ebd., S. 41)

Auch wenn obige Formulierung arg zugespitzt erscheint, kommt negativer Berichterstattungen, z. B. über die ‚faulen Akademiker‘, natürlich eine stark meinungsbildende Funktion zu. Insbesondere für Bevölkerungsgruppen außerhalb des akademischen Feldes werden Probleme von und in Universitäten damit sichtbar. Zum anderen bezieht sich Stichweh auf den Begriff der Wissensgesellschaft. Mit der Pluralisierung der Zugänge zu Wissen und zunehmenden Professionalisierungstendenzen anderer gesellschaftlicher Gruppen bzw. Berufe verlieren die alten Professionen zunehmend ihren exklusiven Status (vgl. ebd., S. 42). Jeder wird schnell zu einem vermeintlichen Experten auf einem Gebiet – man denke beispielsweise an die vielfältigen Möglichkeiten zur zweifelhaften Selbstdiagnose bei medizinischen Fragen im Internet.

Den zweiten Grund für ein sinkendes Vertrauen in die akademische Profession sieht Schimank in der ebenfalls in den 70er Jahren zunehmenden gesellschaftlichen Forderung nach außerwissenschaftlicher Relevanz der akademischen Arbeit an Universitäten. Sowohl die Forschung als auch die Lehre solle sich mehr an den Anforderungen anderer gesellschaftlicher Teilbereiche (z. B. Industrie, Militär) orientieren und entsprechend anwendungsbezogener ausgestaltet werden. Während einige Professorinnen und Professoren „sich diesen wissenschaftspolitischen Herausforderungen gestellt und sie als Chance gesehen [haben]“ (ebd., S. 156), stand die Mehrheit der Wissenschaftlerinnen und Wissenschaftler diesem Versuch der Einflussnahme in grundlegend autonome Entscheidungsbereiche ihrer Arbeit kritisch gegenüber. Diese Abwehrhaltung wiederum wird von staatlicher Seite „als tiefgreifendes Desinteresse an den Bedürfnissen der Gesellschaft gedeutet. Derart misstrauisch, lesen die staatlichen Akteure Autonomie als Verantwortungslosigkeit“ (ebd.).

Verschärft wurden diese beiden Entwicklungen durch eine staatliche Finanzknappheit, die die staatlichen Akteure zu einer möglichst effizienten Mittelverwendung zwang. Evaluationen der Lehr- und Forschungsleistung sollen dabei Qualität und Relevanz der wissenschaftlichen Arbeit in den Blick nehmen und als Entscheidungsgrundlage für die Mittelverteilung zwischen und innerhalb der Universitäten dienen (ebd., S. 156–157). Zudem befördert nach Schimank der Generationenwechsel innerhalb der Professorenschaft die Versuche einer Neugestaltung des Governance-Regimes der Universitäten

und dem damit verbundenen Ziel der Schwächung der Autonomie der akademischen Profession (vgl. auch Kapitel 2.1): Im Rahmen der W-Besoldung (vgl. Kapitel 2.2.3) werden neue Professorinnen und Professoren in Form eines leistungsbezogenen Vergütungssystems bezahlt, Mittel für die Grundausstattung der neuen Professuren werden nur noch auf Zeit vergeben und zudem wird von den „staatlichen Stellen [gehofft], dass die neue Generation von Professoren im neuen Governance-Regime sozialisiert werden wird und sich dementsprechend nicht mehr so stark mit der professionellen Autonomie identifizieren wird wie die nun ausscheidende Generation“ (ebd., S. 157).

Sicherlich lässt sich darüber streiten, ob die dargestellten Entwicklungen alle in dem vermuteten Maße zur Schwächung der akademischen Profession beigetragen haben bzw. beitragen. Unbestreitbar ist aber, dass die Versuche der Einflussnahme in Form von staatlichen und organisationsinternen Steuerungsverfahren zugenommen haben (vgl. Kapitel 2.1 und 2.2) und sich die akademische Profession einem neuen Legitimationsdruck ausgesetzt sieht. Legitimieren soll sich diese nun wie bereits angemerkt über Leistung bzw. Leistungstransparenz (vgl. Daniel et al. 1998, S. 14–15). Dies bedeutet aber keineswegs, dass die akademische Profession ihre Autonomie vollständig zugunsten der Organisation Universität oder dem Staat eingebüßt hätte. Wie sich in den beiden vorangegangenen Kapiteln bereits zeigte, ergeben sich aus den Spezifika der Organisation Universität und denen der akademischen Profession eine Reihe von Gründen, die staatliche oder organisationsinterne Steuerungsversuche des Forschungs- oder Lehrverhaltens zumindest erschweren und der akademischen Profession – trotz des beschriebenen Vertrauensverlustes – weiterhin in hohem Maße Autonomie garantieren (vgl. im Folgenden Kapitel 3.1 und 3.2).

Hier ist zunächst die lose Kopplung zwischen den einzelnen Organisationseinheiten zu erwähnen, die eine systematische und zentrale Steuerung von einer oberen Hierarchieebene erschweren. So sind z. B. in den Naturwissenschaften andere Schwerpunkte in Forschung (z. B. andere Publikationskultur) und Lehre (z. B. höhere Internationalität) zu setzen als in den Geisteswissenschaften. Auch innerhalb einer Fakultät gibt es fachspezifische Differenzen und selbst zwischen den Lehrstühlen eines Faches können sich, beispielsweise im Hinblick auf die Ausrichtung der Forschung (z. B. Grundlagenforschung vs. anwendungsbezogene Forschung), starke Unterschiede ergeben, die eine jeweils individuelle Betrachtungsweise je Organisationseinheit erfordern (würden). Mit dieser losen Kopplung verknüpft ist das jeweils spezifische Fach- oder Expertenwissen jedes Faches und jeder Professorin sowie jedes Professors. Außenstehende oder hierarchisch übergeordnete Steuerungsinstanzen haben dadurch immer mit einer Wissensasymmetrie zu kämpfen, die die Kontrolle der wissenschaftlichen Arbeit – vor allem im

Hinblick auf qualitative, weniger auf quantitative Aspekte – schwierig und die Formulierung konkreter Handlungsanweisungen bzw. Vorgaben für die Arbeit in Forschung und Lehre nahezu unmöglich macht.

Die starke Stellung der akademischen Profession wird auch im Rahmen des Profibürokratieansatzes deutlich, in welchem die Professorinnen und Professoren den stark ausgeprägten operativen Kern der Organisation Universität bilden. Dieser operative Kern erbringt die zentralen Leistungen in Forschung und Lehre und hat, auch aufgrund der engen Verknüpfung zur organisationalen Mittellinie, neben formeller Macht in Gremien und ähnlichem auch große informelle Macht inne. Trotz der Reformbestrebungen im Governance-Regime der Universitäten (vgl. Kapitel 2.1) können den Steuerungsversuchen der Hochschulleitung damit nach wie vor von starker Kollegialität geprägte Interessen (vgl. Schimank 2005, S. 149) des operativen Kerns gegenüberstehen.

Zudem erschweren drei Merkmale von Universitäten, die bereits in Kapitel 3.1 als Kennzeichen von organisierten Anarchien herausgearbeitet wurden, die hochschulinternen Steuerungsversuche: die nicht eindeutigen Zielorientierungen von Universitäten, deren unklare Technologien sowie die fluiden Teilnehmerschaften in Entscheidungssituationen in Gremien und anderen Verhandlungen. Eine gezielte Steuerung wird also erstens dadurch erschwert, dass die Priorisierung von Zielen unklar sein kann bzw. Zielkonflikte auftreten können (z. B. Zeitaufwand für Lehre vs. Forschung). Zweitens ist unklar bzw. unsicher, auf welche Weise bzw. ob das gewünschte Ziel erreicht werden kann (z. B. Erfolgsunsicherheit bei Beantragung von Fördergeldern). Drittens können – sofern keine persönlichen Vereinbarungen abgeschlossen werden – die Zuständigkeiten aufgrund wechselnder Ansprechpartner intransparent sein und zudem bei den Adressaten der Steuerungsversuche möglicherweise mit einer Verantwortungsdiffusion einhergehen, so dass die Steuerungsversuche ins Leere laufen.

Neben den genannten Gründen ist abschließend noch eine bereits in Kapitel 3.2 thematisierte Besonderheit der akademischen Profession zu erwähnen, welche maßgeblich die tendenzielle Steuerungsresistenz der akademischen Profession mitbegründet: der hohe Stellenwert der Scientific Community als zentraler Bezugspunkt der Professionsangehörigen und als die Produktionsstätte der wissenschaftlichen ‚Währung‘ Reputation. Die zugrundeliegende Annahme hierbei lautet, dass sich die Professorinnen und Professoren primär an ihrer jeweiligen fachspezifischen Scientific Community orientieren. Sie sind dabei bestrebt, in der Scientific Community möglichst sichtbar zu sein und entsprechende Reputation zu gewinnen. Staatliche oder organisationsinterne Steuerungsversuche wären daher für die akademische Profession nur von nachrangiger Bedeutung und ihnen wird – je nachdem wie stark sich die Ziele der Steuerungsversuche mit den Anforderungen der Scientific Community decken – im eigenen Handeln mal mehr und mal weniger intensiv nachgegangen werden.

Die vorangegangenen Ausführungen zur Organisation Universität, der akademischen Profession und dem Verhältnis von Steuerung und akademischer Autonomie zeigen auf, dass trotz Reformen im Governance-Regime der Hochschulen und trotz einer zu beobachtenden Schwächung der Stellung der akademischen Profession (Vertrauensverlust) staatlichen und organisationsinternen Steuerungsversuchen nach wie vor eine hohe Autonomie der Wissenschaft gegenübersteht. Ein Anreiz für die Erfüllung der Kriterien der Steuerungs- und Anreizverfahren, welche an die Professorinnen und Professoren herangetragen werden (z. B. Steigerung der Promotionsquote, Steigerung der Zitationsquote etc.), kann aber – neben einem zusätzlichen monetären Ertrag – die in diesem Teilkapitel beschriebene Legitimationsfunktion sein, welche das verloren gegangene Vertrauen in die akademische Profession ersetzt. Zudem gibt es empirische Hinweise, dass die Steuerungsversuche insbesondere dann erfolgversprechend sind, wenn die betroffenen Akteure bei der Implementierung der Steuerungsinstrumente beteiligt werden (vgl. Minssen et al. 2003, S. 121). Hier spielt es sicherlich eine Rolle, dass die Beteiligung von Professorinnen und Professoren als Fachvertreter bei der Implementierung der Steuerungsinstrumente dazu führt, dass die Ausgestaltung von diesen in der Folge mehr der Fachkultur der entsprechenden Scientific Community entspricht. In der Summe ist aber dennoch davon auszugehen, dass aufgrund der beschriebenen Spezifika der akademischen Profession, insbesondere der dominanten Rolle der Scientific Community, die Wirkung von staatlichen und organisationsinternen Steuerungs- und Anreizverfahren deutlich eingeschränkt ist.

Vor dem Hintergrund der hier vorgenommenen organisations- und professionstheoretischen Überlegungen soll nun im nächsten Kapitel das theoretische Modell auf der Mikroebene der handelnden Akteure, also der Professorinnen und Professoren, herausgearbeitet werden. Hierzu werden zunächst die motivationstheoretischen Zugänge erläutert und im Anschluss unter Rückgriff auf die Erkenntnisse dieses Kapitels die Forschungshypothesen abgeleitet.

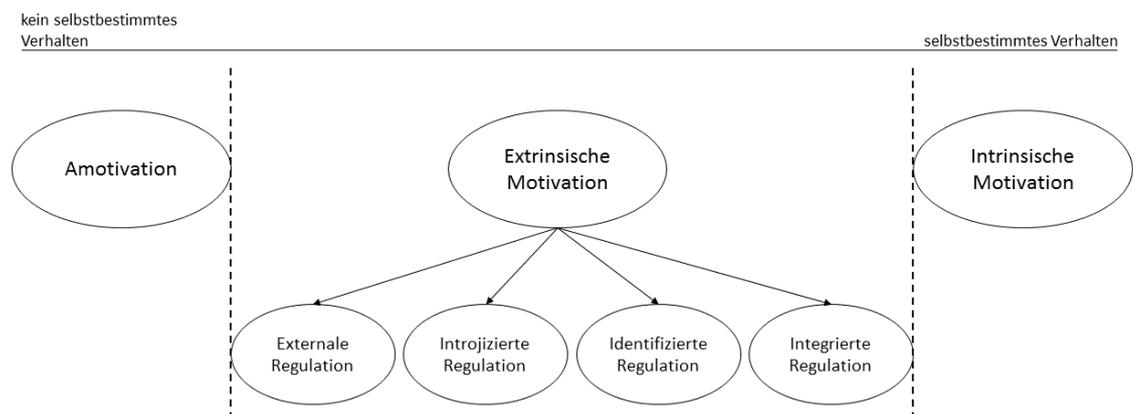
4. Theoretisches Modell und Ableitung der Forschungshypothesen

4.1 Motivationstheoretischer Rahmen

4.1.1 Die Selbstbestimmungstheorie der Motivation

Den Ausgangspunkt für das nun zu entwickelnde theoretische Modell bildet die *Selbstbestimmungstheorie der Motivation* von Deci und Ryan (1993; Ryan und Deci 2000; Gagné und Deci 2005). Unter dem Blickwinkel dieser Theorie lässt sich die motivationale Grundlage von Handeln als Kontinuum begreifen, welches von (stark) kontrollierten Formen der Motivation zu (rein) autonomen Formen der Motivation reicht. Im Gegensatz zu anderen motivationstheoretischen Ansätzen wird also nicht nur dichotomisch zwischen extrinsischer Motivation auf der einen und intrinsischer Motivation auf der anderen Seite unterschieden, sondern extrinsisch motivierte Handlungen werden nochmals anhand des Grades ihrer (wahrgenommenen) Kontrolliertheit bzw. Autonomie differenziert (vgl. Deci und Ryan 1993, S. 225). Autonom motiviertes Handeln impliziert dabei einen hohen Grad an (wahrgenommener) Selbstbestimmung und Wahlfreiheit (vgl. Gagné und Deci 2005, S. 333–334), während „(...) being controlled involves acting with a sense of pressure, a sense of *having to* [Hervorhebung im Original] engage in actions“ (ebd.). Das Kontinuum der unterschiedlichen Typen mit den Endpolen der Amotivation und der intrinsischen Motivation ist in Abbildung 1 dargestellt.

Abbildung 1: *Motivationstypen nach der Selbstbestimmungstheorie der Motivation*



Quelle: Darstellung nach Ryan und Deci (2000, S. 72) und Gagné und Deci (2005, S. 336).

Entscheidend ist nun, dass extrinsisch motivierte Verhaltensweisen durchaus als selbstbestimmtes Verhalten erlebt werden können:

Other types of extrinsic motivation [andere als *Externale Regulation*; Anm. d. Verf.] result when a behavioral regulation and the value associated with it have been internalized. Internalization is defined as people taking in values, attitudes, or regulatory structures, such that the external regulation of a behavior is transformed into an internal regulation and thus no longer requires the presence of an external contingency (thus, I work even when the boss is not watching). (...) The more fully it has been internalized, the more autonomous will be the subsequent, extrinsically motivated behavior. (Gagné und Deci 2005, S. 334)

Die Motivationstypen lassen sich damit anhand der Art ihrer Verhaltensregulation (*regulatory styles*) differenzieren (Ryan und Deci 2000, S. 72), also danach, inwieweit externe Werte, Einstellungen und Regulationsmechanismen von einer Person internalisiert wurden. Internalisierung wird dabei als Oberbegriff verstanden, der auf *Introjektion*, *Identifikation* und *Integration* als drei zu unterscheidende Prozesse verweist (Gagné und Deci 2005, S. 334). Die diesen Internalisierungsprozessen zugeordneten sowie die anderen in Abbildung 1 dargestellten Motivationstypen werden nun im Folgenden näher erläutert:

Der linke Pol des Kontinuums, der Zustand der *Amotivation*, bezieht sich auf Handlungen, die mit keinen erkennbaren Intentionen verknüpft sind und die in diesem Sinne auch nicht als motiviert interpretiert werden können (Deci und Ryan 1993, S. 224). Deci und Ryan nennen hier beispielsweise Verhaltensweisen wie Dösen oder ein unkontrollierter Wutausbruch (ebd.). Solche Handlungen erfüllen natürlich auch ihren physiologischen oder psychologischen Zweck für eine Person, aber Deci und Ryan „(...) bezeichnen sie nicht als motiviert, weil sie nicht durch intentionale Prozesse gesteuert werden“ (ebd.). Demgegenüber findet sich auf der rechten Seite des Kontinuums die *intrinsische Motivation*, welche sich auf Verhalten bezieht, „(...) which is propelled by people’s interest in the activity itself“ (Gagné und Deci 2005, S. 334). Eine schon aus dem vorangegangenen Kapitel bekannte Definition intrinsischer Motivation konkretisiert: „We say a person is intrinsically motivated to perform an activity if there is no apparent reward except the activity itself or the feelings which result from the activity“ (Deci 1972, S. 217). Intrinsisch motiviertes Verhalten wird also aufgrund der (Freude mit der) Aktivität an sich ausgeführt und bedarf daher keiner externalen Anreize oder Regulierungen. In diesem Sinne stellt es den Prototyp selbstbestimmten bzw. autonomen Verhaltens dar (Deci und Ryan 1993, S. 226).

Zwischen diesen beiden Polen finden sich nun die vier Formen extrinsisch motivierten Verhaltens, die sich hinsichtlich des Grades der Verhaltensregulation zwischen externaler Kontrolle und Selbstbestimmung (Autonomie) unterscheiden lassen. Die *externale Regulation* kennzeichnet dabei den ‚klassischen‘ Typ extrinsisch motivierten Handelns (Gagné und Deci 2005, S. 334): Die Handlung wird primär aufgrund äußerer Anreize

(Belohnung) oder aufgrund (der Androhung) von Sanktionen ausgeführt. Sie weist damit den höchsten Grad der Kontrolle und den geringsten Grad der Selbstbestimmung bzw. Autonomie auf (vgl. ebd.).

Eine erste Form internalisierter Verhaltensregulation stellt die *introjizierte Regulation* dar. Ein auf diese Weise reguliertes Verhalten resultiert aus einem inneren Antrieb und ist nicht auf äußere Handlungsanstöße angewiesen (Deci und Ryan 1993, S. 227). Eine Handlung wird ausgeführt oder unterlassen, um Schuldgefühle zu vermeiden oder weil sie für das Selbstwertgefühl und die Selbstachtung relevant ist (vgl. ebd.). Obwohl der Antrieb für die Handlung also ‚von innen‘ kommt, sind solche Verhaltensweisen dennoch mehr fremd- als selbstbestimmt. Die internalisierten Regulationsmechanismen sind nicht akzeptierter Teil des Selbst, sondern werden als kontrollierend wahrgenommen: „Die introjizierte Regulierung beschreibt somit eine Form von Motivation, bei der die Verhaltensweisen durch innere Kräfte *kontrolliert* oder erzwungen werden, die außerhalb des Kernbereichs des individuellen Selbst liegen“ (ebd., S. 227–228).

Im Gegensatz zur introjizierten Regulation ist die *identifizierte Regulation* dadurch gekennzeichnet, dass die damit verbundenen Verhaltensweisen in hohem Maße mit den persönlichen Zielen und Werten kongruent und im Selbstkonzept integriert sind (ebd.): „Man tut etwas nicht einfach deshalb, weil man das Gefühl hat, es tun zu sollen [= introjizierte Regulation; Anm. d. Verf.], sondern weil man es für wichtig hält“ (ebd., S. 228). Entsprechend wird das Ausführen einer extrinsisch motivierten Handlung in Form der identifizierten Regulation als mehr selbstbestimmt denn external aufgezwungen oder kontrolliert erlebt (vgl. Gagné und Deci 2005, S. 334).

Die *integrierte Regulation* schließlich ist von allen Formen der Verhaltensregulation durch den höchsten Grad der Internalisierung gekennzeichnet und verweist damit auf extrinsisch motivierte Handlungen, die in höchstem Maße als selbstbestimmt wahrgenommen werden: „With integrated regulation, people have a full sense that the behavior is an integral part of who they are, that it emanates from their sense of self and is thus self-determined“ (ebd., S. 335). Personen führen Verhaltensweisen, die auf solche Weise motiviert sind, also (freiwillig) in völliger Übereinstimmung mit ihren eigenen inneren Zielen, Werten oder Normen aus und erleben die Handlung dadurch als vollkommen selbstbestimmt. Diese Form der Motivation unterscheidet sich dennoch von der intrinsischen Motivation und zwar dahingehend, dass „(...) intrinsisch motivierte Verhaltensweisen autotelischer Natur sind, während integriertes (extrinsisches) Verhalten eine instrumentelle Funktion besitzt, aber freiwillig ausgeführt wird, weil das individuelle Selbst das Handlungsergebnis hoch bewertet“ (Deci und Ryan 1993, S. 228).

Zusammenfassend kann also festgehalten werden, dass sich im Rahmen der Selbstbestimmungstheorie der Motivation neben amotivierten und intrinsisch motivierten Handlungen vier Typen extrinsisch motivierter Handlungen unterscheiden lassen, die einen

zunehmenden Grad an Selbstbestimmtheit aufweisen. Verhaltensweisen, die external oder introjiziert reguliert sind, sind dabei durch keine bzw. nur einen geringen Grad an wahrgenommener Selbstbestimmung und ein hohes Maß an wahrgenommener Kontrolle gekennzeichnet, während identifiziert und integriert regulierte Handlungen als (vollkommen) autonom motiviert angesehen werden können.

Im Hinblick auf die motivationale Grundlage von Handeln ist für die Selbstbestimmungstheorie noch ein weiterer Aspekt entscheidend. Es handelt sich hierbei um die Annahme dreier psychologischer Grundbedürfnisse: dem Bedürfnis nach Kompetenz bzw. Wirksamkeit, nach Autonomie bzw. Selbstbestimmung und dem Bedürfnis nach sozialer Eingebundenheit. Diese drei Grundbedürfnisse sind nun sowohl für intrinsische als auch extrinsische Formen der Motivation von zentraler Bedeutung (ebd., S. 229). Deci und Ryan (1993) beschreiben die motivierende Wirkung dieser Bedürfnisse wie folgt:

Wir gehen also davon aus, daß der Mensch die angeborene motivationale Tendenz hat, sich mit anderen Personen in einem sozialen Milieu verbunden zu fühlen, in diesem Milieu effektiv zu wirken (zu funktionieren) und sich dabei persönlich autonom und initiativ zu erfahren. (ebd.)

Die drei postulierten psychologischen Grundbedürfnisse liefern der Selbstbestimmungstheorie also die Erklärung dafür, „(...) *warum* [Hervorhebung im Original] bestimmte Handlungsziele motivierend sind“ (ebd.). Die Befriedigung dieser Bedürfnisse ist dabei auch notwendige Bedingung für die Internalisierungsprozesse, die mit den oben beschriebenen extrinsischen Motivationstypen verknüpft sind:

SDT [Self-Determination Theory; Anm. d. Verf.] postulates that when people experience satisfaction of the needs for relatedness and competence with respect to a behavior, they will tend to internalize its value and regulation, but the degree of satisfaction of the need for autonomy is what distinguishes whether identification or integration, rather than just introjection, will occur. (Gagné und Deci 2005, S. 337)

Je mehr eine Handlung also mit dem Gefühl der sozialen Eingebundenheit und einem Kompetenzerleben einhergeht, desto eher werden die mit der Handlung verknüpften Werte, Normen oder andere Handlungsregulierungen (z. B. auch Gesetze) verinnerlicht. Der Grad, in welchem sich der Handelnde dabei als autonomer Akteur erlebt, also das Bedürfnis nach Autonomie befriedigt wird, bestimmt schließlich, ob die Internalisierung auf der eher externalen Ebene der introjizierten Regulation verbleibt oder zu autonomen Formen der Motivation (identifizierte oder integrierte Regulation) führt.

Für die intrinsische Motivation sind nun insbesondere das Kompetenz- und das Autonomieerleben von besonderer Bedeutung. Intrinsisch motiviert handeln kann nur, wer sich bei der Ausführung der entsprechenden Handlung als ausreichend kompetent und vor allem selbstbestimmt erlebt. Die (erfolgreiche) Durchführung intrinsisch motivierter Verhaltensweisen trägt wiederum zum Entstehen von Gefühlen der Kompetenz und der Autonomie bei (vgl. Deci und Ryan 1993, S. 230).

Aufgrund der Abhängigkeit intrinsisch motivierten Handelns von einem Kompetenz- und Autonomieerleben ist dieses auch auf entsprechende Umweltbedingungen angewiesen, die ein selbstbestimmtes und als kompetent erlebtes Handeln erst ermöglichen (vgl. Deci und Ryan 1993, S. 230). So sind die Gewährung ausreichend individueller Handlungsfreiräume und Wahlfreiheiten sowie kompetenzförderndes Verhalten durch die Umwelt (z. B. informatives Feedback) positiv mit intrinsisch motiviertem Verhalten assoziiert und können die intrinsische Motivation sogar noch steigern. Die positive Wirkung von autonomie- und kompetenzförderndem Verhalten auf die intrinsische Motivation konnte bereits in zahlreichen Labor- und Feldexperimenten bestätigt werden (vgl. ebd., S. 230–232; Deci et al. 1999). Andererseits können Umweltbedingungen, die für das Autonomie- und Kompetenzerleben hinderlich sind, wie z. B. die Kontrolle von Handlungen oder konkrete Handlungsanweisungen, negative Auswirkungen auf die intrinsische Motivation haben. Dieser Aspekt wird nun im nächsten Teilkapitel im Rahmen der *Motivation Crowding Theory* näher beleuchtet.

4.1.2 Motivation Crowding Theory

Die Überlegungen der Selbstbestimmungstheorie von Deci und Ryan wurden von der ökonomischen Forschung mit der *Motivation Crowding Theory* aufgegriffen und fortgeführt (Frey 1997; Frey und Jegen 2001). Die *Motivation Crowding Theory* geht zunächst grundlegend von der Annahme aus, dass externe Einflussnahmen auf das Handeln von Personen in Form von monetären Belohnungen oder mit negativen Sanktionen verknüpften Regulierungen die intrinsische Motivation sowohl positiv (crowding-in), negativ (crowding-out) oder auch gar nicht beeinflussen können (Frey und Jegen 2001, S. 592). In Übereinstimmung mit den Prämissen der Selbstbestimmungstheorie (vgl. Kapitel 4.1.1) kommt es zu einem Verdrängungseffekt der intrinsischen Motivation durch externe Einflussnahmen, wenn diese als kontrollierend empfunden werden. Die für das Entstehen und die Aufrechterhaltung von intrinsischer Motivation notwendigen psychologischen Grundbedürfnisse des Autonomie- und Kompetenzerlebens können durch als kontrollierend empfundene externale Regulierungen eingeschränkt werden. Die intrinsische Motivation zur Ausübung einer Tätigkeit kann sich dadurch entsprechend reduzieren (vgl. Frey und Jegen 2001, S. 594–595). Eine als unterstützend wahrgenommene

(Arbeits-)Umgebung kann dagegen durch eine Stärkung des Autonomie- und Kompetenzerlebens zu einer Erhöhung der intrinsischen Motivation beitragen (ebd., S. 595).

Da Formen der extrinsischen und intrinsischen Motivation in der Regel gemeinsam auftreten (vgl. ebd., S. 591), sind die Auswirkungen von Verdrängungs- oder Verstärkungseffekten der intrinsischen Motivation auf die tatsächlich resultierende Leistung jeweils differenziert unter Berücksichtigung der parallel wirkenden motivationalen Effekte der externen Einflussnahme zu betrachten. Das heißt mit anderen Worten, dass nicht nur die Wirkung der intrinsischen Motivation auf die Leistungserbringung, sondern gleichzeitig auch die Wirkung der externen Anreize oder Sanktionen auf die erbrachte Leistung zu analysieren ist. Hieraus ergeben sich nach der *Motivation Crowding Theory* bei der Anwendung externer Anreize oder Sanktionen die folgenden möglichen Effekte auf die Leistungserbringung (vgl. im Folgenden ebd., S. 593):

Zunächst kann davon ausgegangen werden, dass (monetäre) Belohnungen oder Sanktionen als extrinsische Form der Motivation wirksam werden und so zu einer Steigerung der Leistung beitragen (relativer Preiseffekt bzw. Disziplinierungseffekt). Wird die externe Einflussnahme nicht als kontrollierend wahrgenommen, das heißt nicht als Bedrohung der eigenen Handlungsautonomie und Kompetenz erlebt, bleibt die intrinsische Motivation dabei weitgehend unbeeinflusst, so dass in der Summe ein leistungssteigernder Effekt angenommen werden kann. Wenn die externe Einflussnahme als unterstützend für das eigene Handeln wahrgenommen wird, kann diese im Sinne des bereits erwähnten *crowding-in*-Effektes sogar mit einer Erhöhung der intrinsischen Motivation des Handelnden einhergehen. In diesem Fall würde sich die leistungssteigernde Wirkung der externen Anreize bzw. Sanktionen noch erhöhen, da neben dem Disziplinierungseffekt noch eine gesteigerte intrinsische Motivation zur Leistungssteigerung beiträgt. Anders verhält es sich, wenn die externe Einflussnahme als kontrollierend empfunden wird, also als external oder introjiziert regulierte Form extrinsischer Motivation wirksam wird (vgl. Kapitel 4.1.1). Hierdurch kann es zu einem Verdrängungseffekt (*crowding-out*) der intrinsischen Motivation kommen. Durch den Disziplinierungseffekt der externen Anreize oder Sanktionen zeigt sich dann zwar weiterhin eine leistungssteigernde Wirkung, gleichzeitig führt die Verringerung der intrinsischen Motivation aber zu einem sinkenden Leistungsniveau. Die Stärke des Disziplinierungseffektes und des Verdrängungseffektes bestimmen jeweils, ob es in einer solchen Situation in der Summe zu einer Leistungssteigerung, zu einer Leistungsminderung oder aber zu keiner sichtbaren Leistungsveränderung kommt. Dabei ist zu beachten, dass sich auch bei keinen sichtbaren Veränderungen in der erbrachten Leistung die motivationalen Grundlagen des Handelns deutlich verschoben haben können: Beispielsweise kann eine starke Verringerung der intrinsischen Motivation und eine daraus resultierende sinkende Leistungsbereitschaft

durch den Disziplinierungseffekt der externen Anreize (z. B. erhöhte Leistungsbereitschaft aufgrund hoher monetärer Boni) kompensiert worden sein. Das sichtbare Leistungsniveau bleibt also unverändert, während die Handlungsmotivation statt auf autonomen Formen der Motivation nun überwiegend auf externalen Formen der Motivation gründet.

Ohne nun vertiefend weitere psychologische Aspekte des Verdrängungseffektes zu diskutieren, soll hier dennoch darauf hingewiesen werden, dass die Verdrängung autonomer Formen der Motivation zugunsten von external regulierten Formen der Motivation – trotz gleichbleibendem Leistungsniveau – negative Auswirkungen auf die Arbeitszufriedenheit und das generelle Wohlbefinden haben kann (vgl. Gagné und Deci 2005, S. 346–347). Auch wenn die Leistung durch externe Anreize oder Sanktionen nicht negativ oder sogar positiv beeinflusst wird, können also nichtintendierte negative Effekte auf die Arbeitszufriedenheit und die psychische Gesundheit auftreten, die sich langfristig wiederum negativ auf die Gesamtbilanz der Leistungserbringung (z. B. Fehltage) niederschlagen können.

Der Verdrängungseffekt der intrinsischen Motivation wurde vor allem in psychologischen Untersuchungen bereits vielfach empirisch überprüft. In einer Metaanalyse von 128 experimentellen Studien konnten Deci et al. (1999) nachweisen, dass „(...) tangible rewards tend to have a substantially negative effect on intrinsic motivation“ (ebd., S. 658–659). Negative Effekte auf die intrinsische Motivation zeigten sich dagegen nicht, wenn die Belohnung unerwartet vergeben wurde oder in keiner direkten Beziehung zur ausgeführten Tätigkeit stand (ebd., S. 656). Zudem ergaben die Analysen in Übereinstimmung mit den theoretischen Annahmen der Selbstbestimmungstheorie und der *Motivation Crowding Theory*, dass die Art und Weise, wie externe Anreizsysteme ausgestaltet und kommuniziert werden (Wahrnehmung als kontrollierend vs. unterstützend), einen eher positiven oder einen eher negativen Effekt auf die intrinsische Motivation haben kann (vgl. ebd., S. 656–657). Für den universitären Bereich deuten die Ergebnisse einer Studie von Jacobsen und Andersen (2014) zur Wirkung universitärer Steuerungssysteme (command systems) auf den Publikationsoutput von Wissenschaftlerinnen und Wissenschaftlern an dänischen Hochschulen auf übereinstimmende Ergebnisse hin: Die Autoren beobachteten einen positiven Zusammenhang zwischen der Anwendung von als unterstützend wahrgenommenen Steuerungssystemen und dem individuellen Publikationsoutput der Untersuchungseinheiten (ebd., S. 17).²⁷ In einer weiteren Studie mit dänischen Lehrerinnen und Lehrern konnten Jacobsen et al. (2014) zeigen, dass verbindlich

²⁷ Wie die Autoren selbst anmerken, kann der vermutete dahinterstehende Verstärkungseffekt der intrinsischen Motivation (crowding-in) damit aber nicht empirisch bestätigt werden, da in der Studie die intrinsische Motivation selbst nicht erfasst wurde, sondern lediglich die Wahrnehmung der Steuerungssysteme und der Publikationsoutput (vgl. Jacobsen und Andersen 2014, S. 17).

vorgegebene Lehrpläne negativ mit der intrinsischen Motivation der Lehrkräfte korrelieren, wenn diese als kontrollierend wahrgenommen werden (ebd., S. 801). Auch im wirtschaftswissenschaftlichen Bereich finden sich eine Reihe von Labor- und Feldstudien, die explizit oder implizit eine empirische Bestätigung der Annahmen der *Motivation Crowding Theory* erlauben (vgl. für eine Übersicht Frey und Jegen 2001, S. 598–606).

Insgesamt kann der Verdrängungseffekt der intrinsischen Motivation als empirisch belegt angesehen werden. Die verschiedenen Studien hierzu machen aber auch deutlich, dass es – in Übereinstimmung mit der Theorie – wesentlich von der Ausgestaltung der Anreiz- oder Sanktionssysteme bzw. der Wahrnehmung von diesen durch die betroffenen Akteure abhängt, ob eine Verdrängung intrinsischer Motivation stattfindet oder ob die externe Einflussnahme keine oder sogar eine positive Wirkung auf die intrinsische Motivation hat. Wie erwähnt ist hier insbesondere entscheidend, ob die Einflussnahme als kontrollierend oder unterstützend wahrgenommen wird. In diesem Zusammenhang spielen Aspekte wie ausreichende Kommunikations- (Information) und Partizipationsmöglichkeiten bei der Implementierung von Anreizsystemen sowie deren faire und transparente Ausgestaltung eine wichtige Rolle (vgl. Frey und Osterloh 1997, S. 313–315).

Wie bereits in Kapitel 4.1.1 dargestellt wurde, kann die intrinsische Motivation als Prototyp autonomer Formen der Motivation verstanden werden. Andere Formen autonom motivierten Handelns lassen sich anhand des Internalisierungsgrades von extrinsischen Regulationsmechanismen (identifizierte oder integrierte Regulation) unterscheiden. Im nächsten Kapitel wird nun anhand des Konzepts der *Public Service Motivation* eine solche weitere Form autonomer Motivation dargestellt, bevor aufbauend auf den theoretischen Vorüberlegungen die Forschungshypothesen abgeleitet werden.

4.1.3 Das Konzept der Public Service Motivation

Das Konzept der *Public Service Motivation* (PSM) entstammt der Public Management Forschung und geht zunächst grundlegend von der Annahme aus, dass der Tätigkeit von Personen im öffentlichen Sektor bzw. staatlichen Organisationen spezielle Motivlagen gemeinsam sind, die sich beispielsweise in der Privatwirtschaft so nicht wiederfinden lassen (Perry et al. 2010, S. 681). Die Unterscheidung zwischen privaten und öffentlichen Organisationen ist allerdings – je nach herangezogenem Kriterium (z. B. Organisationsziel, Träger, Gewinnorientierung) – nicht immer ganz eindeutig. Für die PSM ist als Distinktionskriterium insbesondere das Organisationsziel bzw. die von der Organisation erbrachte Leistung von Bedeutung, welche bei öffentlichen Organisationen einen grundlegenden gesellschaftlichen Wert oder eine Funktion für den sozialen Wohlstand einer Gesellschaft besitzen soll (vgl. Vandenabeele 2008, S. 1091–1092). Der den Universitäten erteilte Lehr- und Forschungsauftrag lässt sich ohne Zweifel als ein solch ge-

meinwohlorientiertes Organisationsziel verstehen. Weitere in diesem Sinne typische Organisationen sind beispielsweise Krankenhäuser, Schulen oder auch andere Bereiche des öffentlichen Dienstes wie die Polizei. Perry et al. (2010) betonen allerdings, dass das Auftreten von PSM nicht ausschließlich auf die oben genannten Organisationstypen bzw. den öffentlichen Sektor beschränkt ist. Sie gehen aber davon aus, dass PSM insbesondere im öffentlichen Sektor stärker verbreitet ist als in anderen gesellschaftlichen Bereichen (vgl. ebd., S. 682).

Trotz zunehmender Rezeption der PSM im internationalen wissenschaftlichen Diskurs (vgl. Bozeman und Su 2015, S. 700–701) bleibt das dahinter stehende theoretische Konzept allerdings recht vage. Dies liegt nicht zuletzt an der Vielzahl unterschiedlicher Definitionsversuche, welche oft keine ausreichende Trennschärfe zu anderen theoretischen Ansätzen (z. B. Altruismus, prosoziales Verhalten) aufweisen (vgl. ebd., S. 701–704; Perry et al. 2010, S. 681–682). Für die Zwecke dieser Arbeit wird der Begriff der PSM nun im nachfolgend beschriebenen Sinne verstanden:

PSM ist nach Perry und Wise (1990) zunächst definiert als „(...) an individual's predisposition to respond to motives grounded primarily or uniquely in public institutions" (ebd., S. 368). Diese recht allgemein gehaltene Begriffsbestimmung wird hier um die Definition von Vandenberg (2007) erweitert, welcher PSM versteht als „(...) the beliefs, values and attitudes that go beyond self-interest and organizational interest, that concern the interest of a larger political entity and that motivate individuals to act accordingly whenever appropriate" (ebd., S. 547). Dabei unterscheidet sich die PSM von der intrinsischen Motivation dadurch, dass PSM im Gegensatz zur rein intrinsisch motivierten Handlung durch die Orientierung an einen Zweck oder ein ‚höheres‘ Ziel gekennzeichnet ist (Perry et al. 2010, S. 682). Im Sinne der Selbstbestimmungstheorie (vgl. Kapitel 4.1.1) kann PSM als (stark) internalisierte extrinsische, und damit autonome Form der Motivation verstanden werden (vgl. Jacobsen 2012, S. 29–30). Das mit der PSM verknüpfte Ziel wird weder aus reinem Eigeninteresse oder aufgrund irgendwelcher institutioneller Zwänge verfolgt, sondern die dahinter stehenden (für das Gemeinwohl relevanten) Werte und Zielsetzungen sind soweit internalisiert, dass keine externe Regulation des Verhaltens notwendig wird und dieses überwiegend autonom motiviert abläuft.

Auch für die PSM als eine autonome Form der Motivation spielen, den Annahmen der Selbstbestimmungstheorie folgend, die drei psychologischen Grundbedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit eine wichtige Rolle. Nur wenn die Umweltbedingungen die Befriedigung dieser drei Bedürfnisse ermöglichen, kann die Handlungsmotivation im Sinne der PSM internalisiert und aufrecht gehalten werden (vgl. ebd., S. 33). Im Gegensatz zur intrinsischen Motivation, die überwiegend auf das Autonomie- und Kompetenzerleben angewiesen ist, ist für die PSM auch das Bedürfnis nach sozialer Eingebundenheit von Bedeutung. Jacobsen (2012) erläutert dies wie folgt:

The need for relatedness [soziale Eingebundenheit; Anm. d. Verf.] matters for public service motivation because it is carried by the wish to help other people, but also because it is institutionally grounded in a set of common norms and values. Thus, the feeling of serving a larger community is inherent in public service motivation. (ebd., S. 33–34)

Da das Auftreten und die Aufrechterhaltung der PSM von der Befriedigung der drei psychologischen Grundbedürfnisse abhängig ist, kann diese, wie andere autonome Formen der Motivation auch, vom Verdrängungseffekt der Motivation betroffen sein (vgl. Kapitel 4.1.2; Jacobsen 2012, S. 32–33). Dies wäre entsprechend der *Motivation Crowding Theory* immer dann der Fall, wenn externe Einflussnahmen, weil diese z. B. als kontrollierend empfunden werden, das Autonomie- oder Kompetenzerleben des Handelnden beschränken (vgl. Frey und Jegen 2001, S. 594–595). Empirische Hinweise hierzu liefert die bereits erwähnte Studie von Jacobsen et al. (2014) zur Wirkung verbindlich vorgegebener Lehrpläne auf die Motivation von dänischen Lehrkräften: Je mehr die Einführung der verbindlichen Lehrpläne als kontrollierend wahrgenommen wurde, desto niedrigere Skalenwerte ergaben sich nicht nur hinsichtlich der intrinsischen Motivation der Lehrkräfte, sondern auch hinsichtlich der PSM (vgl. ebd., S. 801).

Die auch für diese Arbeit relevante Annahme, dass PSM – bei entsprechenden förderlichen Umwelt- bzw. Arbeitsbedingungen – positiv mit der Arbeitsleistung korreliert (vgl. Perry et al. 2010, S. 684; Perry und Wise 1990, S. 370–371), wurde in verschiedenen Studien überprüft (vgl. für einen Überblick Perry et al. 2010, S. 684–685). In Untersuchungen mit Angestellten im öffentlichen Dienst verschiedener Ländern (unter anderem USA, Korea, Schweiz) konnte der positive Zusammenhang zwischen PSM (vgl. zu Möglichkeiten und Versuchen der Operationalisierung Perry et al. 2010, S. 682–683) und Leistung (zumeist erfasst in Form einer Selbstauskunft in schriftlichen Befragungen) überwiegend bestätigt werden (vgl. ebd.). In einzelnen Studien (Bright 2007; Vandenabeele 2009) wurden auch mögliche Mediatorvariablen wie z. B. die Arbeitszufriedenheit oder der *Person-Organization Fit* in den Blick genommen. Hierbei ergeben sich allerdings teils widersprüchliche Ergebnisse, die noch weiteren Forschungsbedarf offenbaren (vgl. Perry et al. 2010, S. 684). Insgesamt aber deuten die empirischen Befunde zur PSM trotz der bereits erwähnten theoretischen Unschärfen des Konzepts darauf hin, dass Personen, die in Organisationen mit mehr gemeinwohlorientierten Organisationszielen – also zumeist öffentliche Organisationen – tätig sind, eine spezifische Form der Arbeitsmotivation teilen, die positiv mit der Arbeitsleistung verknüpft ist.

PSM wurde hier definiert als „(...) an individual’s predisposition to respond to motives grounded primarily or uniquely in public institutions” (Perry und Wise 1990, S. 368) sowie als „(...) the beliefs, values and attitudes that go beyond self-interest and organizational

interest, that concern the interest of a larger political entity and that motivate individuals to act accordingly whenever appropriate” (Vandenabeele 2007, S. 547). Ruft man sich die professionstheoretischen Ausführungen aus Kapitel 3.2 in Erinnerung, in welchen die Kollektivorientierung an einem gesellschaftlichen Zentralwert wie Gesundheit, Gerechtigkeit oder Wahrheit als Bestimmungsmerkmal einer Profession benannt wurde, kann die PSM nun als motivationstheoretisches Pendant zur professionstheoretischen Kollektivorientierung aufgefasst werden. Wie bereits ausgeführt wurde (vgl. Kapitel 3.2), ist für Parsons die Unterscheidung zwischen der Kollektivorientierung der Professionen und der Selbstorientierung der Privatwirtschaft allerdings mehr durch die institutionelle Ebene als durch die individuelle Handlungsmotivation begründet. Er lehnt also die Annahme ab, dass beispielsweise alle Ärzte per se als Person altruistisch motiviert wären und demgegenüber alle privatwirtschaftlichen Unternehmer rein egoistisch motiviert handeln würden. Auf den ersten Blick steht diese Annahme im Widerspruch zum Konzept der PSM, da bei diesem bestimmten Personengruppen eine gemeinsame Form der Handlungsmotivation unterstellt wird. Wenn man aber berücksichtigt, dass ein zentrales Merkmal der PSM ja gerade darin besteht, dass diese vor allem in spezifischen gesellschaftlichen Organisationen (oder auch Institutionen) auftritt, nähert man sich Parsons Sichtweise der institutionell begründeten Orientierungsdifferenzen wieder an: Die Personen, die im Sinne der PSM motiviert sind, arbeiten in Tätigkeitsfeldern, welche in ihren Zielsetzungen meist auf einen gesellschaftlichen Zentralwert bzw. auf das Gemeinwohl hin ausgerichtet sind. Die PSM stellt im Sinne der Selbstbestimmungstheorie der Motivation dann nichts anderes dar als die internalisierten Zielsetzungen und Werte dieser Tätigkeitsfelder, also der Organisation oder Institution, der die Personen jeweils angehören und die damit deren Handlungsmotivation (mit)bestimmen.

Mit diesem Rückgriff auf Parsons werden die allgemeinen motivationstheoretischen Überlegungen abgeschlossen. Im folgenden Kapitel werden nun die dargestellten Theorieansätze in Bezug zur Forschungsfrage dieser Arbeit gesetzt und unter Einbezug organisations- und professionstheoretischer Aspekte (vgl. Kapitel 3) das theoretische Modell spezifiziert. Aus diesem werden schließlich die zu testenden Forschungshypothesen abgeleitet.

4.2 Verknüpfung der theoretischen Ansätze und Ableitung der Forschungshypothesen

In der Diskussion um das sich wandelnde Governance-Regime der Hochschulen in Kapitel 2 dieser Arbeit wurde deutlich, dass sich sozialwissenschaftliche Professorinnen und Professoren an Universitäten (auch) in Deutschland mit der neuen Steuerungslogik

des New Public Managements konfrontiert sehen. Eines der zentralen Elemente des New Public Managements stellt dabei der Governance-Mechanismus des Wettbewerbs bzw. des Konkurrenzdrucks dar. Dieser verzeichnet in den letzten Jahren einen sichtbaren Bedeutungszuwachs, der sich beispielsweise in der zunehmenden Relevanz von Hochschulrankings, dem Wettbewerb zwischen Universitäten im Rahmen der Exzellenzinitiative oder eben auch durch die Einführung indikatorenbasierter Steuerungs- und Anreizverfahren äußert. Der hochschulinterne Einsatz dieser Verfahren (vgl. Kapitel 2.2) bzw. der dabei verwendeten Forschungsindikatoren (vgl. Kapitel 2.3) ist nun Gegenstand der empirischen Analysen dieser Arbeit. Hierbei wird der Frage nachgegangen, inwieweit der Einsatz der Verfahren in den Sozialwissenschaften mittlerweile verbreitet ist und wie diese jeweils ausgestaltet werden, das heißt, welche Forschungsindikatoren mit welcher Gewichtung verwendet werden.

Neben diesen rein deskriptiven Analysen liegt der Fokus des Forschungsinteresses auf der Überprüfung möglicher Zusammenhänge zwischen dem Einsatz der Forschungsindikatoren in den Steuerungs- und Anreizverfahren und dem Forschungsverhalten der sozialwissenschaftlichen Professorinnen und Professoren. Hierzu werden jeweils die in den Verfahren verwendeten Forschungsindikatoren und die dazu korrespondierenden Forschungstätigkeiten (Stellenwert der Forschungstätigkeit in der eigenen Forschungsarbeit als Proxy für das Forschungsverhalten; vgl. Kapitel 5.2.1) in den Blick genommen und geprüft, welche Zusammenhänge sich hinsichtlich des Einsatzes bzw. Nicht-Einsatzes der Indikatoren ergeben. Zur Analyse werden hierbei die folgenden gängigen Forschungsindikatoren (vgl. Kapitel 2.3.2) herangezogen:

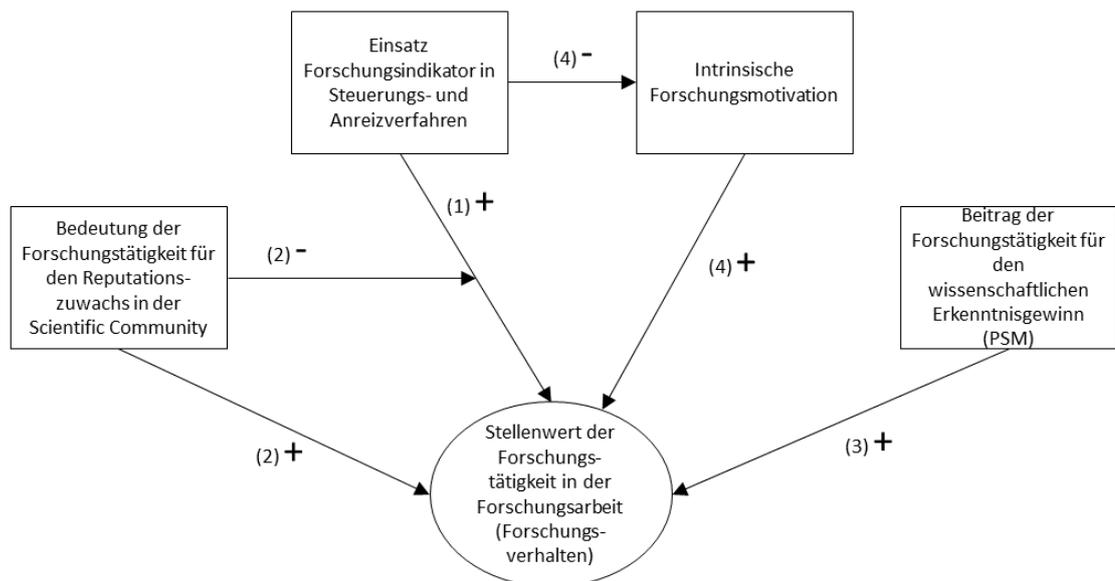
- Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren
- Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren
- Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden etc.)
- Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals
- Anzahl betreuter Promotionen
- Höhe der eingeworbenen Drittmittel der DFG
- Höhe der sonstigen eingeworbenen Drittmittel

Daneben soll unter Rückgriff auf die professionstheoretischen Überlegungen aus Kapitel 3 die Rolle der fachspezifischen Scientific Community beleuchtet werden. Hierzu wird überprüft, ob ein Zusammenhang zwischen der (eingeschätzten) Bedeutung einer Forschungstätigkeit (bzw. deren Output) für den Reputationszuwachs innerhalb der Scientific Community und dem Stellenwert der Tätigkeit in der Forschungsarbeit der Professorinnen und Professoren nachzuweisen ist.

Eine weitere in das theoretische Modell integrierte Annahme betrifft die vermutete Kollektivorientierung der akademischen Profession am ‚Zentralwert‘ des wissenschaftlichen Erkenntnisgewinns, welche als internalisierte Form der Motivation (PSM) in Bezug auf die einzelnen Forschungstätigkeiten wirksam werden kann. Zudem soll die Rolle der intrinsischen Forschungsmotivation und der von der *Motivation Crowding Theory* postulierte mögliche Verdrängungseffekt der intrinsischen Motivation empirisch überprüft werden.

Das theoretische Gesamtmodell mit den vermuteten positiven und negativen Zusammenhängen zwischen den einzelnen Elementen ist in Abbildung 2 dargestellt und wird nun im Folgenden im Detail erläutert. Danach werden darauf aufbauend die konkreten Forschungshypothesen formuliert.

Abbildung 2: *Theoretisches Modell*



Anmerkung. Das Modell ist tätigkeitsspezifisch (z.B. Forschungstätigkeit *Publizieren in Zeitschriften mit Peer-Review-Verfahren* oder *Betreuung von Promotionen*) zu verstehen. Die unabhängige Variable *Einsatz Forschungsindikator in Steuerungs- und Anreizverfahren* bezieht sich damit auf den mit der jeweiligen Tätigkeit korrespondierenden Forschungsindikator (z.B. Indikator *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* oder *Anzahl betreuter Promotionen*). Die intrinsische Forschungsmotivation wird unabhängig von den spezifischen Forschungstätigkeiten generell erfasst.

Quelle: eigene Darstellung

(1)²⁸ Der Einsatz von Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren wird zunächst als Maßnahme verstanden, die als external regulierte Form

²⁸ Diese und die nachfolgenden Nummerierungen mit den zugehörigen Erläuterungen beziehen sich auf die entsprechenden Nummerierungen in Abbildung 2.

der Motivation (vgl. Kapitel 4.1.1) bei den Professorinnen und Professoren wirksam wird und so das Forschungsverhalten beeinflussen kann. Beispielsweise kann innerhalb eines Steuerungsverfahrens durch den Einsatz des Forschungsindikators *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* der entsprechende Publikationsoutput in Form von Leistungszulagen belohnt werden. Dies kann als externer Anreiz dazu führen, dass die Publikationsstrategie der Professorinnen und Professoren entsprechend angepasst und versucht wird, vermehrt Artikel in peer-reviewed Formaten zu platzieren. Dabei muss die äußere Anreizwirkung nicht zwingend nur auf monetären Anreizen beruhen. Eine Anreizwirkung kann auch aus dem in Kapitel 3.3 beschriebenen Vertrauensverlust in die akademische Profession resultieren, der die Professorinnen und Professoren unter Legitimationsdruck setzt. Dieser Legitimationsdruck, ebenfalls verstanden als external regulierte Form der Motivation, kann dazu führen, dass seitens der Professorinnen und Professoren versucht wird, den in den Steuerungs- und Anreizverfahren gesetzten Anforderungen bestmöglich zu entsprechen.

Angenommen wird also eine positive Korrelation zwischen dem Einsatz eines Forschungsindikators in Steuerungs- und Anreizverfahren und dem (eingeschätzten) Stellenwert der korrespondierenden Forschungstätigkeit (z. B. Publizieren in Zeitschriften mit Peer-Review-Verfahren) in der Forschungsarbeit der Professorinnen und Professoren (Forschungsverhalten). Allerdings wird aufgrund der bereits dargestellten Spezifika der Organisation Universität (vgl. die Beschreibung von Universitäten als lose gekoppelte Systeme, Profibürokratien und organisierte Anarchien in Kapitel 3.1) und der daraus resultierenden Steuerungsproblematik (vgl. Kapitel 3.3) davon ausgegangen, dass die Steuerungsmöglichkeiten begrenzter Natur sind und sich damit nur ein schwacher Zusammenhang zwischen dem Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren und dem Forschungsverhalten ergibt. Zudem sind die Besonderheiten der akademischen Profession (vgl. Kapitel 3.2), insbesondere der hohe Grad an gewährter Handlungsautonomie und der hohe Stellenwert der Scientific Community als wesentlicher Bezugspunkt für die Professorinnen und Professoren, zu berücksichtigen. Diese Spezifika der akademischen Profession können die Wirkung hochschulinterner Steuerungsversuche zusätzlich einschränken. Im Folgenden wird dies nun näher erläutert.

(2) Hinsichtlich der fachspezifischen Scientific Community wird angenommen, dass diese als zentraler Bezugspunkt der Professorinnen und Professoren und als Produktionsstätte von Reputation in einem starken Zusammenhang mit dem Forschungsverhalten steht, da der Gewinn von Reputation innerhalb der eigenen Scientific Community als einer der wesentlichen Anreize für das Engagement im wissenschaftlichen Tätigkeitsfeld angesehen werden kann (vgl. Enders 1998, S. 72–73; Kühl 2015b, S. 804; Minssen et al. 2003, S. 36; vgl. Kapitel 3.2). Im Sinne der Selbstbestimmungstheorie der Motivation

wird der Reputationsgewinn in der Scientific Community als eine extrinsische Form der Motivation aufgefasst, die aber zu einem hohen Grad internalisiert ist und als *identifiziert regulierte* Form der Motivation ein hohes Maß an Selbstbestimmtheit bzw. Autonomie aufweist (vgl. Kapitel 4.1.1). Es wird daher davon ausgegangen, dass die mit einem möglichen Reputationsgewinn verknüpften Forschungstätigkeiten auch mit den persönlichen Zielen und Werten der Professorinnen und Professoren übereinstimmen und mit deren Selbstkonzept vereinbar sind (vgl. Deci und Ryan 1993, S. 228). Wenn nun z. B. das Publizieren in *high ranked journals* mit einem Reputationszuwachs in der Scientific Community ‚belohnt‘ wird, wäre entsprechend obiger Annahmen ein erhöhtes Engagement hinsichtlich des Verfassens von Artikeln, welche sich in den betreffenden Zeitschriften veröffentlichen lassen, zu erwarten. Die eingeschätzte Bedeutung einer Forschungstätigkeit für den Reputationszuwachs in der Scientific Community sollte also positiv mit dem Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren korrelieren.

Weiter wird angenommen, dass der Zusammenhang zwischen indikatorenbasierten Steuerungsversuchen und dem Stellenwert der korrespondierenden Forschungstätigkeiten durch die eingeschätzte Bedeutung der jeweiligen Forschungstätigkeit für den Reputationszuwachs in folgender Weise moderiert wird: Wird einer Forschungstätigkeit eine hohe Bedeutung für den Reputationszuwachs in der Scientific Community zugesprochen, orientiert sich das Forschungsverhalten der Professorinnen und Professoren primär an der Logik der Scientific Community und die mit dem möglichen Reputationszuwachs verbundene Forschungstätigkeit nimmt in der eigenen Forschungsarbeit einen entsprechend hohen Stellenwert ein. Ein (zusätzlicher) Steuerungseffekt über den Einsatz eines Forschungsindikators wäre in dieser Situation nur gering ausgeprägt oder überhaupt nicht vorhanden, da sich das Engagement hinsichtlich dieser Forschungstätigkeit (vollständig) über die starke Orientierung an der Scientific Community erklären lassen würde. Mit anderen Worten, wäre der Einsatz von Steuerungsverfahren in einem solchen Fall mehr oder weniger überflüssig, da die Professorinnen und Professoren das gewünschte Verhalten schon unabhängig von hochschulinternen Steuerungsversuchen allein aufgrund ihrer Orientierung an den Maßstäben der Scientific Community zeigen. Spielt eine Forschungstätigkeit für den Reputationszuwachs dagegen nur eine untergeordnete Rolle, kann der Einsatz des Forschungsindikators als externer Anreiz motivational wirksam werden und so zu einem erhöhten Stellenwert der entsprechenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren beitragen. Aber auch in diesem Fall ist von einer nur begrenzten Steuerungswirkung auszugehen, da aufgrund der hohen Relevanz der Scientific Community für die akademische Profes-

sion die begrenzten zeitlichen Ressourcen der Professorinnen und Professoren vermutlich eher für solche Forschungstätigkeiten aufgewendet werden, die mehr den Anforderungen und Erwartungen der Scientific Community entsprechen.

(3) Wie in Kapitel 3.2 dargestellt wurde, kann aus professionstheoretischer Sichtweise die kollektive Orientierung an zentralen gesellschaftlichen Werten wie Gesundheit oder Bildung als ein wesentliches Bestimmungsmerkmal von Professionen aufgefasst werden. Im Zuge der vorangegangenen Ausführungen zum Konzept der *Public Service Motivation* (PSM) wurde ebenfalls bereits gezeigt, dass die PSM als ein auf der Mikroebene des Handelnden angesiedeltes motivationstheoretisches Pendant zur Kollektivorientierung auf der Mesoebene der Profession verstanden werden kann (vgl. Kapitel 4.1.3). Mit dem Blickwinkel der Selbstbestimmungstheorie wird die PSM damit als eine stark internalisierte extrinsische Form der Motivation aufgefasst. Die professionelle Kollektivorientierung an einen gesellschaftlichen Zentralwert ist also soweit verinnerlicht, dass die damit verbundenen Handlungen weitestgehend autonom motiviert ablaufen und keiner externen Anreize bedürfen. Auf dem Selbstbestimmungs-Kontinuum der Motivationstypen von Ryan und Deci (2000) lässt sich die PSM damit je nach Internalisierungsgrad zwischen dem Typus der identifizierten und der integrierten Verhaltensregulation verorten (vgl. ebd., S. 72).

Für die akademische Profession lässt sich nun der wissenschaftliche Erkenntnisgewinn als jener gesellschaftliche Zentralwert begreifen, auf den die Forschungsarbeit der Wissenschaft ausgerichtet ist. Den obigen Annahmen folgend wird davon ausgegangen, dass das Streben nach wissenschaftlichem Erkenntnisgewinn soweit internalisiert ist, dass die Professorinnen und Professoren ‚von sich aus‘ bestrebt sind – also autonom motiviert sind – einen Beitrag zum Fortschreiten des wissenschaftlichen Erkenntnisgewinns zu leisten. Je höher der Beitrag einer Forschungstätigkeit (bzw. deren Ergebnis) für den wissenschaftlichen Erkenntnisgewinn eingeschätzt wird, desto stärker sollte daher die motivationale Wirkung der PSM hinsichtlich dieser Forschungstätigkeit ausgeprägt sein. Je höher wiederum die motivationale Wirkung der PSM ausfällt, desto ein höheres Engagement der Professorinnen und Professoren hinsichtlich dieser Forschungstätigkeit sollte sich zeigen. Wenn beispielsweise eine Professorin das Durchführen eines privatwirtschaftlichen Drittmittelprojektes als wenig förderlich für weiterführende wissenschaftliche Erkenntnisse ansieht, wäre eine nur eingeschränkte autonom wirkende Handlungsmotivation hinsichtlich der Durchführung von diesem und ähnlichen Drittmittelprojekten zu erwarten. Solche Drittmittelforschung sollte in ihrer Forschungsarbeit dann einen entsprechend geringeren Stellenwert einnehmen als bei einem Forscher, der den Drittmittelprojekten einen hohen Beitrag für den Erkenntnisgewinn zuspricht.

Bevor im nächsten Abschnitt die intrinsische Forschungsmotivation in den Blick genommen wird, soll hier nochmals kurz daran erinnert werden, dass die PSM, wie andere autonome Formen der Motivation auch, vom Verdrängungseffekt der Motivation betroffen sein kann (vgl. Kapitel 4.1.3). Da die PSM in dieser Arbeit aber nicht direkt erhoben, sondern mit dem eingeschätzten Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn lediglich die vermutete dahinter stehende kollektive Handlungsorientierung der akademischen Profession erfasst wird, kann der Verdrängungseffekt in Bezug auf die PSM hier nicht überprüft werden. Theoretisch angenommen und empirisch überprüft wird daher lediglich, ob zwischen dem eingeschätzten Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn und dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren eine positive Korrelation besteht.

(4) Die intrinsische Motivation wird im Sinne der Selbstbestimmungstheorie als vollkommen autonome Form der Motivation aufgefasst, die keinerlei externer Anreize bedarf und allein auf dem Interesse an bzw. der Freude mit der Tätigkeit an sich beruht (vgl. Kapitel 4.1.1). Es wird nun von einer positiven Korrelation zwischen der intrinsischen Forschungsmotivation der Professorinnen und Professoren und dem eingeschätzten Stellenwert der Forschungstätigkeiten in der eigenen Forschungsarbeit ausgegangen. Dabei wird die intrinsische Forschungsmotivation aus methodischen Gründen (vgl. Kapitel 5.2.1) nicht differenziert nach den einzelnen Forschungstätigkeiten, sondern für die Forschungsarbeit generell erfasst. Angenommen wird also, dass die intrinsische Forschungsmotivation generell mit allen Aspekten der Forschungsarbeit positiv korreliert ist.

Entsprechend der Annahmen der *Motivation Crowding Theory* wird zudem davon ausgegangen, dass die intrinsische Forschungsmotivation durch den Einsatz indikatorenbasierter Steuerungs- und Anreizverfahren vom Verdrängungseffekt der Motivation betroffen sein kann. Dies wäre nach der *Motivation Crowding Theory* dann der Fall, wenn der Einsatz der Verfahren als kontrollierend wahrgenommen und dadurch das eigene Autonomie- und Kompetenzerleben eingeschränkt wird (vgl. Kapitel 4.1.2). Bei der Beschreibung der Steuerungs- und Anreizverfahren (vgl. Kapitel 2.2) und der hierbei verwendeten Forschungsindikatoren (vgl. Kapitel 2.3) wurde bereits auf die oft kritische Haltung von Professorinnen und Professoren gegenüber den Verfahren und dem generellen Versuch der Messung von Forschungsleistung durch Forschungsindikatoren hingewiesen. Diese eher ablehnende Haltung spiegelt sich teilweise auch in der bisherigen Empirie zur Thematik wider (vgl. Kapitel 2.4). Die teils geringe Akzeptanz der akademischen Profession gegenüber den Verfahren, verbunden mit dem Zweifel an der generellen Eignung von Forschungsindikatoren zur Erfassung der Forschungsleistung (z. B. Vernach-

lässigung qualitativer Aspekte), legen nahe, dass der Einsatz indikatorenbasierter Steuerungs- und Anreizverfahren zumindest teilweise in dem genannten kontrollierenden Sinne wahrgenommen wird. Im Rahmen der Hypothesentestung soll daher auch überprüft werden, ob ein Verdrängungseffekt der Motivation nachgewiesen werden kann. Theoretisch angenommen wird also eine negative Korrelation zwischen dem Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren und der intrinsischen Forschungsmotivation der Professorinnen und Professoren.

Aufbauend auf den bisherigen Überlegungen werden nun im Folgenden die zu testenden Forschungshypothesen abgeleitet. Auch wenn die vorangegangene Darstellung des theoretischen Modells teilweise eindeutige Wirkrichtungen unterstellt hat, werden die Hypothesen dabei bewusst nur als positive oder negative Zusammenhänge zwischen den Variablen formuliert. Der Nachweis von Kausalitäten ist nicht möglich, da die nachfolgenden Analysen lediglich auf einer Querschnitterhebung basieren. Im Rahmen dieser Arbeit können die einzelnen Wirkrichtungen daher nur theoretisch vermutet, aber nicht belegt werden (vgl. zur Kausalität auch das nachfolgende Kapitel 5.1 zum Forschungsdesign der vorliegenden Studie). Statistisch nachweisen lässt sich dagegen die Signifikanz der vermuteten Korrelationen zwischen den Variablen. Diese zu testenden Zusammenhänge lassen sich in Form der folgenden sechs Hypothesen zusammenfassen:

H1: Der Einsatz eines Forschungsindikators in hochschulinternen Steuerungs- und Anreizverfahren korreliert positiv mit dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

H2.1: Je höher die eingeschätzte Bedeutung einer Forschungstätigkeit für den Reputationszuwachs in der Scientific Community ist, desto höher ist der eingeschätzte Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

H2.2: Die Höhe der eingeschätzten Bedeutung einer Forschungstätigkeit für den Reputationszuwachs in der Scientific Community moderiert den positiven Zusammenhang zwischen dem Einsatz eines Forschungsindikators in hochschulinternen Steuerungs- und Anreizverfahren und dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren in folgender Weise: Die Stärke des Zusammenhangs ist bei einer hoch eingeschätzten Bedeutung für den Reputationszuwachs geringer als bei einer niedrig eingeschätzten Bedeutung für den Reputationszuwachs.

H3: Je höher der eingeschätzte Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn ist, desto höher ist der eingeschätzte Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

H4.1: Je höher die intrinsische Forschungsmotivation ist, desto höher ist der eingeschätzte Stellenwert einer Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

H4.2: Der Einsatz eines Forschungsindikators in hochschulinternen Steuerungs- und Anreizverfahren korreliert negativ mit der intrinsischen Forschungsmotivation.

Um die dargestellten Forschungshypothesen zu überprüfen, werden im Folgenden die Ergebnisse der quantitativen Online-Befragung von sozialwissenschaftlichen Professorinnen und Professoren in Deutschland herangezogen. Hierzu werden zunächst im Detail das Forschungsdesign, die Entwicklung des Erhebungsinstrumentes und das Vorgehen bei der Datenerhebung beschrieben. Im Anschluss daran erfolgt die Ergebnisdarstellung der empirischen Analyse.

5. Aufbau und Ablauf der Untersuchung

5.1 Forschungsdesign

Die Studie folgt einer quantitativ-deduktiven Vorgehensweise. Um die abgeleiteten Forschungshypothesen zu überprüfen, wurde eine standardisierte Fragebogenstudie im Querschnitt durchgeführt, wobei die Testung der Hypothesen mittels multivariater Analyseverfahren (multiple Regressionsmodelle) erfolgt. Ein längsschnittliches Forschungsdesign wäre im Hinblick auf das Erkenntnisinteresse der Studie wünschenswert gewesen. Die Umsetzung eines solchen war allerdings nicht zu realisieren, da die hier untersuchten Steuerungsverfahren größtenteils bereits seit einigen Jahren implementiert sind (Bogumil et al. 2013, S. 77–87). Eine Vorher-Nachher-Messung bezüglich der Effekte des Einsatzes der Steuerungs- und Anreizverfahren war daher nicht möglich. Damit sieht sich die vorliegende Studie mit dem Problem konfrontiert, dass Kausalitäten durch das querschnittliche Studiendesign streng genommen nicht nachgewiesen werden können (vgl. Opp 2010; Steinhage und Blossfeld 1999). Die Ergebnisse sollten daher mit entsprechender Vorsicht interpretiert werden. Auch wenn sich diese Studie auf die statistische Prüfung möglicher signifikanter Zusammenhänge zwischen den untersuchten Variablen beschränken muss und dabei die Richtung des Zusammenhangs aus methodischen Gründen unerklärt bleibt, soll hier dennoch nicht gänzlich auf Überlegungen zur Kausalität verzichtet werden. Sofern es das theoretische Modell nahelegt und die Annahmen plausibel erscheinen, wird in den nachfolgenden Kapiteln also auch die (vermutete) Wirkrichtung thematisiert werden, wohlwissend, dass diese Aussagen nur hypothetischen Charakter haben können (vgl. zum Umgang mit derselben Problematik auch Krempkow und Schulz 2012).

Die Studie beschränkt sich in der Untersuchungspopulation auf die Grundgesamtheit der sozialwissenschaftlichen Professorinnen und Professoren an Universitäten in Deutschland. Dies hat zum einen den pragmatischen Grund, dass für diese Population bereits auf eine erste Recherche der Untersuchungseinheiten (Namen, E-Mail) aus einem anderen Forschungsprojekt zurückgegriffen werden konnte (vgl. zur detaillierten Beschreibung der Grundgesamtheit und Datenrecherche Kapitel 5.3). Zum anderen besteht seitens des Autors auch ein persönliches Erkenntnisinteresse aufgrund der eigenen Tätigkeit im sozialwissenschaftlichen Feld und der damit verbundenen inhaltlichen Nähe zu (anderen) sozialwissenschaftlichen Disziplinen. Nicht zuletzt erlaubt die Betrachtung der verschiedenen Fächer und der damit verbundenen unterschiedlichen Fachkulturen eine vergleichende Perspektive innerhalb der Sozialwissenschaften. Im Einzelnen werden unter dem Begriff Sozialwissenschaften hier die Fächer Erziehungswissenschaften inklusive Sonderpädagogik, Politikwissenschaften, Psychologie, Sozialwesen, Soziologie und Wirtschaftswissenschaften verstanden. Diese Definition weicht

zwar von der des Statistischen Bundesamtes ab, welche die Fächer Erziehungswissenschaften einschließlich Sonderpädagogik sowie Psychologie der Fächergruppe Sprach- und Kulturwissenschaften zuordnet (Statistisches Bundesamt 2015, S. 298–299), entspricht aber der Fachsystematik der Deutschen Forschungsgemeinschaft (Deutsche Forschungsgemeinschaft 2012). Angestrebt wurde eine Vollerhebung aller sozialwissenschaftlicher Professorinnen und Professoren. Entsprechend wurden alle Einheiten der Grundgesamtheit zur Befragung eingeladen, sofern dies in Anbetracht von eventuellen Recherchefehlern (z. B. keine Aufnahme in die Stichprobe aufgrund falscher fachlicher Zuordnung) und anderen Fehlerquellen (z. B. keine gültige E-Mail-Adresse) zu realisieren war.

Die Befragung selbst wurde als Online-Befragung konzipiert. Postalischen Befragungen wird zwar (immer noch) eine höhere Rücklaufquote zugeschrieben (Shih und Fan 2008), Online-Befragungen bieten jedoch den nicht zu unterschätzenden Vorteil einer dynamischen Fragebogenführung. Während bei klassischen Paper & Pencil-Befragungen komplexere Filterführungen mit teils umständlichen Ausfüllhinweisen und Sprüngen im Fragebogen verbunden sind, lassen sich diese bei Online-Fragebögen komfortabel durch das Aus- und Einblenden bestimmter Fragen und dem automatischen Überspringen ganzer Seiten umsetzen. Dies kann sowohl die Datenqualität (keine Fehler beim Ausfüllen möglich) als auch die Teilnahmemotivation während des Ausfüllens des Fragebogens und damit die Beendigungsquote positiv beeinflussen. Für den in dieser Studie verwendeten Fragebogen, bei welchem mit umfangreichen Filterführungen gearbeitet wurde, erwies sich die Durchführung einer Online-Befragung daher als geeignete Lösung. Die Entwicklung des Erhebungsinstrumentes und die Umsetzung als Online-Fragebogen sind Teil des nun folgenden Kapitels.

5.2 Entwicklung des Erhebungsinstrumentes

5.2.1 Operationalisierung der Variablen

Im Folgenden wird zunächst dargestellt, in welcher Form die für das theoretisch hergeleitete Modell bzw. für die Hypothesentests relevanten Variablen erhoben wurden. Danach folgt eine kurze Erläuterung zu den übrigen im Fragebogen erhobenen Variablen. Der genaue Wortlaut der einzelnen Fragen und Antwortoptionen kann in der finalen Version des Fragebogens, welcher im Anhang I dieser Arbeit zu finden ist, nachgelesen werden. In diesem Kapitel wird sich daher auf die generelle Beschreibung der Fragetypen und der verwendeten Antwortoptionen bzw. Skalen beschränkt.

Die **unabhängige Variable** des theoretischen Modells stellt der **Einsatz von Forschungsindikatoren in hochschulinternen Steuerungs- oder Anreizverfahren** dar. Hierzu wurde für eine Reihe von Forschungsindikatoren erhoben, ob diese generell an

der Fakultät der Befragten in Steuerungs- oder Anreizverfahren zum Einsatz kommen (Antwortoptionen: *Indikator wird eingesetzt, Indikator wird nicht eingesetzt, weiß nicht*). Die Auswahl der zu erfragenden Forschungsindikatoren erfolgte dabei nach intensiver Literaturrecherche vor allem auf Basis anderer relevanter Studien zum Themenbereich der indikatorenbasierten Steuerung bzw. der Messung von Forschungsleistung (Bogumil und Heinze 2009b; Jaeger et al. 2006; Krempkow et al. 2013; Schröder 2004; Sieweke et al. 2014; Wollersheim et al. 2014). Zudem hatten die Befragten die Möglichkeit, weitere Indikatoren in einem Freitextfeld zu ergänzen.²⁹ Aus der indikatorenspezifischen Abfrage resultiert – nach entsprechender Recodierung – je abgefragtem Indikator eine dichotome Variable mit den Ausprägungen $1 = \text{Indikator wird eingesetzt}$ und $0 = \text{Indikator wird nicht eingesetzt}$.

Für weitere deskriptive Analysen wurde zudem mittels einer fünfstufigen Ratingskala erhoben, wie stark nach Einschätzung der befragten Professorinnen und Professoren die jeweiligen Indikatoren innerhalb von hochschulinternen Steuerungsverfahren insgesamt gewichtet werden (1 *sehr geringe Gewichtung* – 5 *sehr hohe Gewichtung*). Die hierbei verwendete Rating-Skala sowie alle anderen im Fragebogen verwendeten Ratingskalen können die Einschätzungen der Befragten streng genommen nur auf ordinalem Skalenniveau erfassen, da die Abstände zwischen den Antwortmöglichkeiten von den Befragten nicht zwingend als äquidistant interpretiert werden müssen. Eine Prämisse für intervallskalierte Daten wäre somit verletzt. Auf Ordinalskalenniveau sind aber verschiedene mathematische Operationen wie die Mittelwertbildung und die Anwendung darauf basierender statistischer Verfahren nicht zulässig (vgl. Diekmann 2004, S. 250–255). Um in dieser Arbeit die Anwendbarkeit verschiedener Auswertungsverfahren dennoch gewährleisten zu können, wird in Übereinstimmung mit der gängigen sozialwissenschaftlichen Forschungspraxis für die genannten Skalen dennoch quasi-metrisches Skalenniveau angenommen. Die Annahme, dass die Befragten die Abstände zwischen den Antwortkategorien als gleich groß interpretieren, kann dabei als gerechtfertigt angesehen werden, wenn, wie im vorliegenden Fall, numerische Skalen mit lediglich verbalisierten Endpolen verwendet werden (vgl. Allerbeck 1978).

Analog zur Abfrage der eingesetzten Forschungsindikatoren wurde erfragt, welche Steuerungs- und Anreizverfahren an der Fakultät der Befragten zum Einsatz kommen (Antwortoptionen: *Verfahren wird eingesetzt, Verfahren wird nicht eingesetzt, weiß nicht*). Auch hier gründete die Auswahl der abzufragenden Verfahren auf entsprechender Literatur (hier vor allem Bogumil und Heinze 2009b). Erfragt wurde der Einsatz der

²⁹ Eine „sonstige“-Kategorie mit Freitextfeld stand bei allen Fragen zur Auswahl, wo dies inhaltlich sinnvoll erschien (siehe auch den Fragebogen im Anhang). Dies wird daher im weiteren Verlauf dieses Kapitels nicht jedes Mal eigens erwähnt.

folgenden fünf Verfahren: *Formelgebundene/indikatorenbasierte Mittelzuweisung zwischen Hochschulleitung und Fakultät bzw. zwischen Fakultät und Institut/Lehrstuhl, Zielvereinbarungen zwischen Hochschulleitung und Fakultät bzw. zwischen Fakultät und Institut/Lehrstuhl sowie Benchmarking verschiedener Lehrstühle innerhalb der Fakultät.* Der Einsatz der beiden Verfahren *Leistungszulagen im Rahmen der W-Besoldung sowie Bleibe- und Berufungsverhandlungen* wurde nicht erfragt, da diese bei annähernd allen (W-besoldeten) Professorinnen und Professoren zur Anwendung kommen.

Für die einzelnen eingesetzten Steuerungs- und Anreizverfahren wurde im Anschluss ermittelt, welche Indikatoren mit welcher Gewichtung in diesen zum Einsatz kommen. In Übereinstimmung mit der Gesamtabfrage am Anfang des Fragebogens wurde zur Einschätzung der Gewichtung der einzelnen Indikatoren in den jeweiligen Verfahren wieder eine fünfstufige Ratingskala herangezogen. Die Abfrage erfolgte dabei mit Hilfe verschiedener Ein- und Ausblendbedingungen, um den Befragten nur die für sie relevanten Verfahren und Indikatoren anzuzeigen. Die zugrundeliegende Fragebogenlogik wird im Detail noch im zugehörigen Teilkapitel (Kapitel 5.2.2) erläutert.

Die zweite zentrale theoretisch vermutete Einflussgröße auf das Forschungsverhalten der Professorinnen und Professoren stellt der **Reputationsgewinn** in der fachspezifischen **Scientific Community** dar, welcher mit den jeweiligen Forschungstätigkeiten bzw. den daraus resultierenden Outputs verknüpft ist. Um zu erfassen, inwieweit die einzelnen Forschungstätigkeiten nach Einschätzung der Professorinnen und Professoren zum Reputationszuwachs in der Scientific Community beitragen, wurde anhand einer fünfstufigen Ratingskala mit verbalisierten Endpolen (*1 sehr gering* und *5 sehr hoch*) die eingeschätzte Bedeutung der verschiedenen Forschungstätigkeiten bzw. derer Outputs (z. B. *Publikationen in Zeitschriften mit Peer-Review-Verfahren, Anzahl der Zitationen, Höhe der eingeworbenen Drittmittel der DFG etc.*) für eben diesen Reputationszuwachs erfragt.

Zur Erhebung der dritten zentralen theoretischen Einflussgröße, dem eingeschätzten **Beitrag für den wissenschaftlichen Erkenntnisgewinn** der jeweiligen Forschungstätigkeiten bzw. der entsprechenden Outputs, wurde in gleicher Weise verfahren und die Einschätzung der Befragten mit Hilfe einer fünfstufigen Ratingskala abgefragt.

Im Gegensatz zu den drei vorangegangenen Variablen wurde die vierte zentrale theoretisch begründete Einflussgröße, die **intrinsische Forschungsmotivation**, nicht für jeden einzelnen Aspekt der Forschungsarbeit separat abgefragt, sondern mit Hilfe einer Subskala der *Research Motivation Scale* (Deemer et al. 2010) insgesamt erfasst. Die Erhebung der intrinsischen Motivation hinsichtlich einzelner verschiedener Forschungstätigkeiten erscheint nicht zielführend, da beispielsweise die Unterscheidung zwischen dem *Publizieren in Zeitschriften mit Peer-Review-Verfahren* und dem *Verfassen anderer Publikationen* in Bezug auf die zugrundeliegende intrinsische Motivation

wenig sinnvoll ist: Beiden Tätigkeiten liegt der Schreibprozess zu Grunde und auch wenn sich Aufbau und Umfang beider Publikationsformate unterscheiden können, wird hier nicht davon ausgegangen, dass die zugrundeliegende intrinsische Motivation differenziert betrachtet bzw. erhoben werden kann. Vielmehr wird angenommen, dass die intrinsische Forschungsmotivation generell positiv mit dem Stellenwert, den die verschiedenen Forschungstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren einnehmen, korreliert (vgl. Kapitel 4.2).

Zur Erfassung des Konstruktes der intrinsischen Motivation existieren zwar eine Reihe von teils gut bewährten Instrumenten, die aber meist auf einen bestimmten Forschungskontext (z. B. das *Intrinsic Motivation Inventory* (Ryan 1982) für Laborexperimente, die *Work Tasks Motivation Scale for Teachers* (Fernet et al. 2008) oder die *Academic Motivation Scale* (Vallerand et al. 1992) zur Erfassung der Lehrmotivation) zugeschnitten sind. Die Skalen lassen sich daher nicht problemlos auf andere Forschungsfragen übertragen. Mit der *Research Motivation Scale* liegt eine Skala vor, die sich auf die akademische Forschungsmotivation bezieht und daher für den Kontext dieser Arbeit am geeignetsten erscheint. Die englischsprachige Originalskala (Deemer et al. 2010) besteht ursprünglich aus 20 Items, die ein dreifaktorielles Konstrukt der Forschungsmotivation (*Intrinsic Reward*, *Failure Avoidance* und *Extrinsic Reward*) repräsentieren sollen und auf einer fünfstufigen Likert-Skala³⁰ gemessen werden. In der Studie von Deemer et al. (2010) zeigte dabei die Subskala *Intrinsic Reward* mit einem Wert von Cronbachs Alpha von $\alpha = .90$ die höchste interne Konsistenz. Auch im Hinblick auf ergänzende Analysen von Deemer et al. (2012) zur *Research Motivation Scale* kann von einer generellen Eignung der Skala zur Erfassung der Forschungsmotivation ausgegangen werden.

In dieser Arbeit wurde zur Erhebung der intrinsischen Forschungsmotivation lediglich die neun Items umfassende Subskala *Intrinsic Reward* genutzt, da gemäß den theoretischen Überlegungen aus Kapitel 4.2 nur die intrinsische Motivation der Befragten Gegenstand des Forschungsinteresses war. Die Übersetzung der Items erfolgte mit der Methode der Rückübersetzung.³¹ Die fünfstufige Skala wurde beibehalten, allerdings wurde analog zu den anderen im Fragebogen verwendeten Skalen auf die Verbalisierung der Mittelkategorien verzichtet und nur die Endpole (1 *stimme überhaupt nicht zu*, 5 *stimme voll und ganz zu*) verbalisiert. Die Eignung der Skala für die vorliegende Studie wurde im Pretest (vgl. Kapitel 5.2.3) sowie im Rahmen der deskriptiven Analysen der Daten der Hauptbefragung (vgl. Kapitel 6.1.7) überprüft.

³⁰ 1 *strongly disagree*, 2 *disagree*, 3 *neutral*, 4 *agree* und 5 *strongly agree*

³¹ Die Items werden dabei zuerst in die Zielsprache und dann von einer anderen Person wieder zurück in die ursprüngliche Sprache übersetzt. Ein Abgleich der Rückübersetzung mit der Ursprungsversion erlaubt zumindest gravierende Übersetzungsfehler aufzudecken. Nichtsdestotrotz bleibt die Übersetzung von Befragungsinstrumenten fehlerbehaftet (vgl. Haas 2009).

Das Forschungsverhalten als **abhängige Variable** soll über den **subjektiv eingeschätzten Stellenwert**, den die abgefragten **Forschungstätigkeiten** in der Forschungsarbeit der Professorinnen und Professoren einnehmen, erfasst werden. Da manche Forschungstätigkeiten grundsätzlich mehr Zeit als andere beanspruchen (z. B. benötigt das Verfassen einer Monografie in der Regel mehr Zeit als das Verfassen eines Zeitschriftenaufsatzes) wurde bewusst auf eine Abfrage der absolut aufgebrauchten Zeit je Forschungstätigkeit verzichtet. Stattdessen wurde betont, dass die Abfrage auf den subjektiven Stellenwert der Forschungstätigkeiten in der eigenen Forschungsarbeit abzielt – also darauf, wie wichtig es den Befragten war, sich in den abgefragten Tätigkeitsfeldern zu engagieren. Als Referenzzeitraum für die Einschätzung des eigenen Forschungsverhaltens wurden drei Jahre gewählt. Damit soll der Einfluss von kürzeren Phasen mit speziellen Tätigkeitsschwerpunkten, beispielsweise Projektabschlussphasen oder Aufenthalte als Gastwissenschaftler, abgeschwächt und so das durchschnittliche, 'alltägliche' Forschungsverhalten erfasst werden. Die Einschätzung des Stellenwertes der verschiedenen Forschungstätigkeiten in der Forschungsarbeit der letzten drei Jahre erfolgte wieder mit Hilfe einer fünfstufigen Ratingskala (1 *sehr geringer Stellenwert* – 5 *sehr hoher Stellenwert*). Im Gegensatz zu den beiden Variablen *Bedeutung für den Reputationszuwachs in der Scientific Community* und *Beitrag für den wissenschaftlichen Erkenntnisgewinn* wurde bei der abhängigen Variablen das Item (*Anzahl der Zitationen*) nicht abgefragt, da für dieses Item keine entsprechende sinnvoll abzufragende direkte Forschungstätigkeit vorliegt.³² Für mögliche weiterführende Analysen ergänzt wurden dagegen die Items *Publizieren als Koautor / in Mehrautorenschaft*, *Betreiben von Grundlagenforschung* und *Betreiben von anwendungsbezogener Forschung*.³³

³² Die Tätigkeit des Zitierens an sich ist Bestandteil fast jeder wissenschaftlichen Publikationstätigkeit und kann nicht losgelöst von dieser betrachtet werden. Die Abfrage, welchen Stellenwert das Zitieren in der eigenen Forschungsarbeit eingenommen hat, würde daher eher erfassen, welchen Stellenwert das Publizieren in der Forschungsarbeit eingenommen hat.

³³ Hintergrund hierfür waren Annahmen hinsichtlich möglicher Nebeneffekte bzw. nichtintendierter Effekte durch den Einsatz neuer Steuerungsverfahren. So gibt es Überlegungen, dass anwendungsferne Grundlagenforschung zugunsten anwendungsbezogener Forschung verdrängt werden könnte (vgl. Schimank 2008, S. 632). Dem Item *Publizieren als Koautor / in Mehrautorenschaft* liegt die Annahme zugrunde, dass durch dieses Publikationsverhalten der Publikationsoutput und die Zitationsquote strategisch erhöht werden können. Der Einsatz von bibliometrischen bzw. auf dem Publikationsoutput basierenden Indikatoren in Steuerungsverfahren könnte daher zu einer Erhöhung von in Mehrautorenschaft verfassten Publikationen führen (vgl. Kieser 2013; Krempkow et al. 2013, S. 152–154). Da für diese drei Items keine korrespondierende unabhängige (Einsatz Forschungsindikator) bzw. keine korrespondierenden Kontrollvariablen (Reputation und Erkenntnisgewinn) erhoben wurden, wird auf multivariate Analysen hinsichtlich der Variablen *Publizieren als Koautor / in Mehrautorenschaft* und *Betreiben von Grundlagenforschung* bzw. *anwendungsbezogener Forschung* verzichtet. Bivariate Zusammenhangsanalysen wurden dennoch durchgeführt und werden in Kapitel 6.2.1 dargestellt. Die Analysen haben aber mehr explorativen Charakter.

Schließlich wurden im Fragebogen noch eine Reihe weiterer Variablen abgefragt, die entweder für deskriptive Analysen von Interesse waren, für die Filterführung im Fragebogen benötigt wurden, zum Abgleich der Stichprobe mit der Grundgesamtheit herangezogen wurden oder als Kontrollvariablen für die noch folgenden multivariaten Analysen dienen sollten. Im Einzelnen sind dies die folgenden Variablen:

- Fachzugehörigkeit (nominalskaliert)
- Position (nominalskaliert)
- Geschlecht (nominalskaliert)
- Alter, erfasst über das Geburtsjahr (metrisch)
- Qualifikation (Habilitation und / oder Juniorprofessur) zur Professur (nominalskaliert)
- Tätigkeitsdauer als Professorin oder Professor in Jahren (metrisch)
- Bundesland der Universität, an der die Befragten tätig sind (nominalskaliert)
- Trägerschaft der Universität (nominalskaliert)

Nachdem nun die Operationalisierung aller relevanten Variablen erläutert wurde, folgt im nächsten Kapitel eine knappe Darstellung der Umsetzung des erarbeiteten Fragenkatalogs in Form des Online-Fragebogens und der zugehörigen Fragebogenlogik.

5.2.2 Programmierung des Online-Fragebogens und Fragebogenlogik

Der Fragebogen wurde mit dem Umfragetool *Unipark EFS Survey* als Online-Version umgesetzt. Mit dieser Software war es möglich, insbesondere die zahlreichen selektiven Einblendbedingungen und Filterregeln zu definieren. Diese erlauben es, die Anzeige der Fragen und abgefragten Items auf die vorangegangenen Angaben der befragten Person anzupassen. Im Gegensatz zur Papierbefragung kann so die Länge des Fragebogens durch den Ausschluss bzw. das Überspringen nicht relevanter Fragen zum Teil deutlich reduziert werden. Im Folgenden werden nun die wichtigsten Einblendbedingungen und die Filterführung im Fragebogen näher erläutert, so dass in den später folgenden statistischen Analysen nachvollzogen werden kann, worauf die teils unterschiedlich hohen Fallzahlen je Frage zurückzuführen sind (z. B. Missings, da die Frage für den Befragten nicht relevant war und daher nicht abgefragt wurde).

Am Anfang des Fragebogens sollen die Befragten angeben, welche Forschungsindikatoren an ihrer Fakultät im Rahmen von Steuerungs- und Anreizverfahren Verwendung finden. In der Folgefrage nach der Gewichtung der Indikatoren in den Verfahren insgesamt werden dann nur diejenigen Indikatoren eingeblendet, bei denen die Befragten in der vorangegangenen Frage angaben, dass diese eingesetzt werden.

Die fünf folgenden Fragen bis einschließlich der Frage nach den an der Fakultät eingesetzten Steuerungs- und Anreizverfahren werden allen Befragten angezeigt. Die Auswahl der eingesetzten Verfahren (Antwortoption „Verfahren wird eingesetzt“) bestimmt nun, welche der folgenden Detailfragen zu den Steuerungs- und Anreizverfahren für die Befragten sichtbar sind: Gibt eine Teilnehmerin oder ein Teilnehmer beispielsweise an, dass an der Fakultät indikatorenbasierte Mittelzuweisungsverfahren eingesetzt werden, wird in der Folge die Frage eingeblendet, welche Indikatoren mit welcher Gewichtung in diesem Verfahren Verwendung finden. Die als Antwortoption zur Auswahl stehenden Indikatoren werden dabei durch die generelle Frage nach eingesetzten Forschungsindikatoren am Anfang des Fragebogens definiert.

Nicht abgefragt wird, ob die Verfahren *Leistungszulagen im Rahmen der W-Besoldung* sowie *Bleibe- und Berufungsverhandlungen* eingesetzt werden. Stattdessen wird die Detailfrage zu den eingesetzten Indikatoren im Rahmen der Leistungszulagen der W-Besoldung nur denjenigen Befragten gestellt, die bei der Frage nach der aktuellen Position zu Beginn des Fragebogens angaben, eine W-Professur inne zu haben. Die Detailfrage zu den eingesetzten Indikatoren in Bleibe- und Berufungsverhandlungen wird allen Befragten gestellt, sofern sie bei der generellen Frage nach eingesetzten Indikatoren am Anfang des Fragebogens mindestens einen Indikator ausgewählt haben.

Neben den hier erläuterten Besonderheiten wurden in der Online-Version noch weitere Ein- und Ausblendbedingungen definiert (z. B. Einblenden der unter „Sonstige“ eingetragenen Indikatoren in den Folgefragen), die für das Verständnis der Fragebogenlogik aber nicht zwingend erforderlich sind und auf welche daher hier nicht weiter eingegangen wird.

Auf das Erzwingen von Antworten (Antwortpflicht) wurde im Fragebogen verzichtet. Die Befragten konnten also Fragen, die sie nicht beantworten konnten oder wollten, überspringen. Dies hat natürlich zur Folge, dass in jedem Datensatz bei einzelnen Fragen fehlende Werte auftreten können. Im Hinblick auf den ‚Befragungskomfort‘ (Fragen, die nicht beantwortet werden können oder nicht beantwortet werden wollen, müssen nicht jeweils durch das Anklicken einer „keine Angabe“-Kategorie bearbeitet werden) und der damit verbundenen Teilnahmemotivation wurde diese Form der Fragebogengestaltung aber der Variante mit durchgehender Antwortpflicht vorgezogen.

Inwieweit sich der entwickelte Fragebogen inhaltlich und im Hinblick auf seine Funktionalität als Erhebungsinstrument eignet, wurde in einem umfangreichen Pretest ermittelt. Dessen Durchführung und Ergebnisse sind Gegenstand des nächsten Kapitels.

5.2.3 Durchführung und Ergebnisse des Pretests

Der Pretest des Fragebogens wurde von Juli bis August 2015 durchgeführt. Hierzu wurden alle hauptberuflichen Professorinnen und Professoren des Faches Physik an staatlichen deutschen Universitäten, bei welchen eine gültige E-Mail-Adresse ermittelt werden konnte, befragt. Die Recherche der relevanten Professorinnen und Professoren erfolgte über die Webseiten der einzelnen Universitäten, wobei sich zur eindeutigen Abgrenzung des Faches Physik von anderen Fächern an der Fächergruppen-Systematik der Personalstatistik des Statistischen Bundesamtes orientiert wurde (Statistisches Bundesamt 2013, S. 296–297). Insgesamt ergab sich so eine Bruttostichprobengröße von $N = 1368$. Die Einladung zur Befragung erfolgte per E-Mail-Anschreiben mit personalisierter Anrede, aber – zur Sicherstellung der Anonymität der Befragten – einem anonymen Fragebogenlink. Nach zwei Wochen erhielten alle Befragten ein Erinnerungsschreiben mit erneuter Bitte um die Teilnahme an der Befragung.

In der Feldzeit vom 15.07. bis einschließlich 09.08.2015 nahmen insgesamt 201 Personen an der Befragung teil, von denen 156 den Fragebogen abgeschlossen haben. Die Rücklaufquote liegt damit bei 14.7 %. Von den zur Befragung eingeladenen Professorinnen und Professoren haben 11.4 % den Fragebogen vollständig beantwortet. Die Beendigungsquote nach Aufruf des Fragebogens beträgt entsprechend 77.6 %. In Anbetracht der stetig wachsenden Zahl an wissenschaftlichen und nicht-wissenschaftlichen Online-Befragungen (Zerback et al. 2009, S. 15) und einer möglicherweise damit verbundenen abnehmenden Teilnahmemotivation, sind diese Werte als durchaus zufriedenstellend einzustufen.

Der Fragebogen selbst wurde gegenüber der Version der Hauptbefragung für den Pretest leicht abgeändert. Neben kleineren Textunterschieden (z. B. *Physik* statt *Sozialwissenschaften*) wurde insbesondere die Frage nach dem Studienbereich bzw. -fach, in welchem die Befragten tätig sind, dem Fach Physik angepasst (z. B. Experimentelle Physik, Festkörperphysik etc.). Bei der Frage nach dem Stellenwert verschiedener Forschungstätigkeiten in der eigenen Forschungsarbeit wurde zudem das Item *Publizieren als Koautor / in Mehrautorenschaft* nicht abgefragt, da in der Physik das Publizieren mit (inter-)nationalen Koautoren als Standard angesehen werden kann (Schleich 2009, S. 91).

Hinsichtlich der Länge des Fragebogens ergaben sich beim Pretest keine Probleme. Die mittlere Bearbeitungszeit in Sekunden betrug $M = 690$ ($SD = 571$), also etwa elfeinhalb Minuten. Der Median liegt bei einer Bearbeitungszeit von ca. achteinhalb Minuten. Im Anschreiben zur Befragung wurde eine zu erwartende Ausfülldauer von 10 bis 15 Minuten angekündigt, was somit in den meisten Fällen der tatsächlichen Ausfülldauer entsprach. Die recht hohe Streuung bei der benötigten Bearbeitungszeit ist auf die Fragebogenlogik bzw. die Filter im Fragebogen zurückzuführen. Gab eine Befragte oder ein

Befragter beispielsweise am Anfang des Fragebogens an, dass keine der abgefragten Indikatoren in Steuerungs- oder Anreizverfahren zum Einsatz kommen, wurde der überwiegende Teil der darauf folgenden Fragen übersprungen oder ausgeblendet. Je mehr Indikatoren nach Angaben der Befragten in verschiedenen Steuerungs- und Anreizverfahren zum Einsatz kommen, desto mehr zusätzliche Fragen wurden dagegen selektiv eingeblendet, so dass sich die Bearbeitungsdauer in diesen Fällen entsprechend erhöhte. Aufgrund der hohen Streuung in der Bearbeitungszeit und der stark rechtsschiefen Verteilung wurde daher für die Hauptbefragung die Angabe zur erwartenden Ausfülldauer von 10 bis 15 Minuten beibehalten, auch wenn in einigen Fällen die Ausfülldauer (deutlich) unter dem angegebenen Zeitintervall lag.

Die Analyse der Abbrüche während des Ausfüllens zeigte keine auffälligen Häufungen auf bestimmten Seiten des Fragebogens. Erwartungsgemäß gab es die meisten Abbrüche am Anfang des Fragebogens auf der Startseite (5.2 %) sowie auf der auf die Eröffnungsfragen nach Fachbereich und Position folgenden dritten Fragebogenseite (8.0 %). Auch die bereits erwähnte recht akzeptable Beendigungsquote von 77.6 % legte keine notwendigen Änderungen an der Fragebogenstruktur nahe, so dass die Fragenreihenfolge für die Hauptbefragung beibehalten wurde.

Im Pretest fiel allerdings auf, dass die Freitextfelder der „Sonstigen“-Kategorien häufig dafür genutzt wurden, lehrbezogene Indikatoren bzw. lehrbezogene Aspekte der wissenschaftlichen Tätigkeit zu ergänzen. So wurden z. B. Lehrevaluationen oder die Anzahl betreuter Bachelor- und Masterarbeiten als zusätzliche Indikatoren genannt, die in indikatorenbasierten Steuerungsverfahren Verwendung finden. In der ursprünglichen Version des Fragebogens war bei der Frage nach den in Steuerungs- und Anreizverfahren verwendeten Indikatoren, sowie deren Gewichtung in diesen, lediglich von Indikatoren und nicht von *Forschungsindikatoren* die Rede. Die Fragen waren in dieser Form daher inhaltlich nicht eindeutig auf Forschungsindikatoren eingegrenzt. Für die Hauptbefragung wurde der Fragebogen daher angepasst und bei den entsprechenden Fragen von Forschungsindikatoren gesprochen.

Es wurden allerdings auch bei anderen Fragen, die in der Formulierung eindeutiger auf die Aspekte der Forschungsarbeit eingegrenzt waren (z. B. „Wie hoch schätzen Sie die Bedeutung der folgenden Aspekte der Forschungsarbeit für den Reputationszuwachs von Professorinnen und Professoren innerhalb Ihrer Scientific Community ein?“), lehrbezogene Aspekte genannt. Daher kann davon ausgegangen werden, dass – neben Verständnisproblemen oder ungenauem Lesen – auch bewusst die Lehre von einzelnen Befragten thematisiert wurde. Das Ausklammern der Lehre in der vorliegenden Arbeit wurde möglicherweise als falsch bzw. nicht zielführend angesehen. So merkte ein Befragter am Ende des Fragebogens an, dass ein wesentlicher Teil seiner wissenschaftlichen Arbeit unter anderem aus lehrbezogenen Tätigkeiten besteht, welche auch eine

wichtige Steuerungs- und Anreizfunktion haben würden. Ebenso ist es denkbar, dass bestimmte genannte Aspekte, wie z. B. die Betreuung von Bachelor- oder Masterarbeiten, aus Sicht mancher Befragter auch zur Forschungsarbeit gezählt werden können und eine klare Abgrenzung zur Lehre, ganz im Sinne des humboldtschen Bildungsideals, gar nicht immer möglich ist. Beide Punkte sind sicherlich diskussionswürdige Einwände. Da in dieser Arbeit aber primär Zusammenhangsanalysen mit dem Fokus auf einzelnen ausgewählten Indikatoren bzw. Forschungstätigkeiten durchgeführt werden sollen, ist eine erschöpfende Betrachtung aller relevanten Indikatoren und Aspekte der wissenschaftlichen Arbeit zwar wünschenswert, aber nicht zwingend notwendig.

Erste deskriptive Analysen der Pretestdaten ergaben, dass alle abgefragten Indikatoren sowie Steuerungs- und Anreizverfahren Verwendung finden und ausreichend unterschiedlich stark besetzte Gruppen (eingesetzt vs. nicht eingesetzt) vorliegen. Ausnahmen bilden hier der Indikator *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* sowie das Steuerungsverfahren *Benchmarking verschiedener Lehrstühle innerhalb der Fakultät*, bei welchen jeweils nur 9.6 % ($n = 15$) bzw. 6.0 % ($n = 9$) der Befragten angaben, dass diese eingesetzt werden. Für die Hauptbefragung wurden dennoch beide Items beibehalten, da für die Zielgruppe der Sozialwissenschaften zum einen erwartet wurde, dass Publikationen ohne Peer-Review-Verfahren (zumindest noch) einen höheren Stellenwert besitzen. Zum anderen wurde davon ausgegangen, dass bei großen sozialwissenschaftlichen Fächern mit einer oft hohen Anzahl an Lehrstühlen innerhalb einer Fakultät (z. B. Wirtschaftswissenschaften), das Benchmarking von Lehrstühlen als Steuerungsverfahren öfter zum Einsatz kommt, als dies im Fach Physik der Fall ist.

Die Antworten zu den Fragen nach der Bedeutung der verschiedenen Aspekte der Forschungsarbeit für den Reputationszuwachs in der Scientific Community, nach dem Beitrag der Aspekte für den wissenschaftlichen Erkenntnisgewinn sowie nach dem Stellenwert von diesen in der eigenen Forschungsarbeit zeigen eine ausreichend hohe Varianz. Allerdings ergeben sich bei den meisten der abgefragten Items signifikante Abweichungen von der Normalverteilung.

Für die neun Items der auf Deutsch übersetzten Version der *Research Motivation Scale* ergibt sich mit $\alpha = .88$ ein ausreichend hoher Wert von Cronbachs Alpha. Die korrigierten Item-Skala-Korrelationen (Trennschärfen) fallen mit Werten zwischen .63 und .75 ebenfalls ausreichend hoch aus, während die mittlere Inter-Item-Korrelation mit .51 etwas über dem angestrebten Akzeptanzbereich von 0.2 bis 0.4 liegt (Bortz und Döring 2003, S. 219–220). Auffällig ist zudem, dass alle Items eine stark linksschiefe Verteilung aufweisen (Skala von *1 stimme überhaupt nicht zu* bis *5 stimme voll und ganz zu*). Den einzelnen Aussagen wurde demnach tendenziell (voll und ganz) zugestimmt, während die niedrigen Skalenwerte nur schwach besetzt sind. Dieses Ergebnis erscheint aber

durchaus plausibel, da der akademischen Profession grundsätzlich ein hohes Maß an intrinsischer Motivation zugesprochen wird (Kieser 2010, S. 349; Schimank 2005, S. 152), was sich hier auch empirisch widerspiegelt. Als weiteres Ergebnis des Pretests bleibt damit festzuhalten, dass die verwendete Skala insgesamt als geeignetes Erhebungsinstrument angesehen werden kann und daher in dieser Form für die Durchführung der Hauptbefragung beibehalten wurde.

Abschließend sei noch darauf hingewiesen, dass zahlreiche, teils emotionale Reaktionen der Befragten des Pretests auf die Befragung erfolgten. Hierbei wird deutlich, wie stark die Thematik teilweise polarisiert, wobei der überwiegende Teil der Rückmeldungen eine sehr kritische Haltung gegenüber universitären Steuerungsversuchen oder Verfahren der Leistungsmessung offenbarte. Beispielhaft sollen hier drei Anmerkungen im Originalwortlaut wiedergegeben werden:

(...) Welch jämmerliche Gestalt im Rektorat will mir da irgendwelche Vorschriften machen? Wenn die Gängelung noch weitergeht, als sie in den letzten 15 Jahren seit Edelgard Bulmahn, CHE & Co schon gegangen ist, halte ich meine Pflichtvorlesungen ab, mache Forschung noch mit 1-2 Doktoranden, und dann kann sich die Unileitung mit ihren Lichtgestalten der Ruf, den sie herbei indizieren, herbei erzwingen, herbei bewerten will, in die Haare schmieren. Wie lautet der alte Spruch von Siemens aus den 1980er Jahren: Qualität kann man nicht erprüfen, Qualität muß man produzieren! Alle Stellschrauben, alle Mittel, alle Indikatoren, bei denen irgendwelche Erbsenzähler aus der Verwaltung und aus dem Rektorat versuchen, Wissenschaftler auf eine einzige Zahl, z.B. den h-Index, schön Excel-Tabellen-gerecht, zu reduzieren, sind grandios zum Scheitern verurteilt. Gebt den Wissenschaftlern wieder eine anständige Grundausstattung, gebt Ihnen wieder Freiheit, anstatt permanent Reformen durchs Dorf zu treiben, und die Indikatoren stellen sich als Konsequenz von selber ein!

Ich halte die meisten dieser anreizschaffenden Indikatoren - oder wie immer Sie es nennen mögen - für wahnsinnig. Das geht alles einher mit der Bildung à la Bolognese und der Ökonomisierung der Universitäten. Mit Humboldt hat das nicht mehr viel zu tun. Es wäre gut gewesen, wenn Sie stärker erfragt hätten, welches Wissenschaftssystem wünschenswert ist. Sie bleiben zu sehr bei der Erfragung des Ist. So bleibt mir für meine Fundamentalkritik nur dieser Kommentarkasten. So hoffe ich sehr, dass Ihre Recherche nicht zu einer weiteren Perfektionierung der indikatorengestützten Drangsalierung der Professoren führt. (...)

Nach Ausfüllen der Befragung frage ich mich, was denn das Ganze soll. Der Aufwand, der für die Indikatorenbasierte Steuerungs- und Anreizverfahren... und deren Erforschung betrieben wird, ist ein Irrsinn. Würde man die Zeit und das Geld einsparen, hätten alle mehr Zeit und Geld und wir würden mehr Zeit und mehr Geld für die Forschung verwenden! Die Zeit, die in Gremien und deren Vorbereitung verbraten wird ist, gemessen an deren Nutzen, eine klassische Fehlinvestition, die in einem weniger regulierten Umfeld aus Effizienzgründen niemals aufgewendet würde. Warum dies auch noch erforscht werden muss, ist mir ein Rätsel. Funktionieren wird diese Art von Steuerung zum Glück nie. Dazu sind Wissenschaftler zu Recht zu eigensinnig.

Mit diesen Kommentaren schließt die Beschreibung des Pretests und es folgt nun die Darstellung der Datenerhebung im Rahmen der Hauptbefragung.

5.3 Datenerhebung

5.3.1 Recherche der Stichprobe und Durchführung der Befragung

Die Grundgesamtheit der Studie bilden alle zum Erhebungszeitpunkt an staatlichen deutschen Universitäten (ohne Hochschulen für angewandte Wissenschaften) hauptberuflich tätigen Professorinnen und Professoren der folgenden sozialwissenschaftlichen Fächer: Erziehungswissenschaften inklusive Sonderpädagogik, Politikwissenschaften, Psychologie, Sozialwesen, Soziologie und Wirtschaftswissenschaften. Juniorprofessuren und Vertretungsprofessuren wurden für die Studie nicht berücksichtigt.

Zur Recherche der Stichprobe konnte auf bereits vorhandene E-Mail-Listen zurückgegriffen werden. Diese wurden im Rahmen des vom BMBF geförderten Forschungsprojektes „Qualität von sozialwissenschaftlichen Studiengängen und Reputation von Studienstandorten aus Sicht von Professorinnen und Professoren (QualRep)“ erstellt, welches von 2013 bis 2016 unter der Leitung von Professor Schmidt-Hertha am Institut für Erziehungswissenschaften der Eberhard Karls Universität Tübingen durchgeführt wurde (Schmidt-Hertha et al. 2015). Die Recherche erfolgte hierbei zum einen mittels der elektronischen Version des Hochschullehrerverzeichnisses des Deutschen Hochschulverbandes (Deutscher Hochschulverband 2013), zum anderen wurden die relevanten Professorinnen und Professoren manuell auf den Webseiten der Universitäten recherchiert bzw. die Einträge des Hochschullehrerverzeichnisses mit diesen abgeglichen und gegebenenfalls komplettiert. Da zwischen der Erstellung der Listen im Projekt *QualRep* und der hier vorgestellten Hauptbefragung eine zeitliche Differenz von etwa einem Jahr lag, wurden die vorhandenen Listen nochmals mittels einer Internetrecherche geprüft und überarbeitet. So konnten beispielsweise zwischenzeitlich emeritierte Professorinnen und Professoren identifiziert oder Neuberufungen erfasst werden. Zudem wurde im Rahmen der Nachrecherche für alle in den Listen aufgeführten Personen das Geschlecht ergänzt, um eine personalisierte Einladung zu ermöglichen.

Insgesamt ergab sich so eine Bruttostichprobe von 3714 Personen mit gültigen E-Mail-Adressen. Die Verteilung der Bruttostichprobe auf die einzelnen Fächer ist in Tabelle 1 dargestellt. Der Abgleich mit der Personalstatistik des Statistischen Bundesamtes zeigt, dass bei den Fächern Erziehungswissenschaften, Sozialwesen, Soziologie und Wirtschaftswissenschaften teils deutlich weniger Professorinnen und Professoren recherchiert werden konnten, als von der amtlichen Hochschulstatistik erfasst sind (Undercoverage). Zu berücksichtigen ist hier allerdings zum einen, dass in der Hochschulstatistik auch Juniorprofessuren, Gastprofessuren sowie Professuren an Theologischen Hochschulen enthalten sind. Zum anderen beruht die Personalstatistik des Statistischen Bundesamtes auf Selbstauskünften der Hochschulen, so dass auch in dieser Verzerrungen in den Daten nicht ausgeschlossen werden können. Die Differenz lässt sich damit

aber sicherlich nicht vollständig erklären. Vielmehr ist davon auszugehen, dass im Rahmen der Recherche ein Teil der relevanten Professorinnen und Professoren nicht erfasst werden konnte, weil sowohl das Hochschullehrerverzeichnis keinen Anspruch auf Vollständigkeit erheben kann, als auch die internetbasierte Recherche zwangsläufig lückenhaft bleiben muss (z. B. weil Webseiten nicht immer auf dem aktuellsten Stand oder unvollständig sind, Professuren nicht als solche gekennzeichnet werden etc.). Da das beschriebene Undercoverage aber auf keine bestimmte Gruppe von Professorinnen und Professoren beschränkt ist, ist bei den betroffenen Fächern keine systematische Verzerrung der Bruttostichprobe zu befürchten.

Tabelle 1: *Fächerverteilung der Bruttostichprobe*

Fach	N Bruttostichprobe	N Hochschulstatistik (inkl. Juniorprofessuren)	Anteil N Bruttostichprobe an N Hochschulstatistik in %
Erziehungswissenschaften / Sonderpädagogik	722	911	79.3
Politikwissenschaften	367	353	104.0
Psychologie	603	648	93.1
Sozialwesen	52	88	59.1
Soziologie / Sozialwissenschaften	387	488	79.3
Wirtschaftswissenschaften	1583	2033	77.9
Gesamt	3714	4521	82.1

Quelle: Statistisches Bundesamt 2015, S. 95; eigene Berechnungen

Für das Fach Psychologie ergibt sich im Gegensatz zu den oben genannten Fächern eine deutlich höhere Übereinstimmung mit der Personalstatistik. Ob dies auf eine vollständigere Erfassung im Rahmen der Recherche (z. B. aktuellere Webseiten) oder auf eine zu niedrige Anzahl innerhalb der amtlichen Personalstatistik zurückzuführen ist, lässt sich an dieser Stelle nicht klären.

In den Politikwissenschaften übersteigt die Anzahl der recherchierten Professuren die der in der Hochschulstatistik erfassten Professuren sogar. Dies ist wohl darauf zurückzuführen, dass die Einstufung einiger Professuren in eine bestimmte fachliche Disziplin teils schwer fällt bzw. nicht immer eindeutig möglich ist. Dies ist insbesondere dann der Fall, wenn die Fächer Politikwissenschaften und Soziologie einem gemeinsamen Institut zugeordnet sind. So wurden vermutlich Professuren, die in der Personalstatistik des Statistischen Bundesamtes als sozialwissenschaftliche (= Soziologie, Sozialkunde, Sozialwissenschaften allgemein und Demographie) Professuren erfasst sind, in der eigenen

Recherche als politikwissenschaftliche Professuren eingestuft. Für die im Rahmen dieser Studie durchgeführten Befragung bzw. deren Auswertung ist die Frage nach der genauen fachlichen Zuordnung innerhalb der Bruttostichprobe allerdings ohnehin zweitrangig, da die Professorinnen und Professoren am Anfang des Fragebogens gebeten wurden, selbst eine fachliche Eingruppierung vorzunehmen.

Die in der Bruttostichprobe erfassten Professorinnen und Professoren wurden Mitte November 2015 per personalisiertem E-Mail-Anschreiben (siehe Anhang II) zur Teilnahme an der Befragung eingeladen. Die Weiterleitung zum Online-Fragebogen erfolgte wie bereits beim Pretest durch einen anonymisierten Link. Nach zwei Wochen wurde ein Reminder versendet und nach einer weiteren Woche die Befragung beendet. Die gesamte Feldzeit betrug somit etwas mehr als drei Wochen (11.11.2015 bis einschließlich 06.12.2015). Ein detaillierter Blick auf den Rücklauf und die Repräsentativität der Stichprobe folgt im nächsten Kapitel.

5.3.2 Rücklaufquote und Repräsentativität

Im Vergleich zum Pretest konnte die Rücklaufquote bei der Hauptbefragung nochmals deutlich gesteigert werden. Insgesamt nahmen 880 Personen an der Befragung teil, von denen 670 den Fragebogen komplett abgeschlossen haben. 91 Teilnehmerinnen und Teilnehmer mussten allerdings aus der weiteren Analyse ausgeschlossen werden, da sie nicht zu der der Stichprobe zugrundeliegenden Grundgesamtheit gezählt werden konnten. Entweder waren diese gar nicht in einem sozialwissenschaftlichen Fach tätig ($n = 3$), machten keine Angabe zur Fachzugehörigkeit ($n = 23$) oder waren nicht (mehr) hauptberuflich als Professorin oder Professor tätig ($n = 48$) bzw. machten auch hier keine Angabe ($n = 24$). Weitere 60 Personen haben den Fragebogen direkt auf der Startseite (Einleitungstext ohne Fragen) abgebrochen, so dass eine Fächer- bzw. Positionszuordnung auch hier nicht möglich war und damit nicht überprüft werden konnte, ob diese Teilnehmerinnen und Teilnehmer zur Grundgesamtheit gezählt werden können. Abzüglich dieser Fragebögen reduziert sich der Rücklauf damit auf 729 Teilnehmerinnen und Teilnehmer bzw. 624 vollständig abgeschlossene Fragebögen, bei denen eine eindeutige Zuordnung zu einem Fach möglich war und eine hauptberufliche Tätigkeit als Professorin oder Professor vorlag. Die Rücklauf- und Beendigungsquoten für die gesamte Bruttostichprobe sowie für die einzelnen Fächer sind in der folgenden Tabelle 2 aufgeführt.

Tabelle 2: Rücklauf- und Beendigungsquote je Fach

	Bruttostich- probe	Rücklauf (n) und -quote (in %)	Rücklauf (n) und -quote (in %) ab- geschlossene Fragebögen	Beendi- gungs- quote (in %)
Erziehungswissenschaf- ten / Sonderpädagogik	722	176 (24.4)	146 (20.2)	83.0
Politikwissenschaften	367	66 (18.0)	55 (15.0)	83.3
Psychologie	603	127 (21.1)	115 (19.1)	90.6
Sozialwesen	52	14 (27.0)	12 (23.1)	85.7
Soziologie / Sozialwis- senschaften	387	85 (22.0)	73 (18.9)	85.9
Wirtschaftswissenschaf- ten	1583	254 (16.1)	217 (13.7)	85.4
Sonstiges sozialwissen- schaftliches Fach	/	7	6	85.7
Auswertbare Fragebö- gen (Nettorücklauf)	3714	729 (19.6)	624 (16.8)	85.6
Fragebögen gesamt (Bruttorücklauf)	3714	880 (23.7)	670 (18.0)	76.1

Anmerkung. Der Nettorücklauf entspricht allen auswertbaren Fragebögen (sozialwissenschaftliches Fach und hauptberuflich als Professorin oder Professor tätig). Der Bruttorücklauf bezieht sich dagegen auf alle teilnehmenden Personen, also auch auf jene, die nicht zur angestrebten Grundgesamtheit gezählt werden können (z. B. aufgrund fehlender Angaben zu Fach oder Position). Die Beendigungsquote fällt beim Bruttorücklauf mit 76.1 % deutlich geringer als beim Nettorücklauf (85.6 %) aus, da bei der Berechnung des Bruttorücklaufs jene 60 Befragten, die nur die Startseite des Fragebogens aufgerufen haben, mitberücksichtigt werden.

Insgesamt zeigen sich in Anbetracht der bereits thematisierten wachsenden Anzahl an Online-Befragungen und einer damit möglicherweise verbundenen abnehmenden Teilnahmebereitschaft sowie im Vergleich mit ähnlichen Befragungen (vgl. Minssen et al. 2003, S. 43–44) sehr zufriedenstellende Werte. Am höchsten fallen die Rücklaufquoten (Fragebogen abgeschlossen) in den Fächern Erziehungswissenschaften und Sonderpädagogik (20.2 %) sowie Sozialwesen (23.1 %) aus, während in den Wirtschaftswissenschaften mit nur 13.7 % der geringste Rücklauf zu verzeichnen ist. Die Beendigungsquoten liegen zwischen 83.0 % und 85.9 % und sind damit als sehr zufriedenstellend zu bewerten. In der Psychologie wurde der Fragebogen sogar von 90.6 % der teilnehmenden Personen vollständig ausgefüllt.

Die mittlere Bearbeitungsdauer des Fragebogens in Sekunden lag bei allen relevanten Teilnehmerinnen und Teilnehmern bei $M = 558$ ($SD = 338$), also etwa neuneinhalb Minuten. Der Median liegt bei acht Minuten. Die durchschnittliche Bearbeitungszeit fällt

gegenüber dem Pretest damit um etwa zwei Minuten niedriger aus. Dies ist eventuell darauf zurückzuführen, dass sozialwissenschaftliche Professorinnen und Professoren ‚geübter‘ beim Ausfüllen von Fragebögen sind, da diese in den Sozialwissenschaften eines der gängigsten Erhebungsinstrumente darstellen. Hinsichtlich des Abbrecherverhaltens zeigen sich wie bereits beim Pretest keine Auffälligkeiten. Die meisten Abbrüche erfolgten auf der Startseite (7.7 %) bzw. nach den Fragen zu Fach und Position auf der dritten Fragebogenseite (4.2 %).

Zusammenfassend lässt sich zunächst festhalten, dass in der Hauptbefragung hinsichtlich des Rücklaufs und der Beendigungsquote sehr zufriedenstellende Werte erzielt werden konnten. Für die im nächsten Kapitel folgenden deskriptiven Auswertungen ist aber natürlich auch von Interesse, wie generalisierbar die dargestellten Ergebnisse sind, das heißt als wie repräsentativ für die Grundgesamtheit die Stichprobe angesehen werden kann. Auch wenn der Begriff der Repräsentativität und deren Stellenwert als Gütekriterium einer Stichprobe in der Stichprobentheorie nicht unumstritten ist (vgl. Lippe und Kladroba 2002), soll hier daher dennoch kurz auf die zentralen Verteilungsmerkmale der Stichprobe eingegangen werden. Leider liegen die hierfür notwendigen Vergleichskennzahlen seitens des Statistischen Bundesamtes überwiegend auf zu hohem Aggregationsniveau vor. So werden in den relevanten Fachserien beispielsweise die Altersstruktur oder das Bundesland, in welchem die Professorinnen und Professoren tätig sind, nur auf Ebene der Fächergruppe berichtet (z. B. Fächergruppe Rechts-, Wirtschafts- und Sozialwissenschaften). Ein Abgleich auf Ebene der Fächergruppe ist aber nicht zielführend, da zum einen in der Fächergruppe Rechts-, Wirtschafts- und Sozialwissenschaften auch die Rechts- und Verwaltungswissenschaften enthalten sind. Zum anderen sind, wie bereits erwähnt, in der Systematik der Personalstatistik des Statistischen Bundesamtes die Fächer Psychologie, Erziehungswissenschaften und Sonderpädagogik der Fächergruppe Sprach- und Kulturwissenschaften zugeordnet (Statistisches Bundesamt 2015, S. 298–299). Daher muss sich hier auf die Geschlechterverteilung und die prozentuale Verteilung der Professorinnen und Professoren auf die einzelnen sozialwissenschaftlichen Fächer beschränkt werden. Für diese liegen jeweils Kennzahlen auf der Fächerebene vor.

Die Geschlechterverteilung je Fach der Nettostichprobe und in der amtlichen Statistik ist in Tabelle 3 dargestellt. Der Anteil an Frauen und Männern in der Nettostichprobe entspricht überwiegend der Verteilung in der Hochschulstatistik. Die Differenzbeträge bezüglich des Anteils an Frauen je Fach liegen zwischen 1.7 (Politikwissenschaften) und 6.1 (Erziehungswissenschaften) Prozentpunkten. Hinsichtlich des Geschlechts kann also von einer angemessenen Repräsentativität der Stichprobe ausgegangen werden.

Tabelle 3: *Geschlechterverteilung der Stichprobe*

		EZW	POL	PSY	SW	SOZ	WIWI	<i>gesamt</i>
Nettostichprobe	m	88	35	76	6	47	168	420
mit Angabe zu	w	56	16	36	5	24	40	177
Geschlecht	<i>n</i>	144	51	112	11	71	208	597
	Anteil w (in %)	38.9	31.4	32.1	45.5	33.8	19.2	29.6
Hochschulstatis-	m	501	248	407	45	301	1686	3188
tik (inkl. Junior-	w	410	105	241	43	187	347	1333
professuren)	<i>n</i>	911	353	648	88	488	2033	4521
	Anteil w (in %)	45.0	29.7	37.2	48.9	38.3	17.1	29.5

Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SW = Sozialwesen, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften

Quelle: Statistisches Bundesamt 2015, S. 94–95; eigene Berechnungen

Die Fächerverteilung der Bruttostichprobe gegenüber der Verteilung in der Hochschulstatistik wurde bereits in Tabelle 1 dargestellt, wobei bei einigen Fächern teils deutliche Abweichungen zur Hochschulstatistik erkennbar waren (Undercoverage in der Bruttostichprobe). Betrachtet man nun die Verteilung der Professorinnen und Professoren auf die einzelnen Fächer in der Nettostichprobe in Tabelle 4, zeigen sich nur mehr geringfügige Differenzen zur prozentualen Fächerverteilung in der Hochschulstatistik.

Tabelle 4: *Fächerverteilung der Stichprobe*

		EZW	POL	PSY	SW	SOZ	WIWI	<i>gesamt</i>
Nettostichprobe	<i>n</i>	176	66	127	14	85	254	729
	Anteil an gesamt (in %)	24.1	9.1	17.4	1.9	11.7	34.8	100
Hochschulstatis-	<i>n</i>	911	353	648	88	488	2033	4521
tik (inkl. Junior-	Anteil an gesamt (in %)	20.2	7.8	14.3	1.9	10.8	45.0	100
professuren)								

Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SW = Sozialwesen, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften. In der Kategorie *gesamt* der Nettostichprobe sind auch *n* = 7 *sonstige Sozialwissenschaften* enthalten.

Quelle: Statistisches Bundesamt 2015, S. 94–95; eigene Berechnungen

Lediglich in den Wirtschaftswissenschaften ergibt sich eine Differenz von 10.2 Prozentpunkten – der Anteil an wirtschaftswissenschaftlichen Professorinnen und Professoren bei den in dieser Studie betrachteten sozialwissenschaftlichen Fächern liegt nach Angaben der Hochschulstatistik bei 45.0 %, während in der Nettostichprobe nur 34.8 % der teilnehmenden Personen den Wirtschaftswissenschaften zugeordnet werden können. Dies deckt sich auch mit der geringeren Rücklaufquote in den Wirtschaftswissenschaften (vgl. Tabelle 2).

Das teils deutlich vorhandene Problem des Undercoverage in der Bruttostichprobe konnte in der Nettostichprobe vermutlich auch dadurch kompensiert werden, dass Recherchefehler bei der Bruttostichprobe (z. B. falsche Zuordnung von Personen zu den Politikwissenschaften statt zur Soziologie) durch die fachliche Selbsteinstufung im Fragebogen korrigiert werden konnten.

Die dargestellten Ergebnisse lassen sich dahingehend zusammenfassen, dass zumindest für die zentralen Merkmale Geschlecht und fachliche Zugehörigkeit keine starken Verzerrungen der Stichprobe zu erkennen sind. Die Verteilungen entsprechen überwiegend denen in der Grundgesamtheit der hauptberuflich tätigen sozialwissenschaftlichen Professorinnen und Professoren. Dies ist zwar kein Garant für die Repräsentativität der Stichprobe, aber doch ein Hinweis darauf, dass grobe systematische Fehlerquellen bei der Datenerhebung ausgeschlossen werden können. Wie sich die Verteilung der Stichprobe nun hinsichtlich anderer interessierender Merkmale darstellt, ist Gegenstand der nachfolgenden deskriptiven Analysen.

6. Ergebnisdarstellung

6.1 Deskriptive Analysen im Fächervergleich

6.1.1 Stichprobenbeschreibung

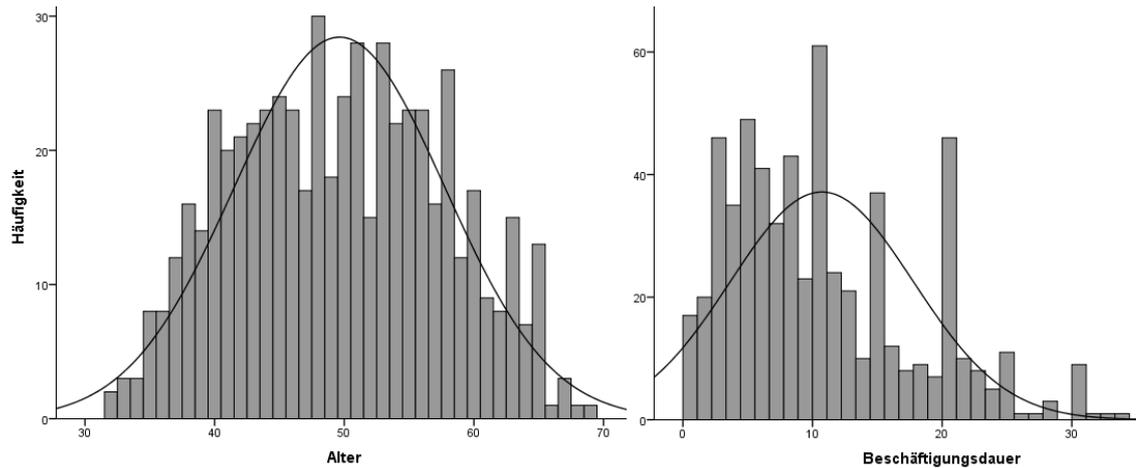
Die Verteilung der Nettostichprobe im Hinblick auf die Fachzugehörigkeit und das Geschlecht wurde bereits im vorangegangenen Kapitel dargestellt. Im Folgenden werden als weitere beschreibende Variablen zunächst das Alter und die Beschäftigungsdauer sowie die Besoldungsart und der Qualifizierungsweg zur Professur näher betrachtet. Abgeschlossen wird die Stichprobenbeschreibung mit Kennzahlen zur regionalen Verortung der Universitäten der Befragten. Die Beschreibung der Verteilung der interessierenden Variablen erfolgt dabei sowohl fachspezifisch³⁴ als auch für die Gesamtstichprobe. Für die deskriptiven und weiter nachfolgenden Analysen werden alle verwertbaren Fragebögen herangezogen, die der angestrebten Zielpopulation (hauptberuflich tätige sozialwissenschaftliche Professorinnen und Professoren) entsprechen. Da die Angaben im Fragebogen nicht verpflichtend waren (vgl. Kapitel 5.2.2) und auch nicht vollständig ausgefüllte Fragebögen (Nettorücklauf insgesamt $n = 729$, hiervon abgeschlossen $n = 624$) mitberücksichtigt werden, divergiert die Fallzahl der gültigen Angaben je nach Frage. Die Fallzahlen werden daher jeweils mit aufgeführt.

Alter und Beschäftigungsdauer

Gültige Angaben zum Alter (Geburtsjahr) liegen von 579 und gültige Angaben zur Beschäftigungsdauer von 592 Befragten vor. Die Häufigkeitsverteilungen beider Variablen sind in Abbildung 3 dargestellt. Das mittlere Alter der Befragten beträgt $M = 49.63$ ($SD = 8.12$), das Minimum liegt bei 32 und das Maximum bei 69 Jahren. Die Altersverteilung in der Stichprobe ist symmetrisch (Schiefe = 0.07), weicht aber im Hinblick auf die flachere Verteilung leicht von einer Normalverteilung ab (Kurtosis = -0.82). Der Altersdurchschnitt der Frauen ($M = 48.49$, $SD = 7.82$) liegt nur geringfügig unter dem der Männer ($M = 50.14$, $SD = 8.24$, $t(571) = 2.22$, $p = .027$). Auch im Fächervergleich zeigen sich nur geringe Differenzen hinsichtlich des Alters der Befragten. Den niedrigsten Altersdurchschnitt in der Stichprobe weisen Professorinnen und Professoren aus den Wirtschaftswissenschaften auf ($M = 48.50$, $SD = 8.27$), während Befragte aus den Erziehungswissenschaften mit $M = 50.91$ ($SD = 7.60$) im Mittel die älteste Gruppe bilden.

³⁴ Die Ergebnisse für Sozialwesen werden aufgrund der geringen Fallzahl von $n = 14$ nicht einzeln aufgeführt, sind aber in der Gesamtbetrachtung der sozialwissenschaftlichen Fächer enthalten. Ebenso wird mit den $n = 7$ sonstigen Sozialwissenschaften verfahren.

Abbildung 3: Histogramm Alter und Beschäftigungsdauer



Quelle: eigene Berechnungen

Die durchschnittliche Beschäftigungsdauer zum Zeitpunkt der Befragung beträgt $M = 10.71$ ($SD = 7.06$) Jahre. Die Verteilung der Beschäftigungsdauer weicht aufgrund der starken Rechtsschiefe (Schiefe = 0.83) signifikant von einer Normalverteilung ab. Die Angaben reichen von einem halben Jahr bis zu 34 Jahren, wobei sich durch die rechtsschiefe Verteilung 77.5 % der Befragten auf den Zeitraum zwischen einem halben und 15 Jahren verteilen. Zwischen den einzelnen Fächern ergeben sich keine signifikanten Unterschiede. Die Werte für die durchschnittliche Beschäftigungsdauer reichen von $M = 9.66$ ($SD = 5.71$) in den Politikwissenschaften bis zu $M = 11.65$ ($SD = 7.26$) in den Wirtschaftswissenschaften.

Wenig überraschend ergibt sich zwischen den Variablen Alter und Beschäftigungsdauer ein starker linearer Zusammenhang ($r(563) = .82$, $p < .001$). Auch wenn dieser Zusammenhang inhaltlich redundant ist, kann er zumindest als Hinweis für die Qualität der gewonnenen Daten gewertet werden. Eine Häufung von bewusst oder unbewusst getätigten Falschangaben erscheint aufgrund der Stärke des erwartbaren Zusammenhangs daher eher unwahrscheinlich.³⁵

Besoldungsart und Qualifizierungsweg zur Professur

Mehr als drei Viertel (78.5 %) der 729 befragten Professorinnen und Professoren sind in den Besoldungsgruppen W2 und W3 eingruppiert, während die übrigen 21.5 % noch nach der C-Besoldung vergütet werden. Im Fächervergleich zeigt sich ein ähnliches Bild. Hier liegt der Anteil der nach der W-Besoldung bezahlten Professorinnen und Professoren zwischen 73 und 86 %. Aufgrund der deutlich kleineren Gruppengröße der C-besol-

³⁵ Diese Schlussfolgerung ist natürlich nur dann gültig, wenn die Falschangaben nicht in der Form erfolgten, dass zwar falsche, aber bewusst plausible Werte für die Variablen Alter und Beschäftigungsdauer eingetragen wurden. Dies wird aber an dieser Stelle nicht angenommen.

deten Befragten ist eine vergleichende Perspektive bzw. eine Kontrolle auf die Besoldungsart in den noch nachfolgenden Analysen nur in der Gesamtbetrachtung (W-Besoldung: $n = 572$; C-Besoldung: $n = 157$) und für die Wirtschaftswissenschaften (W-Besoldung: $n = 186$; C-Besoldung: $n = 68$) sowie mit Einschränkungen für die Erziehungswissenschaften (W-Besoldung: $n = 141$; C-Besoldung: $n = 35$) sinnvoll durchführbar.

Betrachtet man die Frage nach dem Qualifizierungsweg zur Professur, zeigt sich, dass die Juniorprofessur als ein noch recht neuer möglicher Karriereweg³⁶ von nur 10.8 % ($n = 68$) der Befragten genannt wurde. Entsprechend der späteren Einführung der Juniorprofessur ergibt der t-Test für unabhängige Stichproben bei ungleichen Varianzen für diejenigen Befragten, die angaben, sich über die Juniorprofessur qualifiziert zu haben, einen signifikant ($t(121) = 12.46$, $p < .001$) niedrigeren Altersdurchschnitt ($M = 42.15$, $SD = 4.76$) als bei den übrigen Befragten ($M = 50.61$, $SD = 7.96$). Von den 68 Befragten, die eine Juniorprofessur innehatten, haben sich 11 Befragte zusätzlich über die Habilitation für eine Professur qualifiziert. Dies deutet darauf hin, dass die klassische Habilitation auch bei den Nachwuchswissenschaftlerinnen und -wissenschaftlern noch einen hohen Stellenwert hat. Insgesamt liegt der Anteil an habilitierten Professorinnen und Professoren bei den Befragten, die Angaben zum Qualifizierungsweg machten, bei 71.9 % ($n = 451$). Ähnliche Werte ergeben sich im Fächervergleich (68.7 % bis 76.4 %). Neben der Juniorprofessur und der Habilitation werden von 17.1 % ($n = 107$) der Befragten auch habilitationsäquivalente Leistungen als Qualifikationskriterium genannt.

Regionale Verortung

585 Professorinnen und Professoren gaben an, in welchem Bundesland sich die Universität befindet, an der sie zum Befragungszeitpunkt beschäftigt waren. Fast die Hälfte der Befragten (47.7 %) ist demnach an Universitäten in den drei bevölkerungsreichsten Bundesländern Nordrhein-Westfalen (22.1 %), Baden-Württemberg (13.8 %) und Bayern (11.8 %) beschäftigt. In absteigender Reihenfolge folgen Niedersachsen (10.1 %), Hessen (9.4 %), Sachsen (6.3 %), Rheinland-Pfalz (4.6 %), Berlin (4.4 %) und Hamburg (4.1 %). Die restlichen 13.3 % verteilen sich auf die übrigen sieben Bundesländer.

Auch wenn ein exakter Abgleich mit der Verteilung in der Grundgesamtheit aufgrund unterschiedlicher Aggregationsniveaus in der Hochschulstatistik nicht möglich war (vgl. Kapitel 5.3.2), scheint die regionale Stichprobenverteilung im Hinblick auf die Anzahl und die Größe der jeweiligen sozialwissenschaftlichen Hochschulstandorte in den einzelnen

³⁶ Die Einführung erfolgte im Jahr 2002 im Zuge der 5. Novellierung des Hochschulrahmengesetzes (vgl. hierzu und zur weiteren Entwicklung der Juniorprofessur Federkeil und Buch 2007).

Bundesländern damit tendenziell den zu erwartenden Größenverhältnissen zu entsprechen (vgl. zu verschiedenen Personalstatistiken nach Bundesländern Statistisches Bundesamt 2015).

6.1.2 Einsatz von Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren

Bevor näher auf erste Ergebnisse zu den in Steuerungs- und Anreizverfahren eingesetzten Forschungsindikatoren eingegangen wird, soll im Vorfeld darauf hingewiesen werden, dass die Einschätzungen der Professorinnen und Professoren nicht der konkreten Situation an ihrer Fakultät entsprechen müssen. Minssen und Wilkesmann (2003) weisen in diesem Zusammenhang auf eine hohe Unkenntnis im Hinblick auf die in Mittelverteilungsmodellen verwendeten Parameter hin. Dies treffe sowohl auf die Leitungsebene an Hochschulen (Dekane) als auch – in verschärfter Form – auf die Professorinnen und Professoren zu (ebd., S. 132–133). Wenn, wie in der vorliegenden Arbeit, Zusammenhänge zwischen dem Einsatz indikatorenbasierter Steuerungs- und Anreizverfahren und dem individuellen Forschungsverhalten untersucht werden sollen, ist aber gerade von Interesse, wie die Professorinnen und Professoren den Einsatz und die Ausgestaltung dieser Verfahren selbst wahrnehmen. Es geht hier also primär um die wahrgenommene und nicht um die konkrete Ausgestaltung der Steuerungsinstrumente, da für das Forschungsverhalten der Professorinnen und Professoren deren eigener Kenntnisstand über den Einsatz und die Ausgestaltung der Verfahren handlungsleitend ist. Bei der Interpretation der folgenden Daten ist also stets zu beachten, dass diese die Einschätzung der Befragten widerspiegeln, welche nicht zwangsläufig mit der konkreten Situation an einer Fakultät deckungsgleich sein muss.

Weiter ist zu beachten, dass die vorliegenden Daten hierarchisch strukturiert sein können. Dies ist dann der Fall, wenn mehrere Befragte Mitglieder derselben Fakultät (bzw. Universität) sind. Dadurch ergeben sich unterschiedlich große Gruppen von Befragten, die sich nach Fakultätszugehörigkeit differenzieren lassen. Die Angaben der Befragten zum Einsatz von Forschungsindikatoren beziehen sich dann – je nach Stichprobenverteilung in unterschiedlich starkem Ausmaß – auf jeweils dieselbe Fakultät.³⁷ Die anteiligen Werte zur Einsatzhäufigkeit von Forschungsindikatoren in der Gesamtstichprobe können daher nicht als Repräsentant für die Einsatzhäufigkeit der Indikatoren auf Ebene der Fakultäten bzw. Universitäten verstanden werden. Vielmehr sind die Daten

³⁷ Aus Gründen der Anonymität musste im Fragebogen auf die Abfrage der Universität, an der die Befragten tätig sind, verzichtet werden (z. B. wäre in Verbindung mit dem abgefragten Fachbereich eine eindeutige Zuordnung möglich, wenn an einer Universität in einem Fach nur eine Professur besetzt ist). Dadurch ist eine Kontrolle der Antworthäufigkeiten im Hinblick auf die Verteilung der Befragten nach Universitätsstandorten bzw. Fakultäten nicht möglich.

so zu interpretieren, als dass sie anzeigen, welcher Anteil der befragten sozialwissenschaftlichen Professorinnen und Professoren in ihrer Wahrnehmung im Arbeitsalltag mit dem Einsatz verschiedener Forschungsindikatoren konfrontiert ist. Akzeptiert man die Repräsentativitätsannahme der Stichprobe, lassen die Daten damit auch Rückschlüsse auf die Grundgesamtheit zu und erlauben erste Aussagen darüber, wie hoch der Anteil an sozialwissenschaftlichen Professorinnen und Professoren insgesamt ist, der vom Einsatz eines oder mehrerer Forschungsindikatoren in Steuerungs- und Anreizverfahren ‚betroffen‘ ist.

Die gültigen Angaben zur Frage, welche Forschungsindikatoren im Rahmen von hochschulinternen Steuerungs- und Anreizverfahren insgesamt eingesetzt werden, differieren je nach abgefragtem Item.³⁸ Das Maximum liegt bei $n = 678$ gültigen Angaben beim Item *Höhe der sonstigen eingeworbenen Drittmittel* und das Minimum bei $n = 592$ Angaben beim Item *Anzahl der Zitationen / Zitationsindizes*.³⁹ In Tabelle 5 sind die absoluten und prozentualen Antworthäufigkeiten sowie die Fallzahl je Item aufgeführt. Eine grafische Übersicht liefert Abbildung 4.

Zunächst kann festgehalten werden, dass sich in Übereinstimmung mit Minssen und Wilkesmann (2003) auch in dieser Studie zeigt, dass ein nicht unerheblicher Teil der Befragten (5.7 % - 15.2 %) bei einzelnen Indikatoren keine Kenntnis darüber hat, ob diese in indikatorenbasierten Steuerungs- oder Anreizverfahren zur Anwendung kommen. Dies betrifft – mit Ausnahme von Publikationen mit Peer-Review-Verfahren – vor allem die publikations- und zitationsbasierten Indikatoren. Insgesamt konnten dennoch von einem Großteil der Befragten Angaben dazu gemacht werden, ob ein Indikator in den Verfahren eingesetzt wird.

Mit Abstand am häufigsten findet dabei nach Einschätzung der Befragten die Drittmittelhöhe als Forschungsindikator Verwendung in hochschulinternen Steuerungs- und Anreizverfahren. Dabei zeigen sich zunächst keine Differenzen zwischen DFG-Drittmitteln und sonstigen Drittmitteln, die jeweils von einem gleich hohen Anteil der Befragten (80.2 %) genannt wurden. Dies kann entweder dahingehend interpretiert werden, dass in den indikatorenbasierten Verfahren nicht zwischen DFG-Drittmitteln und sonstigen

³⁸ Für den exakten Wortlaut der Fragen und der im Folgenden dargestellten Items siehe den Fragebogen im Anhang I.

³⁹ Im Hinblick auf die deutlich geringeren gültigen Angaben beim Item *Anzahl der Zitationen / Zitationsindizes* ist davon auszugehen, dass, wenn ein Indikator nach Einschätzung der Befragten nicht eingesetzt wurde oder keine Informationen hierzu bekannt waren, öfters keine der vorgegebenen Antwortkategorien ausgewählt wurde. Die Befragten hätten demnach in diesen Fällen teils keine der beiden Kategorien „Indikator wird nicht eingesetzt“ oder „weiß nicht“ angekreuzt und nur beim Zutreffen der Option „Indikator wird eingesetzt“ eine Angabe vorgenommen. Diese Annahme kann anhand der vorliegenden Daten aber nicht überprüft werden, so dass nicht beantwortete Items als ungültige Werte behandelt werden und sich die Auswertung nur auf die tatsächlich getätigten gültigen Angaben („Indikator wird eingesetzt“, „Indikator wird nicht eingesetzt“, „weiß nicht“) bezieht.

Drittmitteln unterschieden wird, oder dass neben DFG-Drittmitteln mit der gleichen Häufigkeit auch sonstige Drittmittel als Indikator eingesetzt werden.

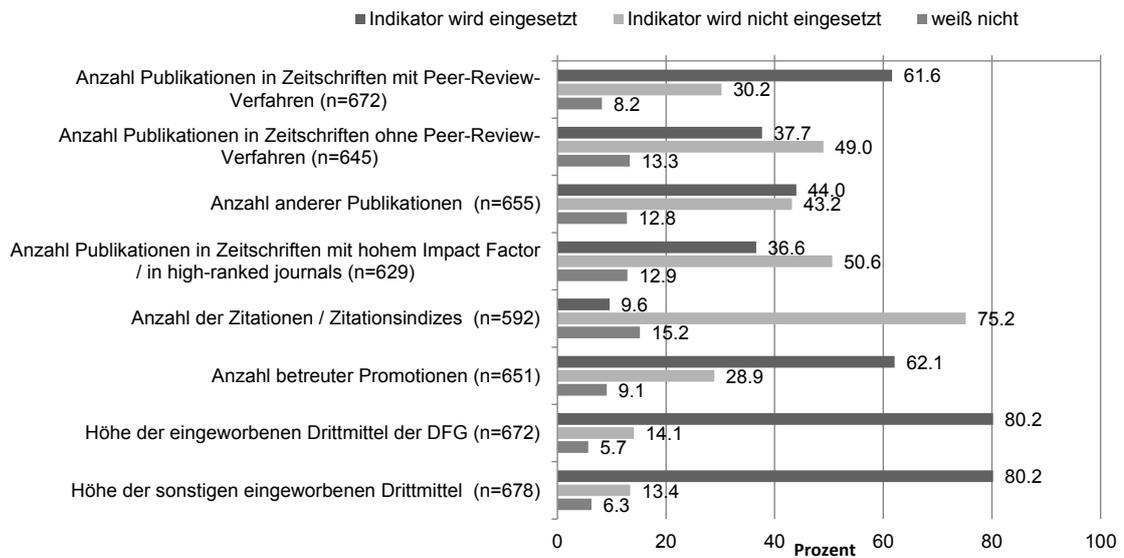
Tabelle 5: *Eingesetzte Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren*

	Indikator wird eingesetzt		Indikator wird nicht eingesetzt		weiß nicht		gesamt <i>n</i>
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	414	61.6	203	30.2	55	8.2	672
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	243	37.7	316	49.0	86	13.3	645
Anzahl anderer Publikationen	288	44.0	283	43.2	84	12.8	655
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	230	36.6	318	50.6	81	12.9	629
Anzahl der Zitationen / Zitationsindizes	57	9.6	445	75.2	90	15.2	592
Anzahl betreuter Promotionen	404	62.1	188	28.9	59	9.1	651
Höhe der eingeworbenen Drittmittel der DFG	539	80.2	95	14.1	38	5.7	672
Höhe der sonstigen eingeworbenen Drittmittel	544	80.2	91	13.4	43	6.3	678

Quelle: eigene Berechnungen

An zweiter Stelle folgen die Anzahl von Publikationen mit Peer-Review-Verfahren sowie die Anzahl betreuter Promotionen, die von 61.1 % bzw. 62.1 % der befragten Professorinnen und Professoren genannt wurden. Mit 44 % gaben deutlich weniger Befragte an, dass andere Publikationen in Form von Monografien, Beiträgen in Sammelbänden etc. als Indikator verwendet werden.⁴⁰

⁴⁰ Im Hinblick auf dieses Item muss erwähnt werden, dass 14 Befragte die Aggregation unterschiedlicher Publikationsformate zum Item *Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden etc.)* (zu Recht) kritisierten. In Bezug auf die hier dargestellte Frage nach den verwendeten Forschungsindikatoren ist dies noch weniger problematisch. Bei anderen Fragen, wie z. B. der nach dem Stellenwert des Verfassens „anderer Publikationen“ in der eigenen Forschungsarbeit, ergeben sich durch die Mehrdimensionalität des Items allerdings Probleme. Die

Abbildung 4: *Eingesetzte Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren*

Quelle: eigene Berechnungen

Für die Anzahl der Publikationen ohne Peer-Review-Verfahren gab etwas mehr als ein Drittel der Befragten (37.7 %) an, dass diese als Forschungsindikator eingesetzt wird. Gleiches gilt für die Anzahl von Publikationen in Zeitschriften mit hohem Impact Factor bzw. auf Basis anderer Kennzahlen hoch geranker Journals (36.6 %). In nur sehr begrenztem Ausmaß finden nach Einschätzung der Befragten die Anzahl der Zitationen oder darauf basierende Zitationskennzahlen als Forschungsindikatoren Verwendung (9.6 %).

Ergänzend zu den vorgegebenen Indikatoren hatten die Befragten die Möglichkeit, in einem Freitextfeld weitere relevante Indikatoren einzutragen. Wie bereits beim Pretest wurde das Freitextfeld von einigen Befragten dazu genutzt, Indikatoren, welche die Lehre betreffen, zu ergänzen. Insgesamt wurden von $n = 47$ Befragten lehrbezogene Indikatoren, insbesondere Lehrevaluationen und die Anzahl betreuter Abschlussarbei-

Befragten sind bei der Beantwortung der Frage (Einschätzung auf einer Ratingskala) dann entweder gezwungen, sich für ein Publikationsformat zu entscheiden (z. B. Monografien), oder sie bilden einen ‚gedanklichen Mittelwert‘ über alle relevanten anderen Publikationen (Monografien, Reviews, Buchkapitel, ...) und legen diesen ihren Angaben zugrunde. In beiden Fällen ist mit verzerrten Werten zu rechnen, sofern die verschiedenen Publikationsformate von den Befragten nicht einheitlich bewertet werden. Hiervon kann aber bei diesem Item nicht ausgegangen werden, da z. B. eine Monografie und ein Beitrag in einem Sammelband weder vom Arbeitsaufwand noch von der eingeschätzten ‚Wertigkeit‘ innerhalb der fachspezifischen Scientific Community als gleich angesehen werden können. Im Pretest stellte dies noch kein Problem dar, da in der Physik Zeitschriftenartikel ohnehin das vorherrschende Publikationsformat sind und andere Formate nur sehr wenig Berücksichtigung finden. Aufgrund grundlegend anderer Publikationskulturen in den Sozialwissenschaften ist bei diesen allerdings mit mehr oder weniger großen Verzerrungen zu rechnen. Im Folgenden wird das Item *Andere Publikationen* dennoch mit in der Analyse berücksichtigt. Die Ergebnisse sind allerdings mit entsprechender Vorsicht zu interpretieren.

ten, genannt. Wie bereits ausgeführt wurde (vgl. Kapitel 5.2.3), können die recht zahlreichen Nennungen von lehrbezogenen Indikatoren entweder auf einer falsch verstandenen Fragestellung oder auf der Intention der Befragten, die Bedeutung der Lehre im Kontext von hochschulinternen Steuerungsversuchen hervorzuheben, beruhen.

Mit absteigender Häufigkeit wurden neben den lehrbezogenen Indikatoren das Engagement in der Selbstverwaltung und Gremienarbeit ($n = 27$), Vorträge auf Tagungen und ähnlichen Veranstaltungen ($n = 10$), internationale Kooperationen und internationale Sichtbarkeit ($n = 7$) sowie die Einwerbung von bzw. die Mitwirkung in Verbundforschungsprojekten ($n = 6$) genannt.

Für die sozialwissenschaftlichen Fächer insgesamt lässt sich somit festhalten, dass in der Wahrnehmung der befragten Professorinnen und Professoren Drittmittel, Promotionen und Publikationen in peer-reviewed Zeitschriften die drei zentralen Forschungsindikatoren sind, mit denen sie in ihrer Forschungsarbeit im Rahmen von indikatorenbasierten Steuerungs- und Anreizverfahren konfrontiert werden. Dies deckt sich mit der bereits in Kapitel 2.3.2 diskutierten hohen Verbreitung der genannten Indikatoren in verschiedenen forschungsevaluativen Verfahren und Hochschulrankings.

Inwieweit sich die Einschätzungen der Professorinnen und Professoren in fächervergleichender Perspektive unterscheiden, ist in Tabelle 6 dargestellt. Für eine bessere Vergleichbarkeit zwischen den Teilstichproben werden nur die gültigen Angaben zu den Antwortkategorien „Indikator wird eingesetzt“ und „Indikator wird nicht eingesetzt“ berücksichtigt. Die Antwortkategorie „weiß nicht“ wird als missing value behandelt. Die Werte in Tabelle 6 beziehen sich dabei auf die gültigen Nennungen zur Antwortkategorie „Indikator wird eingesetzt“. Diese addieren sich zusammen mit den Angaben bei der Kategorie „Indikator wird nicht eingesetzt“ entsprechend zu 100 %.

Zur Prüfung der Signifikanz der Differenzen zwischen den Einschätzungen der Befragten der einzelnen Fächer wurde jeweils ein Chi-Quadrat-Test durchgeführt. Dieser ergibt lediglich für die Indikatoren *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* ($\chi^2(4, n = 602) = 3.10, p = .541$) sowie *Anzahl der Zitationen / Zitationsindizes* ($\chi^2(4, n = 494) = 0.88, p = .927$) keine signifikante Abweichung von der Nullhypothese der Unabhängigkeit der untersuchten Variablen. Über alle Fächer hinweg geben jeweils etwa zwei Drittel der Befragten an, dass die Anzahl begutachteter Publikationen an ihrer Fakultät als Forschungsindikator eingesetzt wird. Der Stellenwert von Publikationen mit Peer-Review-Verfahren als Forschungsindikator scheint damit in allen betrachteten Fächern (mittlerweile) hoch angesiedelt zu sein. Die Differenzen beim Indikator *Anzahl der Zitationen / Zitationsindizes* entziehen sich aufgrund der teils sehr geringen Gruppengrößen einer sinnvollen Interpretation.

Bei allen übrigen abgefragten Indikatoren ergeben sich bezüglich des wahrgenommenen Einsatzes in Steuerungs- und Anreizverfahren höchst signifikante Unterschiede

zwischen den Fächern. Bei Publikationen, die in Zeitschriften ohne Peer-Review-Verfahren veröffentlicht werden und in der Publikationspraxis der Psychologie und der Wirtschaftswissenschaften nur von untergeordneter Bedeutung sind (vgl. Fiedler 2009; Leininger 2009), zeigt sich erwartungsgemäß auch in der Wahrnehmung der Befragten der beiden genannten Fächer nur ein eher mäßiger Einsatz dieses Publikationsformats als Forschungsindikator (37.1 % bzw. 27.7 %). Ähnliches gilt für den Indikator *Anzahl anderer Publikationen*. Dieser wird insbesondere von wirtschaftswissenschaftlichen Professorinnen und Professoren selten (34.0 %) genannt, während der Anteil in der Psychologie fast bei der Hälfte der Befragten liegt, aber mit 48.5 % im Fächervergleich dennoch nur den zweitniedrigsten Wert erreicht.

Tabelle 6: *Eingesetzte Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren im Fächervergleich*

	EZW		POL		PSY		SOZ		WIWI	
	<i>n</i>	%								
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	85	62.5	37	67.3	80	71.4	53	71.6	148	65.8
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	72	56.7	26	52.0	36	37.1	43	62.3	56	27.7
Anzahl anderer Publikationen	80	62.0	31	58.5	49	48.5	48	67.6	69	34.0
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	39	33.9	14	30.4	46	43.8	20	33.3	106	50.2
Anzahl der Zitationen / Zitationsindizes	10	9.1	6	12.5	11	11.3	7	13.5	22	11.8
Anzahl betreuter Promotionen	90	68.7	34	68.0	90	79.6	51	76.1	122	57.3
Höhe der eingeworbenen Drittmittel der DFG	121	85.2	50	89.3	110	94.8	70	92.1	169	75.1
Höhe der sonstigen eingeworbenen Drittmittel	129	87.2	49	86.0	106	94.6	68	90.7	173	77.2

Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften. Die absoluten (*n*) und relativen (%) Werte beziehen sich auf die gültigen Angaben bei der Antwortkategorie „Indikator wird eingesetzt“.

Quelle: eigene Berechnungen

Beim Item *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals* kehrt sich das Bild um: Am häufigsten wird von Befragten aus der Psychologie (43.8 %) und den Wirtschaftswissenschaften (50.2 %) angegeben, dass dieses Publikationsformat als Forschungsindikator Verwendung findet. Bei den übrigen Fächern sind es nur (knapp) ein Drittel.

Der Anteil der Nennungen beim Indikator *Anzahl betreuter Promotionen* schwankt je nach Fach zwischen 68.0 % und 79.6 %. Deutlich seltener wird dieser Indikator nach Einschätzung der wirtschaftswissenschaftlichen Professorinnen und Professoren herangezogen (57.3 %). Auch bei den Drittmitteln (sowohl DFG als auch sonstige) liegt der Anteil der Nennungen in den Wirtschaftswissenschaften mit 75.1 % bzw. 77.2 % teils deutlich unter den Werten der anderen Fächer, die von 85.2 % bis 94.8 % reichen. Insbesondere die Psychologie sticht hier mit Werten von knapp 95 % hervor. Fast alle Befragte dieses Fachbereiches gaben demnach an, dass Drittmittel in hochschulinternen Steuerungs- und Anreizverfahren als Forschungsindikator eingesetzt werden.

Die fächervergleichende Perspektive macht deutlich, dass sich Unterschiede hinsichtlich des Einsatzes von Forschungsindikatoren vor allem zwischen den Fächern Psychologie und Wirtschaftswissenschaften auf der einen und den Erziehungswissenschaften, den Politikwissenschaften sowie der Soziologie auf der anderen Seite zeigen. Hierfür sind vermutlich vor allem Unterschiede in den Publikationskulturen der einzelnen Fächer verantwortlich (vgl. ebd.). Ein grundlegend anderer Stellenwert von Promotionen (z. B. Anteil an berufsbegleitenden, externen Promotionen) und Drittmitteln (z. B. Anteil an Auftragsforschung) in der Forschungspraxis der Wirtschaftswissenschaften könnte als Erklärung für die abweichenden Einschätzungen der wirtschaftswissenschaftlichen Professorinnen und Professoren bei diesen beiden Indikatoren herangezogen werden.

Die Teilnehmerinnen und Teilnehmer der Befragung wurden auch danach gefragt, wie stark die in den Steuerungs- und Anreizverfahren eingesetzten Indikatoren ihrer Einschätzung nach insgesamt gewichtet werden. Es geht also darum, welche Relevanz die Befragten den einzelnen Indikatoren innerhalb von Steuerungs- und Anreizverfahren zuschreiben. Die Einschätzung erfolgte dabei auf einer fünfstufigen numerischen Ratingskala mit den verbalisierten Endpolen „sehr geringe Gewichtung“ und „sehr hohe Gewichtung“. In Tabelle 7 sind die durchschnittlichen Werte je Indikator sowie weitere Verteilungsparameter dargestellt.

Die drittmittelbasierten Indikatoren und die Anzahl von Publikationen in peer-reviewed Zeitschriften, welche nach Einschätzung der Befragten mit am häufigsten in Steuerungs- und Anreizverfahren Verwendung finden, werden in der Wahrnehmung der Befragten auch stark gewichtet. Insbesondere DFG-Drittmittel scheinen einen besonders hohen Stellenwert zu haben ($M = 4.22$, $SD = 0.99$), während die Gewichtung der sonstigen Drittmittel etwas niedriger eingeschätzt wird ($M = 3.87$, $SD = 1.03$). Dem ebenfalls häufig

eingesetzten Indikator *Anzahl betreuter Promotionen* wird dagegen nur eine geringere Relevanz innerhalb der Steuerungs- und Anreizverfahren zugesprochen ($M = 3.09$, $SD = 1.04$). Sofern diese als Indikator eingesetzt wird, stellt die Anzahl von Publikationen in hoch gerankten Zeitschriften nach Einschätzung der Befragten einen weiteren hoch gewichteten Indikator dar ($M = 3.78$, $SD = 1.15$).

Tabelle 7: Gewichtung der Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren insgesamt

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	388	3.68	1.11	-0.58	-0.30
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	224	2.74	0.95	0.04	-0.11
Anzahl anderer Publikationen	266	2.78	1.03	0.12	-0.38
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	210	3.78	1.15	-0.60	-0.55
Anzahl der Zitationen / Zitationsindizes	51	3.14	1.25	-0.01	-0.98
Anzahl betreuter Promotionen	374	3.09	1.04	-0.05	-0.55
Höhe der eingeworbenen Drittmittel der DFG	515	4.22	0.99	-1.21	0.98
Höhe der sonstigen eingeworbenen Drittmittel	512	3.87	1.03	-0.77	0.15

Anmerkung. Die Werte basieren auf einer fünfstufigen Ratingskala mit den Endpolen „sehr geringe Gewichtung“ (1) und „sehr hohe Gewichtung“ (5). Es wurde nur die Gewichtung derjenigen Indikatoren erfragt, die nach Angabe der Befragten auch eingesetzt werden (vgl. Fragebogen im Anhang).

Quelle: eigene Berechnungen

Hinsichtlich der Verteilung der Variablen ist anzumerken, dass die Variablen *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren*, *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals*, *Höhe der sonstigen eingeworbenen Drittmittel* und insbesondere die Variable *Höhe der eingeworbenen Drittmittel der DFG* eine (starke) Linksschiefe aufweisen (Schiefe zwischen -0.58 und -1.21). Die Befragten, die angegeben haben, dass die genannten Indikatoren in Steuerungs- und Anreizverfahren zum Einsatz kommen, gaben also auch mehrheitlich an, dass diese (sehr) stark gewichtet werden.

Signifikante Unterschiede zwischen einzelnen Fächern zeigen sich nach Durchführung einer einfaktoriellen Varianzanalyse bei den Indikatoren *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* ($F(4, 209) = 4.01$, $p = .004$) und *Anzahl anderer Publikationen* ($F(4, 250) = 5.52$, $p < .001$) sowie bei den Drittmittelindikatoren

Höhe der eingeworbenen Drittmittel der DFG ($F(4, 491) = 3.16, p = .014$) und Höhe der sonstigen eingeworbenen Drittmittel ($F(4, 489) = 4.65, p = .001$).⁴¹ Befragte aus den Erziehungswissenschaften schätzen demzufolge die Gewichtung von Publikationen ohne Peer-Review-Verfahren in den Steuerungs- und Anreizverfahren signifikant höher ein ($M = 3.03, SD = 0.82, n = 67$) als Befragte aus der Psychologie ($M = 2.35, SD = 0.92, n = 34$) und den Wirtschaftswissenschaften ($M = 2.53, SD = 1.07, n = 51$).⁴² Ein ähnliches Bild zeigt sich hinsichtlich der *anderen Publikationen*, denen von den erziehungswissenschaftlichen Professorinnen und Professoren ein höheres Gewicht zugesprochen wird ($M = 3.15, SD = 0.90, n = 74$) als von den Wirtschaftswissenschaftlern ($M = 2.53, SD = 1.07, n = 51$).

Auch bei den DFG-Drittmitteln ergibt sich eine signifikante Differenz in der Einschätzung der Gewichtung zwischen den Erziehungswissenschaften und den Wirtschaftswissenschaften. Wieder wird die Gewichtung in den Erziehungswissenschaften ($M = 4.42, SD = 0.89, n = 114$) höher eingestuft als in den Wirtschaftswissenschaften ($M = 4.02, SD = 1.05, n = 162$), wobei die mittlere Differenz mit einem Wert von 0,4 niedriger ausfällt als bei den anderen Indikatoren. Analoges gilt für die sonstigen Drittmittel, bei denen neben den Befragten aus den Erziehungswissenschaften ($M = 4.04, SD = 0.94, n = 120$) auch die Befragten aus der Soziologie ($M = 4.06, SD = 1.02, n = 62$) den darauf basierenden Indikatoren eine höhere Gewichtung zuschreiben, als dies Befragte aus den Wirtschaftswissenschaften ($M = 3.60, SD = 1.07, n = 166$) tun.

Die fächervergleichende Perspektive zu den Einschätzungen der Gewichtungen der Indikatoren innerhalb von Steuerungs- und Anreizverfahren spiegelt im Kern die vorangegangene Analyse zur Einsatzhäufigkeit der Indikatoren wider. Beispielsweise werden diejenigen Indikatoren, die nach Einschätzung der erziehungswissenschaftlichen Befragten hoch und nach Einschätzung von Befragten aus den Wirtschaftswissenschaften eher gering gewichtet werden (*Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren, Anzahl anderer Publikationen, Höhe der DFG und sonstigen eingeworbenen Drittmittel*), auch häufiger (Erziehungswissenschaften) bzw. seltener (Wirtschaftswissenschaften) eingesetzt. In den Wirtschaftswissenschaften und – mit Ausnahme der Drittmittelindikatoren – auch in der Psychologie haben die genannten Indikatoren also anscheinend absolut (eingeschätzter Einsatz der Indikatoren) und relativ (Einschätzung der

⁴¹ Bei Verletzung der Homogenitätsannahme der Varianzen wurden die Ergebnisse der ANOVA mit robusten Testverfahren überprüft (Welch und Brown-Forsythe-Test). Die Ergebnisse der ANOVA konnten in allen Fällen bestätigt werden. Berichtet werden daher in diesem und in den folgenden Kapiteln stets die Ergebniskennzahlen der ANOVA.

⁴² Die hier und im Folgenden berichteten Unterschiede zwischen den Fächern wurden mittels Post-Hoc-Mehrfachvergleichen auf Signifikanz geprüft. Hierzu wurde bei vorliegender Varianzgleichheit die Scheffé-Prozedur herangezogen und bei Verletzung der Homogenitätsannahme der Varianzen (betraf die Variablen *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* und *Höhe der sonstigen eingeworbenen Drittmittel*) der Games-Howell-Test verwendet. Alle berichteten Differenzen sind auf dem 5%-Niveau signifikant.

Gewichtung) einen geringeren Stellenwert als z. B. in den Erziehungswissenschaften. Im Gegensatz dazu wird die Gewichtung der übrigen Indikatoren (*Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren, Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals, Anzahl der Zitationen / Zitationsindizes, Anzahl betreuter Promotionen*) über alle Fächer hinweg etwa gleich hoch eingeschätzt. Insbesondere die publikationsbasierten Indikatoren *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* und *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals* werden nach Einschätzung der Befragten durchgehend hoch gewichtet. Ein hoher Stellenwert von begutachteten Zeitschriftenartikeln und eine generelle ‚Impact-Factor-Orientierung‘, wie sie in den Publikationskulturen der Wirtschaftswissenschaften und der Psychologie bereits mehr als in anderen sozialwissenschaftlichen Fächern verankert ist, scheint sich also bei den übrigen Fächern zumindest auch auf der Ebene der in Steuerungs- und Anreizverfahren verwendeten Indikatoren zu zeigen. Dabei spielt die Anzahl der peer-reviewed Publikationen in den Fächern Erziehungswissenschaften, Politikwissenschaften und Soziologie sowohl in Bezug auf die eingeschätzte Einsatzhäufigkeit als auch in Bezug auf die Höhe der Gewichtung als Indikator eine wichtige Rolle, während der ‚Impact-Indikator‘ in den genannten Fächern absolut noch eine geringe Bedeutung hat. Wird dieser allerdings als Indikator eingesetzt, geschieht dies anscheinend meist auch mit einer hohen Gewichtung.

Inwiefern die Einsatzhäufigkeit und die Gewichtung der Indikatoren die tatsächliche Relevanz dieser Publikationsformate in den Fächern selbst widerspiegelt, wird sich an späterer Stelle der Arbeit zeigen, wenn die Einschätzungen der Befragten zur Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs in der Scientific Community betrachtet werden (vgl. Kapitel 6.1.4). Im Folgenden werden aber zunächst der Einsatz und die Ausgestaltung verschiedener indikatorenbasierter Steuerungs- und Anreizverfahren aus der Sicht der Professorinnen und Professoren näher beleuchtet.

6.1.3 Einsatz und Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren

Auch bezüglich des Einsatzes und der Ausgestaltung hochschulinterner Steuerungs- und Anreizverfahren ist nochmals darauf hinzuweisen, dass die im Folgenden dargestellten Ergebnisse nur die subjektive Einschätzung der Befragten widerspiegeln und zudem hierarchisch strukturiert sein können. Sie müssen daher nicht der tatsächlichen anteiligen Einsatzhäufigkeit und Ausgestaltung der Verfahren an den Fakultäten entsprechen (vgl. hierzu auch die Ausführungen zu Beginn des vorangegangenen Kapitels). Die Ergebnisse zeigen also vielmehr an, welcher Anteil der befragten Professorinnen

und Professoren nach eigener Einschätzung im Arbeitsalltag mit dem Einsatz von indikatorenbasierten Steuerungs- und Anreizverfahren konfrontiert ist.

In Tabelle 8 sowie Abbildung 5 sind die Einschätzungen der Befragten zum Einsatz der einzelnen Steuerungs- und Anreizverfahren dargestellt.

Tabelle 8: *Eingesetzte indikatorenbasierte Steuerungs- und Anreizverfahren*

	Verfahren wird eingesetzt		Verfahren wird nicht eingesetzt		weiß nicht		gesamt <i>n</i>
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
Indikatorenbasierte Mittelzuweisung zwischen Hochschulleitung und Fakultät	364	58.1	153	24.4	110	17.5	627
Indikatorenbasierte Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl	338	54.0	232	37.1	56	8.9	626
Zielvereinbarungen zwischen Hochschulleitung und Fakultät	345	55.3	173	27.7	106	17.0	624
Zielvereinbarungen zwischen Fakultät und Institut/Lehrstuhl	159	25.5	412	66.1	52	8.3	623
Benchmarking verschiedener Lehrstühle innerhalb der Fakultät	79	12.8	443	71.7	96	15.5	618

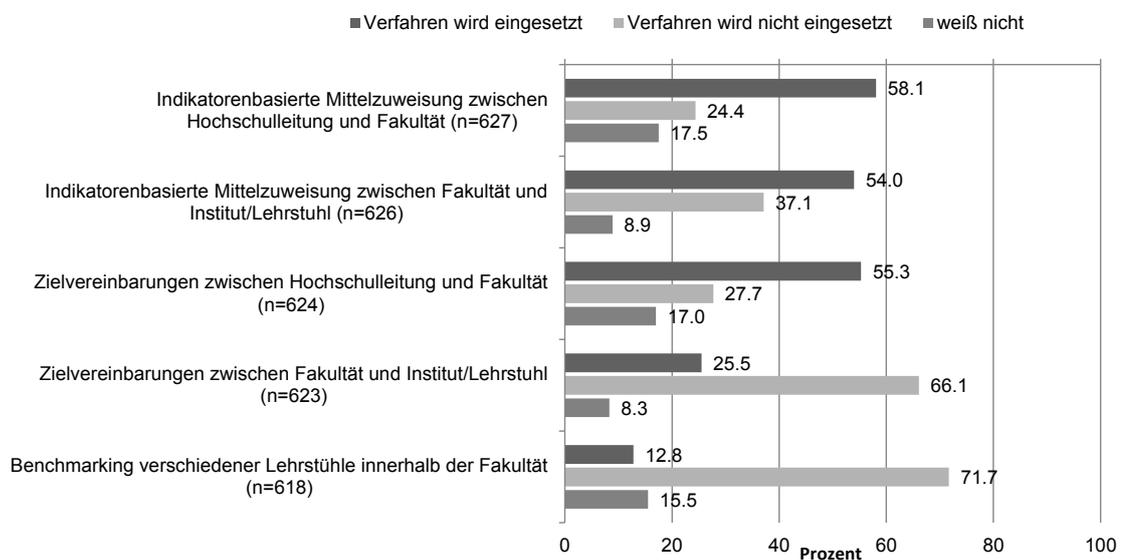
Quelle: eigene Berechnungen

Für Bleibe- und Berufungsverhandlungen sowie die Leistungszulagen im Rahmen der W-Besoldung, welche in dieser Arbeit auch als mögliche indikatorenbasierte Steuerungsinstrumente verstanden werden, wurde aufgrund der Sonderstellung dieser beiden Verfahren keine Abfrage zu deren Einsatz durchgeführt (vgl. zur Fragebogenlogik Kapitel 5.2.2). Bei der später folgenden Beschreibung zur Ausgestaltung der Verfahren werden diese dann wieder mitberücksichtigt.

Wenig überraschend fällt zunächst auf, dass über den Einsatz fakultätsinterner Verfahren mehr Kenntnis besteht als über den Einsatz von Verfahren auf der Ebene Hochschulleitung–Fakultät. Der Anteil der „weiß nicht“-Nennungen ist bei den Verfahren auf der Ebene Hochschulleitung–Fakultät etwa doppelt so hoch. Eine Ausnahme bildet hier das *Benchmarking verschiedener Lehrstühle*, bei dem über 15 % der Befragten angaben, dass sie keine Kenntnis über den Einsatz des Verfahrens haben.

Hinsichtlich der Einsatzhäufigkeit gaben etwas mehr als die Hälfte der Befragten an, dass eine indikatorenbasierte Mittelzuweisung sowohl auf der Ebene Hochschulleitung–Fakultät als auch auf der Ebene Fakultät–Institut/Lehrstuhl existiert. Auch Zielvereinbarungen zwischen der Hochschulleitung und der Fakultät werden nach Ansicht von etwa 55 % der Befragten abgeschlossen. Fakultätsinterne Zielvereinbarungen sind nur nach Einschätzung eines Viertels der Befragten üblich. Ein Benchmarking von Lehrstühlen innerhalb einer Fakultät anhand verschiedener Forschungsindikatoren findet aus Sicht der Professorinnen und Professoren nur vereinzelt statt (12.8 %).

Abbildung 5: *Eingesetzte indikatorenbasierte Steuerungs- und Anreizverfahren*



Quelle: eigene Berechnungen

Auch bei der Frage nach den eingesetzten Steuerungs- und Anreizverfahren war es möglich, in einem Freitextfeld weitere Verfahren zu ergänzen. Relativ viele Befragte ($n = 35$) haben dabei mit Bezug auf Zielvereinbarungen die Ebene Hochschulleitung–Lehrstuhl ergänzt. Diese Angaben beziehen sich vermutlich zu einem großen Teil auf Vereinbarungen im Rahmen der Verhandlungen zu den Leistungszulagen der W-Besoldung. Hierfür spricht, dass bei den übrigen Anmerkungen zu dieser Frage in den meisten Fällen explizit die Leistungszulagen im Rahmen der W-Besoldung genannt wurden.

Im Fächervergleich zeigen sich hinsichtlich der Einsatzhäufigkeit der verschiedenen Verfahren nur geringe Unterschiede.⁴³ Signifikante Abweichungen ergeben sich lediglich

⁴³ Für eine bessere Vergleichbarkeit zwischen den Teilstichproben werden wieder nur die gültigen Angaben zu den Antwortkategorien „Verfahren wird eingesetzt“ und „Verfahren wird nicht eingesetzt“ berücksichtigt. Die Angaben zur Antwortkategorie „weiß nicht“ werden als missing value behandelt.

bei den Verfahren der *indikatorenbasierten Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl* ($\chi^2(4, n = 555) = 10.52, p = .032$) sowie beim *Benchmarking verschiedener Lehrstühle innerhalb einer Fakultät* ($\chi^2(4, n = 511) = 28.37, p < .001$). Hinsichtlich der indikatorenbasierten Mittelzuweisung auf der Ebene Fakultät–Institut/Lehrstuhl gaben Befragte aus den Erziehungswissenschaften deutlich seltener als diejenigen aus der Psychologie an, dass dieses Verfahren zum Einsatz kommt. Während es bei den Erziehungswissenschaftlern nur etwa 48 % sind, liegt der Anteil bei den Psychologen mit etwa 68 % um gut 20 Prozentpunkte höher. Die Werte der anderen Fächer bewegen sich hier zwischen 58 % und 61 %. Das Benchmarking von Lehrstühlen einer Fakultät scheint nur in den Wirtschaftswissenschaften eine (geringe) Rolle zu spielen. Ein Viertel ($n = 50$) der Wirtschaftswissenschaftler, welche die Frage beantwortet haben, gab an, dass das Verfahren eingesetzt wird. Bei den anderen Fächern gibt es dagegen nur vereinzelte Nennungen (zwischen $n = 5$ und $n = 11$).

Im Folgenden wird nun die Ausgestaltung der verschiedenen Steuerungs- und Anreizverfahren in den Blick genommen. Hierzu wird dargestellt, welche Indikatoren nach Einschätzung der Befragten in den einzelnen Verfahren mit welcher Häufigkeit und Gewichtung⁴⁴ Verwendung finden sowie mögliche Auffälligkeiten diskutiert. Eine Übersicht aller Verfahren und Indikatoren (eingeschätzte Einsatzhäufigkeit und Gewichtung) findet sich in Tabelle 9.

⁴⁴ Zur Berechnung der prozentualen Anteile der Einsatzhäufigkeit eines Indikators in den einzelnen Verfahren (dichotome Variable: Indikator wird eingesetzt / nicht eingesetzt) wurden dabei die folgenden Fälle berücksichtigt (vgl. hierzu auch die Beschreibung der Fragebogenlogik in Kapitel 5.2.2 und den Fragebogen im Anhang):

- Indikator wird eingesetzt: Alle Befragten, die angaben, dass das betrachtete Steuerungs- oder Anreizverfahren eingesetzt wird und eine Angabe zur Gewichtung des Indikators in dem Verfahren gemacht haben.
- Indikator wird nicht eingesetzt: Alle Befragten, die angaben, dass das betrachtete Steuerungs- oder Anreizverfahren eingesetzt wird und angegeben haben, dass der Indikator in Steuerungs- und Anreizverfahren generell oder im betrachteten Verfahren nicht zum Einsatz kommt.

Relevant ist also jeweils die Teilstichprobe derjenigen Befragten, die angaben, dass ein Steuerungs- oder Anreizverfahren zum Einsatz kommt. Ausnahmen bilden hier die Verfahren *Leistungszulagen im Rahmen der W-Besoldung* (Teilstichprobe: alle Befragten mit W-Besoldung), sowie *Bleibe- und Berufungsverhandlungen* (alle Befragten). Die Kategorie „weiß nicht“ bei den Fragen nach der Gewichtung der Indikatoren in den einzelnen Verfahren wird als missing value behandelt.

Tabelle 9: *Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren (eingesetzte Indikatoren und deren Gewichtung)*

	Mittelzuweisung zwischen Hochschulleitung und Fakultät			Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl			Zielvereinbarungen zwischen Hochschulleitung und Fakultät			Zielvereinbarungen zwischen Fakultät und Institut/Lehrstuhl		
	<i>n</i>	Indikator wird eingesetzt (%)	<i>M (SD)</i> <i>n</i>	<i>n</i>	Indikator wird eingesetzt (%)	<i>M (SD)</i> <i>n</i>	<i>n</i>	Indikator wird eingesetzt (%)	<i>M (SD)</i> <i>n</i>	<i>n</i>	Indikator wird eingesetzt (%)	<i>M (SD)</i> <i>n</i>
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	265	41.9	3.53 (1.39) 111	273	59.0	3.83 (1.11) 161	222	39.6	3.17 (1.44) 88	119	63.0	3.76 (1.14) 75
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	260	25.0	2.68 (1.06) 65	258	34.9	2.76 (0.99) 90	229	20.1	2.50 (1.09) 46	110	39.1	2.84 (0.90) 43
Anzahl anderer Publikationen	262	30.5	2.78 (1.18) 80	260	44.2	2.81 (1.03) 115	229	25.3	2.72 (1.20) 58	113	43.4	3.00 (1.08) 49
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	266	21.4	3.42 (1.48) 57	266	34.6	3.92 (1.21) 92	226	18.6	3.48 (1.38) 42	107	35.5	3.74 (0.95) 38
Anzahl der Zitationen / Zitationsindizes	273	4.8	2.85 (1.57) 13	252	3.6	3.22 (1.20) 9	244	4.9	3.00 (1.41) 12	104	6.7	2.43 (0.54) 7
Anzahl betreuter Promotionen	263	62.4	3.40 (1.07) 164	273	67.8	3.44 (1.06) 185	208	48.6	3.37 (1.23) 101	114	55.3	3.44 (1.22) 63
Höhe der eingeworbenen Drittmittel der DFG	271	84.1	4.46 (0.86) 228	278	82.4	4.25 (0.92) 229	222	78.4	4.53 (0.82) 174	115	76.5	4.39 (1.02) 88
Höhe der sonstigen eingeworbenen Drittmittel	264	83.7	4.15 (0.92) 221	273	82.8	3.95 (0.99) 226	218	77.5	4.30 (0.89) 169	118	75.4	4.08 (1.00) 89

Fortsetzung Tabelle 9: *Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren (eingesetzte Indikatoren und deren Gewichtung)*

	Benchmarking verschiedener Lehrstühle innerhalb der Fakultät			Leistungszulagen im Rahmen der W-Besoldung			Bleibe- und Berufungsverhandlungen		
	<i>n</i>	Indikator wird eingesetzt (%)	<i>M (SD)</i> <i>n</i>	<i>n</i>	Indikator wird eingesetzt (%)	<i>M (SD)</i> <i>n</i>	<i>n</i>	Indikator wird eingesetzt (%)	<i>M (SD)</i> <i>n</i>
Anzahl Publikationen in Zeitschriften mit Peer- Review-Verfahren	61	75.4	4.07 (0.95) 46	400	51.7	3.82 (1.18) 207	448	50.9	4.04 (1.05) 228
Anzahl Publikationen in Zeitschriften ohne Peer- Review-Verfahren	57	29.8	3.18 (1.07) 17	380	30.0	2.91 (1.04) 114	452	26.8	2.83 (1.05) 121
Anzahl anderer Publika- tionen	64	40.6	3.15 (1.08) 26	380	35.3	2.96 (1.18) 134	445	32.6	2.86 (1.12) 145
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high- ranked journals	61	47.5	4.07 (1.03) 29	380	25.8	3.78 (1.26) 98	455	28.8	4.08 (1.15) 131
Anzahl der Zitationen / Zitationsindizes	59	11.9	4.29 (0.95) 7	385	5.5	3.52 (1.37) 21	484	6.8	3.36 (1.27) 33
Anzahl betreuter Promo- tionen	60	66.7	3.35 (1.15) 40	371	43.7	3.12 (1.33) 162	413	47.5	2.75 (1.27) 196
Höhe der eingeworbe- nen Drittmittel der DFG	62	85.5	4.15 (1.01) 53	399	72.4	4.47 (0.91) 289	444	75.7	4.57 (0.79) 336
Höhe der sonstigen ein- geworbenen Drittmittel	61	85.2	3.85 (1.09) 52	393	73.0	4.24 (0.97) 287	432	75.9	4.25 (0.90) 328

Anmerkung. Die Spalte *n* zeigt jeweils die Anzahl der Befragten an, die angegeben haben, dass das jeweilige Verfahren zum Einsatz kommt und Angaben zum Einsatz des jeweiligen Indikators gemacht haben. Die Prozentangaben in der Spalte „Indikator wird eingesetzt“ beziehen sich jeweils anteilig auf diese Teilgesamtheit. Die Mittelwerte beziehen sich auf die Gewichtung der Indikatoren und basieren auf einer fünfstufigen Ratingskala mit den Endpolen „sehr geringe Gewichtung“ (1) und „sehr hohe Gewichtung“ (5).

Quelle: eigene Berechnungen

Bei der Interpretation der dargestellten prozentualen Häufigkeiten und der Mittelwerte sind die bei einzelnen Verfahren (z. B. Benchmarking) und Indikatoren (z. B. Zitationen) teils geringen Fallzahlen zu berücksichtigen.

Über alle Verfahren hinweg zeigt sich die hohe Bedeutung der Drittmittelindikatoren, welche nach Einschätzung der Befragten durchgängig am häufigsten zum Einsatz kommen und auch meist am stärksten gewichtet werden.⁴⁵ Deutliche Unterschiede ergeben sich dagegen beim Indikator *Anzahl betreuter Promotionen*, der hinsichtlich der Einsatzhäufigkeit vor allem in den Verfahren der indikatorenbasierten Mittelzuweisungen eine wichtige Rolle spielt.

Betrachtet man die Einschätzungen der Befragten zu den Verfahren der indikatorenbasierten Mittelzuweisungen und den Zielvereinbarungen, fällt grundsätzlich auf, dass auf der fakultätsinternen Ebene (Fakultät–Institut/Lehrstuhl) die publikationsbasierten Indikatoren deutlich öfter zum Einsatz kommen als auf der Ebene Hochschulleitung–Fakultät. Beispielsweise gaben hinsichtlich des Indikators *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals* etwa 35 % der Befragten an, dass dieser in indikatorenbasierten Mittelzuweisungsverfahren zwischen der Fakultät und dem Institut bzw. Lehrstuhl eingesetzt wird, während dies bei der indikatorenbasierten Mittelzuweisung zwischen Hochschulleitung und Fakultät nur ca. 21 % waren. Eine noch deutlichere Differenz ergibt sich bei den Zielvereinbarungen, bei denen auf der Ebene Hochschulleitung–Fakultät nur 19 % der Befragten angaben, dass der Indikator zum Einsatz kommt. Auf der Ebene Fakultät–Institut/Lehrstuhl sind es dagegen 36 %.

Bei den *Leistungszulagen im Rahmen der W-Besoldung* sowie den *Bleibe- und Berufungsverhandlungen* ergibt sich hinsichtlich der Einsatzhäufigkeit der Indikatoren ein relativ ähnliches Bild. In beiden Verfahren finden nach Einschätzung der Befragten drittmittelbasierte Indikatoren am häufigsten Verwendung, gefolgt von begutachteten Publikationen und der Anzahl betreuter Promotionen. Andere publikationsbasierte Indikatoren werden nach Einschätzung von einem Viertel bis etwa einem Drittel der Befragten eingesetzt. Die Anzahl der Zitationen bzw. andere zitationsbasierte Kennzahlen spielen, wie in den anderen Verfahren auch, eine nur sehr untergeordnete Rolle.

Wie bereits bei der Analyse des generellen Einsatzes von Forschungsindikatoren (vgl. Kapitel 6.1.2) zeigen sich im Fächervergleich, insbesondere hinsichtlich der Einschätzungen der wirtschaftswissenschaftlichen Professorinnen und Professoren, signifikante Abweichungen. Demnach werden in den Wirtschaftswissenschaften Publikationen in hoch gerankten Zeitschriften innerhalb der indikatorenbasierten Mittelzuweisung zwi-

⁴⁵ Beim Verfahren *Benchmarking verschiedener Lehrstühle innerhalb der Fakultät* liegen bei einzelnen Indikatoren zwar höhere Mittelwerte der Gewichtung als bei den Drittmittelindikatoren vor, hier sind aber wie bereits erwähnt die teils geringen Fallzahlen zu berücksichtigen.

schen Fakultät und Institut/Lehrstuhl (49.1 %) sowie in Bleibe- und Berufungsverhandlungen (37.2 %) deutlich häufiger eingesetzt, wobei hier insbesondere die Differenz zu den Erziehungswissenschaften hoch ausfällt (14.0 bzw. 18.8 %).⁴⁶ Gleiches gilt für den Einsatz von begutachteten Publikationen als Indikator innerhalb der indikatorenbasierten Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl (67.9 % in den Wirtschaftswissenschaften gegenüber 36.4 % in den Erziehungswissenschaften).

Dagegen findet die Anzahl betreuter Promotionen in den Leistungszulagen im Rahmen der W-Besoldung (26.8 %) sowie in Bleibe- und Berufungsverhandlungen (33.1 %) in den Wirtschaftswissenschaften gegenüber den übrigen Fächern (Werte bei den Leistungszulagen zwischen 44.7 % und 54.3 % und bei den Bleibe- und Berufungsverhandlungen zwischen 47.8 % und 64.5 %) deutlich seltener Verwendung als Indikator. Ebenso werden nach Einschätzung der wirtschaftswissenschaftlichen Befragten die Anzahl von nicht-begutachteten Publikationen (13.4 % gegenüber 25.0 % bis 49.1 % bei den übrigen Fächern) sowie andere Publikationen (18.1 % gegenüber 32.1 % bis 46.9 % bei den übrigen Fächern) signifikant seltener als Indikatoren innerhalb von Bleibe- und Berufungsverhandlungen eingesetzt. Ein ähnliches Bild ergibt sich hinsichtlich dieser beiden Indikatoren bei den Leistungszulagen im Rahmen der W-Besoldung. Neben den Befragten aus den Wirtschaftswissenschaften gaben hier auch die Befragten aus der Psychologie deutlich seltener an, dass nicht-begutachtete Publikationen (Wirtschaftswissenschaften 14.5 %, Psychologie 19.2% gegenüber 40.5 % bis 48.9 % bei den übrigen Fächern) bzw. andere Publikationen (Wirtschaftswissenschaften 16.7 %, Psychologie 26.8% gegenüber 45.9 % bis 56.2 % bei den übrigen Fächern) als Indikatoren eingesetzt werden.

Die fächervergleichende Analyse lässt sich mit der Beobachtung abschließen, dass die hohe generelle Bedeutung des Drittmittelindikators in der Psychologie (vgl. Kapitel 6.1.2) auch in der Detailbetrachtung der einzelnen Verfahren bestätigt wird, während die Einschätzungen der Befragten für die Wirtschaftswissenschaften wieder eine nur geringere Bedeutung drittmittelbasierter Indikatoren nahelegen. Sowohl in den Zielvereinbarungen zwischen der Hochschulleitung und Fakultät, innerhalb der Leistungszulagen im Rahmen der W-Besoldung als auch in Bleibe- und Berufungsverhandlungen werden die Höhe der eingeworbenen DFG-Drittmittel und der sonstigen Drittmittel nach Einschätzung der Befragten in der Psychologie signifikant häufiger (zwischen 83.6 % und 92.4 %) als in den Wirtschaftswissenschaften (zwischen 59.3 % und 64.9 %) eingesetzt. Innerhalb der indikatorenbasierten Mittelzuweisung zwischen Hochschulleitung und Fakultät betrifft dies nur die DFG-Drittmittel (92.6 % in der Psychologie gegenüber 73.8 % in den

⁴⁶ Diesen und den im Folgenden berichteten Ergebnissen liegen Chi-Quadrat-Tests zugrunde (Unabhängige Variable: Fach; abhängige Variable: Einsatz Indikator ja/nein), die jeweils signifikante Abweichungen von der Nullhypothese der Unabhängigkeit der untersuchten Variablen ergaben (p zwischen $<.001$ und $.041$).

Wirtschaftswissenschaften), während sich bei den sonstigen Drittmitteln keine signifikanten Unterschiede ergeben.

Zusammenfassend lässt sich zunächst festhalten, dass alle der betrachteten Steuerungs- und Anreizverfahren in den hier untersuchten sozialwissenschaftlichen Fächern – mit unterschiedlicher Intensität – eingesetzt werden. Eine Ausnahme bildet das *Benchmarking verschiedener Lehrstühle einer Fakultät*, welches überwiegend in den Wirtschaftswissenschaften zum Einsatz kommt und in den anderen Fächern so gut wie keine Rolle spielt. Von einer flächendeckenden Implementierung der Verfahren – gemessen an dem Anteil der von den Verfahren direkt betroffenen Professorinnen und Professoren – kann anhand der Ergebnisse der Befragung nicht gesprochen werden: Der höchste Wert wird beim Verfahren der *indikatorenbasierten Mittelzuweisung zwischen Hochschulleitung und Fakultät* erreicht, bei welchem ca. 58 % der Befragten angaben, dass das Verfahren zum Einsatz kommt.

Hinsichtlich der in den einzelnen Verfahren verwendeten Indikatoren und deren jeweiliger Gewichtung zeigen sich teils deutliche Unterschiede. Die Anzahl betreuter Promotionen wird beispielsweise vor allem in den Verfahren der *indikatorenbasierten Mittelzuweisung*, sowohl auf der Ebene Hochschulleitung–Fakultät als auch zwischen der Fakultät und dem Institut/Lehrstuhl, häufig als Indikator verwendet (62.4 % bzw. 67.8 %). Publikationsbasierte Indikatoren werden dagegen insbesondere auf der fakultätsinternen Ebene in der indikatorenbasierten Mittelzuweisung und in Zielvereinbarungen eingesetzt.

Im Fächervergleich zeigen sich hinsichtlich der Ausgestaltung der Verfahren verschiedene Schwerpunktsetzungen, die überwiegend die Ergebnisse der Betrachtung des generellen Einsatzes von Forschungsindikatoren in indikatorenbasierten Verfahren (vgl. Kapitel 6.1.2) widerspiegeln. Die Fächer lassen sich demnach wieder in die Gruppen Erziehungswissenschaften, Politikwissenschaften und Soziologie auf der einen sowie Psychologie und Wirtschaftswissenschaften auf der anderen Seite einteilen. Gemeinsamkeiten ergeben sich bei der Psychologie und den Wirtschaftswissenschaften dabei vor allem hinsichtlich des Einsatzes und der Gewichtung publikationsbasierter Indikatoren, während sich in Bezug auf die Verwendung von Drittmittelindikatoren zwischen diesen beiden Fächern sogar die größten Unterschiede zeigen.

Inwiefern sich ähnliche fachliche Differenzen im Hinblick auf die Einschätzung der Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs in der fachspezifischen Scientific Community zeigen, ist unter anderem Gegenstand des folgenden Kapitels.

6.1.4 Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs in der Scientific Community

Zunächst werden wieder die Kennzahlen für die gesamte Stichprobe dargestellt, bevor die fachspezifischen Unterschiede in den Blick genommen werden. Für die Gesamtstichprobe zeigen sich dabei hinsichtlich der eingeschätzten Bedeutung der abgefragten Aspekte der Forschungsarbeit für den Reputationszuwachs in der Scientific Community teils deutliche Unterschiede. Die Mittelwerte und andere Verteilungsparameter sind in Tabelle 10 aufgeführt.

Tabelle 10: *Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs*

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>
Publikationen in Zeitschriften mit Peer-Review-Verfahren	656	4.42	0.76	-1.41	2.24
Publikationen in Zeitschriften ohne Peer-Review-Verfahren	643	2.28	1.04	0.34	-0.74
Andere Publikationen	648	2.71	1.08	0.15	-0.61
Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	651	4.29	1.01	-1.45	1.46
Anzahl der Zitationen / Zitationsindizes	643	3.30	1.20	-0.31	-0.76
Anzahl betreuter Promotionen	639	2.46	1.00	0.41	-0.26
Höhe der eingeworbenen Drittmittel der DFG	656	4.09	1.01	-0.97	0.27
Höhe der sonstigen eingeworbenen Drittmittel	653	3.60	1.04	-0.46	-0.35

Anmerkung. Die Werte basieren auf einer fünfstufigen Ratingskala mit den Endpolen „sehr gering“ (1) und „sehr hoch“ (5).

Quelle: eigene Berechnungen

Die höchste Bedeutung wird klar *peer-reviewed Publikationen* ($M = 4.42$, $SD = 0.76$) sowie *Publikationen in hoch gerankten Zeitschriften* ($M = 4.29$, $SD = 1.01$) zugesprochen – das Gutachterverfahren sowie (zitationsbasierte) Zeitschriftenrankings werden anscheinend trotz aller Kritik (vgl. Kapitel 2.3.2) als Qualitätskriterium für eine Publikation gewertet. Veröffentlichungen in entsprechenden Zeitschriften sind daher nach Einschätzung der Befragten mit einem Reputationszuwachs in der Scientific Community verknüpft. Die durchgängig hoch eingeschätzte Bedeutung zeigt sich auch in der stark links-schiefen Verteilung bei diesen beiden Variablen. Knapp 90 % bzw. 81 % der Befragten stufen die Bedeutung von peer-reviewed Publikationen bzw. von Publikationen in hoch gerankten Zeitschriften auf der fünfstufigen Ratingskala mit den Werten 4 oder 5 als (sehr) hoch ein.

Nicht ganz so stark, aber immer noch deutlich, fällt die Linksschiefe bei der Variablen *Höhe der eingeworbenen Drittmittel der DFG* aus. Etwa 75 % der Befragten sprechen diesem Aspekt eine (sehr) hohe Bedeutung für den Reputationszuwachs zu. Der *Höhe der sonstigen eingeworbenen Drittmittel* wird gegenüber den DFG-Drittmitteln mit $M = 3.60$ ($SD = 1.04$) eine deutlich geringere Bedeutung zugesprochen. Die Verteilung nähert sich einer Normalverteilung an, weicht aber nach Überprüfung mit dem Kolmogorov-Smirnov-Test noch immer signifikant ($p < .001$) von dieser ab. Die höhere Bedeutung von DFG-Drittmitteln für den Reputationszuwachs ist vermutlich zum einen auf die (gewachsene) Reputation der DFG selbst und der durch diese geförderten Projekte zurückzuführen. Zum anderen geht der Vergabe von DFG-Drittmitteln in der Regel ein anerkanntes Gutachterverfahren voraus, was bei anderen Drittmitteln (z. B. aus der Privatwirtschaft) nicht immer der Fall ist. Wie bereits bei den Publikationen ist zu vermuten, dass das Gutachterverfahren in der Scientific Community als Qualitätskriterium gewertet wird, welches mit erhöhtem Reputationsgewinn verbunden ist.

Die *Anzahl der Zitationen bzw. Zitationsindizes*, welche als Indikator in Steuerungs- und Anreizverfahren nur selten Verwendung finden (vgl. Kapitel 6.1.2), spielen für die Reputation in der Scientific Community durchaus eine bedeutende Rolle. Der Mittelwert liegt bei $M = 3.30$ ($SD = 1.20$), wobei die Verteilung eine leichte Linksschiefe aufweist. Weniger Bedeutung wird dagegen *nicht-begutachteten Publikationen* ($M = 2.28$, $SD = 1.04$) sowie den *anderen Publikationen* ($M = 2.71$, $SD = 1.08$) beigemessen. Auch die *Anzahl der betreuten Promotionen* ist für den Reputationszuwachs – zumindest in der fächerübergreifenden Betrachtung – mit einem Wert von $M = 2.46$ ($SD = 1.00$) nur von geringerer Bedeutung. Dies erscheint nachvollziehbar, da der Betreuungsschlüssel oder auch der Inhalt der betreuten Arbeiten den wenigsten Kolleginnen und Kollegen in der Scientific Community bekannt und auch für die meisten von geringerem Interesse sein dürfte. Inwiefern sich eine Professorin oder ein Professor in der wissenschaftlichen Nachwuchsförderung engagiert, ist vermutlich eher auf der hochschulinternen Ebene von Interesse und mit entsprechender Anerkennung verknüpft.

Auch bei der Frage nach der Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs hatten die Befragten die Möglichkeit, eigene Aspekte zu ergänzen. Am häufigsten ($n = 13$) wurde hierbei auf Leistungen in der universitären Selbstverwaltung, in der Gremienarbeit (z. B. DFG-Gutachter) und auf besondere Funktionen bzw. besonderes Engagement innerhalb der Scientific Community (z. B. Engagement in Fachgesellschaften, Sprecher von Forschungsnetzwerken, Herausgeberschaften) verwiesen. Die Präsenz sowie Beiträge auf Tagungen, Konferenzen oder Kongressen wurde von $n = 8$ Befragten als reputationsträchtiges Kriterium ergänzt. In absteigender Reihenfolge wurden weiter Preise und Ehrungen ($n = 4$), die internationale Ausrichtung ($n = 3$) sowie die Qualität der Forschung generell ($n = 3$) erwähnt. Wie auch schon

bei den vorangegangenen Fragen wurden zudem vereinzelt ($n = 5$) lehrbezogene Aspekte (Lehrleistung, Studierende etc.) genannt.

In der fachspezifischen Perspektive werden teils deutliche Differenzen zwischen den Einschätzungen der Professorinnen und Professoren sichtbar. Die Mittelwerte je Fach sind in Tabelle 11 und Abbildung 6 aufgeführt.

Tabelle 11: *Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs im Fächervergleich*

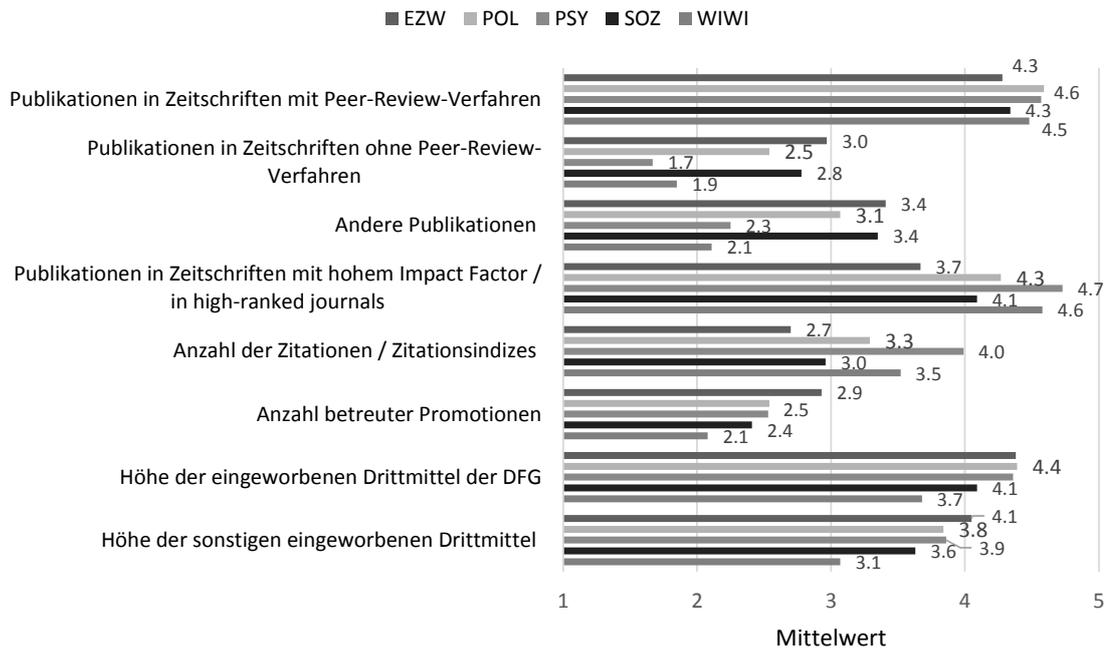
	EZW	POL	PSY	SOZ	WIWI
	<i>M (SD)</i>				
Publikationen in Zeitschriften mit Peer-Review-Verfahren	4.28 (0.79)	4.59 (0.63)	4.57 (0.65)	4.34 (0.86)	4.48 (0.73)
Publikationen in Zeitschriften ohne Peer-Review-Verfahren	2.97 (0.89)	2.54 (1.00)	1.67 (0.82)	2.78 (1.02)	1.85 (0.81)
Andere Publikationen	3.41 (0.85)	3.07 (0.87)	2.25 (0.88)	3.35 (0.94)	2.11 (0.91)
Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	3.67 (1.22)	4.27 (1.00)	4.73 (0.58)	4.09 (0.92)	4.58 (0.84)
Anzahl der Zitationen / Zitationsindizes	2.70 (1.27)	3.29 (1.14)	3.99 (0.93)	2.96 (1.03)	3.52 (1.11)
Anzahl betreuter Promotionen	2.93 (1.02)	2.54 (0.91)	2.53 (0.86)	2.41 (0.96)	2.08 (0.93)
Höhe der eingeworbenen Drittmittel der DFG	4.38 (0.97)	4.39 (0.82)	4.36 (0.84)	4.09 (0.94)	3.68 (1.04)
Höhe der sonstigen eingeworbenen Drittmittel	4.05 (0.88)	3.84 (0.92)	3.86 (0.92)	3.63 (1.03)	3.07 (1.01)

Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften; n je Fach (min – max): $n_{EZW} = 149-155$, $n_{POL} = 54-56$, $n_{PSY} = 117-119$, $n_{SOZ} = 74-76$, $n_{WIWI} = 226-233$.

Quelle: eigene Berechnungen

Zur Überprüfung der Signifikanz der Mittelwertunterschiede wurden erneut einfaktorielle Varianzanalysen gerechnet, welche für alle abgefragten Items signifikante Gruppenunterschiede anzeigten und im Folgenden im Detail berichtet werden.⁴⁷

⁴⁷ Um im Detail zu prüfen, auf welche Gruppenunterschiede die signifikanten Ergebnisse der Varianzanalysen beruhen, wurden im Anschluss wieder Post-Hoc-Mehrfachvergleiche durchgeführt (Scheffé-Prozedur bzw. bei Verletzung der Homogenitätsannahme der Varianzen der Games-Howell-Test). Alle berichteten Differenzen sind auf dem 5%-Niveau signifikant.

Abbildung 6: *Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs im Fächervergleich*

Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften; n je Fach (min – max): $n_{EZW} = 149\text{--}155$, $n_{POL} = 54\text{--}56$, $n_{PSY} = 117\text{--}119$, $n_{SOZ} = 74\text{--}76$, $n_{WIWI} = 226\text{--}233$.

Quelle: eigene Berechnungen

Am geringsten fallen die Unterschiede bei den *Publikationen in Zeitschriften mit Peer-Review-Verfahren* aus, deren Bedeutung für den Reputationszuwachs über alle Fächer hinweg sehr hoch eingeschätzt wird. Die durchgeführte Varianzanalyse ergab dennoch einen signifikanten Mittelwertunterschied ($F(4, 633) = 3.87$, $p = .004$), der auf der Differenz zwischen den Einschätzungen der Befragten aus den Erziehungswissenschaften ($M = 4.28$, $SD = 0.79$) und aus der Psychologie ($M = 4.57$, $SD = 0.65$) beruht.

Bei den für den Reputationszuwachs insgesamt ebenfalls hoch eingeschätzten *Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals* zeigen sich dagegen bereits deutlichere Fachunterschiede ($F(4, 629) = 30.11$, $p < .001$). Insbesondere in den Erziehungswissenschaften ($M = 3.67$, $SD = 1.22$), aber auch in der Soziologie ($M = 4.09$, $SD = 0.92$), wird die Bedeutung dieser Publikationsformate geringer eingeschätzt als in den anderen Fächern. Zu beachten ist allerdings, dass die Mittelwerte bei diesen beiden Fächern dennoch recht hoch ausfallen. Publikationen in hoch gerankten Zeitschriften oder in Zeitschriften mit hohem Impact Factor sind also auch in den Erziehungswissenschaften und der Soziologie von hoher Bedeutung für den Reputationszuwachs, auch wenn die Relevanz nicht ganz so hoch eingeschätzt wird wie in der Psychologie ($M = 4.73$, $SD = 0.58$) oder den Wirtschaftswissenschaften ($M = 4.58$, $SD = 0.84$).

Die unterschiedlichen Publikationskulturen der Fächer machen sich auch bei den Einschätzungen zu den *nicht-begutachteten* und den *anderen Publikationen* bemerkbar, welche zwischen den Befragten der verschiedenen Fächer stark differieren ($F(4, 620) = 59.13, p < .001$ bzw. $F(4, 625) = 69.13, p < .001$). Wieder ist eine Zweiteilung zwischen den Fächern Psychologie und Wirtschaftswissenschaften auf der einen und den übrigen Fächern auf der anderen Seite zu erkennen. Erstere schätzen die Bedeutung dieser beiden Publikationstypen für den Reputationszuwachs deutlich geringer ein als die Befragten der anderen Fächer. Insgesamt liegt der eingeschätzte Stellenwert dieser Publikationstypen für den Reputationszuwachs aber über alle Fächer hinweg unter dem der begutachteten Publikationen und der Publikationen in hochgerankten Zeitschriften.

Die fachliche Zweiteilung setzt sich auch bei den *Zitationen* bzw. den *Zitationsindizes* fort ($F(4, 621) = 25.90, p < .001$). Die Bedeutung von diesen wird insbesondere in der Psychologie ($M = 3.99, SD = 0.93$) und den Wirtschaftswissenschaften ($M = 3.52, SD = 1.11$) hoch eingeschätzt, während in den Erziehungswissenschaften ($M = 2.70, SD = 1.27$) und der Soziologie ($M = 2.96, SD = 1.03$) die Relevanz für den Reputationszuwachs geringer eingestuft wird. Die Politikwissenschaften liegen mit einem Wert von $M = 3.29 (SD = 1.14)$ zwischen den beiden Gruppierungen.

Bei der *Anzahl der betreuten Promotionen* fallen die Fachunterschiede weniger stark aus ($F(4, 617) = 18.98, p < .001$). Hervorzuheben sind hier der signifikant höhere Wert in den Erziehungswissenschaften ($M = 2.93, SD = 1.02$) und der signifikant niedrigere Wert in den Wirtschaftswissenschaften ($M = 2.08, SD = 0.93$). In den Erziehungswissenschaften scheint also ein erhöhtes Engagement in der Nachwuchsförderung mehr als in anderen Fächern mit einer höheren Anerkennung durch Fachkollegen verbunden zu sein.

Hinsichtlich der Bedeutung von erfolgreichen Drittmittelinwerbungen für die Reputation in der Scientific Community ergeben sich im Fächervergleich für *DFG-Drittmittel* und die *sonstigen Drittmittel* ähnliche Ergebnisse. Die signifikanten Gruppenunterschiede ($F(4, 633) = 17.87, p < .001$ bzw. $F(4, 630) = 29.37, p < .001$) sind dabei auf die deutlich geringere Bedeutung von Drittmitteln für den Reputationszuwachs in den Wirtschaftswissenschaften ($M = 3.68, SD = 1.04$ bzw. $M = 3.07, SD = 1.01$) zurückzuführen. Zudem wird der Stellenwert von DFG-Drittmitteln über alle Fächer hinweg höher eingeschätzt als der von sonstigen Drittmitteln.

Insgesamt machen die Ergebnisse die hohe Bedeutung von begutachteten Publikationen und Veröffentlichungen in hoch gerankten Zeitschriften sowie von erfolgreichen Drittmittelinwerbungen für den Reputationszuwachs in der Scientific Community deutlich. Obwohl sich zwischen den Fächern jeweils signifikante Unterschiede in der eingeschätzten Bedeutung zeigen, liegen die Werte bei diesen Forschungstätigkeiten fast

durchgehend auf hohem Niveau. Bei den anderen publikationsbasierten Items sowie den Zitationen und Zitationsindizes ergeben sich dagegen teils erhebliche Differenzen in den Einschätzungen der Befragten, welche auch die verschiedenen Publikationskulturen der einzelnen Fächer widerspiegeln (vgl. zu den Publikationskulturen verschiedener Fächer Alexander von Humboldt Stiftung 2009). Die Anzahl der betreuten Promotionen ist für den Reputationszuwachs sowohl in der Gesamtbetrachtung als auch in den einzelnen Fächern im Vergleich zu anderen Aspekten der Forschungsarbeit von eher geringerer Bedeutung. Allerdings zeigen sich auch hier signifikante Unterschiede zwischen den Fächern.

Ob eine Forschungstätigkeit mit hohem Reputationsgewinn verknüpft ist, sagt noch nichts darüber aus, ob diese auch einen Beitrag für das der Forschungsarbeit zugrunde liegende Ziel – dem wissenschaftlichen Erkenntnisgewinn – leistet. Publikationen in hoch gerankten Zeitschriften mögen beispielsweise mit hohem Reputationsgewinn belohnt werden, doch daraus muss nicht zwangsläufig folgen, dass in diesen gerade diejenigen Publikationen erscheinen, die für das Fortschreiten des Erkenntnisgewinns von besonderer Bedeutung sind (vgl. hierzu auch die Diskussion möglicher nichtintendierter Effekte indikatorenbasierter Leistungsmessung in Kapitel 2.3.2, insbesondere die Befürchtung vermehrter Mainstreamforschung durch eine zu starke Fokussierung auf Zeitschriftenrankings). Im nächsten Kapitel wird nun die Perspektive auf eben diesen wissenschaftlichen Erkenntnisgewinn als Ziel der Forschungsarbeit gerichtet. Hierfür werden die Einschätzungen der Befragten zum Beitrag der einzelnen Aspekte der Forschungsarbeit für den wissenschaftlichen Erkenntnisgewinn näher analysiert.

6.1.5 Beitrag verschiedener Aspekte der Forschungsarbeit für den wissenschaftlichen Erkenntnisgewinn

In der Gesamtbetrachtung (vgl. Tabelle 12) fällt zunächst auf, dass auch für den wissenschaftlichen Erkenntnisgewinn der Beitrag von *begutachteten Publikationen* ($M = 4.02$, $SD = 0.96$) und *Publikationen in hochgerankten Zeitschriften* ($M = 3.67$, $SD = 1.01$) höher eingeschätzt wird als der von *nicht-begutachteten* ($M = 2.77$, $SD = 1.11$) und *anderen Publikationen* ($M = 3.11$, $SD = 1.08$). Die Differenzen fallen allerdings nicht so hoch aus wie bei der eingeschätzten Bedeutung der verschiedenen Publikationsformate für den Reputationszuwachs. Der Beitrag von *Zitationen* für den Erkenntnisgewinn im wissenschaftlichen Forschungsprozess wird mit einem Wert von $M = 2.84$ ($SD = 1.22$) relativ niedrig eingeschätzt. Wenn man davon ausgeht, dass Zitationen durch die Bezugnahme auf andere Publikationen und der damit einhergehenden Verknüpfung und Bündelung verschiedener Forschungsergebnisse ein zentrales Element der wissenschaftlichen Forschungsarbeit darstellen, erscheint dieses Ergebnis überraschend. Möglicherweise spielt hier eine Rolle, dass nicht alle Zitationen Veröffentlichungen betreffen, die

für den Erkenntnisgewinn als besonders relevant anzusehen sind, sondern oft auch weniger wichtige Arbeiten oder sogar für das bearbeitete Themenfeld nicht relevante Arbeiten (Stichwort Selbstzitation und Zitationskartelle) zitiert werden. Angenommen werden könnte also, dass aufgrund eines teilweise ‚inflationären‘ Gebrauchs von Zitationen in wissenschaftlichen Publikationen, die wirklich relevanten Zitationen in der Masse von anderen Zitationen untergehen und der Beitrag für den Erkenntnisgewinn im Mittel eher entsprechend gering eingeschätzt wird.

Tabelle 12: *Beitrag verschiedener Aspekte der Forschungsarbeit für den Erkenntnisgewinn*

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>
Publikationen in Zeitschriften mit Peer-Review-Verfahren	625	4.02	0.96	-0.92	0.57
Publikationen in Zeitschriften ohne Peer-Review-Verfahren	620	2.77	1.11	0.08	-0.70
Andere Publikationen	620	3.11	1.08	-0.12	-0.53
Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	622	3.67	1.01	-0.65	-0.17
Zitationen	605	2.84	1.22	0.02	-0.98
Betreuung von Promotionen	600	3.04	1.11	-0.04	-0.66
Durchführung von Drittmittelprojekten der DFG	626	3.55	1.11	-0.47	-0.44
Durchführung von sonstigen Drittmittelprojekten	626	3.24	1.11	-0.26	-0.64

Anmerkung. Die Werte basieren auf einer fünfstufigen Ratingskala mit den Endpolen „sehr gering“ (1) und „sehr hoch“ (5).

Quelle: eigene Berechnungen

Die Einschätzung der Befragten hinsichtlich der *Betreuung von Promotionen* liegt mit einem Wert von $M = 3.04$ ($SD = 1.11$) im mittleren Bereich der Skala. Mit einer Schiefe von -0.04 ergibt sich eine nahezu vollkommen symmetrische Verteilung der Antworthäufigkeiten um den Wert 3 der Skala.

Die Durchführung von Drittmittelprojekten wird insgesamt als durchaus bedeutsam für den wissenschaftlichen Erkenntnisgewinn angesehen, wobei auch hier wieder die Relevanz von *DFG-Drittmitteln* ($M = 3.55$, $SD = 1.11$) höher als die der *sonstigen Drittmittel* ($M = 3.24$, $SD = 1.11$) eingestuft wird. Dies erscheint nachvollziehbar, da die sonstigen Drittmittel z. B. auch industrielle Auftragsforschung umfassen können, deren Beitrag für den wissenschaftlichen Erkenntnisgewinn sicherlich als geringer anzusehen ist als der von Projekten, die z. B. im Bereich der Grundlagenforschung angesiedelt sind.

Als weitere für den wissenschaftlichen Erkenntnisgewinn relevante Aspekte der Forschungsarbeit wurden von den Befragten generell die inhaltliche Qualität der Forschung ($n = 4$), z. B. die Qualität von Publikationen unabhängig vom Format, sowie – damit wohl verbunden – ausreichend Zeit für die Forschungsarbeit ($n = 4$) genannt. Auch der Austausch mit Kollegen auf Tagungen und Konferenzen ($n = 4$) sowie der internationale Austausch ($n = 3$) wurden von mehreren Befragten erwähnt.

In Tabelle 13 und Abbildung 7 sind nun die Einschätzungen der Befragten im Fächervergleich dargestellt. Die Mittelwertunterschiede wurden wieder mit Hilfe von einfaktoriellen Varianzanalysen auf Signifikanz geprüft und im Anschluss Post-Hoc-Mehrfachvergleiche (Scheffé-Prozedur bzw. Games-Howell-Test) durchgeführt.

Tabelle 13: *Beitrag verschiedener Aspekte der Forschungsarbeit für den Erkenntnisgewinn im Fächervergleich*

	EZW	POL	PSY	SOZ	WIWI
	<i>M (SD)</i>				
Publikationen in Zeitschriften mit Peer-Review-Verfahren	3.92 (1.03)	3.94 (0.86)	4.30 (0.95)	3.97 (0.94)	4.02 (0.89)
Publikationen in Zeitschriften ohne Peer-Review-Verfahren	3.30 (1.00)	3.08 (1.12)	2.33 (1.00)	3.09 (1.10)	2.40 (1.02)
Andere Publikationen	3.64 (0.90)	3.71 (0.97)	2.64 (0.97)	3.48 (1.07)	2.69 (1.03)
Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals	3.24 (1.22)	3.65 (1.15)	3.85 (0.95)	3.53 (0.99)	3.99 (0.98)
Zitationen	2.51 (1.18)	2.86 (1.02)	2.97 (1.20)	2.72 (1.20)	3.08 (1.26)
Betreuung von Promotionen	3.37 (1.08)	3.07 (0.97)	3.03 (1.09)	3.04 (1.20)	2.79 (1.09)
Durchführung von Drittmittelprojekten der DFG	3.98 (1.05)	3.62 (1.16)	3.82 (1.01)	3.53 (1.08)	3.09 (1.04)
Durchführung von sonstigen Drittmittelprojekten	3.73 (0.99)	3.35 (1.12)	3.43 (1.07)	3.25 (1.08)	2.74 (1.05)

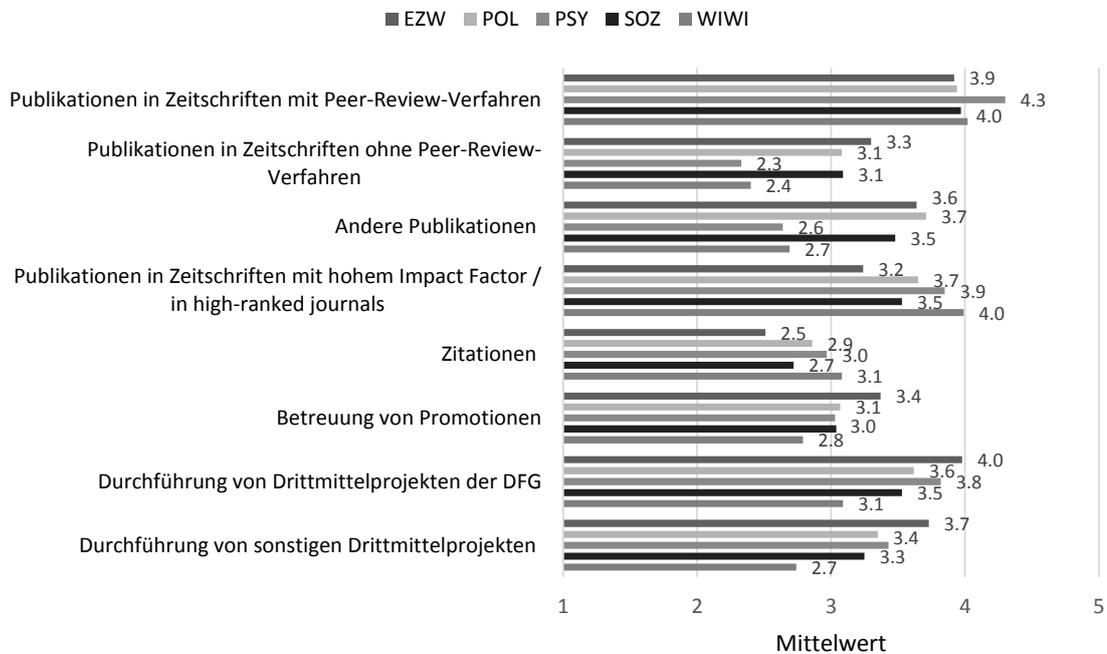
Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften; n je Fach (min – max): $n_{EZW} = 146-150$, $n_{POL} = 44-53$, $n_{PSY} = 112-115$, $n_{SOZ} = 69-75$, $n_{WIWI} = 211-219$.

Quelle: eigene Berechnungen

Zwischen den Fächern ergeben sich bei allen abgefragten Items signifikante Unterschiede. Diese fallen bei den *peer-reviewed Publikationen* relativ moderat aus ($F(4, 604) = 2.98$, $p = .019$). Ein signifikanter Mittelwertunterschied zeigt sich hier lediglich zwischen den Einschätzungen der Erziehungswissenschaftler ($M = 3.92$, $SD = 1.03$)

und der Psychologen ($M = 4.30, SD = 0.95$). Insgesamt wird diesen Publikationen aber über alle Fächer hinweg eine hohe Relevanz für den wissenschaftlichen Erkenntnisgewinn zugesprochen.

Abbildung 7: *Beitrag verschiedener Aspekte der Forschungsarbeit für den Erkenntnisgewinn im Fächervergleich*



Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften; n je Fach (min – max): $n_{EZW} = 146-150, n_{POL} = 44-53, n_{PSY} = 112-115, n_{SOZ} = 69-75, n_{WIWI} = 211-219$.

Quelle: eigene Berechnungen

Deutlichere Unterschiede – und wieder die bereits bekannte fachliche Zweiteilung von Psychologie und Wirtschaftswissenschaften auf der einen und den übrigen Fächern auf der anderen Seite – zeigen sich bei den *Publikationen in Zeitschriften ohne Peer-Review-Verfahren* ($F(4, 599) = 25.06, p < .001$). Dass der eingeschätzte Beitrag für den Erkenntnisgewinn in der Psychologie ($M = 2.33, SD = 1.00$) und den Wirtschaftswissenschaften ($M = 2.40, SD = 1.02$) deutlich geringer ausfällt, ist angesichts der bereits beschriebenen Unterschiede in den Publikationskulturen (starke Fokussierung auf peer-reviewed Zeitschriften und geringere Bedeutung anderer Publikationsformate) wenig verwunderlich. Ähnlich verhält es sich bei den *anderen Publikationen* ($F(4, 599) = 34.03, p < .001$), denen in den Erziehungswissenschaften ($M = 3.64, SD = 0.90$), den Politikwissenschaften ($M = 3.71, SD = 0.97$) und der Soziologie ($M = 3.48, SD = 1.07$) eine recht hohe Relevanz für den wissenschaftlichen Erkenntnisgewinn zugesprochen wird,

während die Psychologie und die Wirtschaftswissenschaften mit Werten von $M = 2.64$ ($SD = 0.97$) und $M = 2.69$ ($SD = 1.03$) wieder deutlich zurück liegen.

Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals wird über alle Fächer hinweg ein relativ hoher Beitrag für den Erkenntnisgewinn zugesprochen, auch wenn der Wert in den Erziehungswissenschaften mit $M = 3.24$ ($SD = 1.22$) etwas geringer ausfällt. In der Psychologie ($M = 3.85$, $SD = 0.95$) und den Wirtschaftswissenschaften ($M = 3.99$, $SD = 0.98$) werden dagegen wieder die höchsten Werte erzielt ($F(4, 602) = 12.20$, $p < .001$). Bei den *Zitationen* folgen die Einschätzungen derselben fachlichen Reihenfolge. Mit Werten zwischen $M = 2.51$ ($SD = 1.18$) und $M = 3.08$ ($SD = 1.26$) liegen diese nur auf deutlich niedrigerem Niveau ($F(4, 585) = 5.15$, $p < .001$).

Der *Betreuung von Promotionen* wird vor allem in den Erziehungswissenschaften ein hoher Beitrag für den wissenschaftlichen Erkenntnisgewinn zugesprochen ($M = 3.37$, $SD = 1.08$), während die Wirtschaftswissenschaften mit $M = 2.79$ ($SD = 1.09$) den niedrigsten Wert erreichen ($F(4, 580) = 6.05$, $p < .001$). Auch bei der Beurteilung des Beitrags von *Drittmittelprojekten der DFG* und *sonstigen Drittmittelprojekten* fallen die Werte in den Wirtschaftswissenschaften mit $M = 3.09$ ($SD = 1.04$) bzw. $M = 2.74$ ($SD = 1.09$) am niedrigsten aus, was vermutlich durch die bereits angesprochene stärkere Verbreitung von industrieller Auftragsforschung in diesem Fach begründet ist. Die Werte bei den übrigen Fächern liegen sowohl bei den *Drittmittelprojekten der DFG* ($F(4, 605) = 18.53$, $p < .001$) als auch bei den *sonstigen Drittmitteln* ($F(4, 605) = 21.63$, $p < .001$) signifikant über denen der Wirtschaftswissenschaften.

Für dieses Teilkapitel kann resümiert werden, dass die Befragungsergebnisse zum einen den überraschenden Befund liefern, dass Zitationen nach Einschätzung der befragten Professorinnen und Professoren einen nur eher mäßigen Beitrag für den wissenschaftlichen Erkenntnisgewinn leisten. Zum anderen zeigt sich, dass über alle Fächer hinweg und insbesondere in der Psychologie und den Wirtschaftswissenschaften begutachteten Publikationen eine höhere Relevanz für den Erkenntnisgewinn als anderen Publikationsformaten attestiert wird. Eine naheliegende Erklärung hierfür ist, dass das akademische Gutachterverfahren anscheinend trotz mancher Kritik als Qualitätskriterium anerkannt wird, indem es durch seine Selektionsfunktion einen Beitrag für den wissenschaftlichen Erkenntnisgewinn leistet. Eine alternative Erklärung hierfür wäre, dass aufgrund einer zunehmenden – auch von außen forcierten – Orientierung an Peer-Review-Formaten, primär versucht wird, Publikationen in eben solchen Zeitschriften zu platzieren. Für andere Veröffentlichungsformate eingereicht oder anderweitig publiziert werden dann nur noch solche Publikationen, die durch das Gutachterverfahren ausselektiert wurden bzw. bei denen sich die Autorin oder der Autor von vornherein keine Aus-

sicht auf Erfolg im Gutachterverfahren versprochen hat. Das heißt, nicht primär die Validität des Begutachtungsverfahrens wäre für die Qualität der in den entsprechenden Zeitschriften erscheinenden Aufsätze verantwortlich, sondern schlicht die Tatsache, dass im Vorfeld durch die Autorinnen und Autoren vorselektiert wird. Die für den wissenschaftlichen Erkenntnisgewinn bedeutsamen Publikationen finden sich dann entsprechend überproportional in peer-reviewed Publikationen wieder.

Nach den Einschätzungen der Befragten zur Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs in der Scientific Community und deren Beitrag für den wissenschaftlichen Erkenntnisgewinn soll es im nächsten Kapitel nun um den konkreten Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren gehen.

6.1.6 Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit

Um das konkrete Forschungsverhalten der befragten Professorinnen und Professoren in den Blick zu bekommen, wurde im Fragebogen danach gefragt, welchen subjektiven Stellenwert verschiedene Forschungstätigkeiten in der Forschungsarbeit der letzten drei Jahre hatten. Der aufgebrauchte absolute Zeitaufwand für die einzelnen Tätigkeiten wurde bewusst nicht erfragt, da manche Tätigkeiten (z. B. Verfassen eines Drittmittelantrages) generell mehr Zeit als andere Tätigkeiten (z. B. Verfassen eines kurzen Artikels) in Anspruch nehmen. Die aufgewendete Zeit kann daher nicht als Indikator für die subjektive Relevanz der einzelnen Tätigkeiten in der Forschungsarbeit angesehen werden (vgl. zur Operationalisierung Kapitel 5.2.1). Die nachfolgend präsentierten Werte sind daher nicht als aufgebrauchter Arbeitsaufwand je Tätigkeit zu interpretieren, sondern – entsprechend der Fragenformulierung – als der Stellenwert, den die Professorinnen und Professoren diesen Tätigkeiten in den letzten drei Jahren ihrer Forschungsarbeit aus eigener Sicht beigemessen haben.

In Tabelle 14 finden sich die Verteilungskennzahlen je Item für die auswertbare Gesamtstichprobe. Dem *Publizieren in Zeitschriften mit Peer-Review-Verfahren* wird mit einem Wert von $M = 3.97$ ($SD = 1.13$) im Mittel der höchste Stellenwert zugeschrieben. Bei einer stark linksschiefen Verteilung wählten etwa 72 % der Befragten die Skalenwerte 4 oder 5 (sehr hoher Stellenwert). Dass aber durchaus nicht alle Professorinnen und Professoren auf die Veröffentlichungen von begutachteten Publikationen fokussiert sind, zeigt sich daran, dass immerhin 13.5 % ($n = 86$) der Befragten diesem Publikationsformat nur einen geringen Stellenwert (Skalenwerte 1 oder 2) in ihrer Forschungsarbeit zuschreiben. Für die Mehrheit der Professorinnen und Professoren ist das Verfassen von Artikeln zur Veröffentlichung in begutachteten Zeitschriften aber – wenig überraschend – ein wesentlicher Aspekt der Forschungsarbeit.

Tabelle 14: Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>
Publizieren in Zeitschriften mit Peer-Review-Verfahren	635	3.97	1.13	-0.94	-0.02
Publizieren in Zeitschriften ohne Peer-Review-Verfahren	630	2.56	1.22	0.18	-1.11
Verfassen anderer Publikationen	633	3.09	1.27	-0.20	-0.99
Publizieren in Zeitschriften mit hohem Impact Factor / in high-ranked journals	630	3.28	1.40	-0.29	-1.18
Publizieren als Koautor / in Mehrautorenschaft	628	3.83	1.03	-0.65	-0.22
Betreuung von Promotionen	621	3.74	1.01	-0.53	-0.27
Einwerben von Drittmitteln der DFG	631	3.00	1.36	-0.02	-1.19
Einwerben von sonstigen Drittmitteln	628	3.29	1.40	-0.33	-1.19
Betreiben von Grundlagenforschung	621	3.70	1.25	-0.64	-0.66
Betreiben von anwendungsbezogener Forschung	621	3.51	1.30	-0.59	-0.76

Anmerkung. Die Werte basieren auf einer fünfstufigen Ratingskala mit den Endpolen „sehr geringer Stellenwert“ (1) und „sehr hoher Stellenwert“ (5).

Quelle: eigene Berechnungen

Dem *Publizieren in Zeitschriften ohne Peer-Review-Verfahren* wird dagegen im Mittel der geringste Stellenwert zugeschrieben ($M = 2.56$, $SD = 1.22$). Für nur etwa 5 % ($n = 30$) der Befragten hat diese Tätigkeit nach eigenen Angaben einen sehr hohen Stellenwert, während die übrigen 95 % relativ gleichmäßig über die anderen Skalenpunkte streuen (21.7–26.3 %). Der subjektive Stellenwert des Publizierens in nicht-begutachteten Zeitschriften wird also erwartungsgemäß insgesamt deutlich niedriger als der des Publizierens in begutachteten Zeitschriften eingeschätzt.

Das *Verfassen anderer Publikationen* sowie das *Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals* werden mit Werten von $M = 3.09$ ($SD = 1.27$) bzw. $M = 3.28$ ($SD = 1.40$) in ihrer Relevanz für die eigene Forschungsarbeit durchschnittlich wieder höher eingestuft. In beiden Fällen streuen die Werte relativ stark. Der höhere Wert bei den Publikationen in hochgerankten Zeitschriften beruht dabei vor allem auf dem deutlich höheren Anteil an Befragten, die diesen einen sehr hohen Stellenwert (Skalenpunkt 5) zuschreiben: Während es hier 25.4 % sind, sahen hinsichtlich des Verfassens anderer Publikationen nur 13.9 % der Befragten einen sehr hohen Stellenwert in ihrer Forschungsarbeit.

Veröffentlichungen gemeinsam mit einem oder mehreren Autorinnen und Autoren scheinen eher die Regel als die Ausnahme zu sein. Dem *Publizieren als Koautor bzw. in Mehrautorenschaft* wird mit einem Wert von $M = 3.83$ ($SD = 1.03$) insgesamt ein hoher

Stellenwert zugeschrieben. Nur 11.5 % der Befragten gaben an, dass diese Tätigkeit für sie einen (sehr) geringen Stellenwert in der Forschungsarbeit der letzten drei Jahre hatte. Die *Betreuung von Promotionen* hat ebenfalls einen festen Stellenwert in der Forschungsarbeit der Professorinnen und Professoren. Bei einem durchschnittlichen Wert von $M = 3.74$ ($SD = 1.01$) sieht die Mehrheit der Befragten eine hohe Relevanz dieser Tätigkeit in ihrer Arbeit. Für wieder nur einen kleinen Teil von 11.6 % ist der Stellenwert dagegen (sehr) gering.

Hinsichtlich des Stellenwerts der *Einwerbung von Drittmitteln* ergeben sich leichte Unterschiede zwischen DFG-Drittmitteln und den sonstigen Drittmitteln. Während sich die Angaben bei den DFG-Drittmitteln relativ gleichmäßig auf die einzelnen Skalenpunkte verteilen, wird dem Einwerben von sonstigen Drittmitteln von über der Hälfte der Befragten (52.1 %) ein (sehr) hoher Stellenwert in der eigenen Forschungsarbeit zugesprochen. Es ergeben sich so durchschnittliche Werte von $M = 3.00$ ($SD = 1.36$) für das Einwerben von DFG-Drittmitteln und von $M = 3.29$ ($SD = 1.40$) für die sonstigen Drittmittel. Auch wenn die Durchführung von Drittmittelprojekten der DFG sowohl hinsichtlich des Reputationszuwachses in der Scientific Community (vgl. Kapitel 6.1.4) als auch hinsichtlich des Beitrags für den wissenschaftlichen Erkenntnisgewinn (vgl. Kapitel 6.1.5) als bedeutender eingeschätzt wird, hat das Einwerben von Drittmitteln anderer Mittelgeber in der Forschungsarbeit der befragten Professorinnen und Professoren insgesamt im Mittel einen leicht höheren Stellenwert. Dies lässt sich schlicht damit begründen, dass die Fördermöglichkeiten durch die DFG begrenzt sind und nicht immer eine Passung zwischen dem eigenen Forschungsfeld und den Fördermöglichkeiten der DFG besteht. Die Professorinnen und Professoren greifen zur Finanzierung ihrer Forschungsarbeit daher konsequenterweise (auch) auf ein breites Angebot alternativer Mittelgeber zurück. Bieten diese mehr und erfolversprechendere Fördermöglichkeiten, steigt damit auch der Stellenwert, den das Einwerben von Drittmitteln anderer Mittelgeber in der eigenen Forschungsarbeit hat.

Betrachtet man den subjektiven Stellenwert von *Grundlagenforschung* gegenüber *anwendungsbezogener Forschung* in der Forschungsarbeit der Professorinnen und Professoren, zeigt sich mit einem Wert von $M = 3.70$ ($SD = 1.25$) eine leichte Priorisierung der Grundlagenforschung gegenüber der anwendungsbezogenen Forschung mit einem Wert von $M = 3.51$ ($SD = 1.30$). Bei beiden Items ergibt sich dabei eine deutlich linkschiefe Verteilung.

Am Ende des Frageblocks hatten die Befragten wieder die Möglichkeit, anhand der halboffenen „sonstige“-Kategorie weitere Forschungstätigkeiten zu ergänzen. Sechs Nennungen bezogen sich dabei auf die Arbeit in internationalen Kooperationen und der internationalen Vernetzung der Forschungsarbeit. Der Austausch mit Kollegen sowie die

Teilnahme an Tagungen und Konferenzen wurden von jeweils drei Befragten thematisiert. Administrative Tätigkeiten benannten zwei Befragte und zwei weitere Teilnehmende wiesen darauf hin, dass generell das Nachdenken über Forschungsprobleme und das Verstehen von Sachzusammenhängen einen hohen Stellenwert in ihrem Forschungsalltag einnehmen.

Zwischen den Fächern ergeben sich hinsichtlich aller abgefragten Forschungstätigkeiten signifikante Differenzen. Die im Anschluss an die einfaktoriellen Varianzanalysen durchgeführten Post-Hoc-Mehrfachvergleiche (Scheffé-Prozedur bzw. Games-Howell-Test) zeigen dabei teilweise deutliche Mittelwertunterschiede zwischen einzelnen Fächern. Die Mittelwerte je Forschungstätigkeit und Fach sind in Tabelle 15 und Abbildung 8 aufgeführt.

Tabelle 15: *Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit im Fächervergleich*

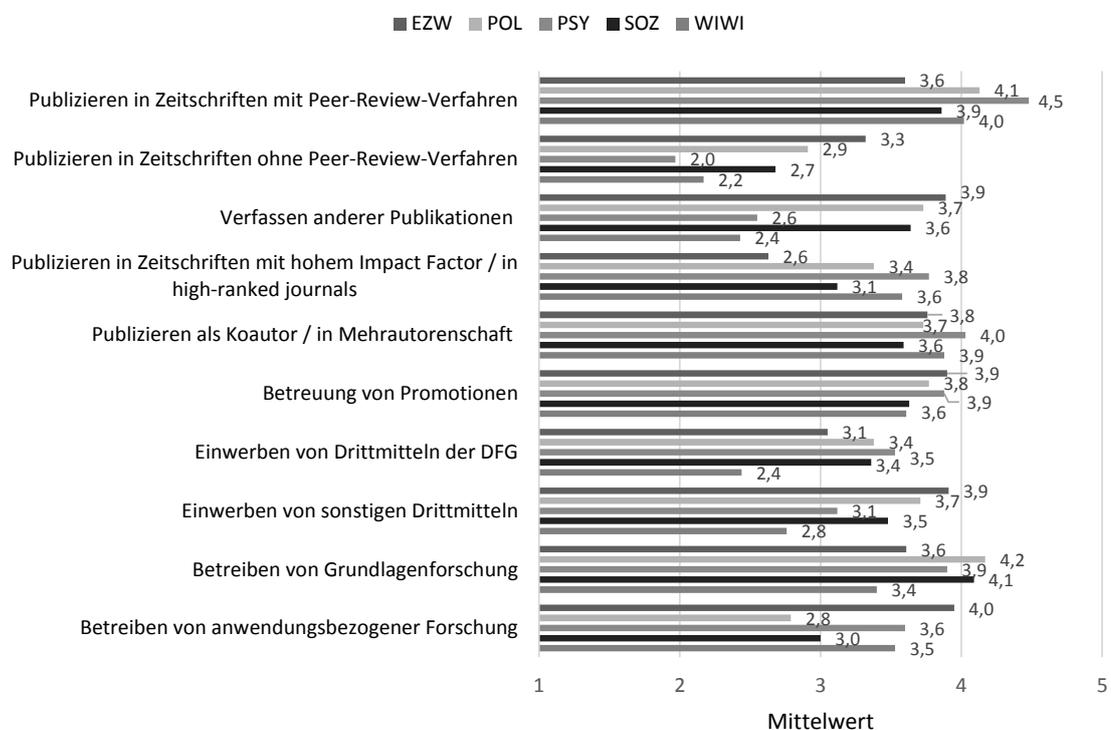
	EZW	POL	PSY	SOZ	WIWI
	<i>M (SD)</i>				
Publizieren in Zeitschriften mit Peer-Review-Verfahren	3.60 (1.21)	4.13 (0.94)	4.48 (0.73)	3.86 (1.16)	4.02 (1.13)
Publizieren in Zeitschriften ohne Peer-Review-Verfahren	3.32 (1.10)	2.91 (1.22)	1.97 (1.07)	2.68 (1.24)	2.17 (1.06)
Verfassen anderer Publikationen	3.89 (0.90)	3.73 (1.04)	2.55 (1.16)	3.64 (1.07)	2.43 (1.18)
Publizieren in Zeitschriften mit hohem Impact Factor / in high-ranked journals	2.63 (1.37)	3.38 (1.42)	3.77 (1.11)	3.12 (1.34)	3.58 (1.36)
Publizieren als Koautor / in Mehrautorenschaft	3.76 (0.97)	3.73 (1.10)	4.03 (0.92)	3.59 (1.02)	3.88 (1.10)
Betreuung von Promotionen	3.90 (0.92)	3.77 (1.08)	3.88 (0.97)	3.63 (0.91)	3.61 (1.08)
Einwerben von Drittmitteln der DFG	3.05 (1.36)	3.38 (1.39)	3.53 (1.30)	3.36 (1.22)	2.44 (1.20)
Einwerben von sonstigen Drittmitteln	3.91 (1.20)	3.71 (1.36)	3.12 (1.42)	3.48 (1.27)	2.76 (1.37)
Betreiben von Grundlagenforschung	3.61 (1.26)	4.17 (0.94)	3.90 (1.31)	4.09 (1.08)	3.40 (1.28)
Betreiben von anwendungsbezogener Forschung	3.95 (1.16)	2.79 (1.34)	3.60 (1.33)	3.00 (1.35)	3.53 (1.23)

Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften; *n* je Fach (min – max): $n_{EZW} = 143-148$, $n_{POL} = 52-55$, $n_{PSY} = 116-117$, $n_{SOZ} = 74-76$, $n_{WIWI} = 217-222$.

Quelle: eigene Berechnungen

Das *Publizieren in Zeitschriften mit Peer-Review-Verfahren* ($F(4, 613) = 11.44, p < .001$) hat erwartungsgemäß insbesondere in der Psychologie einen sehr hohen Stellenwert in der Forschungsarbeit der Professorinnen und Professoren. Mit Ausnahme der Politikwissenschaften, bei denen mit $M = 4.13$ ($SD = 0.94$) der Stellenwert ebenfalls sehr hoch eingeschätzt wird, fällt der Wert in der Psychologie mit $M = 4.48$ ($SD = 0.73$) signifikant höher aus als bei den anderen untersuchten Fächern. Auch in den Wirtschaftswissenschaften ergibt sich mit $M = 4.02$ ($SD = 1.13$) ein sehr hoher Wert. Dieser ist signifikant höher als in den Erziehungswissenschaften, bei denen die Relevanz mit $M = 3.60$ ($SD = 1.21$) im Fächervergleich am geringsten eingestuft wird, dabei aber immer noch auf einem hohen Niveau liegt.

Abbildung 8: Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit im Fächervergleich



Anmerkung. EZW = Erziehungswissenschaften / Sonderpädagogik, POL = Politikwissenschaften, PSY = Psychologie, SOZ = Soziologie / Sozialwissenschaften, WIWI = Wirtschaftswissenschaften; n je Fach (min – max): $n_{EZW} = 143-148$, $n_{POL} = 52-55$, $n_{PSY} = 116-117$, $n_{SOZ} = 74-76$, $n_{WIWI} = 217-222$.

Quelle: eigene Berechnungen

Die untergeordnete Bedeutung von *Zeitschriften ohne Peer-Review-Verfahren* in den Publikationskulturen der Psychologie und der Wirtschaftswissenschaften spiegelt sich deutlich im Stellenwert dieses Publikationsformats in der Forschungsarbeit der befragten Professorinnen und Professoren wider. Mit Werten von $M = 1.97$ ($SD = 1.07$) und

$M = 2.17$ ($SD = 1.06$) wird der Stellenwert dieser Forschungstätigkeit im Mittel eher gering eingeschätzt und liegt signifikant unter den Werten der übrigen untersuchten Fächer ($F(4, 608) = 34.21, p < .001$). Auffällig ist zudem der recht hohe Stellenwert von Publikationen in nicht-begutachteten Zeitschriften in den Erziehungswissenschaften ($M = 3.32, SD = 1.10$). Hier scheinen ausreichend anerkannte Fachzeitschriften vorzuliegen, in denen die Veröffentlichung eines Beitrages auch ohne Peer-Review-Verfahren für die Professorinnen und Professoren attraktiv erscheint.

Unabhängig von Publikationen in Zeitschriften spielt das *Verfassen anderer Publikationen* sowohl in den Erziehungswissenschaften ($M = 3.89, SD = 0.90$) als auch in den Politikwissenschaften ($M = 3.73, SD = 1.04$) und der Soziologie ($M = 3.64, SD = 1.07$) eine relativ große Rolle. Die fachliche Zweiteilung zwischen diesen Fächern auf der einen sowie der Psychologie und den Wirtschaftswissenschaften auf der anderen Seite setzt sich dabei auch hier fort und zeigt sich in den signifikant niedrigeren Werten bei den letztgenannten ($F(4, 611) = 57.54, p < .001$).

Hinsichtlich des *Publizierens in Zeitschriften mit hohem Impact Factor / in high ranked journals* ergibt sich wieder eine Sonderstellung der Erziehungswissenschaften, in welchen der Stellenwert mit einem Wert von $M = 2.63$ ($SD = 1.37$) signifikant niedriger als in den übrigen Fächern eingeschätzt wird ($F(4, 610) = 16.06, p < .001$). Eine Ausnahme bildet lediglich die Soziologie, in welcher die Orientierung am Impact-Factor und Zeitschriftenrankings ebenfalls etwas geringer ausgeprägt zu sein scheint ($M = 3.12, SD = 1.34$).

Das *Publizieren in Ko- oder Mehrautorenschaft* hat über alle Fächer hinweg einen hohen Stellenwert in der Forschungsarbeit der Professorinnen und Professoren. Lediglich zwischen den Fächern Psychologie ($M = 4.03, SD = 0.92$) und Soziologie ($M = 3.59, SD = 1.02$) ergibt sich eine signifikante Differenz ($F(4, 606) = 2.64, p = .033$). Ähnliches gilt für die *Betreuung von Promotionen*, bei der die fachspezifischen Werte ebenfalls durchgängig auf hohem Niveau sind und sich ein signifikanter Unterschied ($F(4, 599) = 2.69, p = .030$) nur zwischen den Erziehungswissenschaften ($M = 3.90, SD = 0.92$) und den Wirtschaftswissenschaften ($M = 3.61, SD = 1.08$) nachweisen lässt.

Der Stellenwert, den das *Einwerben von Drittmitteln* in der Forschungsarbeit der befragten Professorinnen und Professoren einnimmt, differiert dagegen wieder deutlich zwischen den verschiedenen Fächern. Bei den *DFG-Drittmitteln* ergibt sich ein signifikant ($F(4, 609) = 18.46, p < .001$) niedrigerer Wert in den Wirtschaftswissenschaften, der mit $M = 2.44$ ($SD = 1.20$) auch absolut betrachtet recht gering ausfällt. Bei den *sonstigen Drittmitteln* liegt der Wert bei den Wirtschaftswissenschaften ($M = 2.76, SD = 1.37$) ebenfalls deutlich unter dem Niveau der anderen Fächer ($F(4, 608) = 19.13, p < .001$). Auffällig sind zudem die recht hohen Werte von $M = 3.91$ ($SD = 1.20$) in den Erziehungswissenschaften und von $M = 3.71$ ($SD = 1.36$) in den Politikwissenschaften. Dies lässt

sich möglicherweise damit erklären, dass in diesen beiden Fächern – mehr als das bei den anderen Fächern der Fall ist – neben der DFG alternative und attraktive Möglichkeiten zur Drittmittelinwerbung bestehen (z. B. BMBF, EU, Auftragsforschung).

Höchst signifikante Unterschiede zwischen den Fächern ergeben sich auch hinsichtlich des Stellenwerts von *Grundlagenforschung* ($F(4, 600) = 8.06, p < .001$) und *anwendungsbezogener Forschung* ($F(4, 600) = 11.92, p < .001$). Während der Stellenwert der Grundlagenforschung in den Politikwissenschaften ($M = 4.17, SD = 0.94$), der Psychologie ($M = 3.90, SD = 1.31$) und der Soziologie ($M = 4.09, SD = 1.08$) am höchsten eingeschätzt wird, fallen die Werte in den Erziehungswissenschaften ($M = 3.61, SD = 1.26$) und insbesondere in den Wirtschaftswissenschaften ($M = 3.40, SD = 1.28$) deutlich niedriger aus. In Bezug auf den Stellenwert der anwendungsbezogenen Forschung zeigt sich erwartungsgemäß ein umgekehrtes Bild: In den Politikwissenschaften und der Soziologie wird der Stellenwert mit Werten von $M = 2.79 (SD = 1.34)$ und $M = 3.00 (SD = 1.35)$ am geringsten eingestuft, während die Werte bei den übrigen Fächern entsprechend höher ausfallen. Zudem zeigt sich, dass der Stellenwert von anwendungsbezogener Forschung in den Erziehungswissenschaften und den Wirtschaftswissenschaften etwas höher eingeschätzt wird als der Stellenwert der Grundlagenforschung. Welche Relevanz Grundlagenforschung bzw. anwendungsbezogene Forschung in der eigenen Forschungsarbeit einnimmt, ist natürlich auch in hohem Maße von den innerhalb einer Disziplin behandelten Fragestellungen und den damit verbundenen Forschungsthemen abhängig, also davon, welche Inhalte generell innerhalb einer Disziplin behandelt werden (können). So bieten die Erziehungs- und die Wirtschaftswissenschaften durch ihren teils engen Praxisbezug sicherlich mehr Möglichkeiten für die Bearbeitung anwendungsbezogener Forschungsthemen, als dies bei anderen Fächern der Fall ist.

Insgesamt zeigten sich auch bei der Analyse des Stellenwerts der verschiedenen Forschungstätigkeiten in der Forschungsarbeit der befragten Professorinnen und Professoren durchgehend signifikante Fachunterschiede. Die fachliche Zweiteilung mit der Psychologie und den Wirtschaftswissenschaften auf der einen und den übrigen Fächern auf der anderen Seite lässt sich bei den publikationsbasierten Tätigkeiten mit Einschränkungen auch hier beobachten. Insbesondere bei den Wirtschaftswissenschaften ergaben sich auch bei anderen Tätigkeitsbereichen, vor allem hinsichtlich der Drittmittelinwerbung, starke Abweichungen zu den anderen Fächern.

Von den einzelnen Forschungstätigkeiten und deren eingeschätzten Stellenwert in der Forschungsarbeit der Professorinnen und Professoren wird nun der Bogen zum individuellen Stellenwert der Forschungsarbeit generell gespannt, indem im Folgenden die intrinsische Forschungsmotivation der befragten Professorinnen und Professoren in den Blick genommen wird.

6.1.7 Intrinsische Forschungsmotivation

Die intrinsische Forschungsmotivation wurde unter Verwendung einer Subskala der *Research Motivation Scale* von Deemer et al. (2010; 2012) erfasst (vgl. Kapitel 5.2.1). Im Folgenden werden zunächst die deskriptiven Verteilungskennzahlen der gültigen Gesamtstichprobe für alle Items der Skala berichtet. Für die weiteren Analysen wurde mit Hilfe der verwendeten Skala ein Index gebildet, dessen Berechnung und Verteilungskennzahlen im Anschluss dargestellt werden. Abschließend werden die Werte des gebildeten Index wieder in fächervergleichender Perspektive betrachtet.

In Tabelle 16 sind die Verteilungsparameter und die gültigen Fallzahlen je Item aufgeführt.

Tabelle 16: *Intrinsische Forschungsmotivation (Itemwerte)*

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Kurtosis</i>
<i>I1</i> : Wenn ich forsche, fühle ich mich zufrieden.	640	4.44	0.70	-1.05	0.54
<i>I2</i> : Ich forsche aus Freude an der Sache.	642	4.55	0.67	-1.49	2.06
<i>I3</i> : Ich fühle mich rundum wohl, wenn ich in Forschungsarbeit involviert bin.	639	4.26	0.83	-1.01	0.70
<i>I4</i> : Mir gefällt es sehr durch Forschung neue Erkenntnisse zu gewinnen.	639	4.68	0.58	-1.96	4.76
<i>I5</i> : Ich habe das Bedürfnis wissenschaftliche Sachverhalte zu verstehen.	641	4.69	0.54	-1.91	5.01
<i>I6</i> : Wenn ich neue Erkenntnisse in meinem Forschungsfeld gewonnen habe, empfinde ich große Freude.	640	4.64	0.63	-1.75	2.80
<i>I7</i> : Die Forschungsarbeit an sich macht mir Spaß.	636	4.57	0.62	-1.29	1.09
<i>I8</i> : Ich betreibe Forschung um ihrer selbst willen.	630	3.81	1.28	-0.84	-0.39
<i>I9</i> : Die Zeit vergeht wie im Flug, wenn ich mit Forschungstätigkeiten beschäftigt bin.	636	4.17	0.98	-1.01	0.30

Anmerkung. Die Werte basieren auf einer fünfstufigen Ratingskala mit den Endpolen „stimme überhaupt nicht zu“ (1) und „stimme voll und ganz zu“ (5).

Quelle: eigene Berechnungen

Auffällig sind die durchgehend hohen Mittelwerte (zwischen $M = 3.81$ und $M = 4.69$), die mit jeweils stark linksschiefen Verteilungen einhergehen. Bei fast allen Items hat dementsprechend eine deutliche Mehrheit der Befragten den vorgegebenen Aussagen (voll und ganz) zugestimmt. Deutlich weniger stark fällt die Zustimmung mit $M = 3.81$

($SD = 1.28$) beim Item *I8* aus. Hier wählten ‚nur‘ etwa 65 % der Befragten die Skalenergebnisse 4 oder 5 (stimme voll und ganz zu). Auch bei den Items *I3* und *I9* ergeben sich im Mittel etwas geringere Zustimmungswerte, welche mit Werten von $M = 4.26$ ($SD = 0.83$) und $M = 4.17$ ($SD = 0.98$) aber dennoch auf sehr hohem Niveau liegen. Obwohl der akademischen Profession generell ein hoher Grad an intrinsischer (Forschungs-)Motivation nachgesagt wird (Kieser 2010, S. 349; Schimank 2005, S. 152), und somit entsprechend hohe Zustimmungswerte zu den einzelnen Aussagen zu erwarten waren, ist die Zustimmungstendenz in diesem hohen Ausmaß doch bemerkenswert.⁴⁸

Zur Bildung des Index wurden die Items *I1* bis *I9* zunächst einer Reliabilitätsanalyse unterzogen. Als Reliabilitätskriterium wurde das hierfür gängige *Cronbachs Alpha* verwendet (vgl. Diekmann 2004, S. 220–221). Dieses erreicht unter Einbezug aller Items einen Wert von 0.88 (0.90 für standardisierte Items) und fällt damit ähnlich hoch aus wie bei der englischen Originalskala, bei welcher der Wert von Cronbachs Alpha bei 0.90 liegt (Deemer et al. 2012, S. 186). Die interne Konsistenz der Skala kann also zunächst als sehr gut eingeschätzt werden.

Betrachtet man die Veränderung des Wertes von Cronbachs Alpha unter Ausschluss einzelner Items, zeigt sich, dass bis auf Item *I8* („Ich betreibe Forschung um ihrer selbst willen.“) alle Items zu einer Erhöhung von Cronbachs Alpha beitragen. Schließt man dagegen Item *I8* aus der Indexbildung aus, erhöht sich der Wert von Cronbachs Alpha von 0.88 auf 0.90. Auch der recht niedrige Trennschärfekoeffizient von 0.49 – bei Werten zwischen 0.62 und 0.74 bei den übrigen Items – spricht für einen Ausschluss von Item *I8* (vgl. Diekmann 2004, S. 212). Aufgrund der Ergebnisse der Reliabilitätsanalysen wurde sich daher dafür entschieden, Item *I8* bei der Indexbildung nicht zu berücksichtigen. Die eher schlechte Passung des Items *I8* in der deutschen Version der Skala könnte auf Übersetzungsschwierigkeiten und damit verbundenen Bedeutungsverschiebungen zurückzuführen sein. Die Aussage *I enjoy doing research for its own sake* im englischen Originalwortlaut korrespondiert demnach vermutlich nur unzureichend mit der hier vorgenommenen Übersetzung *Ich betreibe Forschung um ihrer selbst willen*.

Die angenommene eindimensionale Struktur der Subskala zur intrinsischen Motivation wurde von Deemer et al. mit Hilfe einer konfirmatorischen Faktorenanalyse überprüft und bestätigt (2010). Zumindest die faktorielle Validität (vgl. Bühner 2011, S. 64) der Skala kann damit als gegeben angesehen werden, während andere Aspekte der Validität der Skala (z. B. Inhalts- oder Kriteriumsvalidität) hier ungeprüft bleiben müssen (vgl. zu Möglichkeiten und Problemen der Überprüfbarkeit der Validität einer Skala Diekmann

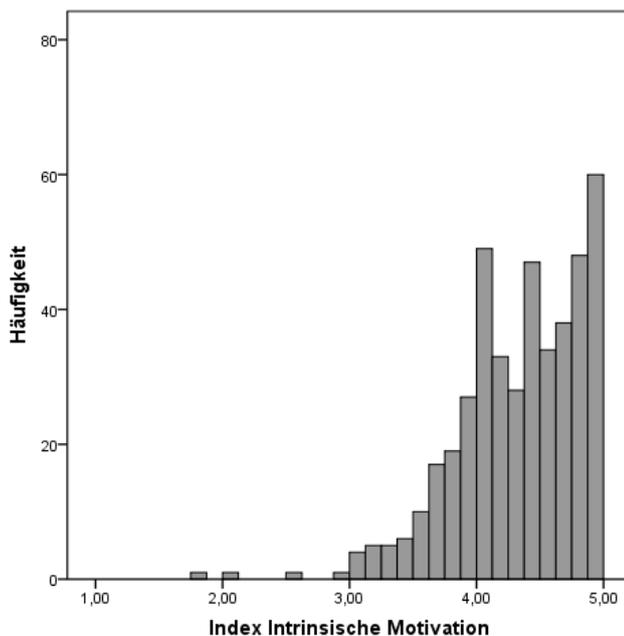
⁴⁸ Im Sinne einer Erhöhung der Antwortvarianz wäre zu prüfen, ob eine Änderung der Skalbreite auf eine sechs-, sieben- oder gegebenenfalls sogar zehnstufige Skala zu reliablen differenzierteren Ergebnissen führt.

2004, S. 223–227). Insgesamt kann aufgrund der dargestellten Ergebnisse dennoch von einer grundsätzlichen Eignung der Skala zur Bildung eines Index ausgegangen werden.

Bei der nachfolgenden Indexkonstruktion wurde sich für die Bildung eines Mittelwertindex mit dem Wertebereich von 1 bis 5 entschieden. Da die Angaben im Fragebogen nicht verpflichtend waren, kam es auch bei den acht für die Indexbildung relevanten Items zu vereinzelt fehlenden Werten. Dies trifft auf insgesamt 23 Datensätze zu.⁴⁹ Um für die angestrebten multivariaten Analysen dennoch möglichst viele Fälle einbeziehen zu können, wurden für die Indexbildung alle Fälle berücksichtigt, bei denen für mindestens sechs Items gültige Werte vorliegen. Dies betrifft insgesamt 11 Datensätze, so dass die Anzahl an auszuschließenden Datensätzen auf 12 reduziert werden konnte.

Auch für den aus $n = 639$ gültigen Fällen gebildeten Index *Intrinsische Forschungsmotivation* ergibt sich entsprechend der hohen Mittelwerte der Einzelitems ein sehr hoher Mittelwert von $M = 4.50$ ($SD = 0.53$). Die wieder stark linksschiefe Verteilung ist deutlich in dem in Abbildung 9 dargestellten Histogramm zu erkennen.

Abbildung 9: Histogramm Index *Intrinsische Forschungsmotivation*



Quelle: eigene Berechnungen

Zur Ermittlung möglicher Unterschiede im Fächervergleich wurde erneut eine einfaktorielle Varianzanalyse durchgeführt ($F(4, 616) = 3.62, p = .006$). Die Post-Hoc-Mehrfachvergleiche ergaben dabei einen signifikant höheren Indexwert bei Befragten der Politik-

⁴⁹ Nicht berücksichtigt sind hierbei Datensätze, bei denen die Teilnehmer den Fragebogen bereits vor der Abfrage der intrinsischen Motivation abgebrochen haben.

wissenschaften ($M = 4.70$, $SD = 0.39$, $n = 54$) als bei Befragten der Erziehungswissenschaften ($n = 148$), der Soziologie ($n = 75$) und der Wirtschaftswissenschaften ($n = 227$), bei welchen der Mittelwert des Index um jeweils etwa 0.24 Punkte niedriger liegt. Der Wert in der Psychologie beträgt $M = 4.60$ (0.49) und ist damit nur geringfügig niedriger als in den Politikwissenschaften. Insgesamt fallen die Indexwerte aber über alle Fächer hinweg sehr hoch aus und die zwar signifikanten, aber eher geringfügigen Differenzen sollten nicht überinterpretiert werden.

Mit der Darstellung des Index zur intrinsischen Motivation der Befragten sind die fächervergleichenden deskriptiven Analysen nahezu abgeschlossen. Ergänzt werden diese nun noch durch eine zusammenfassende Beschreibung der zahlreichen Anmerkungen und Kommentare, welche die Befragten im Freitextfeld am Ende des Fragebogens eingetragen oder per E-Mail versendet haben.

6.1.8 Anmerkungen und Kommentare der Befragten

Bevor näher auf die Anmerkungen und Kommentare der Befragten im Fragebogen eingegangen wird, sollen die teils ausführlichen direkten Rückmeldungen per E-Mail nicht unerwähnt bleiben. Aus Gründen der Anonymitätswahrung werden diese im Folgenden aber nur in einem kurzen thematischen Überblick wiedergegeben.

Sechs zur Befragung eingeladene Professorinnen und Professoren äußerten grundsätzliche Kritik an universitären Steuerungsversuchen, an einer zu starken Orientierung an Forschungsindikatoren und an dem Versuch, Forschungsleistung zu messen. Eine fortschreitende Ökonomisierung der Hochschulen wurde von diesen generell mit großer Skepsis gesehen. Aufgrund dieser Skepsis waren die meisten dieser Professorinnen und Professoren auch nicht bereit, an der Befragung teilzunehmen. Es wurde entweder befürchtet, dass die vorliegende Studie dem Zweck dienen könnte, die indikatorenbasierte Steuerung von Hochschulen noch weiter zu forcieren bzw. zu optimieren, oder aber, dass die Studie – obwohl sie als neutral wahrgenommen wurde – der Thematik allein durch deren Untersuchung unnötige Aufmerksamkeit verschafft.

Aufgrund dieser Rückmeldungen könnte angenommen werden, dass es in der Stichprobe zu einer systematischen Verzerrung der Untersuchungspopulation kommt. Befragte, die Verfahren der Leistungsmessung und indikatorenbasierten Steuerungsverfahren besonders kritisch gegenüber stehen, würden demnach öfter von vornherein die Teilnahme an der Befragung verweigern und wären in der Stichprobe entsprechend unterrepräsentiert. Da sich – wie sich bei der Analyse der Anmerkungen der Befragten noch zeigen wird – unter den Teilnehmerinnen und Teilnehmern der Befragung aber auch eine Vielzahl an Personen befindet, die der angesprochenen Thematik mit einer ähnlich großen Skepsis begegnen, kann die Annahme des Verzerrungseffektes entkräf-

tet werden. Denkbar wäre zudem auch, dass sich die Teilnahmemotivation genau gegenteilig darstellt und eben diejenigen Personen, die dem Thema kritisch (und damit oft auch emotionaler) gegenüberstehen, ein erhöhtes Interesse haben, an der Befragung teilzunehmen. Aufgrund dieser möglichen Selbstselektion wären in diesem Fall die ‚Kritiker‘ in der Stichprobe sogar überrepräsentiert. Beide Annahmen, also sowohl die These der Über- als auch der Unterrepräsentation der ‚Kritiker‘, können anhand der in dieser Arbeit vorliegenden Daten aber nicht überprüft werden.

Neben den teils sehr emotionalen kritischen Rückmeldungen gab es auch durchaus positive Rückmeldungen von den Befragten. Insgesamt 14 Teilnehmerinnen und Teilnehmer betonten in ihrer E-Mail die Relevanz und Aktualität des untersuchten Themas bzw. der Studie. Zehn Personen bekundeten dabei ihr Interesse an den Ergebnissen der Untersuchung.

Weitere E-Mail-Rückmeldungen bezogen sich auf andere Gründe der Nichtteilnahme (z. B. Zeitnot, keine Motivation aufgrund der Vielzahl an Anfragen, emeritiert etc.), wünschten ein erfolgreiches Projekt oder informierten im Zuge der Reminderversendung darüber, dass bereits teilgenommen wurde. Welche Anmerkungen und Kommentare die Befragten im Fragebogen selbst abgaben, wird nun im Folgenden dargestellt.

Insgesamt nutzten 101 Professorinnen und Professoren die Möglichkeit für (anonyme) Rückmeldungen anhand der offenen Frage am Ende des Fragebogens. Die Rückmeldungen lassen sich dabei überwiegend drei inhaltlichen Kategorien zuordnen: Anmerkungen und Kritik zum Untersuchungsgegenstand der vorliegenden Studie bzw. damit verbundener Themen (Leistungsmessung in der Wissenschaft, Orientierung an Indikatoren, New Public Management etc.) mit 41 Nennungen, Anmerkungen zum konkreten Einsatz und der Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren an der Hochschule der Befragten und damit verbundener Besonderheiten mit 72 Nennungen und schließlich Anmerkungen zum Fragebogen selbst mit 35 Nennungen.⁵⁰ Fünf weitere Rückmeldungen betrafen Erfolgswünsche für das Projekt ($n = 2$) sowie Informationen zum aktuellen Beschäftigungsstatus ($n = 3$).

Was sich schon bei den E-Mail-Rückmeldungen zeigte, setzt sich auch bei den im Fragebogen abgegebenen Anmerkungen fort: Dem Versuch einer Quantifizierung von Forschungsleistung und dem damit verbundenen Einsatz von indikatorenbasierten Steuerungs- und Anreizverfahren steht eine Vielzahl der Befragten kritisch gegenüber.⁵¹

⁵⁰ In der Summe ergibt sich hier ein Wert über 101, da sich viele Befragte in ihren Rückmeldungen auf mehrere der beschriebenen inhaltlichen Kategorien bezogen haben.

⁵¹ Insgesamt wurden zu dieser Thematik zwar ‚nur‘ 39 Nennungen gezählt, es ist aber davon auszugehen, dass der Anteil der Befragten mit einer kritischen Haltung deutlich höher liegt, da bei Befragungen offene Fragen aufgrund des erhöhten zeitlichen Aufwandes erfahrungsgemäß nur von einem kleinen Teil der Befragten überhaupt beantwortet werden.

Es wird bezweifelt, ob die Arbeit der wissenschaftlichen Profession überhaupt von außen gesteuert werden kann oder soll:⁵²

Die Frage welche Anreize für Forschung die richtigen sind, kann ich eigentlich überhaupt nicht verstehen. Entweder ist man zum Forschen geboren oder man soll es lassen! Sich mit der Frage zu beschäftigen, welche monetäre Anreizhöhe wird einen Wissenschaftler dazu bringen, ordentlich zu arbeiten, deutet darauf hin, dass das wissenschaftliche Denken in einer Gesellschaft eigentlich schon gestorben ist.

In obiger Aussage schwingt bereits ein weiterer oft genannter Aspekt mit, nämlich die Annahme der grundsätzlich intrinsischen Motivation von Wissenschaftlerinnen und Wissenschaftlern. Externe Steuerungs- und Anreizversuche seien aufgrund dieser Besonderheit der akademischen Profession daher unnötig bzw. unwirksam:

(...) Meine Erwartung ist, dass das intrinsische Forschungsinteresse ein viel wichtigerer Motivator ist als irgendwelche Anreiz- und Controllingsysteme, zumal wenn sie gar nicht mit den Kriterien korrespondieren, die in der Community relevant sind (was vielfach der Fall ist). Den nicht unerheblichen Aufwand kann man sich weitgehend sparen. Insbesondere jüngere Kollegen stehen unter hohem Drittmitteldruck, obwohl in der Wirtschaftswissenschaft (VWL) Drittmittelinwerbungen praktisch keinen prädiktiven Wert für Forschungsproduktivität (gemessen am qualitativen Ergebnis) haben. (...)

Einige Befragte sehen den Einsatz von indikatorenbasierten Steuerungs- und Anreizverfahren nicht nur als unnötig, sondern sogar als kontraproduktiv an, weil es durch extrinsische Anreize und Vorgaben zu einem Verdrängungseffekt der intrinsischen Motivation kommen könne (vgl. hierzu auch Kapitel 4.1.2). Neben einem möglichen motivationalen Verdrängungseffekt wird zudem die Gefahr gesehen, dass eine falsche Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren (z. B. zu starker Fokus auf Quantität als auf Qualität) nichtintendierte Effekte auf das Forschungsverhalten der Professorinnen und Professoren hat und der Forschung im Speziellen sowie der Wissenschaft im Allgemeinen mehr schadet als nützt:

Es ist schon Aufmerksamkeit erzeugend, dass es eine solche Erhebung gibt. Einfach irritierend, Menschen, die sich der Wissenschaft als Berufung verschrieben haben, so einzustufen, wie Arbeitnehmer, die durch eine angeblich objektivierbare Leistung schlichte materielle Gratifikationen erhalten. Vertretermentalität in der Wissenschaft. Wenn man berufen wird, muss man als bis dahin gereifte wissenschaftliche Persönlichkeit auch Verantwortung und Maß verinnerlicht haben, so dass dann die wissenschaftliche Leistung emergiert und glänzen kann, die endogen gesteuert und verantwortet emergiert. Falls das System, wie hier in den Fragen repräsentiert, zum alleinigen relevanten Steuerungsinstrumentarium verabsolutiert wird, schreiben demnächst wohl noch mehr Kolleginnen und Kollegen ohne Überzeugung und Liebe zum Fach ihre Nachweis- und Pflichtartikel pro Monat - leider allzu häufig dann noch mehr ohne Sinn und Verstand. Hauptsache Datensatz besorgen, irgendwie einsetzen ins Auswertungsprogramm und irgendetwas empirisch angeblich Bedeutsames und Neues in die Welt setzen. Viele kleine unverbundene Fragmente werden geschaffen, die in den Regalen der Archive vor sich dahin vergammeln und nicht mehr in gesellschaftliche Entwurfskonzepte ganzheitlich eingefügt werden. So beraubt sich die Wissenschaft ihrer bisherigen Funktion, auch Visionen, Utopien und Kritik zu entwickeln. Solche werden letztlich von der Wissenschaft nicht mehr berücksichtigt. Nur noch Wenige bieten mehr Orientierung stiftende Modelle und Ansätze, an denen man sich in der wissenschaftlich konstru-

⁵² Im Folgenden werden zu den verschiedenen inhaltlichen Kategorien jeweils beispielhaft einzelne, besonders prägnante Aussagen im Originalwortlaut wiedergegeben.

tiv reiben kann. Eine auch über Steuerungsinstrumente generierte Oberflächlichkeit und Marketing werden zur Selbsterosion der Wissenschaftsethik und der Wissenschaft beitragen. Solche Geister, die man so gerufen hat, bekommt man nicht mehr los.

Oder auch:

Diese ganzen Steuerungs- und Anreizverfahren zerstören mehr und mehr die intrinsische Motivation. Nicht zu vernachlässigen ist die Gefahr, die von fehlerhaften Steuerungs- und Anreizverfahren ausgeht, denn perverse Anreize führen zu perversen Ergebnissen. Doch wenn die perversen Ergebnisse erst mal vorliegen, dann ist es zu spät. (...) Es wird nur noch gezählt. Richtige Qualität interessiert niemanden mehr. Es ist unerträglich. In Berufungsverfahren etc. Aber es ist ja auch nicht verwunderlich, können Wissenschaftsmanager, die die Anreize setzen nur zählen. Publikationen und Euros.

Die falsche Ausgestaltung der Steuerungs- und Anreizverfahren wird auch in den drei folgenden Aussagen konkret thematisiert:

Die Anreizsysteme sind sehr stark auf Drittmittel ausgerichtet. Das ist grundsätzlich gut, weil es sich um einen objektiven Indikator handelt. Leider werden die Qualität der Arbeiten, der wissenschaftliche Erkenntnisgewinn (!), das Standing innerhalb der Scientific Community u.ä. praktisch überhaupt nicht berücksichtigt (verständlich, weil das objektiv schwer messbar ist).

Aktuell beobachte ich eine Fehlsteuerung dahingehend, dass der Versuch, die frühere leistungsunabhängige Mittel-/Besoldungsvergabe durch die Anwendung von Leistungsindikatoren zu beenden, teilweise die falschen Anreize setzt. In meiner Hochschule sind internationale Zeitschriftenbeiträge mit peer-review (neben Drittmitteln) die Königsklasse. Dass es innerhalb dieser Königsklasse auch viele Vertreter mit sehr geringen Qualitätsmaßstäben gibt, wird nicht berücksichtigt. In der Tendenz strebt die Mehrheit daher nach möglichst vielen Publikationen in der Kategorie international und peer-review. Der eigentliche Erkenntnisgewinn, das Streben nach wissenschaftlichem Fortschritt sind allenfalls Mittel zum Zweck und werden häufig völlig außen vor gelassen. Mit 5 Publikationen in internationalen Journals schlechter Qualität stehe ich im System meiner Hochschule besser da als mit 2 Top-Publikationen in der Science.

Ich fühle ein gewaltiges Unbehagen hinsichtlich des enormen Publikationsoutputs in meinem Fach, an dem ich mich jedoch trotz allem rege beteilige. Ich würde mich freuen, wenn in Zukunft Qualität statt Quantität eher im Mittelpunkt von Anerkennungsverfahren stünde. Denn meine Artikel in sog. peer-review-Zeitschriften sind beileibe nicht die besten gewesen... Die wirklich großen Würfe finden sich eher abseits des Mainstreams, in kleinen aber feinen Blättern. (...)

Noch einen Schritt weiter gehen Befürchtungen, dass die Steuerungs- und Anreizverfahren manipulatives und wissenschaftliches Fehlverhalten fördern könnten:

Es wäre wünschenswert, wenn andere Indikatoren genutzt würden. Diese tragen nicht zu Qualität der Forschung bei! Replikationskrise, Betrugsfälle und co. werde hiermit m.E. nur gefördert!

Die Annahme, dass Professoren wegen ein paar Hundert Euro pro Jahr höherer Mittelausstattung der Professur mehr leisten, halte ich, gelinde gesagt, für hinterfragenswert. Die leistungsbezogene Mittelverteilung wirkt sich nicht als Leistungsanreiz aus (zumal sie in der W-Besoldung heute ja überwiegend konsumiert wird) und ist in der Fakultät sehr umstritten. Ich gehe im Übrigen davon aus, dass getäuscht, getrickst und betrogen wird. Der bürokratische Aufwand für die Erfassung ist hoch, dass Verteil-Verfahren im Detail intransparent. (...)

Die Aussagen der Befragten machen deutlich, wie kritisch und dabei teils höchst emotional der Thematik der vorliegenden Studie begegnet wird. Neutrale oder gar positive Anmerkungen zum Einsatz von indikatorenbasierten Steuerungs- und Anreizverfahren

finden sich in den Kommentaren nur vereinzelt wieder. So merkte ein Teilnehmer das Folgende an:

(...) Aber zu Recht wächst der Druck die Organisation als solche anhand der Indikatoren zu benchmarken - und zwar von den Trägern - und letztendlich auch durch die Community selbst.

Wodurch der wachsende Druck dabei begründet sein könnte (z. B. zu geringer Output im internationalen Vergleich, Vertrauensverlust in die Profession etc.), bleibt allerdings offen. Hinsichtlich der bei den Kommentaren stark überwiegenden kritischen Äußerungen ist anzumerken, dass die Motivation zum Verfassen zusätzlicher Anmerkungen vermutlich höher ausfällt, wenn einer Thematik kritisch gegenübergestellt wird – dies gilt insbesondere dann, wenn auch emotionale Aspekte eine Rolle spielen. Steht man indikatorenbasierten Steuerungs- und Anreizverfahren dagegen neutral gegenüber oder befürwortet diese sogar, fällt die Motivation zum Verfassen zusätzlicher Anmerkungen, so die Annahme, entsprechend geringer aus. Nichtsdestotrotz zeigen die zahlreichen Anmerkungen, sowohl jene per E-Mail als auch jene im Fragebogen, dass indikatorenbasierte Steuerungs- und Anreizversuche im Rahmen der aktuellen hochschulpolitischen Entwicklungen einen unter den Professorinnen und Professoren höchst streitbaren Gegenstand darstellen.

Die zweite große inhaltliche Kategorie der Kommentare im Fragebogen bilden die Anmerkungen zum konkreten Einsatz und der Ausgestaltung indikatorenbasierter Steuerungs- und Anreizverfahren an der Hochschule der Befragten und den damit verbundenen Besonderheiten mit 72 Nennungen. Diese beziehen sich in den meisten Fällen auf ganz konkrete Hinweise zur Ausgestaltung (z. B. kein Einsatz indikatorenbasierter Steuerungsverfahren, Verwendung von lehrbezogenen Indikatoren, individuelle Aushandlung relevanter Indikatoren, Anteil der leistungsbezogen verausgabten Mittel etc.), die jeweils die einzelnen Universitäten bzw. Fakultäten betreffen. Für die hier vorgenommene Analyse sind diese Einzelfalldarstellungen von geringerem Interesse, so dass auf eine detaillierte Zusammenstellung der Aussagen verzichtet wird.

Die letzte Kategorie von Aussagen betrifft die Anmerkungen zum Fragebogen mit 35 Nennungen. Hier wurde zum einen wieder darauf hingewiesen, dass im Fragebogen das Thema (Anreizverfahren für die) Lehre ausgeklammert wurde (vgl. hierzu die Ausführungen in Kapitel 5.2.3 und 6.1.2). Zum anderen wurden einzelne methodische (z. B. Gefahr der sozialen Erwünschtheit bei einzelnen Fragen, Mehrdimensionalität beim Item *Andere Publikationen*, zu wenig Möglichkeiten, weitere Indikatoren zu ergänzen) und inhaltliche (z. B. Anzahl Promotionen weniger als Indikator für Forschung als für Nachwuchsförderung, fehlende Berücksichtigung von Praxiskontakten als Teil des Erkenntnisgewinns, Fragen in der differenzierten Form nicht zu beantworten) Aspekte thematisiert. Die Kommentare der Befragten variieren auch bei den Anmerkungen zum Fragebogen inhaltlich

sehr stark, so dass eine weitere thematische Gruppierung schwer fällt. Auf eine Auflistung aller Einzelaussagen wird hier verzichtet.

Die vielen Rückmeldungen der Befragten⁵³ heben die Relevanz des Themas in der hochschulpolitischen Debatte hervor und geben zudem erste Hinweise auf mögliche Zusammenhänge zwischen dem Einsatz indikatorenbasierter Steuerungs- und Anreizverfahren und dem Forschungsverhalten von Professorinnen und Professoren. So wurde deutlich, dass die Wirksamkeit dieser Verfahren teils als sehr gering bzw. nicht existent eingeschätzt wird und negative, nichtintendierte Effekte wie die Verdrängung der intrinsischen Motivation befürchtet werden. Externe Anreize scheinen also diesen Aussagen zufolge für die wenigsten Professorinnen und Professoren handlungsleitend für die Ausgestaltung ihrer Forschungsarbeit zu sein.

Inwieweit sich diese und weitere Annahmen in den erhobenen Daten empirisch widerspiegeln, wird sich nun in den beiden folgenden Teilkapiteln zeigen, in denen die forschungsleitenden Hypothesen dieser Arbeit in den Blick genommen werden sollen. Hierzu werden in Kapitel 6.2.1 zunächst bivariate Analysen vorgeschaltet, um einen ersten Überblick über die Zusammenhangsstruktur der interessierenden Variablen zu bekommen. In Kapitel 6.2.2 folgt dann der eigentliche Test der Forschungshypothesen mittels multipler Regressionsmodelle.

6.2 Überprüfung der Forschungshypothesen

6.2.1 Bivariate Analysen

Bei den nun folgenden bivariaten Zusammenhangsanalysen wird der Fokus vor allem auf die für die späteren Hypothesentests relevanten Variablen gelegt. Das heißt, primär werden Zusammenhänge zwischen dem eingeschätzten Stellenwert einzelner Forschungstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren (*AV*)⁵⁴ und den zentralen theoretisch hergeleiteten Einflussfaktoren dargestellt (vgl. Kapitel 4.2). Im Einzelnen sind dies:

- der Einsatz von spezifischen Forschungsindikatoren in hochschulinternen Steuerungs- und Anreizverfahren, die die jeweilige Forschungstätigkeit betreffen (*UV*)
- die eingeschätzte Bedeutung der jeweiligen Forschungstätigkeit bzw. deren Output für den Reputationszuwachs in der Scientific Community (*KV_REP*)

⁵³ Für die überwiegend konstruktiven Anregungen und Kritikpunkte sowie die positiven Rückmeldungen möchte ich mich an dieser Stelle bei den Teilnehmerinnen und Teilnehmern der Befragung bedanken.

⁵⁴ In diesem Kapitel werden zur besseren Lesbarkeit die folgenden Kürzel verwendet: *AV* = abhängige Variable, *UV* = unabhängige Variable, *KV* = Kontrollvariable.

- der eingeschätzte Stellenwert der jeweiligen Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn (*KV_ERK*)
- das Ausmaß der intrinsischen Motivation der Befragten (*KV_IM*)

Daneben werden im Anschluss noch mögliche Zusammenhänge zwischen der intrinsischen Motivation der Befragten und dem Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren (vgl. Kapitel 4.2, *H4.2*) sowie zwischen dem Forschungsverhalten und weiteren Kontrollvariablen (z. B. Besoldungsart, Alter) näher betrachtet.

Im Fragebogen wurde zudem erhoben, welchen Stellenwert das Publizieren als Koautor bzw. in Mehrautorenschaft sowie das Betreiben von Grundlagenforschung und anwendungsbezogener Forschung in der Forschungsarbeit der Professorinnen und Professoren hat (vgl. Kapitel 5.2.1). Diese drei Variablen werden allerdings nicht Teil der multivariaten Hypothesentestung in Kapitel 6.2.2 sein, da für diese Forschungstätigkeiten keine korrespondierenden Forschungsindikatoren (z. B. wie der Indikator *Höhe der eingeworbenen Drittmittel* für die Forschungstätigkeit *Einwerben von Drittmitteln*) und damit keine direkt entsprechenden unabhängigen Variablen vorliegen. Dennoch wurden auch für diese Forschungstätigkeiten bivariate Zusammenhangsanalysen mit verschiedenen interessierenden Variablen durchgeführt, deren Ergebnisse in diesem Kapitel präsentiert werden. Möglicherweise ergeben sich aus diesen Analysen Impulse für die inhaltliche Ergebnisinterpretation zu den Hypothesentests der anderen Variablen.

Die Analyse der bivariaten Zusammenhänge erfolgt in Form von einfachen Korrelationen mittels des *Bravais-Pearson-Korrelationskoeffizienten*. Neben intervallskalierten (oder dichotomen) Variablen und der Annahme eines linearen Zusammenhangs zwischen den zu analysierenden Variablen wird für die Signifikanzprüfung der Korrelation zudem die bivariate Normalverteilung der Variablen vorausgesetzt. Auch wenn die Signifikanzprüfung von Korrelationen sehr robust gegenüber Verletzungen der Verteilungsannahme ist (vgl. hierzu und zu den Voraussetzungen allgemein Bortz 1999, S. 205), wird hier aufgrund der teils stark schiefen Verteilungen der zu analysierenden Variablen (vgl. Kapitel 6.1) und dem nur quasi-metrischen Skalenniveau ergänzend *Spearman's Rho* als alternativer, nichtparametrischer Korrelationskoeffizient berichtet. Dieser basiert auf Rängen und setzt im Gegensatz zum Bravais-Pearson-Korrelationskoeffizienten lediglich einen monotonen und keinen linearen Zusammenhang zwischen den Variablen voraus. Zudem ist Spearman's Rho auch bei nur ordinalskalierten Variablen anwendbar, so dass das quasi-metrische Skalenniveau als unproblematisch angesehen werden kann. Aufgrund der Berechnung auf Basis von Rängen ist Spearman's Rho außerdem robust gegenüber Ausreißern (vgl. zu Spearman's Rho Fahrmeir et al. 2003, S. 139–143).

In den nun folgenden Tabellen sind die Werte des Bravais-Pearson-Korrelationskoeffizienten (r) sowie von Spearmans Rho (r_s) mit den zugehörigen Signifikanzniveaus jeweils für die Gesamtstichprobe und für die fachspezifischen Substichproben aufgeführt.⁵⁵ Zur Erläuterung der in den Tabellen genutzten Abkürzungen der jeweiligen Variablen sei hier auf die Anmerkungen unter den Tabellen verwiesen.

Neben den in den Tabellen dargestellten Zusammenhangsanalysen wurden zudem Berechnungen mit weiteren Variablen durchgeführt (Gewichtung der Indikatoren in den Steuerungs- und Anreizverfahren insgesamt, Einsatz und Gewichtung der Indikatoren in spezifischen Steuerungs- und Anreizverfahren), die hier aus Gründen der Übersichtlichkeit nicht alle einzeln aufgeführt werden. Wenn sich bei den ergänzenden Analysen aber signifikante Zusammenhänge nachweisen ließen, werden diese im Folgenden berichtet.⁵⁶

Anzumerken ist zudem, dass an dieser Stelle auf inhaltliche Interpretationen der dargestellten Ergebnisse weitgehend verzichtet wird. Diese erfolgen erst im Anschluss an den multivariaten Analyseteil, in welchem die theoretisch hergeleiteten Forschungshypothesen unter Berücksichtigung aller relevanten Variablen in einer ganzheitlichen Perspektive überprüft werden sollen (vgl. Kapitel 6.2.2 und 6.3).

Publizieren in Zeitschriften mit Peer-Review-Verfahren

In Tabelle 17 sind die Ergebnisse der Korrelationsanalysen zur Variablen *Publizieren in Zeitschriften mit Peer-Review-Verfahren (AV_1)* dargestellt. In der Gesamtbetrachtung zeigen sich dabei signifikant positive Korrelationen bei allen Variablen, wobei die beiden herangezogenen Korrelationskoeffizienten sehr ähnliche Werte erreichen. Die stärksten Zusammenhänge mit der Variablen *AV_1* ergeben sich bei den Variablen *KV_REP_1* und *KV_ERK_1*. Eine hohe eingeschätzte Bedeutung von Publikationen in Zeitschriften mit Peer-Review-Verfahren für den Reputationszuwachs in der Scientific Community

⁵⁵ Der Abgleich der Ergebnisse der beiden Koeffizienten soll auch zur Kontrolle möglicher auftretender Verzerrungen durch Verletzungen der Verteilungsannahmen dienen: Wenn sich im bivariaten Fall nur geringe Differenzen zwischen den Werten der beiden Koeffizienten ergeben, kann angenommen werden, dass sich die Verzerrungen aufgrund von Verletzungen der Verteilungsannahmen in Grenzen halten. Entsprechend sollten die Verzerrungen dann auch bei den im nächsten Kapitel folgenden parametrischen multivariaten Analyseverfahren (Multiple Regression) weniger schwer ins Gewicht fallen.

⁵⁶ Dabei ist zu berücksichtigen, dass für die ergänzenden Analysen des Einsatzes der Indikatoren in den einzelnen Steuerungs- und Anreizverfahren (z. B. Zusammenhang zwischen dem Einsatz des Indikators *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* im Verfahren *Benchmarking* und dem Stellenwert des Publizierens von peer-reviewed Artikeln in der Forschungsarbeit), nur die Teilstichprobe derjenigen Befragten relevant ist, die angegeben haben, dass das jeweilige Steuerungs- oder Anreizverfahren zum Einsatz kommt (Ausnahmen bilden die Verfahren der *Leistungszulagen im Rahmen der W-Besoldung* und *Bleibe-/Berufungsverhandlungen*; vgl. zur Fragebogenlogik und Berechnung der Variablen Kapitel 5.2.2 und 6.1.3 sowie Tabelle 9). Hieraus resultieren auch die teils niedrigen Fallzahlen bzw. Freiheitsgrade der dargestellten Korrelationen.

bzw. ein hoch eingeschätzter Beitrag dieses Publikationsformats für den wissenschaftlichen Erkenntnisgewinn korrespondiert demnach mit einem hohen Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren in der Forschungsarbeit der Professorinnen und Professoren. Auch hinsichtlich der intrinsischen Motivation der Befragten ergibt sich eine höchst signifikante Korrelation mit dem eingeschätzten Stellenwert des Publizierens in peer-reviewed Zeitschriften, wobei die Stärke des Zusammenhangs etwas geringer ausfällt.

Tabelle 17: Korrelationen zur Variablen AV_1: Publizieren in Zeitschriften mit Peer-Review-Verfahren

		gesamt	EZW	POL	PSY	SOZ	WIWI
UV_1	<i>r</i>	.097*	.100	-.056	.079	.186	.062
	<i>r_s</i>	.085*	.091	-.052	.116	.122	.318
	<i>n</i>	571	124	48	109	69	208
KV_REP_1	<i>r</i>	.318***	.416***	.491***	.219*	.050	.250***
	<i>r_s</i>	.323***	.419***	.440***	.115	.121	.278***
	<i>n</i>	632	148	54	117	76	220
KV_ERK_1	<i>r</i>	.474***	.316***	.518***	.416***	.596***	.498***
	<i>r_s</i>	.474***	.353***	.560***	.373***	.566***	.476***
	<i>n</i>	617	148	51	115	75	212
KV_IM	<i>r</i>	.238***	.303***	.335*	.160	.212	.189**
	<i>r_s</i>	.242***	.328***	.251	.090	.235*	.180**
	<i>n</i>	631	147	54	117	75	221

Anmerkung. UV_1: Einsatz der Anzahl von Publikationen in Zeitschriften mit Peer-Review-Verfahren als Forschungsindikator in indikatorenbasierten Steuerungs- und Anreizverfahren (dichotom), KV_REP_1: Bedeutung von Publikationen in Zeitschriften mit Peer-Review-Verfahren für den Reputationszuwachs in der Scientific Community, KV_ERK_1: Beitrag von Publikationen in Zeitschriften mit Peer-Review-Verfahren für den wissenschaftlichen Erkenntnisgewinn, KV_IM: Index Intrinsische Motivation; *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig): **p* ≤ .05, ***p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

Die sehr schwache, aber in der Gesamtbetrachtung noch signifikante Korrelation zwischen UV_1 und AV_1 verschwindet in den fachspezifischen Analysen gänzlich. Ein positiver Zusammenhang zwischen dem Einsatz des Indikators Anzahl von Publikationen in Zeitschriften mit Peer-Review-Verfahren in Steuerungs- und Anreizverfahren und dem Stellenwert von peer-reviewed Publikationen in der Forschungsarbeit der Professorinnen und Professoren kann hier also zunächst nicht nachgewiesen werden.

Während sich bei der Variablen KV_ERK_1 für alle Fächer durchgehend höchst signifikante positive Korrelationen ergeben, zeigen sich bei den Variablen KV_REP_1 und KV_IM fachspezifische Unterschiede: Der Zusammenhang zwischen KV_REP_1 und

AV_1 ist in der Psychologie nur schwach ausgeprägt bzw. bei Betrachtung von *Spearman's Rho* nicht mehr auf dem 5%-Niveau signifikant. In der Soziologie ergibt sich ebenso kein signifikanter Zusammenhang zwischen den beiden Variablen. Eine signifikante Korrelation zwischen *KV_IM* und *AV_1* zeigt sich bei Befragten der Erziehungs- und der Wirtschaftswissenschaften. Bei den anderen drei Fächern ergeben sich keine (Psychologie) bzw., je nach herangezogenem Korrelationskoeffizienten, schwache Zusammenhänge auf dem 5%-Niveau (Politikwissenschaften und Soziologie).

Die vertiefende Betrachtung des Einsatzes und der Gewichtung des Indikators *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* in den einzelnen Steuerungs- und Anreizverfahren ergibt die folgenden signifikanten Zusammenhänge mit der Variablen *AV_1*:⁵⁷

Für die Gesamtstichprobe zeigt sich, dass der Einsatz des Indikators sowohl im Verfahren *Benchmarking* ($r(59) = .51, p < .001$) als auch in *Bleibe-/Berufungsverhandlungen* ($r(435) = .14, p = .003$) positiv mit der Variablen *AV_1* korreliert. Ebenso korrespondiert eine höhere Gewichtung des Indikators im Verfahren *Benchmarking* ($r(44) = .50, p < .001$) und in den *Leistungszulagen im Rahmen der W-Besoldung* ($r(205) = .19, p = .008$) mit einem erhöhten Stellenwert der Tätigkeit des Publizierens in Zeitschriften mit Peer-Review-Verfahren in der Forschungsarbeit der Professorinnen und Professoren.

Für die Erziehungswissenschaften bestätigt sich die positive Korrelation zwischen der Variablen *AV_1* und der Gewichtung des Indikators in den *Leistungszulagen der W-Besoldung* ($r(49) = .30, p = .032$). Zudem korreliert die eingeschätzte Gewichtung des Indikators in den Steuerungs- und Anreizverfahren insgesamt positiv mit der Variablen *AV_1* ($r(71) = .37, p = .001$). Die Ergebnisse in der recht großen Substichprobe der Wirtschaftswissenschaften spiegeln nochmals in Teilen die Ergebnisse der Gesamtstichprobe wieder: Sowohl der Einsatz des Indikators in *Bleibe-/Berufungsverhandlungen* ($r(157) = .21, p = .007$) als auch die Höhe der Gewichtung des Indikators im Verfahren *Benchmarking* ($r(34) = .50, p = .002$) korreliert positiv mit der Variablen *AV_1*.

Publizieren in Zeitschriften ohne Peer-Review-Verfahren

Im folgenden Abschnitt wird nun der Stellenwert des *Publizierens in Zeitschriften ohne Peer-Review-Verfahren* (*AV_2*) in der Forschungsarbeit der Professorinnen und Professoren in den Blick genommen. Die Ergebnisse der bivariaten Korrelationen zwischen der

⁵⁷ Berichtet werden hier und im Folgenden nur diejenigen Korrelationen, die sowohl unter Verwendung des Bravais-Pearson-Korrelationskoeffizienten als auch unter Verwendung von Spearman's Rho auf dem 5%-Niveau signifikant sind. Als Maß für die Stärke des Zusammenhangs wird aus Gründen der Übersichtlichkeit lediglich der Wert des Bravais-Pearson-Korrelationskoeffizienten wiedergegeben. Auf eine tabellarische Darstellung der Ergebnisse wird bei der vertiefenden Betrachtung der einzelnen Verfahren (auch bei den folgenden Forschungstätigkeiten) verzichtet.

Variablen *AV_2* und den anderen zentralen Variablen finden sich in Tabelle 18. In der Gesamtstichprobe ergeben sich auch hier für alle Variablen höchst signifikante positive Zusammenhänge mit *AV_2*. Eine Ausnahme bildet allerdings die intrinsische Motivation, bei der sich nur beim Rangkorrelationskoeffizienten r_s ein schwacher negativer Zusammenhang mit *AV_2* zeigt ($r_s(624) = -.10, p = .015$). Dieser lässt sich bei den fachspezifischen Analysen aber nicht reproduzieren, so dass hier vermutlich von einem statistischen Artefakt ausgegangen werden kann.

Tabelle 18: *Korrelationen zur Variablen AV_2: Publizieren in Zeitschriften ohne Peer-Review-Verfahren*

		gesamt	EZW	POL	PSY	SOZ	WIWI
UV_2	r	.257***	.152	.382*	.199	.141	.136
	r_s	.258***	.130	.397**	.212*	.126	.141
	n	513	115	42	94	65	185
KV_REP_2	r	.591***	.442***	.581***	.535***	.448***	.467***
	r_s	.595***	.392***	.564***	.515***	.430***	.453***
	n	615	144	51	115	74	214
KV_ERK_2	r	.559***	.403***	.497***	.530***	.505***	.462***
	r_s	.566***	.392***	.456***	.537***	.512***	.470***
	n	608	148	49	114	74	207
KV_IM	r	-.069	.008	-.128	-.096	-.077	-.073
	r_s	-.097*	.044	-.141	-.133	-.076	-.084
	n	626	147	53	116	75	218

Anmerkung. UV_2: Einsatz der Anzahl von Publikationen in Zeitschriften ohne Peer-Review-Verfahren als Forschungsindikator in indikatorenbasierten Steuerungs- und Anreizverfahren (dichotom), KV_REP_2: Bedeutung von Publikationen in Zeitschriften ohne Peer-Review-Verfahren für den Reputationszuwachs in der Scientific Community, KV_ERK_2: Beitrag von Publikationen in Zeitschriften ohne Peer-Review-Verfahren für den wissenschaftlichen Erkenntnisgewinn, KV_IM: Index *Intrinsische Motivation*; r = Korrelationskoeffizient nach Bravais-Pearson, r_s = Spearman's Rho; Signifikanz (zweiseitig): * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$
Quelle: eigene Berechnungen

Sowohl für die Variable *KV_REP_2* als auch für *KV_ERK_2* zeigen sich auch im Fächervergleich durchgehend höchst signifikante positive Korrelationen, wobei die Werte des Bravais-Pearson-Koeffizienten und des Rangkorrelationskoeffizienten Spearman's Rho nur geringfügig voneinander abweichen. Für die Variable *UV_2* – dem Einsatz der Anzahl von Publikationen ohne Peer-Review-Verfahren als Indikator in Steuerungs- und

Anreizverfahren generell – bleibt die in der Gesamtstichprobe signifikante positive Korrelation nur in den Politikwissenschaften und mit Einschränkungen in der Psychologie⁵⁸ bestehen.

Betrachtet man den Einsatz und die Gewichtung der Anzahl nicht-begutachteter Publikationen als Forschungsindikator in den verschiedenen Verfahren im Detail, ergeben sich eine ganze Reihe signifikanter Korrelationen mit der Variablen AV_2.

Für die Gesamtstichprobe zeigt sich zunächst, dass der Einsatz des Indikators und dessen Gewichtungshöhe für nahezu alle abgefragten Steuerungs- und Anreizverfahren signifikant positiv mit AV_2 korreliert. Lediglich der Einsatz des Indikators in *Zielvereinbarungen zwischen Fakultät und Institut* ($r(108) = .17, p = .068$) sowie im Verfahren *Benchmarking* ($r(55) = .19, p = .150$) korreliert nicht signifikant mit dem Stellenwert, den das Publizieren in nicht-begutachteten Formaten in der Forschungsarbeit der Professorinnen und Professoren einnimmt.

In den Substichproben der einzelnen Fächer finden sich diese Zusammenhänge in dieser Form nur vereinzelt wieder: Signifikante positive Korrelationen ergeben sich in den Politikwissenschaften bezüglich des Einsatzes des Indikators in *Bleibe-/Berufungsverhandlungen* ($r(38) = .47, p = .002$) und in den Verfahren der *indikatorenbasierten Mittelzuweisung zwischen Hochschulleitung und Fakultät* ($r(27) = .51, p = .005$) sowie der *Mittelzuweisung zwischen Fakultät und Institut* ($r(22) = .53, p = .008$). Zumindest bei den beiden letztgenannten Korrelationen sind aber die geringen Fallzahlen von $n = 29$ bzw. $n = 24$ zu beachten. Diese resultieren wie bereits angemerkt aus der starken Ausdifferenzierung der Gesamtstichprobe nach verschiedenen Filterkriterien.

Auch für die Psychologie ergibt sich eine signifikant positive Korrelation zwischen dem Einsatz des Indikators in *Bleibe-/Berufungsverhandlungen* und dem eingeschätzten Stellenwert des Publizierens von nicht-begutachteten Publikationen in der Forschungsarbeit ($r(81) = .24, p = .029$). Ein letzter signifikanter Zusammenhang zeigt sich in den Wirtschaftswissenschaften zwischen dem Einsatz des Indikators in den *Leistungszulagen im Rahmen der W-Besoldung* und der Variablen AV_2 ($r(123) = .22, p = .012$).

Neben den genannten ergaben sich in den einzelnen Fächern noch weitere signifikante Korrelationen hinsichtlich des Einsatzes und der Gewichtung des Indikators in den verschiedenen Verfahren. Aufgrund der teilweise sehr geringen Fallzahlen erlauben diese Ergebnisse aber keine sinnvolle Interpretation und werden daher nicht weiter berichtet.

⁵⁸ In der Psychologie ist nur r_s auf dem 5%-Niveau signifikant, während r mit einem Wert von $p = .054$ knapp darüber liegt. Ähnlich knapp wurde die Signifikanzgrenze in den Wirtschaftswissenschaften überschritten ($p = .066$ (r) bzw. $.055$ (r_s)), was vermutlich die in der Gesamtstichprobe identifizierte höchst signifikante Korrelation erklärt.

Verfassen anderer Publikationen

Im Folgenden sind nun nicht mehr Zeitschriftenpublikationen, sondern andere gängige Publikationsformate (Monografien, Beiträge in Sammelbänden)⁵⁹ Gegenstand der Analyse. Die relevanten Korrelationen hinsichtlich des eingeschätzten Stellenwerts des *Verfassens anderer Publikationen (AV_3)* in der Forschungsarbeit der Professorinnen und Professoren sind in Tabelle 19 aufgeführt.

Tabelle 19: Korrelationen zur Variablen AV_3: Verfassen anderer Publikationen

		gesamt	EZW	POL	PSY	SOZ	WIWI
UV_3	<i>r</i>	.216***	.184*	.315*	.138	.070	.046
	<i>r_s</i>	.219***	.146	.284	.174	.061	.049
	<i>n</i>	527	116	45	98	67	189
KV_REP_3	<i>r</i>	.569***	.196*	.403**	.292***	.433***	.527***
	<i>r_s</i>	.572***	.180*	.487***	.261**	.413***	.511***
	<i>n</i>	622	145	54	116	74	216
KV_ERK_3	<i>r</i>	.589***	.408***	.454***	.502***	.418***	.524***
	<i>r_s</i>	.578***	.382***	.426**	.461***	.439***	.503***
	<i>n</i>	610	145	51	115	75	208
KV_IM	<i>r</i>	-.020	.058	-.046	-.127	-.005	-.003
	<i>r_s</i>	-.002	.100	-.022	-.112	.085	-.022
	<i>n</i>	629	146	54	117	75	220

Anmerkung. UV_3: Einsatz der *Anzahl anderer Publikationen* als Forschungsindikator in indikatorenbasierten Steuerungs- und Anreizverfahren (dichotom), KV_REP_3: Bedeutung von *Anderen Publikationen* für den Reputationszuwachs in der Scientific Community, KV_ERK_3: Beitrag von *Anderen Publikationen* für den wissenschaftlichen Erkenntnisgewinn, KV_IM: Index *Intrinsische Motivation*; *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig): **p* ≤ .05, ***p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

Die stärksten positiven Korrelationen mit der Variablen AV_3 ergeben sich in der Gesamtbetrachtung wieder hinsichtlich der Bedeutung der Publikationsformate für den Reputationszuwachs in der Scientific Community (KV_REP_3) sowie hinsichtlich des Beitrages von diesen für den wissenschaftlichen Erkenntnisgewinn (KV_ERK_3). Während der Zusammenhang zwischen KV_ERK_3 und AV_3 auch im Fächervergleich auf recht hohem Niveau bleibt, zeigen sich bei der Variablen KV_REP_3 deutliche Unterschiede bezüglich der Stärke des Zusammenhangs. Dieser fällt in den Erziehungswissenschaft-

⁵⁹ Auf das Problem der Mehrdimensionalität bei diesem Item wurde bereits an anderer Stelle hingewiesen (vgl. Kapitel 6.1.2).

ten mit $r = .20$ und der Psychologie mit $r = .29$ deutlich geringer als in der Gesamtstichprobe ($r = .57$) aus. Bei beiden Variablen liefern die beiden herangezogenen Korrelationskoeffizienten überwiegend übereinstimmende Ergebnisse.

Die intrinsische Motivation scheint – zumindest in der bivariaten Betrachtung – in keinem Zusammenhang mit der Variablen *AV_3* zu stehen. Sowohl in der Gesamtstichprobe als auch in den fächerspezifischen Substichproben ergeben sich hier keine signifikanten Korrelationen. Gemischte Ergebnisse zeigen sich wieder hinsichtlich des Einsatzes der Anzahl anderer Publikationen (*UV_3*) als Forschungsindikator in Steuerungs- und Anreizverfahren generell. Dieser korreliert in der Gesamtstichprobe positiv mit *AV_3*. Im Fächervergleich zeigen sich (knapp) signifikante Ergebnisse nur noch für die Erziehungs- und die Politikwissenschaften, und diese auch nur beim Bravais-Pearson-Korrelationskoeffizient, während die Überprüfung mittels des nichtparametrischen Rangkorrelationskoeffizienten r_s zu keinem signifikanten Ergebnis führt.

Wenn man nun wieder die einzelnen Steuerungs- und Anreizverfahren und den Einsatz bzw. die Gewichtungshöhe des Indikators *Anzahl anderer Publikationen* in diesen im Detail in den Blick nimmt, ergeben sich in der Gesamtstichprobe signifikante positive Zusammenhänge zwischen *AV_3* und dem Einsatz des Indikators in den *Leistungszulagen im Rahmen der W-Besoldung* ($r(367) = .30, p < .001$) sowie in *Bleibe-/Berufungsverhandlungen* ($r(427) = .23, p < .001$). Zudem korreliert die Höhe der Gewichtung des Indikators in den Verfahren insgesamt ($r(247) = .28, p < .001$) positiv mit der Variablen *AV_3*. Dies gilt mit Ausnahme der Verfahren *Zielvereinbarungen zwischen Hochschulleitung und Fakultät* ($r(56) = .12, p = .39$) und *Benchmarking* ($r(24) = .29, p = .15$; Fallzahl hier nur $n = 26$) auch für jedes Verfahren im Einzelnen.

Die Detailanalyse der einzelnen Verfahren im Fächervergleich ist aufgrund der wieder geringen Fallzahlen nur sehr eingeschränkt möglich. Berichtet werden daher nur diejenigen signifikanten Zusammenhänge, denen eine ausreichend große Fallzahl zugrunde liegt. Hierbei ergibt sich lediglich in den Politikwissenschaften ($r(40) = .35, p = .022$) und der Psychologie ($r(78) = .23, p = .041$) ein signifikant positiver Zusammenhang zwischen dem Einsatz des Indikators *Anzahl anderer Publikationen* in *Bleibe-/Berufungsverhandlungen* und der Variablen *AV_3*. Des Weiteren korreliert in den Wirtschaftswissenschaften die Höhe der Gewichtung des Indikators im Verfahren der *indikatorenbasierten Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl* mit der Variablen *AV_3* ($r(33) = .44, p = .008$). Weitere signifikante Zusammenhänge lassen sich hier – auch aufgrund der geringen Fallzahlen – nicht nachweisen.

Publizieren in Zeitschriften mit hohem Impact Factor / in high-ranked journals

Die Korrelationen zwischen dem Stellenwert des *Publizierens in Zeitschriften mit hohem Impact Factor / in high-ranked journals* (AV_4) in der Forschungsarbeit der Professorinnen und Professoren und den zentralen Untersuchungsvariablen sind in Tabelle 20 aufgeführt.

Tabelle 20: Korrelationen zur Variablen AV_4: Publizieren in Zeitschriften mit hohem Impact Factor / in high-ranked journals

		gesamt	EZW	POL	PSY	SOZ	WIWI
UV_4	<i>r</i>	.213***	.119	.323*	.076	.206	.241***
	<i>r_s</i>	.210***	.116	.312*	.048	.182	.251***
	<i>n</i>	507	103	40	102	58	194
KV_REP_4	<i>r</i>	.508***	.487***	.358**	.489***	.378***	.453***
	<i>r_s</i>	.490***	.479***	.314*	.476***	.362***	.435***
	<i>n</i>	624	144	54	117	75	219
KV_ERK_4	<i>r</i>	.580***	.531***	.510***	.466***	.519***	.603***
	<i>r_s</i>	.584***	.537***	.550***	.476***	.533***	.601***
	<i>n</i>	612	144	51	115	75	212
KV_IM	<i>r</i>	.170***	.171*	.159	.246**	.068	.158*
	<i>r_s</i>	.152***	.119	.118	.154	.074	.143*
	<i>n</i>	626	146	54	117	75	219

Anmerkung. UV_4: Einsatz der Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals als Forschungsindikator in indikatorenbasierten Steuerungs- und Anreizverfahren (dichotom), KV_REP_4: Bedeutung von Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals für den Reputationszuwachs in der Scientific Community, KV_ERK_4: Beitrag von Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals für den wissenschaftlichen Erkenntnisgewinn, KV_IM: Index Intrinsische Motivation; *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig): **p* ≤ .05, ***p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

In der Gesamtstichprobe ergeben sich zunächst für alle Variablen signifikante positive Korrelationen mit der Variablen AV_4. Am stärksten fallen diese bei den Variablen KV_REP_4 und KV_ERK_4 aus. Bei diesen zeigen sich auch im Fächervergleich durchgehend höchst signifikante Zusammenhänge. Der in der Gesamtbetrachtung höchst signifikante Zusammenhang zwischen UV_4 und AV_4 lässt sich in den fachspezifischen Analysen nur für die Politikwissenschaften und die Wirtschaftswissenschaften nachweisen, während sich für die anderen Fächer keine signifikanten Korrelationen ergeben. Auch hinsichtlich der intrinsischen Motivation (KV_IM) sind die Korrelationen mit der Variablen AV_4 im Fächervergleich weniger eindeutig als in der Gesamtstichprobe: Für die

Erziehungswissenschaften und die Psychologie ergibt sich ein signifikant positiver Zusammenhang beim Bravais-Pearson-Korrelationskoeffizienten, der sich unter Verwendung von Spearmans Rho aber nicht bestätigen lässt. Einzig in den Wirtschaftswissenschaften zeigt sich bei beiden Koeffizienten eine eher schwache positive Korrelation zwischen den Variablen *KV_IM* und *AV_4*.

Betrachtet man die einzelnen Steuerungs- und Anreizverfahren wieder isoliert, zeigt sich, dass der Einsatz des Indikators *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals* für alle Verfahren positiv mit der Variablen *AV_4* assoziiert ist. Die stärkste Korrelation ergibt sich dabei beim Verfahren *Benchmarking* ($r(59) = .39, p = .002$), während die Stärke des Zusammenhang bei den *Zielvereinbarungen zwischen Hochschulleitung und Fakultät* ($r(222) = .17, p = .011$) am geringsten ausfällt. Die Höhe der Gewichtung des Indikators korreliert allerdings lediglich bei den Verfahren *Zielvereinbarungen zwischen Fakultät und Institut/Lehrstuhl* ($r(36) = .36, p = .026$), *Leistungszulagen im Rahmen der W-Besoldung* ($r(97) = .30, p = .003$) und *Bleibe-/Berufungsverhandlungen* ($r(128) = .41, p < .001$) signifikant positiv mit der Variablen *AV_4*.

Die Detailbetrachtung der Verfahren in fachspezifischer Perspektive führt in den Wirtschaftswissenschaften zu einem ähnlichen Ergebnis wie in der Gesamtstichprobe. Auch hier korreliert der Einsatz des Indikators *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals* für alle Verfahren auf dem 5%-Niveau signifikant positiv mit der Variablen *AV_4* (r zwischen $.22$ und $.66$ bei Fallzahlen von $n = 36$ bis $n = 165$). Für die übrigen Fächer ergeben die Analysen signifikant positive Zusammenhänge zwischen dem Einsatz des Indikators in *Bleibe-/Berufungsverhandlungen* ($r(35) = .35, p = .035$) und der Variablen *AV_4* (Politikwissenschaften) sowie zwischen dem Einsatz des Indikators in den *Leistungszulagen im Rahmen der W-Besoldung* ($r(39) = .38, p = .016$) und *AV_4* (Soziologie).

Betreuung von Promotionen

Nach Betrachtung der verschiedenen publikationsbasierten Forschungstätigkeiten werden nun die Zusammenhangsanalysen zum eingeschätzten Stellenwert der *Betreuung von Promotionen (AV_5)* in der Forschungsarbeit der Professorinnen und Professoren dargestellt. Die entsprechenden Koeffizienten und Fallzahlen finden sich in Tabelle 21.

Zunächst lässt sich festhalten, dass sich zwischen dem generellen Einsatz des Indikators *Anzahl betreuter Promotionen* in Steuerungs- und Anreizverfahren und dem eingeschätzten Stellenwert der Betreuung von Promotionen in der Forschungsarbeit der Professorinnen und Professoren weder in der Gesamtstichprobe noch in den fachspezifischen Substichproben signifikante Korrelationen nachweisen lassen.

Tabelle 21: Korrelationen zur Variablen AV_5: Betreuung von Promotionen

		gesamt	EZW	POL	PSY	SOZ	WIWI
UV_5	<i>r</i>	.064	.105	.012	.003	.072	.046
	<i>r_s</i>	.059	.103	-.031	.016	.111	.031
	<i>n</i>	536	116	40	108	62	193
KV_REP_5	<i>r</i>	.281***	.280***	.112	.304***	.280*	.224***
	<i>r_s</i>	.274***	.287***	.198	.260**	.284*	.225***
	<i>n</i>	603	139	49	114	75	210
KV_ERK_5	<i>r</i>	.415***	.375***	.449**	.451***	.495***	.349***
	<i>r_s</i>	.409***	.372***	.432***	.462***	.476***	.358***
	<i>n</i>	580	141	41	111	71	201
KV_IM	<i>r</i>	.184***	.301***	.042	.255**	-.086	.165*
	<i>r_s</i>	.187***	.349***	.088	.287**	-.107	.101
	<i>n</i>	617	143	51	116	74	216

Anmerkung. UV_5: Einsatz der Anzahl betreuter Promotionen als Forschungsindikator in indikatorenbasierten Steuerungs- und Anreizverfahren (dichotom), KV_REP_5: Bedeutung der Anzahl betreuter Promotionen für den Reputationszuwachs in der Scientific Community, KV_ERK_5: Beitrag der Betreuung von Promotionen für den wissenschaftlichen Erkenntnisgewinn, KV_IM: Index *Intrinsische Motivation*; *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig): **p* ≤ .05, ***p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

Höchst signifikante positive Zusammenhänge mit der Variablen AV_5 zeigen sich in der Gesamtstichprobe dagegen für die Variablen KV_REP_5, KV_ERK_5 sowie für KV_IM. Die stärkste Korrelation ergibt sich dabei wieder zwischen dem eingeschätzten Beitrag der Betreuung von Promotionen für den wissenschaftlichen Erkenntnisgewinn (KV_ERK_5) und AV_5. Auch in den fachspezifischen Analysen bleiben die signifikanten Korrelationen überwiegend bestehen. Bei der Variablen KV_REP_5 ergibt sich lediglich in den Politikwissenschaften kein signifikanter Zusammenhang. Hinsichtlich der intrinsischen Motivation trifft dies neben den Politikwissenschaften auch für die Soziologie zu. Für die Wirtschaftswissenschaften zeigt zudem nur der Bravais-Pearson-Koeffizient einen signifikanten Zusammenhang auf dem 5%-Niveau an, während Spearmans Rho deutlich über der festgelegten Signifikanzgrenze liegt ($r_s(214) = .10, p = .139$).

Während der Einsatz des Indikators *Anzahl betreuter Promotionen* in den Steuerungs- und Anreizverfahren insgesamt (UV_5) nicht signifikant mit AV_5 korreliert, ergeben sich in der isolierten Betrachtung des Einsatzes des Indikators in den einzelnen Verfahren signifikante – wenn auch nur schwache – positive Zusammenhänge bei den *Leistungszulagen im Rahmen der W-Besoldung* ($r(356) = .11, p = .033$) sowie den *Bleibe- und Berufungsverhandlungen* ($r(393) = .12, p = .013$). Auch die Höhe der Ge-

wichtung des Indikators korreliert in der Gesamtstichprobe sowohl bei den *Leistungszulagen W-Besoldung* ($r(156) = .21, p = .007$) als auch bei den *Bleibe- und Berufungsverhandlungen* ($r(190) = .28, p < .001$) signifikant positiv mit der Variablen *AV_5*.

In den fachspezifischen Detailanalysen ergeben sich teils analoge Ergebnisse. In den Erziehungswissenschaften ($r(48) = .38, p = .006$) und der Psychologie ($r(47) = .33, p = .021$) korreliert die Höhe der Gewichtung des Indikators *Anzahl betreuter Promotionen* in *Bleibe- und Berufungsverhandlungen* signifikant positiv mit der Variablen *AV_5*. In der Soziologie zeigt sich ein signifikanter Zusammenhang zwischen dem Einsatz des Indikators in *Bleibe- und Berufungsverhandlungen* und der Variablen *AV_5* ($r(44) = .35, p = .018$). Ebenso korreliert auch hier die Höhe der Gewichtung des Indikators in *Bleibe- und Berufungsverhandlungen* mit der Variablen *AV_5* ($r(20) = .47, p = .026$), wobei anzumerken ist, dass die Aussagekraft dieses Ergebnisses aufgrund der geringen Fallzahl von $n = 22$ eingeschränkt ist.

Ein weiterer signifikanter Zusammenhang ergibt sich nochmals in der Psychologie, in welcher der Einsatz des Indikators *Anzahl betreuter Promotionen* innerhalb von *Zielvereinbarungen zwischen Hochschulleitung und Fakultät* positiv mit der Variablen *AV_5* korreliert ($r(38) = .33, p = .039$).

Einwerben von Drittmitteln der Deutschen Forschungsgemeinschaft

Im vorliegenden Abschnitt wird nun der Stellenwert, den das *Einwerben von Drittmitteln der DFG (AV_6)* in der Forschungsarbeit der Professorinnen und Professoren hat, näher beleuchtet. Tabelle 22 bietet wieder eine Übersicht der Korrelationskoeffizienten und der Fallzahlen je Untersuchungsvariablen.

In der Gesamtstichprobe ergeben sich für alle vier zentralen Untersuchungsvariablen signifikant positive Zusammenhänge mit der Variablen *AV_6*. Der Zusammenhang zwischen *UV_6* und *AV_6* ist mit $r = .10$ dabei nur sehr schwach ausgeprägt. Im Fächervergleich verschwindet der signifikante Zusammenhang zwischen den beiden Variablen gänzlich. Ebenfalls eher schwach ausgeprägt – allerdings höchst signifikant – ist die Korrelation zwischen *KV_IM* und *AV_6*. In der fachspezifischen Perspektive ergeben sich außer für die Erziehungswissenschaften allerdings wieder nur nicht-signifikante Zusammenhänge. In den Wirtschaftswissenschaften ist zumindest der Korrelationskoeffizient nach Bravais-Pearson auf dem 5%-Niveau signifikant, während Spearmans Rho keine Signifikanz mehr anzeigt ($r_s(218) = .14, p = .085$).

Hinsichtlich des Zusammenhangs zwischen den Variablen *KV_REP_6* und *AV_6* ergeben sich teils deutliche Unterschiede zwischen den Fächern. Am stärksten fällt die Korrelation in den Politikwissenschaften mit $r = .41$ aus, während sich für die Psychologie kein signifikanter Zusammenhang nachweisen lässt. Zwischen den Variablen *KV_ERK_6* und *AV_6* zeigt sich in der Gesamtstichprobe der stärkste Zusammenhang

und auch in den fachspezifischen Analysen ergeben sich durchgehend höchst signifikante Zusammenhänge.

Tabelle 22: Korrelationen zur Variablen AV_6: Einwerben von Drittmitteln der DFG

		gesamt	EZW	POL	PSY	SOZ	WIWI
UV_6	<i>r</i>	.097*	.099	.031	-.116	.128	-.001
	<i>r_s</i>	.096*	.091	.018	-.142	.145	.007
	<i>n</i>	586	128	49	112	72	208
KV_REP_6	<i>r</i>	.303***	.271***	.405**	.148	.285*	.208**
	<i>r_s</i>	.305***	.253**	.433***	.170	.234*	.194**
	<i>n</i>	629	145	54	116	76	221
KV_ERK_6	<i>r</i>	.423***	.446***	.540***	.309***	.383***	.319***
	<i>r_s</i>	.423***	.431***	.559***	.357***	.369***	.300***
	<i>n</i>	615	146	52	114	75	212
KV_IM	<i>r</i>	.169***	.268***	.075	.132	.095	.139*
	<i>r_s</i>	.163***	.237**	.050	.138	.153	.116
	<i>n</i>	627	145	54	116	75	220

Anmerkung. UV_6: Einsatz der Höhe der eingeworbenen Drittmittel der DFG als Forschungsindikator in indikatorenbasierten Steuerungs- und Anreizverfahren (dichotom), KV_REP_6: Bedeutung der Höhe der eingeworbenen Drittmittel der DFG für den Reputationszuwachs in der Scientific Community, KV_ERK_6: Beitrag der Durchführung von Drittmittelprojekten der DFG für den wissenschaftlichen Erkenntnisgewinn, KV_IM: Index Intrinsische Motivation; *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig): **p* ≤ .05, ***p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

Die Detailbetrachtung des Einsatzes und der Gewichtung des Indikators *Höhe der eingeworbenen Drittmittel der DFG* in den verschiedenen Steuerungs- und Anreizverfahren ergibt die folgenden (bei ausreichend hoher Fallzahl) signifikanten Korrelationen mit der Variablen AV_6: Zunächst zeigt sich für die Gesamtstichprobe, dass die Gewichtung des Indikators in den Verfahren insgesamt positiv mit AV_6 korreliert ($r(485) = .12, p = .007$). Hinsichtlich des Einsatzes des Indikators in den einzelnen Verfahren ergeben sich signifikante positive Zusammenhänge zwischen AV_6 und dem Einsatz in *Zielvereinbarungen zwischen Hochschulleitung und Fakultät* ($r(218) = .18, p = .008$), *Zielvereinbarungen zwischen Fakultät und Institut* ($r(113) = .28, p = .002$), *Leistungszulagen im Rahmen der W-Besoldung* ($r(393) = .14, p = .006$) sowie in *Bleibe- und Berufungsverhandlungen* ($r(434) = .16, p = .001$). Bei den *Bleibe- und Berufungsverhandlungen* zeigt sich zudem eine signifikante Korrelation hinsichtlich der Höhe der Gewichtung des Indikators und der Variablen AV_6 ($r(332) = .15, p = .006$).

In der fachspezifischen Perspektive ergeben sich bei der isolierten Betrachtung der einzelnen Verfahren nur vereinzelt signifikante Korrelationen. In der Psychologie korreliert die Höhe der Gewichtung des Indikators im Verfahren der *indikatorenbasierten Mittelzuweisung zwischen Hochschulleitung und Fakultät* signifikant positiv mit *AV_6* ($r(47) = .37, p = .008$). In der Soziologie geht der Einsatz des Indikators in den *Leistungszulagen im Rahmen der W-Besoldung* ($r(44) = .49, p = .001$) bzw. in *Bleibe- und Berufungsverhandlungen* ($r(51) = .38, p = .005$) mit einem höheren eingeschätzten Stellenwert des Einwerbens von DFG-Drittmitteln in der Forschungsarbeit der Professorinnen und Professoren einher.

Einwerben von sonstigen Drittmitteln

Im Hinblick auf die *sonstigen Drittmittel (AV_7)* lässt sich zunächst ein tendenziell ähnliches Bild wie bei den DFG-Drittmitteln erkennen. Wie aus Tabelle 23 ersichtlich wird, zeigt sich insbesondere hinsichtlich der Korrelationen zwischen den Variablen *KV_REP_7* bzw. *KV_ERK_7* und *AV_7* ein vergleichbares Ergebnis mit nahezu durchgehend signifikant positiven Zusammenhängen. Lediglich die Korrelation zwischen *KV_REP_7* und *AV_7* im Fach Psychologie ist – wie auch bei den DFG-Drittmitteln – nicht signifikant.

Tabelle 23: Korrelationen zur Variablen *AV_7: Einwerben von sonstigen Drittmitteln*

		gesamt	EZW	POL	PSY	SOZ	WIWI
UV_7	<i>r</i>	.073	.016	.160	.022	.102	.009
	<i>r_s</i>	.073	-.007	.175	.018	.091	.006
	<i>n</i>	583	134	49	109	70	206
KV_REP_7	<i>r</i>	.330***	.268***	.444***	.152	.280*	.239***
	<i>r_s</i>	.336***	.240**	.469***	.156	.288*	.243***
	<i>n</i>	623	146	53	117	75	217
KV_ERK_7	<i>r</i>	.484***	.434***	.604***	.297***	.412***	.456***
	<i>r_s</i>	.482***	.426***	.626***	.294***	.431***	.457***
	<i>n</i>	613	147	52	114	74	211
KV_IM	<i>r</i>	.041	.079	-.067	.164	.037	-.010
	<i>r_s</i>	.069	.145	.063	.142	.056	.017
	<i>n</i>	624	146	54	117	74	219

Anmerkung. UV_7: Einsatz der Höhe der eingeworbenen sonstigen Drittmittel als Forschungsindikator in indikatorenbasierten Steuerungs- und Anreizverfahren (dichotom), KV_REP_7: Bedeutung der Höhe der eingeworbenen sonstigen Drittmittel für den Reputationszuwachs in der Scientific Community, KV_ERK_7: Beitrag der Durchführung von sonstigen Drittmittelprojekten für den wissenschaftlichen Erkenntnisgewinn, KV_IM: Index Intrinsische Motivation; *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig): * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

Quelle: eigene Berechnungen

Bei den anderen beiden Variablen *UV_7* und *KV_IM* ergeben sich dagegen Unterschiede zu den Analysen der DFG-Drittmittel. Während sich bei diesen in der Gesamtstichprobe ein signifikant positiver Zusammenhang zwischen dem Einsatz des Forschungsindikators (*UV_6*) und der Variablen *AV_6* sowie zwischen dem Maß der intrinsischen Motivation und *AV_6* zeigte, lassen sich diese beiden signifikanten Zusammenhänge bei den sonstigen Drittmitteln nicht mehr beobachten. Sowohl die Korrelation zwischen *UV_7* und *AV_7* als auch zwischen *KV_IM* und *AV_7* ist auf dem 5%-Niveau nicht signifikant. Auch die fachspezifischen Analysen ergeben durchgehend nicht-signifikante Zusammenhänge.

Während der generelle Einsatz der Höhe der eingeworbenen sonstigen Drittmittel als Forschungsindikator in Steuerungs- und Anreizverfahren (*UV_7*) weder in der Gesamtstichprobe noch in den einzelnen Fächern signifikant mit *AV_7* korreliert, zeigen sich bei der isolierten Betrachtung der einzelnen Verfahren sowie bei der Gewichtung des Indikators in diesen, teils signifikante Zusammenhänge.

Für die Gesamtstichprobe ergibt sich zunächst ein signifikant positiver, aber schwacher Zusammenhang zwischen der Gewichtung des Forschungsindikators *Höhe der eingeworbenen sonstigen Drittmittel* in den Verfahren insgesamt und der Variablen *AV_7* ($r(481) = .11, p = .020$). Der Einsatz des Indikators in den Verfahren *Leistungszulagen im Rahmen der W-Besoldung* ($r(383) = .15, p = .004$) sowie *Bleibe- und Berufungsverhandlungen* ($r(421) = .12, p = .015$) korreliert signifikant positiv mit *AV_7*, ebenso wie die Höhe der Gewichtung des Indikators in diesen beiden Verfahren ($r(282) = .14, p = .017$ bzw. $r(324) = .22, p < .001$). Zudem zeigt sich in der Gesamtstichprobe ein signifikant positiver Zusammenhang zwischen der Gewichtung des Indikators im Verfahren *indikatorbasierte Mittelzuweisung zwischen Hochschulleitung und Fakultät* und *AV_7* ($r(217) = .21, p = .002$).

In den fachspezifischen Analysen lassen sich unter Berücksichtigung der teils geringen Fallzahlen lediglich für die Fächer Politikwissenschaften ($r(31) = .49, p = .004$) und Soziologie ($r(39) = .34, p = .032$) signifikante Korrelationen zwischen der Gewichtung des Indikators in *Bleibe- und Berufungsverhandlungen* und der Variablen *AV_7* nachweisen.

Publizieren als Koautor / in Mehrautorenschaft

Einen weiteren im Fragebogen thematisierten Aspekt der Forschungsarbeit stellt das *Publizieren als Koautor bzw. in Mehrautorenschaft* dar. Diese Forschungstätigkeit kann nicht unabhängig von anderen hier untersuchten Forschungstätigkeiten wie z. B. dem Publizieren in Zeitschriften mit Peer-Review-Verfahren oder dem Publizieren in high ranked journals gedacht werden. Das Publizieren zusammen mit anderen Autoren ist schon

immer gängige Praxis der wissenschaftlichen Arbeit und in manchen, vor allem naturwissenschaftlichen Disziplinen, mehr die Regel als die Ausnahme (vgl. zu den verschiedenen Publikationskulturen Alexander von Humboldt Stiftung 2009). Zu erwarten sind daher zum einen positive Zusammenhänge mit dem generellen Stellenwert von einzelnen Publikationstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren (z. B. hoher Stellenwert von peer-reviewed Artikeln korrespondiert mit hohem Stellenwert von Koautorenschaften). Zum anderen sind aufgrund der Möglichkeit, durch das Publizieren in Ko- bzw. Mehrautorenschaft den eigenen Publikationsoutput oder die eigene Zitationsquote (mit möglicherweise geringerem Aufwand als bei Einzelpublikationen) zu erhöhen, auch Zusammenhänge mit möglichen handlungsleitenden Anreizen, die erhöhten Output ‚belohnen‘ (z. B. indikatorenbasierte Steuerungsverfahren), anzunehmen. Im Folgenden werden nun analog zu den vorangegangenen Analysen einige dieser möglichen bivariaten Zusammenhänge dargestellt und auf statistische Signifikanz geprüft.⁶⁰

Betrachtet man den eingeschätzten Stellenwert, den das Publizieren in verschiedenen Publikationsformaten für die Professorinnen und Professoren in der Forschungsarbeit hat, ergeben sich für die Gesamtstichprobe hinsichtlich des Zusammenhangs mit dem eingeschätzten Stellenwert des Publizierens als Koautor bzw. in Mehrautorenschaft die in Tabelle 24 dargestellten Ergebnisse.

Während sich für das Publizieren in peer-reviewed Zeitschriften und in Zeitschriften mit hohem Impact Factor / in high ranked journals erwartungsgemäß signifikant positive Korrelationen nachweisen lassen, zeigt sich für das Verfassen anderer Publikationen und für das Publizieren in Zeitschriften ohne Peer-Review eine schwache negative Korrelation bzw. kein signifikanter Zusammenhang. Die negative Korrelation lässt sich damit erklären, dass unter die *anderen Publikationsformate* auch Monografien fallen, welche oft in Alleinautorenschaft verantwortet werden.

Die fachspezifischen Analysen kommen in den Wirtschaftswissenschaften zu einem übereinstimmenden Ergebnis mit der Gesamtbetrachtung, während in der Psychologie und der Soziologie der signifikante negative Zusammenhang zwischen dem Stellenwert des Verfassens anderer Publikationen und dem Publizieren als Koautor bzw. in Mehrautorenschaft verschwindet. In den Erziehungswissenschaften ergibt sich sogar ein signifikant positiver Zusammenhang zwischen den beiden Variablen. Dies kann sowohl Unterschieden in der Publikationskultur geschuldet (z. B. vermehrt Mehrautorenschaften bei Monografien in den Erziehungswissenschaften) als auch auf die Mehrdimensionalität

⁶⁰ Die Korrelationsanalysen erfolgten analog zu den vorangegangenen Auswertungen mittels des Bravais-Pearson-Korrelationskoeffizienten und Spearmans Rho. Zusammenhänge werden nur dann als signifikant interpretiert, wenn beide Koeffizienten ein signifikantes Ergebnis anzeigten. Im Text werden aus Gründen der Übersichtlichkeit wieder jeweils nur die Kennzahlen für den Bravais-Pearson-Koeffizienten berichtet.

des Items *Andere Publikationen* zurückzuführen sein. In diesem Fall würde es sich lediglich um ein methodisches Artefakt handeln.

Tabelle 24: *Korrelationen zwischen der Variablen „Stellenwert des Publizierens als Koautor / in Mehrautorenschaft“ und dem Stellenwert anderer Publikationstätigkeiten*

		gesamt	EZW	POL	PSY	SOZ	WIWI
Publizieren in Zeitschriften mit Peer-Review	<i>r</i>	.367***	.325***	.195	.182*	.286*	.558***
	<i>r_s</i>	.402***	.348***	.278*	.209*	.285*	.579***
	<i>n</i>	628	146	55	117	76	217
Publizieren in Zeitschriften ohne Peer-Review	<i>r</i>	-.074	.052	-.177	.054	-.061	-.101
	<i>r_s</i>	-.091*	.084	-.103	.051	-.064	-.121
	<i>n</i>	623	146	54	116	76	214
Verfassen anderer Publikationen	<i>r</i>	-.083*	.207***	-.260	.047	.049	-.174*
	<i>r_s</i>	-.089*	.198*	-.189	-.004	.068	-.179**
	<i>n</i>	626	145	55	117	76	216
Publizieren in Zeitschriften mit hohem Impact Factor	<i>r</i>	.367***	.277***	.258	.185*	.417***	.502***
	<i>r_s</i>	.378***	.243**	.286*	.205*	.414***	.520***
	<i>n</i>	624	145	55	117	76	216

Anmerkung. *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig):

p* ≤ .05, *p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

In den Politikwissenschaften ergibt sich unter Verwendung des Bravais-Pearson-Korrelationskoeffizienten überraschenderweise bei keiner der dargestellten Variablen ein signifikanter Zusammenhang. Die alternative Überprüfung mittels Spearmans Rho und die jeweiligen Vorzeichen der Koeffizienten zeigen aber eine ähnliche Tendenz an.

Der Einsatz der verschiedenen publikations- und zitationsbasierten Forschungsindikatoren in Steuerungs- und Anreizverfahren generell korreliert nur vereinzelt signifikant mit dem eingeschätzten Stellenwert der Forschungstätigkeit *Publizieren als Koautor / in Mehrautorenschaft*. In der Gesamtstichprobe ergibt sich ein jeweils signifikant schwacher negativer Zusammenhang mit dem Einsatz der Indikatoren *Anzahl von Publikationen in Zeitschriften ohne Peer-Review-Verfahren* ($r(509) = -.16, p < .001$) sowie der *Anzahl anderer Publikationen* ($r(520) = -.13, p = .002$). Hier könnte wieder eine Rolle spielen, dass bestimmte Publikationsformate (z. B. Monografien) mehr als andere öfter in Alleinautorenschaft verfasst werden. Eine (wirksame) Förderung von solchen Formaten durch externe Anreize würde damit zu einer Verringerung von in Mehrautorenschaft verfassten Publikationsformaten bzw. zu einem geringeren Stellenwert von diesen in der aktuellen Forschungsarbeit führen.

In den fachspezifischen Analysen zeigt sich, dass die ermittelten signifikanten Zusammenhänge in der Gesamtstichprobe im Wesentlichen auf den Substichproben der Politikwissenschaften und der Wirtschaftswissenschaften beruhen, bei denen sich in beiden Fällen dieselben signifikanten Korrelationen zeigten. Bei den übrigen Fächern sind dagegen hinsichtlich des Einsatzes von publikations- und zitationsbasierten Forschungsindikatoren und dem Stellenwert des Publizierens in Mehrautorenschaft keine signifikanten Zusammenhänge zu beobachten.

Betreiben von Grundlagenforschung und anwendungsbezogener Forschung

Wie bereits bei der im vorangegangenen Teilkapitel dargestellten Forschungstätigkeit des Publizierens in Ko- und Mehrautorenschaft werden die folgenden Analysen nach einem etwas anderen Schema als bei den übrigen Variablen (AV_1 bis AV_7) durchgeführt. Da für den abgefragten Stellenwert des Betriebens von Grundlagenforschung und anwendungsbezogener Forschung in der Forschungsarbeit der Professorinnen und Professoren keine direkten Steuerungs- oder Anreizmechanismen bzw. Forschungsindikatoren existieren – wie das z. B. in Bezug auf die Drittmittelinwerbung der Fall ist – sollen hier mögliche indirekte Zusammenhänge in den Blick genommen werden. Die Analyse der Zusammenhänge zwischen den hier dargestellten Variablen hat dabei mehr explorativen Charakter.

Zunächst kann festgehalten werden, dass der eingeschätzte Stellenwert von Grundlagenforschung und anwendungsbezogener Forschung erwartungsgemäß signifikant negativ miteinander korreliert ($r(614) = -.33, p < .001$). Dies gilt sowohl für die Gesamtstichprobe als auch – mit Ausnahme der Politikwissenschaften, bei welchen mit $r(50) = -.25, p = .076$ keine Signifikanz auf dem 5%-Niveau nachgewiesen werden konnte – für alle fachspezifischen Substichproben. Hier scheinen sich insgesamt doch relativ eindeutige Orientierungen in der Forschungsrichtung zu ergeben.

Betrachtet man nun mögliche Zusammenhänge mit dem eingeschätzten Stellenwert der anderen Forschungstätigkeiten, zeigen sich für die Gesamtstichprobe die in Tabelle 25 dargestellten Ergebnisse. Der *Stellenwert von Grundlagenforschung* in der Forschungsarbeit korreliert signifikant positiv mit dem Stellenwert der Forschungstätigkeiten *Publizieren in Zeitschriften mit Peer-Review-Verfahren*, *Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals*, *Publizieren als Koautor / in Mehrautorenschaft* sowie dem *Einwerben von DFG-Drittmitteln*.

Tabelle 25: Korrelationen zwischen dem Stellenwert von Grundlagenforschung bzw. anwendungsbezogener Forschung und dem Stellenwert anderer Forschungstätigkeiten

		Betreiben von Grundlagen- forschung	Betreiben von anwen- dungsbezogener For- schung
Publizieren in Zeitschriften mit Peer-Review-Verfahren	<i>r</i>	.188***	-.034
	<i>r_s</i>	.213***	-.023
	<i>n</i>	621	621
Publizieren in Zeitschriften ohne Peer-Review-Verfahren	<i>r</i>	-.079*	.185***
	<i>r_s</i>	-.091*	.186***
	<i>n</i>	617	616
Verfassen anderer Publika- tionen	<i>r</i>	.034	.119**
	<i>r_s</i>	.045	.132**
	<i>n</i>	620	619
Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals	<i>r</i>	.135**	-.072
	<i>r_s</i>	.148***	-.065
	<i>n</i>	618	618
Publizieren als Koautor / in Mehrautorenschaft	<i>r</i>	.096*	.068
	<i>r_s</i>	.104*	.092*
	<i>n</i>	616	617
Betreuung von Promotionen	<i>r</i>	.056	.133**
	<i>r_s</i>	.071	.161***
	<i>n</i>	609	609
Einwerben von Drittmitteln der DFG	<i>r</i>	.263***	-.037
	<i>r_s</i>	.266***	-.023
	<i>n</i>	619	619
Einwerben von sonstigen Drittmitteln	<i>r</i>	-.088*	.293***
	<i>r_s</i>	-.079	.306***
	<i>n</i>	617	617

Anmerkung. *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig):

p* ≤ .05, *p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

Mit Blick auf den Stellenwert der *anwendungsbezogenen Forschung* ergibt sich dagegen für keinen dieser Zusammenhänge ein signifikantes Ergebnis. Stattdessen finden sich signifikant positive Korrelationen bezüglich des Stellenwertes des *Publizierens in Zeitschriften ohne Peer-Review-Verfahren*, dem *Verfassen anderer Publikationen*, der *Betreuung von Promotionen* sowie dem *Einwerben von sonstigen Drittmitteln*. Für den Stellenwert der Grundlagenforschung zeigen sich zudem signifikante, aber sehr schwache negative Korrelationen hinsichtlich des Publizierens in Zeitschriften ohne Peer-Review-

Verfahren sowie dem Einwerben sonstiger Drittmittel. In Bezug auf die Drittmittelvariablen deckt sich die Stärke und Richtung der ermittelten Korrelationen mit der grundsätzlich stärkeren Fokussierung auf Grundlagenforschung bei DFG-geförderten Projekten (vgl. Fritsch et al. 2007, S. 66; Grözinger und Fromm 2014, S. 114). Zudem scheint das Betreiben von Grundlagenforschung eher mit der Veröffentlichung der Ergebnisse in peer-reviewed bzw. high ranked journals einherzugehen, während mehr anwendungsbezogene orientierte Forscherinnen und Forscher vermehrt (auch) in nicht-begutachteten Zeitschriften und anderen Publikationsformaten veröffentlichen.

In den fachspezifischen Analysen können diese Ergebnisse für die Politikwissenschaften und die Psychologie bestätigt werden. Eine Ausnahme bildet hier die Korrelation zwischen dem Stellenwert von anwendungsbezogener Forschung und der Betreuung von Promotionen, die sich nur für das Fach Wirtschaftswissenschaften nachweisen lässt ($r(212) = .16, p = .022$).

Auf die Erziehungswissenschaften, die Soziologie und die Wirtschaftswissenschaften sind die Ergebnisse der Gesamtstichprobe nur eingeschränkt übertragbar. Bei den Erziehungswissenschaften zeigen sich lediglich hinsichtlich der beiden Drittmittelvariablen übereinstimmende Ergebnisse ($r(140) = .32, p < .001$ zwischen DFG-Drittmittel und Grundlagenforschung und $r(140) = .36, p < .001$ zwischen sonstige Drittmittel und anwendungsbezogene Forschung). In der Soziologie lassen sich die signifikanten Korrelationen für die Variablen *Publizieren in Zeitschriften mit Peer-Review-Verfahren* ($r(73) = .39, p = .001$), *Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals* ($r(73) = .24, p = .039$) sowie dem *Einwerben von DFG-Drittmitteln* ($r(73) = .23, p = .048$) bestätigen. In den Wirtschaftswissenschaften stimmen die Ergebnisse bei den Variablen *Publizieren in Zeitschriften ohne Peer-Review-Verfahren* ($r(214) = -.18, p = .009$ für Grundlagenforschung und $r(213) = .20, p = .003$ für anwendungsbezogene Forschung), *Verfassen anderer Publikationen* ($r(215) = .17, p = .015$), *Einwerben von sonstigen Drittmitteln* ($r(215) = .30, p < .001$) und wie erwähnt bei der *Betreuung von Promotionen* überein. Die Ergebnisse in der Gesamtbetrachtung sind damit wohl zum Teil auf Drittvariableneffekte (insbesondere Fachzugehörigkeit; vgl. zum Stellenwert verschiedener Forschungstätigkeiten je Fach Kapitel 6.1.6) zurückzuführen. Nichtsdestotrotz zeigen diese Resultate im Abgleich mit den fachspezifischen Analysen, dass unterschiedliche Forschungsorientierungen im Sinne von Grundlagenforschung gegenüber anwendungsbezogener Forschung mit jeweils spezifischen Schwerpunktsetzungen in den Publikationstätigkeiten (Formaten) und bei der Drittmittelinwerbung einhergehen.

Im Folgenden werden nun mögliche Zusammenhänge zwischen dem Einsatz verschiedener Forschungsindikatoren und einer eher grundlagen- bzw. anwendungsorientierten Forschung untersucht. Forschungsindikatoren, welche sich direkt auf die Aspekte Grundlagenforschung und anwendungsbezogene Forschung beziehen, existieren wie

bereits erwähnt nicht. Denkbar ist aber, dass indikatorenbasierte Steuerungsversuche indirekt über die verwendeten Indikatoren einen Einfluss auf die Ausrichtung der Forschung bzw. der Forschungsthemen der Wissenschaftlerinnen und Wissenschaftler nehmen. Dies wäre dann der Fall, wenn den in den Steuerungsverfahren verwendeten Indikatoren durch eine bestimmte Ausrichtung der Forschung eher gerecht werden kann. Wird z. B. die Höhe der eingeworbenen DFG-Drittmittel als Indikator herangezogen, verspricht ein auf Grundlagenforschung ausgerichtetes Forschungsprofil möglicherweise mehr Erfolg bei der Drittmittelinwerbung als mehr anwendungsbezogene Forschungsthemen. Ebenso könnten Anreize zur Erhöhung der Publikationen in Zeitschriften mit hohem Impact Factor dazu führen, dass sich die Wahl der Forschungsthemen am momentanen wissenschaftlichen Mainstream orientiert, wenn dieser sich vermeintlich leichter in bestimmten Journals publizieren lässt (vgl. zu den möglichen Auswirkungen auf die Ausrichtung der Forschung Alberts 2013; Kieser 2012; Oancea 2014; Schimank 2015). Die Analyse der Zusammenhänge zwischen dem Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren und dem Stellenwert von Grundlagenforschung bzw. anwendungsbezogener Forschung ist daher durchaus von Interesse.

Betrachtet man nun die Ergebnisse dieser Analysen, zeigen sich in der Gesamtstichprobe allerdings keine und in den fachspezifischen Analysen nur zwei signifikante Zusammenhänge. Lediglich in den Politikwissenschaften korreliert der generelle Einsatz des Indikators *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals* signifikant positiv mit dem Stellenwert der Grundlagenforschung ($r(38) = .37, p = .018$). Zudem ergibt sich in den Wirtschaftswissenschaften eine signifikant negative Korrelation zwischen dem Einsatz der *Anzahl betreuter Promotionen* als Forschungsindikator in den Verfahren insgesamt und dem eingeschätzten Stellenwert der Grundlagenforschung ($r(192) = -.18, p = .013$). Sofern die hier dargestellten rein bivariaten Analysen überhaupt erste interpretative Aussagen erlauben, zeigen die Ergebnisse zunächst also keinen (größeren) Zusammenhang zwischen dem Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren und dem Stellenwert von Grundlagen- bzw. anwendungsbezogener Forschung in der Forschungsarbeit der Professorinnen und Professoren an.

Zum Abschluss dieses Teilkapitels sollen noch kurz mögliche Zusammenhänge mit der intrinsischen Forschungsmotivation der Befragten in den Blick genommen werden. Hierbei ergibt sich für die Gesamtstichprobe eine signifikant positive Korrelation zwischen dem Grad der intrinsischen Motivation der Befragten und der Höhe des Stellenwerts von Grundlagenforschung in ihrer Forschungsarbeit ($r(615) = .31, p < .001$). Dieser Zusammenhang bestätigt sich auch in den einzelnen fachspezifischen Substichproben, wobei die Korrelation in den Politikwissenschaften mit $r(50) = .41, p = .003$ am stärksten und in der Psychologie mit $r(114) = .20, p = .031$ am schwächsten ausfällt. In

der Psychologie zeigt sich zudem auch eine signifikante Korrelation zwischen der intrinsischen Forschungsmotivation und dem eingeschätzten Stellenwert von anwendungsbezogener Forschung ($r(115) = .26, p = .006$).

Im nächsten Teilkapitel wird die intrinsische Forschungsmotivation nun nochmals genauer in den Blick genommen und untersucht, inwieweit sich in den bivariaten Analysen Zusammenhänge zwischen dem Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren und der intrinsischen Motivation der Befragten nachweisen lassen.

Einsatz von Forschungsindikatoren und intrinsische Motivation

Gegenstand der nachfolgenden Analysen sind mögliche Zusammenhänge zwischen der intrinsischen Forschungsmotivation der befragten Professorinnen und Professoren und dem Einsatz der einzelnen Forschungsindikatoren in den Steuerungs- und Anreizverfahren insgesamt. Die zugehörigen Korrelationskoeffizienten und Fallzahlen sind in Tabelle 26 dargestellt.

Ein Blick auf die dargestellten Koeffizienten macht deutlich, dass sich ein signifikant negativer Zusammenhang in der Gesamtstichprobe nur bei den Indikatoren *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* und *Anzahl anderer Publikationen* zeigt. Bei den fachspezifischen Analysen ergibt sich lediglich in der Psychologie ein signifikant negativer Zusammenhang bei genau diesen Indikatoren. Die signifikanten Korrelationen in der Gesamtstichprobe sind also wohl überwiegend auf die Ergebnisse in der Substichprobe der Psychologie zurückzuführen. Der signifikant positive Wert des Bravais-Pearson-Korrelationskoeffizienten beim Indikator *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals* in den Politikwissenschaften kann durch Spearmans Rho nicht bestätigt werden.

Auch wenn sich in den bivariaten Analysen nur vereinzelt signifikante Korrelationen ergeben, gehen die nicht-signifikanten Zusammenhänge – mit Ausnahme der Substichprobe Soziologie – zumindest hinsichtlich des Vorzeichens zu einem großen Teil in die theoretisch erwartete Richtung (vgl. Kapitel 4.2). Dass die Ergebnisse nicht deutlicher ausfallen, ist vermutlich auch auf die stark linksschiefe Verteilung des Index zur intrinsischen Motivation zurückzuführen (vgl. Kapitel 6.1.7).

Tabelle 26: Korrelationen zwischen der Variablen „Einsatz Forschungsindikator in Steuerungs- und Anreizverfahren insgesamt“ und der intrinsischen Forschungsmotivation

		gesamt	EZW	POL	PSY	SOZ	WIWI
Anzahl Publikationen in	r	-.053	-.106	-.066	-.072	.106	-.054
Zeitschriften mit Peer-	r_s	-.048	-.076	-.076	-.067	.137	-.069
Review	n	574	124	47	109	68	212
Anzahl Publikationen in	r	-.129**	-.164	-.160	-.291**	.060	-.112
Zeitschriften ohne	r_s	-.121**	-.136	-.154	-.268**	.105	-.118
Peer-Review	n	519	115	43	94	64	190
Anzahl anderer Publi-	r	-.105*	-.151	-.118	-.261**	.081	-.077
kationen	r_s	-.093*	-.102	-.113	-.237*	.110	-.087
	n	529	117	44	98	66	191
Anzahl Publikationen in	r	.023	-.131	.332*	.061	.111	.059
Zeitschriften mit hohem	r_s	.036	-.051	.310	.031	.135	.048
IF/ high ranked Jour-	n	514	104	40	102	57	200
nals							
Anzahl der Zitationen /	r	.017	.022	.121	.146	.010	-.021
Zitationsindizes	r_s	.007	.006	.150	.129	.087	-.079
	n	468	99	41	94	50	176
Anzahl betreuter Pro-	r	-.024	-.036	-.177	-.040	.154	-.061
motionen	r_s	-.020	-.014	-.240	-.014	.175	-.078
	n	550	120	43	109	62	199
Höhe der eingeworbe-	r	.030	-.021	-.036	-.067	.071	.042
nen Drittmittel der DFG	r_s	.016	-.060	-.077	-.059	.089	.027
	n	592	130	48	113	71	212
Höhe der sonstigen	r	.003	-.048	-.121	-.058	.108	.014
eingeworbenen Dritt-	r_s	-.009	-.087	-.176	-.048	.139	-.001
mittel	n	591	135	48	109	70	211

Anmerkung. r = Korrelationskoeffizient nach Bravais-Pearson, r_s = Spearmans Rho; Signifikanz (zweiseitig):

* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

Quelle: eigene Berechnungen

Weitere Analysen: Alter, Beschäftigungsdauer, Besoldungsart und Geschlecht

Die bivariaten Analysen werden nun mit einem knappen Überblick zu weiteren möglichen Einfluss- bzw. Kontrollvariablen abgeschlossen. Im Einzelnen sind dies das Alter und die Beschäftigungsdauer der Befragten, die Besoldungsart (W- oder C-Besoldung) und das Geschlecht. Im Fokus der Analysen stehen dabei wieder mögliche Zusammenhänge zwischen den genannten Variablen und dem eingeschätzten Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren. Die Ergebnisse der entsprechenden Korrelationsanalysen sind in Tabelle 27 aufgeführt.

Tabelle 27: Korrelationen zwischen dem Stellenwert verschiedener Forschungstätigkeiten und den Variablen Alter, Beschäftigungsdauer, Besoldungsart und Geschlecht

		Alter	Beschäftigungsdauer	Besoldungsart	Geschlecht
Publizieren in Zeitschriften mit Peer-Review-Verfahren	<i>r</i>	-.280***	-.180***	-.136**	-.016
	<i>r_s</i>	-.278***	-.205***	-.127**	-.007
	<i>n</i>	576	589	635	598
Publizieren in Zeitschriften ohne Peer-Review-Verfahren	<i>r</i>	.217***	.120***	.144***	.067
	<i>r_s</i>	.221***	.132***	.147***	.067
	<i>n</i>	571	584	630	593
Verfassen anderer Publikationen	<i>r</i>	.276***	.151***	.131**	.095*
	<i>r_s</i>	.267***	.155***	.129**	.099*
	<i>n</i>	574	587	633	596
Publizieren in Zeitschriften mit hohem Impact Factor / in high-ranked journals	<i>r</i>	-.236***	-.120***	-.074	-.043
	<i>r_s</i>	-.238***	-.157***	-.079*	-.039
	<i>n</i>	574	586	630	595
Publizieren als Koautor / in Mehrautorenschaft	<i>r</i>	-.233***	-.172***	-.095*	.013
	<i>r_s</i>	-.239***	-.181***	-.097*	.010
	<i>n</i>	571	582	628	592
Betreuung von Promotionen	<i>r</i>	.067	.067	.028	.037
	<i>r_s</i>	.061	.081	.022	.041
	<i>n</i>	564	576	621	585
Einwerben von Drittmitteln der DFG	<i>r</i>	-.075	-.111**	-.132**	.160***
	<i>r_s</i>	-.083*	-.129**	-.131**	.160***
	<i>n</i>	573	585	631	595
Einwerben von sonstigen Drittmitteln	<i>r</i>	.078	-.020	-.001	.120**
	<i>r_s</i>	.080	-.016	.006	.117**
	<i>n</i>	572	583	628	592
Betreiben von Grundlagenforschung	<i>r</i>	-.029	-0.33	-.083*	-.013
	<i>r_s</i>	-.025	-.029	-.075	-.002
	<i>n</i>	566	578	621	586
Betreiben von anwendungsbezogener Forschung	<i>r</i>	.068	.070	.102*	.072
	<i>r_s</i>	.066	.050	.098*	.095*
	<i>n</i>	566	577	621	586

Anmerkung. Alter und Beschäftigungsdauer sind metrische Variablen, die Besoldungsart und das Geschlecht dichotom mit den Ausprägungen 0 = männlich / W-Besoldung und 1 = weiblich / C-Besoldung; *r* = Korrelationskoeffizient nach Bravais-Pearson, *r_s* = Spearmans Rho; Signifikanz (zweiseitig): **p* ≤ .05, ***p* ≤ .01, ****p* ≤ .001

Quelle: eigene Berechnungen

Für die Variable *Alter* ergeben sich signifikante Korrelationen bei allen publikationsbasierten Forschungstätigkeiten. Dabei zeigt sich, dass mit steigendem Alter der Stellenwert des Publizierens in peer-reviewed Zeitschriften sowie in hochgerankten Zeitschriften niedriger eingeschätzt wird. Auch das Publizieren als Koautor bzw. in Mehrautorenschaft verliert anscheinend in den höheren Altersgruppen an Relevanz. Dagegen korreliert das Alter signifikant positiv mit den Forschungstätigkeiten des Publizierens in Zeitschriften ohne Peer-Review-Verfahren sowie dem Verfassen anderer Publikationen.

Analoge, allerdings weniger stark ausgeprägte Zusammenhänge zeigen sich hinsichtlich der Beschäftigungsdauer. Da die Beschäftigungsdauer nicht unabhängig vom Alter der Befragten gedacht werden kann, korreliert diese mit $r(563) = .82, p < .001$ erwartungsgemäß ihrerseits sehr stark mit dem Alter. Unter Kontrolle der Variablen *Alter* mittels der Berechnung partieller Korrelationen verschwinden die hier berichteten signifikanten Zusammenhänge teilweise, die Stärke des Zusammenhangs verringert sich oder die Richtung des Zusammenhangs kehrt sich sogar um.⁶¹ Signifikante Korrelationen zeigen sich unter Kontrolle des Alters hinsichtlich des Stellenwerts des Publizierens in Zeitschriften ohne Peer-Review ($r(554) = -.09, p = .037$), des Verfassens anderer Publikationen ($r(557) = -.13, p = .002$) sowie für den Stellenwert des Publizierens in Zeitschriften mit hohem Impact Factor / in high ranked journals ($r(557) = .13, p = .002$). Die Richtung des ursprünglichen Zusammenhangs kehrt sich dabei bei allen drei Variablen um. Professorinnen und Professoren derselben Altersgruppe schätzen mit steigender Beschäftigungsdauer den Stellenwert des Publizierens in Zeitschriften ohne Peer-Review-Verfahren sowie des Verfassens anderer Publikationen in der eigenen Forschungsarbeit entsprechend geringer ein, während das Publizieren in hochgerankten Zeitschriften an Gewicht gewinnt. Zudem ergibt sich eine signifikant negative Korrelation zwischen Beschäftigungsdauer und dem Stellenwert des Einwerbens von DFG-Drittmitteln, die auch unter Kontrolle des Alters bestehen bleibt. Für das Einwerben von sonstigen Drittmitteln ergibt sich dagegen nur dann ein signifikant negativer Zusammenhang mit der Beschäftigungsdauer ($r(555) = -.11, p = .007$), wenn das Alter der Befragten kontrolliert wird. Das Einwerben von Drittmitteln scheint also insbesondere für Professorinnen und Professoren am Anfang der Karrierelaufbahn von erhöhter Relevanz zu sein.

Auch für die dichotome Variable *Besoldungsart* (0 = W-Besoldung / 1 = C-Besoldung) gilt, dass diese stark mit dem Alter ($r(577) = .54, p < .001$) und insbesondere auch mit der Beschäftigungsdauer ($r(590) = .65, p < .001$) korreliert. Mit der ‚alten‘ C-Besoldung gehen also ein höheres Alter und eine längere Beschäftigungsdauer der Professorinnen und Professoren einher. Mögliche Drittvariableneffekte müssen daher auch in dieser

⁶¹ Auch wenn dieses Kapitel lediglich bivariate Analysen zum Gegenstand hat, werden aufgrund der offensichtlichen Abhängigkeiten der Variablen untereinander hier auch knapp mögliche Drittvariableneffekte diskutiert.

Analyse berücksichtigt werden. Ohne Drittvariablenkontrolle zeigen sich zunächst überwiegend übereinstimmende Ergebnisse zu den Korrelationen bei den Variablen *Alter* und *Beschäftigungsdauer*. Zudem ergibt sich eine zwar schwache, aber signifikant positive Korrelation mit dem Stellenwert des Betreibens von anwendungsbezogener Forschung und eine negative Korrelation hinsichtlich des Betreibens von Grundlagenforschung. Die Signifikanz des letztgenannten Zusammenhangs lässt sich unter Verwendung von Spearmans Rho allerdings nicht bestätigen ($r_s(619) = -.08, p = .062$).

Berechnet man die partiellen Korrelationen unter Kontrolle der Beschäftigungsdauer⁶², lassen sich signifikante Zusammenhänge mit der Besoldungsart nur noch für den eingeschätzten Stellenwert des Publizierens in Zeitschriften ohne Peer-Review ($r(581) = .10, p = .012$), des Einwerbens von DFG-Drittmitteln ($r(582) = -.09, p = .031$), des Betreibens von Grundlagenforschung ($r(575) = -.11, p = .010$) sowie des Betreibens von anwendungsbezogener Forschung ($r(574) = .09, p = .033$) nachweisen.

C-besoldete Professorinnen und Professoren schätzen damit tendenziell die Relevanz des Publizierens ohne Peer-Review in der eigenen Forschungsarbeit höher ein und die des Einwerbens von DFG-Drittmitteln geringer. Das Betreiben von Grundlagenforschung hat bei den Befragten mit C-Besoldung zudem einen etwas geringeren Stellenwert als bei denen mit W-Besoldung, während im Hinblick auf die anwendungsbezogene Forschung Gegenteiliges gilt.

In Bezug auf das Geschlecht zeigt sich, dass Frauen den Stellenwert des Verfassens anderer Publikationen sowie des Einwerbens von DFG- und sonstigen Drittmitteln etwas höher einschätzen. Hierbei spielen sicherlich auch Fachspezifika eine Rolle: Der Anteil der Frauen fällt in den Wirtschaftswissenschaften mit 19.2 % am geringsten aus (vgl. Kapitel 5.3.2). Da dies auch das Fach ist, in welchem der Stellenwert der oben genannten Forschungstätigkeiten am geringsten eingeschätzt wird (vgl. Kapitel 6.1.6), lassen sich die geschlechtsspezifischen Unterschiede primär durch die Verteilung der Geschlechter auf die einzelnen Fächer erklären.

Ergebniszusammenfassung der bivariaten Analysen

Abschließend werden an dieser Stelle die wesentlichen Erkenntnisse aus den bivariaten Analysen zusammengefasst, bevor im nächsten Kapitel zu den multivariaten Hypothesentests übergegangen wird.

Zunächst kann festgehalten werden, dass die Ergebnisse der bivariaten Analysen grundsätzlich einige der in den Forschungshypothesen formulierten Zusammenhänge (vgl. Kapitel 4.2) widerspiegeln und damit eine vorläufige Bestätigung der Hypothesen

⁶² Die signifikant positive Korrelation zwischen Alter und Besoldungsart verschwindet unter Kontrolle auf die Beschäftigungsdauer. Daher wurde für die partiellen Korrelationen zwischen Besoldungsart und dem Stellenwert der verschiedenen Forschungstätigkeiten nur auf die Beschäftigungsdauer kontrolliert.

nahelegen. Dies gilt insbesondere für die Zusammenhänge zwischen dem Stellenwert verschiedener Forschungstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren (z. B. Betreuung von Promotionen) und der eingeschätzten Relevanz von diesen für den Reputationszuwachs in der Scientific Community sowie deren eingeschätzten Beitrag für den wissenschaftlichen Erkenntnisgewinn.

Für die intrinsische Motivation und den Einsatz von Forschungsindikatoren in Steuerungs- und Anreizverfahren zeigen sich dagegen in der bivariaten Betrachtung nur vereinzelte signifikante Zusammenhänge mit dem eingeschätzten Stellenwert der einzelnen Forschungstätigkeiten. In Bezug auf den Einsatz von Forschungsindikatoren zeigt sich aber auch, dass in der Detailbetrachtung des Einsatzes der Indikatoren in den einzelnen Verfahren teils signifikante Korrelationen ermittelt werden können, die sich in der Gesamtbetrachtung in dieser Form nicht ergeben. So korreliert beispielsweise der Einsatz des Indikators *Anzahl betreuter Promotionen* in den Steuerungs- und Anreizverfahren insgesamt nicht mit dem eingeschätzten Stellenwert der Betreuung von Promotionen in der Forschungsarbeit. In der isolierten Betrachtung des Einsatzes des Indikators *Anzahl betreuter Promotionen* in den Verfahren *Leistungszulagen im Rahmen der W-Besoldung* sowie *Bleibe- und Berufungsverhandlungen* zeigt sich dagegen ein signifikant positiver Zusammenhang mit dem eingeschätzten Stellenwert der Betreuung von Promotionen. Im Hinblick auf die bereits öfters angesprochene Fallzahlenproblematik bei zunehmendem Grad der Ausdifferenzierung der Untersuchungsvariablen muss in den nachfolgenden multivariaten Analysen aber auf die Verwendung von verfahrensspezifischen Variablen (z. B. *Einsatz Indikator Anzahl betreuter Promotionen in fakultätsinternen Zielvereinbarungen*) verzichtet werden. Als unabhängige Variable wird stattdessen der Einsatz der verschiedenen Forschungsindikatoren in den *Steuerungs- und Anreizverfahren insgesamt* herangezogen.

Weiter zeigten sich bei den bivariaten wie auch schon bei den deskriptiven Analysen für die meisten der untersuchten Variablen teils deutliche Unterschiede zwischen den Fächern. Eine alleinige Betrachtung der Gesamtstichprobe wird diesen Fachunterschieden nicht gerecht und kann, z. B. aufgrund sich gegenseitig neutralisierender Effekte zweier Fächer, zu falschen Schlussfolgerungen verleiten. Daher wird in den nachfolgenden multivariaten Analysen auch die Fachzugehörigkeit kontrolliert.

Als abschließende Erkenntnis aus den bivariaten Analysen ist festzuhalten, dass sich überwiegend übereinstimmende Ergebnisse zwischen dem Korrelationskoeffizienten nach Bravais-Pearson und der parameterfreien Kennzahl Spearmans Rho ergeben. Etwaige Verletzungen von Verteilungsannahmen (vor allem hinsichtlich der teils stark linksschiefen Verteilungen der Variablen) scheinen also zu keinen größeren Verzerrun-

gen der Ergebnisse geführt zu haben. Hieraus lässt sich die vorsichtige Annahme ableiten, dass auch in den nachfolgenden multivariaten Verfahren mögliche Verletzungen der Modellvoraussetzungen nicht zu stark ins Gewicht fallen sollten.

Die vorangegangenen bivariaten Analysen haben zu einem ersten Überblick über die Datenstruktur und Zusammenhänge zwischen einzelnen Variablen verholfen. Über die bivariaten Zusammenhänge hinausgehende Abhängigkeiten der Untersuchungsvariablen untereinander blieben dabei überwiegend unberücksichtigt. Diesem blinden Fleck in den bisherigen Analysen soll nun begegnet werden, indem nachfolgend mit Hilfe von multiplen Regressionsmodellen Drittvariableneffekte kontrolliert und Interaktionseffekte zwischen den Variablen analysiert werden. Im nächsten Kapitel soll also die Zusammenhangsstruktur aller für die Forschungsfragen relevanten Variablen analysiert und damit nicht zuletzt die theoretisch abgeleiteten Forschungshypothesen überprüft werden.

6.2.2 Multivariate Analysen

6.2.2.1 Generelles Vorgehen bei der Berechnung der multiplen Regressionsmodelle

Bevor die Ergebnisse der multivariaten Analysen dargestellt werden, wird im Folgenden ein genereller Überblick über das methodische Vorgehen bei der Berechnung der Regressionsmodelle gegeben und die statistischen Hintergründe bzw. Voraussetzungen für die Anwendung des Verfahrens der multiplen Regression erläutert. Die darauf nachfolgende Darstellung der einzelnen Regressionsmodelle setzt den Fokus dann vor allem auf die Analyse der Ergebnisse.

Modellspezifikation und Berechnung der Regressionsmodelle

Zum Testen der Forschungshypothesen wird je betrachteter Forschungstätigkeit ein eigenes Regressionsmodell gerechnet. Wie bereits angemerkt, erlaubt das Verfahren der multiplen Regression durch die Berücksichtigung mehrerer unabhängiger Variablen die Kontrolle von Drittvariableneffekten. Zudem lässt sich durch die Aufnahme eines Produktterms auch ein wie in Hypothese *H2.2* (vgl. Kapitel 4.2) formulierter Moderatoreffekt modellieren und auf Signifikanz testen. Auch über einen Mediator vermittelte Zusammenhänge zwischen zwei Variablen können mit Hilfe von (mehreren) Regressionsmodellen auf Signifikanz getestet werden (vgl. zum Verfahren der multiplen Regression sowie zur Modellierung von Moderator- und Mediatoreffekten mit dieser Fahrmeir et al. 2003; Urban und Mayerl 2011). Ein solcher (partieller) Mediatoreffekt wird in dem hier entwickelten Modell zwischen dem Einsatz eines Forschungsindikators als Steuerungsinstrument (unabhängige Variable), der intrinsischen Motivation (Mediator) und dem Stellenwert einer Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren (abhängige Variable) vermutet (vgl. Kapitel 4.2 Hypothesen *H4.1* und *H4.2*).

Die Berechnung der Regressionsmodelle erfolgte mit der Software *IBM SPSS Statistics* in der Version 22. Zudem wurde für die Berechnungen das von Hayes (2013) programmierte SPSS-Makro *Process* genutzt.⁶³ *Process* wurde in dieser Arbeit vor allem aus zwei Gründen herangezogen: Erstens vereinfacht es das Testen von Mediatoreffekten, indem alle hierfür notwendigen Regressionen nach einmaliger Aufforderung automatisiert berechnet werden und die Signifikanz des indirekten Effektes der unabhängigen Variablen X über den Mediator Z auf die abhängige Variable Y getestet werden kann (vgl. Urban und Mayerl 2011, S. 303–305). Dies kann in *Process* entweder mit Hilfe des Sobel-Tests erfolgen oder durch Bootstrap-basierte Tests. Nach Hayes (2013, S. 104–105) und Baltés-Götz (2015, S. 17–19) sind Bootstrap-basierte Tests dem Sobel-Test vorzuziehen, da die dem Sobel-Test zugrundeliegenden Annahmen über die Stichprobenverteilung nicht gerechtfertigt seien und sich Signifikanztests vom Sobel-Typ auch in Simulationsstudien nicht bewährt hätten. Bootstrap-basierte Verfahren erlauben dagegen die Berechnung von Konfidenzintervallen unabhängig von der Verteilung der Stichprobe (Baltés-Götz 2015, S. 17). Die dahinterstehende Logik beschreibt Baltés-Götz (2015) wie folgt:

Man behandelt die Stichprobe als *Population*, ermittelt durch Ziehen mit Zurücklegen zahlreiche Sekundärstichproben (mindestens 1000) mit derselben Größe wie Original- bzw. Primärstichprobe, wobei in der Regel etliche Fälle *mehrfach* in einer Sekundärstichprobe vertreten sind. Aus jeder Sekundärstichprobe wird mit den üblichen Methoden ein Schätzer für den interessierenden Parameter gewonnen, so dass man eine empirische Stichprobenkennwerteverteilung erhält. Diese ersetzt die theoretische Stichprobenkennwerteverteilung, die z. B. auf der Normalverteilungsannahme basiert. Aus der empirischen Stichprobenkennwerteverteilung lassen sich Vertrauensintervalle und Testentscheidungen konstruieren, die nicht von der Normalverteilungsannahme abhängen. (ebd., S. 17)

In dieser Arbeit wurde den vorangegangenen Ausführungen folgend der Bootstrap-basierte Test zur Berechnung des Konfidenzintervalls des indirekten Effektes in *Process* gewählt.

Den zweiten Grund zur Nutzung des *Process*-Makros stellt die Möglichkeit dar, robuste Standardfehler⁶⁴ zur Berechnung von Signifikanztests und Vertrauensintervallen zu verwenden. Bei einer hohen Heterogenität der Residualvarianzen (siehe hierzu auch

⁶³ Das Makro steht unter <http://processmacro.org/index.html> (22.09.2016) kostenlos zum Download zur Verfügung. Hier finden sich außerdem Dokumentationen und weitere hilfreiche Materialien zu *Process*.

⁶⁴ In *Process* wird der sogenannte HC3-Schätzer verwendet, der hierfür als am geeignetsten angesehen werden kann (Long und Ervin 2000, S. 222).

die noch nachfolgenden Erläuterungen zu den Modellvoraussetzungen der multiplen Regression) kann durch die Verwendung von robusten Standardfehlern einer daraus resultierenden fehlerhaften Berechnung von Signifikanztests entgegengewirkt werden (vgl. Hayes und Cai 2007).

Hinsichtlich der Spezifikation von Regressionsmodellen ist generell das Problem des sogenannten *underfitting* bzw. dem gegenteiligen *overfitting* zu berücksichtigen. Während ersteres sich auf den Fall bezieht, dass wichtige erklärende Variablen nicht in das Regressionsmodell integriert werden, bezieht sich das Problem des *overfitting* darauf, dass zu viele, nicht relevante erklärende Variablen in das Modell aufgenommen werden. *Underfitting* eines Modells führt zu verzerrten Schätzwerten, während *overfitting* zu ineffizienten Schätzern mit nicht mehr minimaler Varianz führt (Backhaus et al. 2006, S. 85). Wie Urban und Mayerl (2011) anmerken, sollte die Spezifikation eines Regressionsmodells primär auf theoretisch oder analytisch begründbaren Zusammenhängen zwischen den Variablen beruhen. Mögliche Spezifikationsfehler sind mit rein statistischen Methoden nur schwer nachweisbar (ebd., S. 217–222), so dass die Modelle vor allem hinsichtlich der theoretisch vermuteten Variablenzusammenhänge vollständig spezifiziert sein sollten.

Entsprechend dieser Vorgabe erfolgte die Auswahl der in die Regressionsmodelle aufzunehmenden Variablen vorrangig aufgrund der in Kapitel 4 dargestellten theoretischen Vorüberlegungen und den daraus abgeleiteten Forschungshypothesen. Für die einzelnen tätigkeitsspezifischen Regressionsmodelle werden damit die folgenden theoretisch begründeten erklärenden Variablen herangezogen (hier am Beispiel der abhängigen Variablen *Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren*):

- *Abhängige Variable*: Eingeschätzter Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren in der Forschungsarbeit
- *Unabhängige Variablen*:
 - *UV_1*: Einsatz Forschungsindikator *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* in hochschulinternen Steuerungs- und Anreizverfahren generell (dichotom)
 - *KV_REP_1*: Bedeutung von *Publikationen in Zeitschriften mit Peer-Review-Verfahren* für den Reputationszuwachs in der Scientific Community (Rating-Skala)
 - Interaktionseffekt (Produktterm) *UV_1 X KV_REP_1*
 - *KV_ERK_1*: Beitrag *Publikationen in Zeitschriften mit Peer-Review-Verfahren* für den wissenschaftlichen Erkenntnisgewinn (Rating-Skala)
 - *KV_IM*: Intrinsische Forschungsmotivation (Likert-Skala basierter Index)

Die am Interaktionseffekt beteiligte Variable *KV_REP* wird dabei vor Aufnahme in das Modell zentriert. Zwar kann der (mathematische) Nutzen einer Zentrierung vor der Modellierung eines Interaktionseffektes zur Vermeidung von hoher Multikollinearität (vgl. zur Multikollinearität auch die nachfolgenden Erläuterungen zu den Modellvoraussetzungen) angezweifelt werden (vgl. Baltès-Götz 2014, 2015; Hayes 2013; Urban und Mayerl 2011), die Zentrierung kann aber die Interpretation der Regressionsschätzungen erleichtern. Wenn für den in das Modell aufgenommenen Interaktionseffekt keine Signifikanz nachgewiesen werden kann, wird dieser in den folgenden Berechnungen aus den Regressionsmodellen entfernt (zur statistischen Begründung siehe Baltès-Götz 2015, S. 54–55) und entsprechend auf eine Zentrierung der Variablen *KV_REP* verzichtet.

Neben den theoretisch hergeleiteten Regressoren werden noch weitere Variablen, die zwar nicht explizit im theoretischen Modell enthalten sind, bei denen aber Zusammenhänge mit der abhängigen Variablen oder den anderen unabhängigen Variablen vermutet werden können, in die Regressionsmodelle aufgenommen. So können mögliche Drittvariableneffekte anderer relevanter Einflussgrößen kontrolliert werden. Dies sind im vorliegenden Fall neben den soziodemografischen Merkmalen *Alter* (metrisch) und *Geschlecht* (dichotom), die *Beschäftigungsdauer* (metrisch) sowie die *Fachzugehörigkeit* (Dummy-Kodierung mit der Referenzkategorie Wirtschaftswissenschaften).

Hinsichtlich der Fachzugehörigkeit haben sich schon in den vorangegangenen Analysen teils deutliche Unterschiede bei den interessierenden Variablen gezeigt (vgl. Kapitel 6.1 und 6.2.1), so dass eine Kontrolle auf das Fach der befragten Professorinnen und Professoren zwingend notwendig erscheint.

Die Beschäftigungsdauer könnte – unabhängig vom Alter der Professorinnen und Professoren – im Sinne eines wissenschaftlichen (sekundären) Sozialisationsprozesses im Zusammenhang mit dem Forschungsverhalten stehen. So könnten Professorinnen und Professoren, die schon länger im Wissenschaftsbetrieb tätig sind, über eine hohe gewachsene Reputation bzw. einen höheren Grad an Autonomie verfügen, und sich so weniger einem äußeren Zwang ausgesetzt sehen, reihenweise Artikel in peer-reviewed journals zu veröffentlichen (vgl. hierzu auch Frey 2003). Stattdessen widmen sie ihre Zeit z. B. lieber dem Verfassen einer Monografie. Im Gegensatz dazu könnte das Alter als wissenschaftliche Erstsozialisation aufgefasst werden: Ältere Professorinnen und Professoren kamen z. B. in ihrer Studien- oder auch frühen Beschäftigungszeit als Wissenschaftlerinnen und Wissenschaftler mit einer anderen Forschungskultur in Berührung, die ihr eigenes späteres Forschungsverhalten (mit-)prägte. Kieser (2010) merkt beispielsweise in einem Aufsatz an:

Im bereits erwähnten populären *Handelsblatt*-Ranking für Betriebswirtschaftsprofessoren sind die 200 – nach dem Maßstab des *Handelsblatt* – erfolgreichsten BWL-

Professoren aus deutschsprachigen Ländern gelistet. Unter diesen finden sich keine Professoren, die ihre Forschungsergebnisse vorwiegend in Monografien veröffentlichen, mögen diese auch noch so einflussreich sein. Es werden nämlich im [sic] auch im *Handelsblatt*-Ranking, wie in Rankings üblich, nur Zeitschriften-Veröffentlichungen berücksichtigt. Die Folge ist, dass Wissenschaftler weniger Monografien schreiben, was ich persönlich sehr bedauerlich finde, denn als Student habe ich vor allem aus Monografien gelernt. Und wenn man mich jetzt fragt, welche Veröffentlichungen ich für die wichtigsten halte, führe ich fast nur Monografien auf. (ebd., S. 355)

Die Besoldungsart der befragten Professorinnen und Professoren (Besoldungsgruppe W oder C) wurde nicht als weitere Kontrollvariable aufgenommen. Obwohl hier Zusammenhänge mit dem Forschungsverhalten denkbar wären, zeigte sich bei der testweisen Aufnahme dieser Variablen in die Regressionsmodelle keine Verbesserung der Modellgüte. Das korrigierte Bestimmtheitsmaß R^2 als ein Maß für die Varianzaufklärung des Modells (Backhaus et al. 2006, S. 68) verringerte sich sogar durch die Aufnahme der Besoldungsart als Kontrollvariable. Möglicherweise ist dies auch auf die stark unterschiedlich besetzten Gruppen der W- (78.5 %) und C-Besoldeten (21.5 %) in der Stichprobe zurückzuführen. Wie sich bereits in Kapitel 6.1.1 zeigte, wäre eine vergleichende Analyse zwischen beiden Gruppen aufgrund der Stichprobengrößen nur für die Fächer Wirtschaftswissenschaften und mit Einschränkungen für die Erziehungswissenschaften sinnvoll durchführbar gewesen. Bei den folgenden Regressionsmodellen wird daher auf die Aufnahme der Besoldungsart als Kontrollvariable verzichtet.

Neben der möglichst korrekten Spezifikation der Regressionsmodelle ist die Erfüllung der statistischen Modellvoraussetzungen eine notwendige Bedingung, um Verzerrungen bei der Durchführung des Schätzverfahrens zu minimieren. Im Folgenden wird daher noch das allgemeine Vorgehen zur Überprüfung der Modellvoraussetzungen und bei eventuellen Verletzungen von diesen erläutert, bevor im nächsten Teilkapitel die Ergebnisdarstellung der Regressions-schätzungen erfolgt.

Überprüfung der Modellvoraussetzungen und Ausreißerdiagnostik

Zur Überprüfung der Modellvoraussetzungen wurde sich in dieser Arbeit vor allem an den Lehrbüchern bzw. Empfehlungen von Backhaus et al. (2006), Baltés-Götz (2014), Fahrmeir et al. (2003) und Urban und Mayerl (2011) orientiert. In diesen sind für die weiterführende Lektüre auch vertiefende Informationen zu den statistischen Hintergründen der einzelnen Modellvoraussetzungen und zu den Effekten bei Verletzungen von diesen zu finden. In der folgenden Darstellung wird daher auf eine umfassende Beschreibung verzichtet und stattdessen knapp das konkrete Vorgehen in der vorliegenden Stu-

die erläutert. Inwieweit die Modellvoraussetzungen bei den verschiedenen Regressionsmodellen erfüllt oder verletzt sind, wird dann – sofern an dieser Stelle keine generellen Aussagen hierzu getroffen werden – vor allem in den entsprechenden Unterkapiteln thematisiert.

Eine zentrale Voraussetzung, die bei der Anwendung einer multiplen Regression erfüllt sein sollte, betrifft die *Linearität des Zusammenhangs* zwischen der abhängigen und den unabhängigen Variablen. Das heißt, diese Variablen sollten in einer linearen, zumindest aber monotonen Beziehung zueinander stehen. Eine Nichtbeachtung der Linearitätsannahme kann zu einer Verzerrung der Schätzwerte führen. Ob ein linearer bzw. monotoner Zusammenhang zwischen den Variablen besteht, wurde in dieser Arbeit mittels partieller Regressionsdiagramme (Streudiagramme) getestet. Lässt sich in diesen kein monotoner linearer, sondern ein anderer funktionaler Zusammenhang erkennen (z. B. u-förmig), können – theoretisch zu begründende (Baltes-Götz 2014, S. 92) – Variablentransformationen Abhilfe schaffen. Bei allen hier gerechneten Modellen ergab die grafische Analyse mittels der partiellen Regressionsdiagramme zwar nicht immer eindeutig lineare Zusammenhänge, alternative funktionale Zusammenhänge waren aber auch nicht zu beobachten. Variablentransformationen wurde daher nicht durchgeführt. In den vorangegangenen bivariaten Analysen konnten zudem die Ergebnisse des linearen Zusammenhangsmaßes (Bravais-Pearsons r) durch das parameterfreie Maß *Spearman's Rho* überwiegend bestätigt werden (vgl. Kapitel 6.2.1), so dass die Annahme gerechtfertigt erscheint, dass sich die Zusammenhänge zwischen den Variablen für die hier durchgeführten Berechnungen als ausreichend linear darstellen. Die Anmerkungen hinsichtlich der Linearitätsannahme gelten für alle hier berechneten Modelle. In den im nächsten Teilkapitel nachfolgenden Ergebnisdarstellungen wird daher auf weitere Anmerkungen hierzu verzichtet.

Eine zweite zu testende Annahme bezieht sich auf die *Normalverteilung der Residuen*. Eine Verletzung der Normalverteilungsannahme führt zu ungültigen Signifikanztests (F- und t-Tests) des Regressionsmodells. Überprüfen lässt sich die Normalverteilungsannahme mittels grafischer (Histogramm)⁶⁵ oder inferenzstatistischer (Kolmogorov-Smirnov-Test oder Shapiro-Wilk) Tests. Hinsichtlich der Anwendung statistischer Tests zur Überprüfung der Normalverteilungsannahme merkt Baltes-Götz (2014) allerdings an, dass diese insbesondere bei großen Stichproben fast immer die Nullhypothese der Normalverteilung ablehnen (ebd., S. 64). In dieser Arbeit wurde die Annahme daher

⁶⁵ Baltes-Götz (2014) weist darauf hin, dass die in SPSS zum grafischen Test der Normalverteilungsannahme standardmäßig herangezogenen standardisierten Residuen streng genommen weniger geeignet sind, und stattdessen die ausgeschlossenen studentisierten Residuen zu analysieren seien. Entsprechend wurden in dieser Arbeit die ausgeschlossenen studentisierten Residuen zum Test der Normalverteilungsannahme verwendet, auch wenn in der Praxis keine großen Differenzen zu erwarten sind (ebd., S. 61).

lediglich grafisch mit Hilfe von Histogrammen überprüft (vgl. Anhang III). Zudem ist darauf hinzuweisen, dass bei ausreichend hohen Fallzahlen – wie hier je nach Modell zwischen $n = 411$ und $n = 478$ – aufgrund des zentralen Grenzwertsatzes keine Probleme hinsichtlich der Verteilungsannahmen zu erwarten sind (ebd.; Fahrmeir et al. 2003, S. 313–316).

Probleme können sich dagegen in Bezug auf die geforderte *Varianzhomogenität der Residuen* als weitere Modellvoraussetzung ergeben. Weichen die Varianzen der Residuen systematisch voneinander ab, z. B. in Form einer zunehmenden Streuung bei steigenden Werten, sind verzerrte Varianzen der geschätzten Regressionskoeffizienten und daraus resultierende fehlerhafte Berechnungen von Konfidenzintervallen bzw. Signifikanztests mögliche Folgen (Urban und Mayerl 2011, S. 242–243). Getestet wurde die Homoskedastizitätsannahme in dieser Arbeit mit Hilfe von Residuen-Plots (Plot der ausgeschlossenen-studentisierten Residuen gegen die standardisierten prognostizierten Werte) sowie dem Breusch-Pagan-Test⁶⁶. Zeigten sich in den Residuenplots systematische Auffälligkeiten und musste durch den Breusch-Pagan-Test die Nullhypothese der Varianzhomogenität verworfen werden, wurde auf die Verwendung robuster Standardfehler zur Berechnung der Vertrauensintervalle bzw. Signifikanztests des Regressionsmodelles zurückgegriffen. Wie bereits weiter oben beschrieben, konnte hierfür das SPSS-Makro *Process* genutzt werden, welches eine Option zur Verwendung heteroskedastizitätskonsistenter Standardfehler bietet (vgl. zum Problem der Heteroskedastizität und der Verwendung robuster Standardfehler Hayes und Cai 2007).

Die Verletzung einer weiteren Modellvoraussetzung, nämlich der *Unkorreliertheit der Residuen*, stellt in der Praxis meist nur bei im Längsschnitt erhobenen Daten ein Problem dar (Urban und Mayerl 2011, S. 260). Der in SPSS implementierte Durbin-Watson-Test zeigte entsprechend für keines der hier berechneten Modelle Probleme mit korrelierten Residuen an, so dass im weiteren Verlauf dieser Arbeit auf zusätzliche Anmerkungen hinsichtlich dieser Modellvoraussetzung verzichtet wird.

Multikollinearität, das heißt eine starke lineare Abhängigkeit zwischen den einzelnen Regressoren, kann bei multiplen Regressionsmodellen ebenfalls zu Problemen führen. Zu hohe Korrelationen zwischen den unabhängigen Variablen können die Standardfehler der Regressionskoeffizienten stark erhöhen. Die Schätzungen der Parameter können dadurch unpräzise werden (Backhaus et al. 2006, S. 89–90). Das Vorhandensein von Multikollinearität wurde mittels der direkt in SPSS abrufbaren Kollinearitätsdiagnose überprüft. Die hierbei ausgegebenen *VIF*- (*Variance Inflation Factor*) bzw. Toleranzwerte erlauben eine Beurteilung, ob von Multikollinearität ausgegangen werden muss. Baltes-

⁶⁶ Der Breusch-Pagan-Test ist in der SPSS Basis-Version nicht implementiert. Durch Installation der R-Essentials für SPSS lässt sich diese Funktionalität aber hinzufügen. R-Essentials sind verfügbar unter <https://www.ibm.com/developerworks/> (05.10.2016).

Götz (2014) nennt Toleranzwerte kleiner 0.1 als problematisch (ebd., S. 108), während Urban und Mayerl (2011) einen strengeren Grenzwert von 0.20 bis 0.25 vorschlagen (ebd., S. 232). An dieser Stelle kann bereits festgehalten werden, dass sich bei keinem der berechneten Modelle Probleme in Bezug auf Multikollinearität ergaben. Die Toleranzwerte lagen jeweils über dem kritischen Wert von 0.1 und auch der strengere Grenzwert von 0.25 wurde nicht unterschritten.

Neben Verletzungen der Modellvoraussetzungen können Ausreißer bzw. besonders einflussreiche Fälle für Verzerrungen bei der Schätzung der Regressionsparameter verantwortlich sein. Ausreißer können sich dabei hinsichtlich der Residuen oder hinsichtlich der Werte der Regressoren zeigen. Verzerrungen bei ersteren sind dann zu erwarten, wenn ungewöhnlich große Residuen beobachtet werden können. Urban und Mayerl (2011) nennen hier einen Wert von mehr als zwei Standardabweichungen bei den ausgeschlossenen standardisierten Residuen (ebd., S. 86–87), während Gordon (2010) erst einen Wert von mehr als drei Standardabweichungen als problematisch ansieht (ebd., S. 366). Die Ausreißeranalyse bezüglich der Residuen erfolgte in dieser Arbeit anhand der im Regressionsdialog in SPSS anzufordernden Residuenstatistik. Wie auch schon beim Test auf Normalverteilung wurden für die Analyse allerdings statt der standardisierten die studentisierten Residuen verwendet (vgl. Baltés-Götz 2014, S. 101–102).

Zur Ausreißeranalyse der Regressoren wurden die zentrierten Hebelwerte sowie die Kennzahl *Cooks D* herangezogen, welche beide über die Ausreißerstatistik im Regressionsdialog angefordert werden können. Zunächst wurden die Hebelwerte je Fall näher betrachtet. Lagen diese über einem vorher festgelegten kritischen Wert⁶⁷, wurde für diese Fälle weiter der Wert von *Cooks D* überprüft. Für *Cooks D* wurde wiederum ein kritischer Wert festgelegt, ab welchem dem betreffenden Fall ein besonders starker Einfluss attestiert werden kann. Den Empfehlungen einschlägiger Statistik-Literatur folgend (Gordon 2010, S. 367; Hamilton 2006, S. 210) wurde hier ein Grenzwert von $4/n$ gewählt. Fälle mit einem Wert größer dem Grenzwert $4/n$ wurden zudem grafisch mittels partieller Regressionsdiagramme hinsichtlich ihrer Einflussstärke überprüft. Die so ermittelten Ausreißer wurden im Anschluss im Detail analysiert, indem der gesamte Datensatz der betroffenen befragten Person nochmals gesichtet und überprüft wurde. Sofern die Ausreißerwerte nicht auf fehlerhaften Daten beruhen (z. B. offensichtliche Falschangaben) oder andere plausible Gründe für einen Ausschluss der Fälle sprechen (z. B. wenn diese nicht zur Zielpopulation der Stichprobe gezählt werden können), ist die Nichtberücksichtigung von einzelnen Datensätzen allein aufgrund des ermittelten hohen Einflusses als problematisch anzusehen. Es besteht die Gefahr, dass diejenigen Fälle, die nicht zum spezifizierten Modell ‚passen‘, ausgeschlossen werden, während andere einflussreiche

⁶⁷ Dieser lag bei $2k/n$, wobei k die Anzahl der Regressoren und n die Stichprobengröße darstellt (Urban und Mayerl 2011, S. 188).

Fälle, die dem Modell gut entsprechen, beibehalten werden. Die Modellgüte kann auf diese Weise zwar deutlich verbessert werden, allerdings erscheint in einem solchen Fall der Vorwurf der Datenmanipulation durchaus gerechtfertigt. In dieser Arbeit wurden ermittelte Ausreißer daher nur dann aus der Analyse ausgeschlossen, wenn der Ausschluss anhand der Datenstruktur oder stichprobentheoretisch begründet werden konnte.

Nachdem nun mit der Beschreibung des grundlegenden Vorgehens bei der Spezifikation der Regressionsmodelle und der Überprüfung der Modellprämissen der methodische Rahmen für die Berechnung der Regressionsmodelle dargestellt wurde, folgt nun im nächsten Kapitel die Ergebnisdarstellung der durchgeführten Analysen.

6.2.2.2 Ergebnisse der Regressionsanalysen

Die Ergebnisdarstellung der Regressionsanalysen erfolgt wie bereits angesprochen je untersuchter Forschungstätigkeit (z. B. Publizieren in Zeitschriften mit Peer-Review-Verfahren, Publizieren in Zeitschriften ohne Peer-Review-Verfahren etc.). Dabei werden zunächst ausführlich die Ergebnisse der Regressionsschätzungen (Regressionskoeffizienten⁶⁸, Modellgüte) beschrieben. Die inhaltliche Diskussion der Ergebnisse und deren Abgleich mit den in Kapitel 4.2 aufgestellten Forschungshypothesen erfolgt dann im Anschluss an die Ergebnisdarstellung aller gerechneten Modelle in dem eigenständigen Kapitel 6.3.

Modell 1: Publizieren in Zeitschriften mit Peer-Review-Verfahren

Das Regressionsmodell mit der abhängigen Variablen *Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren* in der Forschungsarbeit der Professorinnen und Professoren wurde wie im vorangegangenen Teilkapitel beschrieben spezifiziert – die aufgenommenen Variablen sind in Tabelle 28 aufgeführt. Nach der Berechnung wurde das Modell im Hinblick auf die oben genannten Modellvoraussetzungen hin überprüft. Hierbei ergaben sich hinsichtlich der Linearitäts- und Normalverteilungsannahme sowie in Bezug auf Multikollinearität keine größeren Auffälligkeiten. Allerdings ergab der Breusch-Pagan-Test zur Überprüfung der Homoskedastizitätsannahme eine höchst signifikante Abweichung von der Nullhypothese der Varianzhomogenität ($\chi^2 = 39.57$,

⁶⁸ In den folgenden Ergebnisdarstellungen werden nur die unstandardisierten Regressionskoeffizienten berichtet und interpretiert. Zum einen ist die Berechnung von standardisierten Koeffizienten bei Dummy-Variablen nicht sinnvoll, zum anderen muss die Eignung von standardisierten Regressionskoeffizienten zum Vergleich von Effektstärken generell kritisch gesehen werden (vgl. Urban und Mayerl 2011, S. 103-108). Da das Ziel dieser Arbeit primär der Test der aufgestellten Forschungshypothesen ist, und der detaillierte Vergleich unterschiedlicher Effektstärken eine Exaktheit der berechneten Werte implizieren würde, die die vorliegenden Daten nicht für sich beanspruchen können, wird auf die Darstellung und Interpretation der standardisierten Regressionskoeffizienten verzichtet.

$p < .001$), so dass zur Berechnung der Vertrauensintervalle bzw. Signifikanztests des Regressionsmodelles auf robuste Standardfehler zurückgegriffen wurde.

Die Ausreißerdiagnostik hinsichtlich der Residuen ergab insgesamt 27 Fälle, bei denen die fallweise Diagnose eine Überschreitung des Schwellenwertes von zwei Standardabweichungen anzeigte. Da dieser Schwellenwert bei der Vielzahl an Ausreißerfällen als zu niedrig angesetzt angesehen werden kann, wurde der von Gordon (2010) vorgeschlagene großzügigere Schwellenwert von drei Standardabweichungen (ebd., S. 366) gewählt. Dieser wurde lediglich von zwei Fällen überschritten, die daraufhin im Detail analysiert wurden. Da keine Anzeichen für Falschangaben oder fehlender Zugehörigkeit zur Stichprobe auszumachen waren, wurden die beiden Fälle für die weiteren Analysen beibehalten. In Bezug auf die Regressoren ergab die Ausreißerdiagnostik mit Hilfe der Hebelwerte und *Cooks D* insgesamt 28 Fälle, welche die berechneten Schwellenwerte überschritten. Nach nochmaliger Analyse der betroffenen Datensätze wurden von diesen 28 Fällen vier aus den weiteren Berechnungen ausgeschlossen. Diese stellen gegenüber der angestrebten Stichprobe Spezialfälle dar (außeruniversitäre Forschung, Professur im Ausland (zwei Mal), keine eindeutige Fachzuordnung), bei denen ein Ausschluss aufgrund der inhaltlichen Unterschiede gerechtfertigt erschien. Diese vier Datensätze wurden daher auch bei allen weiteren Regressionsmodellen nicht berücksichtigt. Nach Ausschluss der Fälle wurde das Regressionsmodell erneut berechnet und wieder auf Verletzungen der Modellvoraussetzungen überprüft. Die zuvor festgestellte Heteroskedastizität der Residuen bestätigte sich dabei (Breusch-Pagan: $\chi^2 = 39.64$, $p < .001$), so dass weiterhin robuste Standardfehler zur Berechnung herangezogen wurden. Die Ergebnisse des so berechneten Modells sind in Tabelle 28 darstellt.

Für das Gesamtmodell ergibt sich zunächst mit einem Wert von $F(12, 454) = 20.41$ eine höchst signifikante ($p < .001$) Abweichung von der Nullhypothese, dass alle Regressionskoeffizienten in der Grundgesamtheit Null sind. Betrachtet man die p -Werte zu den einzelnen Regressoren, ergeben sich insgesamt sieben Variablen, die in einem signifikanten Zusammenhang mit der abhängigen Variablen stehen. Demnach geht der Einsatz des Forschungsindikators *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* in Steuerungs- oder Anreizverfahren unter Konstanthaltung aller anderen Regressoren mit einem erhöhten Stellenwert der korrespondierenden Forschungstätigkeit, dem Publizieren in peer-reviewed Zeitschriften, einher ($B = .24$, $p = .011$). Ebenfalls in positivem Zusammenhang mit der abhängigen Variablen stehen die eingeschätzte Bedeutung von Publikationen in Zeitschriften mit Peer-Review für den Reputationszuwachs in der Scientific Community ($B = .29$, $p = .001$), der eingeschätzte Beitrag von diesen für den wissenschaftlichen Erkenntnisgewinn ($B = .39$, $p < .001$) sowie die Höhe der intrinsischen Forschungsmotivation ($B = .25$, $p = .007$). In allen drei Fällen führt eine Erhöhung des Wertes der unabhängigen Variablen unter Kontrolle der übrigen Variablen zu

einer Erhöhung des Wertes der abhängigen Variablen. Auch wenn sich die Effektstärke aufgrund der teils unterschiedlichen Skalierung der Regressoren nicht exakt vergleichen lässt, kann festgehalten werden, dass diese mit Werten zwischen $B = .24$ und $B = .39$ eher mäßig bis mäßig stark ausfällt.

Tabelle 28: *Multiple Regression mit der abhängigen Variablen „Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren“*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
<i>Konstante</i>	3.076	.560	5.49	.000
Einsatz Indikator „Anzahl Publikationen in Zeitschriften mit Peer-Review“ (0 kein Einsatz / 1 Einsatz)	.242	.095	2.54	.011
Bedeutung „Publikationen in Zeitschriften mit Peer-Review“ für Reputationszuwachs (zentriert)	.288	.087	3.32	.001
Interaktion Einsatz Indikator x Bedeutung Reputationszuwachs	-.240	.113	-2.12	.034
Beitrag „Publikationen in Zeitschriften mit Peer-Review“ für wissenschaftlichen Erkenntnisgewinn	.394	.058	6.78	.000
Index Intrinsische Motivation	.249	.091	2.72	.007
Alter	-.042	.010	-4.27	.000
Beschäftigungsdauer	.011	.012	0.90	.367
Geschlecht (0 männlich / 1 weiblich)	-.152	.104	-1.47	.142
Dummy Erziehungswissenschaft	-.240	.130	-1.84	.067
Dummy Politikwissenschaft	.156	.152	1.03	.306
Dummy Psychologie	.296	.112	2.64	.009
Dummy Soziologie	.104	.150	0.70	.486
<i>R</i>²		.345		
Korrigiertes <i>R</i>²		.327		
<i>F</i>(12, 454)		20.41 (<i>p</i> < .001)		
<i>n</i>		467		

Anmerkung. *B*: Unstandardisierter Regressionskoeffizient; *SE*: Standardfehler von *B*; *t*: t-Statistik von *B*; *R*²: Bestimmtheitsmaß; *F*: F-Statistik mit Freiheitsgraden; Referenzkategorie für Fachdummies: Wirtschaftswissenschaften

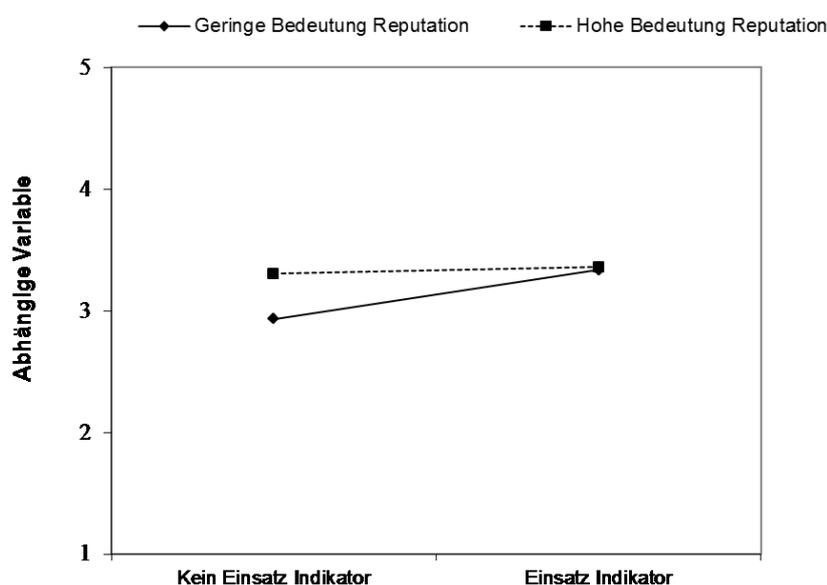
Quelle: eigene Berechnungen

Gegenüber der Referenzkategorie der Wirtschaftswissenschaften fällt der eingeschätzte Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren in der eigenen Forschungsarbeit im Fach Psychologie signifikant höher aus ($B = .30$, $p = .009$). Bei den übrigen Fächern ergeben sich keine signifikanten Differenzen zur Referenzkategorie.

Ebenso scheinen die Beschäftigungsdauer ($B = .01, p = .367$) und das Geschlecht ($B = -.15, p = .142$) keinen Einfluss zu haben, während der eingeschätzte Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren mit zunehmendem Alter (leicht) abnimmt ($B = -.04, p < .001$).

Für die vermutete Interaktion zwischen dem Einsatz des Forschungsindikators und der eingeschätzten Bedeutung von peer-reviewed Publikationen für den Reputationszuwachs ergibt sich ebenfalls ein signifikanter negativer Effekt ($B = -.24, p = .034$). In Abbildung 10 ist die Art des Zusammenhangs zwischen den drei beteiligten Variablen zum besseren Verständnis grafisch dargestellt.

Abbildung 10: *Interaktionseffekt Bedeutung Reputationszuwachs x Einsatz Indikator (Publikationen in Zeitschriften mit Peer-Review)*



Anmerkung. Die Werte für die zwei Ausprägungen „Geringe Bedeutung Reputation“ und „Hohe Bedeutung Reputation“ ergeben sich aus max. $M \pm 1 SD$ der Ursprungsvariablen.

Quelle: Excel template zur Erstellung der Grafik von <http://www.jeremydawson.com/slopes.htm> (06.10.2016); eigene Berechnungen

Der positive Steigungseffekt der unabhängigen Variablen (Einsatz Forschungsindikator) auf die abhängige Variable verschwindet bei zunehmender eingeschätzter Bedeutung von Publikationen in Zeitschriften mit Peer-Review-Verfahren für den Reputationszuwachs in der Scientific Community. Der Zusammenhang zwischen der unabhängigen Variablen (Einsatz Indikator) und der abhängigen Variablen (Stellenwert in der Forschungsarbeit) wird also durch die Drittvariable (Bedeutung Reputationszuwachs) moderiert. Dies wird auch deutlich, wenn man die in *Process* berechneten bedingten Effekte der unabhängigen auf die abhängige Variable unter gegebenen Werten des (zentrierten) Moderators betrachtet. Diese sind in Tabelle 29 aufgeführt.

Bei einem niedrigen Wert des Moderators⁶⁹ (-.74), also einer eher niedrigen eingeschätzten Bedeutung von peer-reviewed Publikationen für den Reputationszuwachs, zeigt sich noch ein höchst signifikanter positiver Effekt ($B = .42, p = .001$) des Einsatzes des Forschungsindikators *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* auf den eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit. Beim Mittelwert des Moderators fällt der positive Effekt bereits deutlich geringer aus ($B = .24, p = .012$), während er bei denjenigen Befragten, welche die Bedeutung dieser Publikationsform für den Reputationszuwachs in der Scientific Community am höchsten einschätzen (.56), gänzlich verschwindet ($B = .11, p = .33$).

Tabelle 29: *Bedingte Effekte des Einsatzes des Indikators „Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren“*

Ausprägung des Moderators (M)	B	SE	t	p
<i>Bedeutung Publikationen in Zeitschriften mit Peer-Review für Reputationszuwachs (zentriert)</i>				
-.744	.420	.131	3.21	.001
.000	.240	.095	2.53	.012
.560	.107	.111	0.97	.334

Anmerkung. B: Unstandardisierter Regressionskoeffizient der Variablen *Einsatz Indikator „Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren“* unter gegebenen Ausprägungen des Moderators M; SE: Standardfehler von B; t: t-Statistik von B; Abhängige Variable: *Eingeschätzter Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren in der eigenen Forschungsarbeit*

Quelle: eigene Berechnungen

Zur Beurteilung der Erklärungskraft des Gesamtmodells wird hier das *korrigierte R²* als Gütemaß herangezogen. Im Gegensatz zum einfachen Bestimmtheitsmaß *R²* ist dieses so konstruiert, dass die zusätzliche Aufnahme von Regressoren in das Modell nur dann zu einer Erhöhung des Wertes der Maßzahl führt, wenn es dadurch tatsächlich zu einer Verbesserung der Erklärungskraft des Modells kommt. Das einfache *R²* steigt dagegen generell mit der Aufnahme jedes weiteren Regressors in das Modell an (Backhaus et al. 2006, S. 68).

Der Wert des korrigierten Bestimmtheitsmaßes beträgt bei Modell 1 $R^2 = 0.33$, das heißt, dass etwa 33 % der Variation der Werte der abhängigen Variablen *Stellenwert des Publizierens in Zeitschriften mit Peer-Review-Verfahren in der eigenen Forschungsarbeit* durch das Regressionsmodell erklärt werden können. Dieser Wert ist im Hinblick auf die

⁶⁹ Die gegebenen Werte des Moderators zur Berechnung der bedingten Effekte ergeben sich aus dem Mittelwert der zentrierten Variablen ($M = 0$) +/- einer Standardabweichung ($SD = 0.74$). Aufgrund der stark linksschiefen Verteilung der Variablen (vgl. Kapitel 6.1.4) liegt das obere Maximum bereits bei einem Wert von 0.56.

in den Sozialwissenschaften generell häufig geringere Varianzaufklärung (vgl. Fromm 2003, S. 15) als durchaus zufriedenstellend zu bewerten.

Neben den tätigkeitsspezifischen Regressionsmodellen wurde zur Überprüfung der Verdrängungshypothese *H4.2* (vgl. Kapitel 4.2) bzw. des vermuteten Mediationseffektes jeweils zusätzlich im Makro *Process* die Berechnung eines Regressionsmodells mit der abhängigen Variablen *Intrinsische Forschungsmotivation* sowie der Bootstrap-basierte Test des indirekten Effektes (über den Mediator vermittelter Effekt der unabhängigen auf die abhängige Variable) angefordert. Hierbei zeigte sich kein signifikanter Effekt des Einsatzes des Forschungsindikators *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* auf die intrinsische Forschungsmotivation ($B = -.05$, $p = .306$). Weitere Analysen zu einem möglichen Mediationseffekt wurden daher nicht durchgeführt.

Modell 2: Publizieren in Zeitschriften ohne Peer-Review-Verfahren

Beim nun folgenden Modell 2 bildet der *Stellenwert des Publizierens in Zeitschriften ohne Peer-Review-Verfahren* die abhängige Variable. Zunächst wurde wie im vorangegangenen Modell ein Produktterm in das Modell aufgenommen, um den Hypothese *H2.2* zugrundeliegenden Interaktionseffekt zu testen. Für diesen konnte allerdings keine signifikante Abweichung von der Nullhypothese nachgewiesen werden ($B = -.12$, $p = .220$). Daher wurde der Produktterm wieder aus dem Modell entfernt und ein weiteres Modell ohne diesen gerechnet. Die folgenden Ausführungen zu den Modellvoraussetzungen sowie die in Tabelle 30 dargestellten Ergebnisse beziehen sich entsprechend auf das Modell ohne Interaktionseffekt.

Hinsichtlich der Modellvoraussetzungen zeigen sich ähnliche Ergebnisse wie bei Modell 1. Mit Ausnahme der Homoskedastizitätsannahme können die Voraussetzungen überwiegend als erfüllt angesehen werden. Aufgrund des höchst signifikanten Ergebnisses des Breusch-Pagan-Tests ($\chi^2 = 10.82$, $p = .001$) wurden daher wieder robuste Standardfehler zur Berechnung des Regressionsmodells genutzt. Auch bei Modell 2 ergab sich im Zuge der Ausreißerdiagnostik eine Reihe von besonders einflussreichen Fällen. Auffälligkeiten bezüglich der Residuen zeigten sich bei zwei Fällen, bei welchen die Werte der ausgeschlossenen studentisierten Residuen über dem Schwellenwert von drei Standardabweichungen lagen. Ausreißer hinsichtlich der Regressorenwerte ergaben sich nach Analyse der Hebelwerte und von *Cooks D* bei 23 Fällen. Insbesondere durch den Ausschluss der Ausreißerfälle bezüglich der Residuen hätte die Güte des vorliegenden Modells verbessert werden können. Da nach Durchsicht der Datensätze aber keine inhaltliche oder stichprobentheoretische Begründung für einen Ausschluss gefunden werden konnte, wurden die Fälle dennoch beibehalten.

Auch für Modell 2 ergibt sich mit $F(11, 407) = 40.04$ für das Gesamtmodell eine höchst signifikante Abweichung von der Nullhypothese, dass alle *B*-Werte gleich Null

sind. Signifikante Effekte lassen sich bei Betrachtung der einzelnen Regressoren allerdings nur für die eingeschätzte Bedeutung von Publikationen ohne Peer-Review für den Reputationszuwachs ($B = .41$, $p < .001$) sowie für deren eingeschätzten Beitrag für den wissenschaftlichen Erkenntnisgewinn ($B = .34$, $p < .001$) nachweisen. Zudem wird der Stellenwert der untersuchten Publikationsform in der eigenen Forschungsarbeit von Professorinnen und Professoren der Erziehungswissenschaften signifikant höher eingeschätzt als von denen der Wirtschaftswissenschaften ($B = .42$, $p = .005$).

Tabelle 30: *Multiple Regression mit der abhängigen Variablen „Stellenwert des Publizierens in Zeitschriften ohne Peer-Review-Verfahren“*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
<i>Konstante</i>	.632	.668	0.95	.345
Einsatz Indikator „Anzahl Publikationen in Zeitschriften ohne Peer-Review“ (0 kein Einsatz / 1 Einsatz)	.063	.107	0.58	.560
Bedeutung „Publikationen in Zeitschriften ohne Peer-Review“ für Reputationszuwachs	.414	.068	6.07	.000
Beitrag „Publikationen in Zeitschriften ohne Peer-Review“ für wissenschaftlichen Erkenntnisgewinn	.338	.057	5.97	.000
Index Intrinsische Motivation	-.102	.091	-1.11	.267
Alter	.004	.012	0.31	.755
Beschäftigungsdauer	.015	.013	1.23	.221
Geschlecht (0 männlich / 1 weiblich)	-.008	.108	-0.70	.944
Dummy Erziehungswissenschaft	.418	.148	2.83	.005
Dummy Politikwissenschaft	.192	.192	1.00	.317
Dummy Psychologie	-.034	.121	-0.28	.780
Dummy Soziologie	-.110	.165	-0.67	.506
<i>R</i>²		.464		
Korrigiertes <i>R</i>²		.450		
<i>F</i>(11, 407)		40.04 (<i>p</i> < .001)		
<i>n</i>		419		

Anmerkung. *B*: Unstandardisierter Regressionskoeffizient; *SE*: Standardfehler von *B*; *t*: t-Statistik von *B*; *R*²: Bestimmtheitsmaß; *F*: F-Statistik mit Freiheitsgraden; Referenzkategorie für Fachdummies: Wirtschaftswissenschaften

Quelle: eigene Berechnungen

Weder für den Einsatz des Forschungsindikators *Anzahl Publikationen ohne Peer-Review-Verfahren* ($B = .06$, $p = .560$) noch für die intrinsische Motivation ($B = -.10$, $p = .267$) lassen sich signifikante Effekte auf die abhängige Variable nachweisen. Die

p -Werte der übrigen Kontrollvariablen (Alter, Beschäftigungsdauer, Geschlecht) liegen ebenfalls über dem Signifikanzniveau von $\alpha = 5\%$. Dennoch weist das Gesamtmodell mit einem korrigierten Bestimmtheitsmaß von $R^2 = .45$ eine recht hohe Varianzaufklärung auf. Die eingeschätzte Bedeutung von Publikationen ohne Peer-Review-Verfahren für den Reputationszuwachs in der Scientific Community sowie der eingeschätzte Beitrag von diesen für den wissenschaftlichen Erkenntnisgewinn scheinen damit recht gute Prädiktoren für den Stellenwert zu sein, den das Verfassen solcher Publikationen in der Forschungsarbeit der Professorinnen und Professoren einnimmt.

Die Verdrängungshypothese *H4.2* bzw. der Mediationseffekt wurde auch für den Einsatz des Indikators *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* überprüft. Ein über die intrinsische Motivation vermittelter Effekt des Einsatzes des Forschungsindikators auf den eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit konnte dabei nicht nachgewiesen werden. Wie bereits aus Tabelle 30 ersichtlich wurde, zeigt sich weder für die unabhängige Variable ($B = .06$, $p = .560$) noch für die intrinsische Motivation ($B = -.10$, $p = .267$) ein signifikanter direkter Effekt auf die abhängige Variable. Auch das berechnete 95 %-Konfidenzintervall des Bootstrap-basierten Tests zeigt mit einem Wertebereich von $[-.012; .068]$ keine Signifikanz des indirekten ($B = .021$) Effektes an.

Allerdings ergibt sich bei der im Zuge der Überprüfung des Mediationseffektes gerechneten Regression der abhängigen Variablen *Intrinsische Motivation* auf die Prädiktoren⁷⁰ ein höchst signifikanter negativer Zusammenhang zwischen dem Einsatz des Forschungsindikators *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* und der intrinsischen Motivation ($B = -.21$, $p < .001$). Die intrinsische Forschungsmotivation fällt also – unter Konstanthaltung der anderen Variablen – bei der Gruppe der Befragten, bei denen der Forschungsindikator *Anzahl Publikationen in Zeitschriften ohne Peer-Review* nicht zum Einsatz kommt, signifikant höher aus als bei denjenigen, bei denen ein solcher Indikator eingesetzt wird.

Modell 3: Verfassen anderer Publikationen

Wie auch für Modell 2 konnte für Modell 3 mit der abhängigen Variablen *Stellenwert des Verfassens anderer Publikationen (Monografien, Beiträge in Sammelbänden etc.)* kein signifikanter Interaktionseffekt zwischen der eingeschätzten Bedeutung dieser Publikationsformen für den Reputationszuwachs und dem Einsatz des Forschungsindikators

⁷⁰ Die verwendeten Prädiktoren sind identisch mit den in Tabelle 30 dargestellten Variablen. Im Zuge der Überprüfung des Mediationseffektes wurde lediglich die abhängige Variable *Stellenwert des Publizierens in Zeitschriften ohne Peer-Review-Verfahren* durch die abhängige Variable *Intrinsische Motivation* ersetzt.

Anzahl anderer Publikationen in Steuerungs- und Anreizverfahren nachgewiesen werden ($B = .04$, $p = .624$). Entsprechend wurde auch hier der Produktterm nach der Signifikanzprüfung aus dem in Tabelle 31 dargestellten Modell entfernt.

Tabelle 31: *Multiple Regression mit der abhängigen Variablen „Stellenwert des Verfassens anderer Publikationen“*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
<i>Konstante</i>	.009	.589	0.02	.988
Einsatz Indikator „Anzahl anderer Publikationen“ (0 kein Einsatz / 1 Einsatz)	.156	.096	1.62	.101
Bedeutung „Andere Publikationen“ für Reputationszuwachs	.200	.060	3.30	.001
Beitrag „Andere Publikationen“ für wissenschaftlichen Erkenntnisgewinn	.427	.054	7.98	.000
Index Intrinsische Motivation	-.037	.088	-0.42	.675
Alter	.020	.011	1.92	.056
Beschäftigungsdauer	.005	.012	0.39	.695
Geschlecht (0 männlich / 1 weiblich)	.101	.105	0.97	.334
Dummy Erziehungswissenschaft	.575	.148	3.89	.000
Dummy Politikwissenschaft	.494	.188	2.64	.009
Dummy Psychologie	.027	.133	0.20	.841
Dummy Soziologie	.384	.169	2.27	.024
<i>R</i>²		.456		
Korrigiertes <i>R</i>²		.442		
<i>F</i>(11, 414)		31.55 (<i>p</i> < .001)		
<i>n</i>		426		

Anmerkung. *B*: Unstandardisierter Regressionskoeffizient; *SE*: Standardfehler von *B*; *t*: t-Statistik von *B*; *R*²: Bestimmtheitsmaß; *F*: F-Statistik mit Freiheitsgraden; Referenzkategorie für Fachdummies: Wirtschaftswissenschaften

Quelle: eigene Berechnungen

Die Überprüfung der Modellvoraussetzungen ergab keine Auffälligkeiten. Auch das in den vorherigen Modellen aufgetretene Heteroskedastizitätsproblem konnte bei Modell 3 nicht festgestellt werden. Auf die Verwendung robuster Standardfehler wurde daher verzichtet. Allerdings konnten im Rahmen der Ausreißerdiagnostik wieder eine Reihe von einflussreichen Fällen identifiziert werden. Dies betraf drei Fälle mit Residuenwerten von mehr als drei Standardabweichungen und 21 Fälle mit auffälligen Hebelwerten bzw. auffälligen Werten von *Cooks D*. Allerdings lagen auch hier keine inhaltlichen Gründe vor,

die einen Ausschluss der Datensätze gerechtfertigt hätten. Die aus diesen Fällen resultierende geringfügig schlechtere Modellgüte wurde daher in Kauf genommen und die Datensätze für die weiteren Analysen beibehalten.

Der globale Overall-F-Test zeigt mit einem Wert von $F(11, 414) = 31.55$ auch für Modell 3 an, dass mindestens ein signifikanter linearer Zusammenhang zwischen der abhängigen Variablen und den Prädiktoren besteht ($p < .001$). Analog zu Modell 2 ergibt sich hinsichtlich der einzelnen theoretisch begründeten Prädiktoren ein signifikanter Steigungseffekt auf die abhängige Variable nur bei der eingeschätzten Bedeutung von *anderen Publikationen* für den Reputationszuwachs in der Scientific Community ($B = .20$, $p = .001$) sowie dem eingeschätzten Beitrag von diesen für den wissenschaftlichen Erkenntnisgewinn ($B = .43$, $p < .001$). Daneben zeigen sich in Bezug auf die Kontrollvariable *Fachzugehörigkeit* gegenüber der Referenzkategorie „Wirtschaftswissenschaften“ signifikant höhere Werte der Steigungsparameter bei den Fächern Erziehungswissenschaft ($B = .58$, $p < .001$), Politikwissenschaft ($B = .49$, $p = .009$) und Soziologie ($B = .38$, $p = .024$). Dies deckt sich mit den bereits in Kapitel 6.1.6 dargestellten fächervergleichenden deskriptiven Analysen, die einen (deutlich) höheren eingeschätzten Stellenwert von anderen Publikationsformen abseits von Zeitschriftenartikel in den Fächern Erziehungswissenschaft, Politikwissenschaft und Soziologie als in der Psychologie und den Wirtschaftswissenschaften beschreiben.

Obwohl auch in Modell 3 nur wenige Prädiktoren signifikant zur Varianzaufklärung der abhängigen Variablen beitragen, fällt die Erklärungskraft des Gesamtmodells mit einem korrigierten Bestimmtheitsmaß von $R^2 = .44$ recht gut aus. Neben der Fachzugehörigkeit scheinen damit die eingeschätzte Bedeutung für den Reputationszuwachs und der eingeschätzte Beitrag für den wissenschaftlichen Erkenntnisgewinn, wie auch beim vorangegangenen Modell, geeignete Prädiktoren für die abhängige Variable darzustellen. Nichtsdestotrotz bleiben in diesem Modell 56 % der Varianz der abhängigen Variablen unerklärt und sind damit auf andere Einflussfaktoren zurückzuführen.

Hinsichtlich des Mediationseffektes bzw. der Verdrängungshypothese *H4.2* lässt sich Ähnliches resümieren wie für Modell 2: Weder zeigt sich ein signifikanter direkter Effekt der intrinsischen Forschungsmotivation ($B = -.04$, $p = .675$), noch führt der Einsatz des Indikators *Anzahl anderer Publikationen* zu einer signifikanten Erhöhung oder Verringerung ($B = .16$, $p = .101$) des Wertes der abhängigen Variablen. Das Konfidenzintervall des Bootstrap-basierten Tests zur Überprüfung des indirekten Effektes schließt entsprechend mit Werten zwischen $[-.012; .068]$ die Null mit ein und zeigt folglich keine Signifikanz des Effektes an. Für den Einsatz des Forschungsindikators *Anzahl anderer Publikationen* ergab die Mediationsanalyse dagegen wieder einen signifikanten negativen Effekt ($B = -.15$, $p = .004$) auf die intrinsische Forschungsmotivation. Ermittelt wurde dieser anhand der Regression der Variablen *Intrinsische Motivation* auf die Prädiktoren aus

Modell 3. Die abhängige Variable aus Modell 3 bleibt also sowohl vom Einsatz des Forschungsindikators als auch vom Ausmaß der intrinsischen Forschungsmotivation unbeeinflusst, während sich zwischen diesen beiden Prädiktoren der beschriebene negative Zusammenhang ergibt.

Modell 4: Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals

In Modell 4 bildet der eingeschätzte *Stellenwert des Publizierens in Zeitschriften mit hohem Impact Factor / in high ranked journals* die abhängige Variable. Wieder konnte für den theoretisch vermuteten Interaktionseffekt zwischen der unabhängigen Variablen *Einsatz Indikator* und der Variablen *Bedeutung für den Reputationszuwachs* keine Signifikanz nachgewiesen werden ($B = -.08$, $p = .543$). Der Produktterm wurde daher auch hier aus dem Modell entfernt.

Die Modellvoraussetzungen zur Berechnung des Regressionsmodells können nach Überprüfung mit den in Kapitel 6.2.2.1 beschriebenen Verfahren überwiegend als erfüllt angesehen werden. Allerdings stellt die geforderte Varianzhomogenität der Residuen wieder ein Problem dar. Der Breusch-Pagan-Test ergab eine signifikante Abweichung von der Nullhypothese der Varianzhomogenität ($\chi^2 = 5.42$, $p = .020$), so dass für die weiteren Analysen wieder die in *Process* implementierten heteroskedastizitätskonsistenten Standardfehler genutzt wurden. Die Ausreißerdiagnostik ergab insgesamt 27 besonders einflussreiche Fälle (drei Ausreißer hinsichtlich der Residuen sowie 24 Ausreißer hinsichtlich der Werte der Regressoren). Auch in Modell 4 wurde aufgrund fehlender inhaltlicher Begründungen auf einen Ausschluss der Ausreißerfälle verzichtet.

Die Ergebnisse der Regressionsanalyse finden sich in Tabelle 32. Für die F-Statistik ergibt sich mit einem Wert von $F(11, 399) = 35.46$ auch bei Modell 4 ein höchst signifikantes Ergebnis ($p < .001$).

Im Gegensatz zu den beiden vorangegangenen Modellen zeigt sich bei der abhängigen Variablen des eingeschätzten Stellenwerts des Publizierens in high ranked journals für den Einsatz des korrespondierenden Forschungsindikators wieder ein höchst signifikanter positiver Steigungseffekt ($B = .38$, $p < .001$). Auch für die Variablen *Bedeutung von Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals für den Reputationszuwachs* ($B = .27$, $p < .001$) sowie deren *Beitrag für den wissenschaftlichen Erkenntnisgewinn* ($B = .49$, $p < .001$) ergeben sich wieder signifikante Regressionskoeffizienten. Wie in Modell 1 lässt sich zudem ein signifikanter leicht negativer Effekt des Alters der Befragten ($B = -.03$, $p = .016$) nachweisen.

Für die intrinsische Motivation, die Beschäftigungsdauer, das Geschlecht und die Fachzugehörigkeit kann dagegen kein signifikanter Effekt nachgewiesen werden. Die Varianzaufklärung des Gesamtmodells erreicht mit einem korrigierten Bestimmtheitsmaß von $R^2 = .40$ dennoch einen sehr zufriedenstellenden Wert.

Tabelle 32: *Multiple Regression mit der abhängigen Variablen „Stellenwert des Publizierens in Zeitschriften mit hohem Impact Factor / in high ranked journals“*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
<i>Konstante</i>	1.006	.737	1.37	.173
Einsatz Indikator „Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals“ (0 kein Einsatz / 1 Einsatz)	.381	.109	3.48	.001
Bedeutung „Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals“ für Reputationszuwachs	.272	.070	3.90	.000
Beitrag „Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals“ für wissenschaftlichen Erkenntnisgewinn	.494	.057	8.69	.000
Index Intrinsische Motivation	.167	.111	1.50	.135
Alter	-.033	.014	-2.43	.016
Beschäftigungsdauer	.006	.016	0.37	.713
Geschlecht (0 männlich / 1 weiblich)	-.133	.120	-1.11	.269
Dummy Erziehungswissenschaft	-.078	.166	-0.47	.636
Dummy Politikwissenschaft	.125	.223	0.56	.575
Dummy Psychologie	.217	.140	1.55	.122
Dummy Soziologie	.196	.202	0.97	.333
<i>R</i>²		.416		
Korrigiertes <i>R</i>²		.400		
<i>F</i>(11, 399)		35.46 (<i>p</i> < .001)		
<i>n</i>		411		

Anmerkung. *B*: Unstandardisierter Regressionskoeffizient; *SE*: Standardfehler von *B*; *t*: t-Statistik von *B*; *R*²: Bestimmtheitsmaß; *F*: F-Statistik mit Freiheitsgraden; Referenzkategorie für Fachdummies: Wirtschaftswissenschaften

Quelle: eigene Berechnungen

Ein über die intrinsische Motivation medierter Zusammenhang zwischen dem Regressor *Einsatz Indikator „Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals“* und der abhängigen Variablen kann auch in Modell 4 nicht nachgewiesen werden. Das mit Hilfe des Bootstrap-basierten Tests ermittelte Vertrauensintervall zeigt mit Werten zwischen [-.009; .038] keine Signifikanz des vermuteten indirekten Effektes an.

Zudem konnte bei der Regression der intrinsischen Motivation auf die unabhängige Variable *Einsatz Indikator „Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high-ranked journals“* kein signifikanter Effekt der unabhängigen Variablen auf die

intrinsische Motivation festgestellt werden ($B = .03$, $p = .582$). Die intrinsische Motivation steht demnach – zumindest hinsichtlich der hier betrachteten Publikationsform – weder in einem Zusammenhang mit dem eingeschätzten Stellenwert des Publizierens in high ranked journals in der Forschungsarbeit der Professorinnen und Professoren noch wird diese vom Einsatz des entsprechenden Forschungsindikators in hochschulinternen Steuerungs- und Anreizverfahren beeinflusst.

Modell 5: Betreuung von Promotionen

Im nun folgenden Modell 5 wird der eingeschätzte Stellenwert der Betreuung von Promotionen in der Forschungsarbeit der Professorinnen und Professoren als abhängige Variable herangezogen. Zunächst wurde wieder ein Modell gerechnet, welches den Produktterm zwischen den Variablen *Einsatz Indikator „Anzahl betreuter Promotionen“* und *Bedeutung „Anzahl betreuter Promotionen“ für den Reputationszuwachs* zur Modellierung des Interaktionseffektes enthielt. Für den Steigungsparameter des Interaktionseffektes ergab sich dabei ein Wert von $B = .164$ mit $p = .078$. Der p -Wert überschreitet damit knapp das festgelegte Signifikanzniveau von $\alpha = 5\%$. Beim Vorliegen einer gerichteten Hypothese wäre die Annahme eines einseitigen Testproblems und damit die Halbierung des p -Wertes gerechtfertigt (Baltes-Götz 2015, S. 26; Bortz und Döring 2003, S. 498–499; Fahrmeir et al. 2003, S. 409–413). Der Interaktionseffekt wäre dann als signifikant zu interpretieren. Dieses Vorgehen ist aber nicht unumstritten und erfordert eine klare theoretische Fundierung, warum eine gegenteilige Richtung des Zusammenhangs im Voraus ausgeschlossen wird. Diese Eindeutigkeit kann das vorliegende theoretische Modell allerdings nicht für sich beanspruchen, so dass das Testproblem weiterhin als zweiseitig behandelt wird. Der Produktterm wurde entsprechend aus Modell 5 entfernt und die Analysen ohne Berücksichtigung des Interaktionseffektes fortgesetzt.

Bei der Überprüfung der Modellvoraussetzungen ergab der Breusch-Pagan-Test auch für Modell 5 eine Verletzung der Annahme der Varianzhomogenität der Residuen ($\chi^2 = 8.20$, $p = .004$), so dass bei den weiteren Analysen wieder auf robuste Standardfehler zurückgegriffen wurde. Hinsichtlich der übrigen Modellvoraussetzungen waren keine gravierenden Verletzungen zu beobachten.

Die Analyse der einflussreichen Fälle führte zur Ermittlung von drei Ausreißern hinsichtlich der Residuen (Werte größer drei Standardabweichungen) und 22 Ausreißern hinsichtlich der Regressorwerte. Auch wenn ein Ausschluss zu einer deutlichen Erhöhung der Erklärungskraft des Modells geführt hätte, konnte dieser auch bei diesen Fällen inhaltlich nicht begründet werden. Auf einen Ausschluss der Fälle wurde entsprechend verzichtet.

In Tabelle 33 sind die Ergebnisse der Berechnungen zu Modell 5 aufgeführt. Der globale Overall-F-Test ergibt einen höchst signifikanten ($p < .001$) Wert von $F(11, 404) = 11.30$.

Tabelle 33: *Multiple Regression mit der abhängigen Variablen „Stellenwert der Betreuung von Promotionen“*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
<i>Konstante</i>	.823	.590	1.39	.164
Einsatz Indikator „Anzahl betreuter Promotionen“ (0 kein Einsatz / 1 Einsatz)	.095	.101	0.94	.348
Bedeutung „Anzahl betreuter Promotionen“ für Reputationszuwachs	.097	.053	1.83	.068
Beitrag „Betreuung von Promotionen“ für wissenschaftlichen Erkenntnisgewinn	.338	.047	7.22	.000
Index Intrinsische Motivation	.242	.095	2.56	.011
Alter	.010	.010	1.00	.316
Beschäftigungsdauer	.000	.012	0.02	.988
Geschlecht (0 männlich / 1 weiblich)	.024	.103	0.24	.814
Dummy Erziehungswissenschaft	-.016	.133	-0.12	.904
Dummy Politikwissenschaft	-.092	.208	-0.44	.658
Dummy Psychologie	.152	.132	1.15	.250
Dummy Soziologie	-.149	.156	-0.96	.339
<i>R</i>²		.218		
Korrigiertes <i>R</i>²		.196		
<i>F</i>(11, 404)		11.30 ($p < .001$)		
<i>n</i>		416		

Anmerkung. *B*: Unstandardisierter Regressionskoeffizient; *SE*: Standardfehler von *B*; *t*: t-Statistik von *B*; *R*²: Bestimmtheitsmaß; *F*: F-Statistik mit Freiheitsgraden; Referenzkategorie für Fachdummies: Wirtschaftswissenschaften

Quelle: eigene Berechnungen

Als signifikante Einflussgrößen lassen sich nach Betrachtung der p -Werte lediglich der eingeschätzte Beitrag der Betreuung von Promotionen für den wissenschaftlichen Erkenntnisgewinn ($B = .34$, $p < .001$) sowie die intrinsische Forschungsmotivation ($B = .24$, $p = .011$) ausmachen. Der bisher bei allen vorangegangenen Modellen signifikante (positive) Zusammenhang zwischen der eingeschätzten Bedeutung der analysierten Forschungstätigkeit für den Reputationszuwachs in der Scientific Community und dem eingeschätzten Stellenwert von dieser in der eigenen Forschungsarbeit ist in Modell 5 nicht

mehr zu beobachten. Mit einem Koeffizienten von $B = .10$ und einem p -Wert von $p = .068$ wird die Signifikanzgrenze knapp überschritten.⁷¹

Auch für alle übrigen Prädiktoren zeigt sich kein signifikanter Effekt. Entsprechend gering fällt die Varianzaufklärung aus. Mit einem korrigierten Bestimmtheitsmaß von $R^2 = .20$ können nur knapp 20 Prozent der Variation der abhängigen Variablen durch das Modell erklärt werden. Für Modell 5 ist daher anzunehmen, dass wesentliche Prädiktoren bei der Spezifikation des Modells unberücksichtigt geblieben sind. Der Stellenwert, den die Betreuung von Promotionen in der Forschungsarbeit einnimmt, kann z. B. schlicht davon abhängen, inwieweit überhaupt die einzelne Professorin oder der einzelne Professor als Betreuer gefragt ist. Die Anzahl der Betreuungsanfragen bzw. der potentiellen Promovierenden könnte also eine entscheidende erklärende Variable darstellen und wiederum von anderen Aspekten wie dem fachlichen Schwerpunkt (z. B. Forschungsthematiken, mit denen sich viel oder wenig potentielle Promovierende beschäftigen) der Professorin oder des Professors abhängen.

Der theoretisch vermutete negative Zusammenhang zwischen dem Einsatz des Forschungsindikators *Anzahl betreuter Promotionen* und der intrinsischen Forschungsmotivation konnte im Zuge der Überprüfung des Mediationseffektes ebenfalls nicht bestätigt werden ($B = -.08$, $p = .17$). Auch ein indirekter Effekt, also ein über die intrinsische Motivation vermittelter Effekt des Einsatzes des Forschungsindikators *Anzahl betreuter Promotionen* auf die abhängige Variable *Stellenwert der Betreuung von Promotionen in der Forschungsarbeit*, konnte nicht nachgewiesen werden. Das über den Bootstrap-basierten Test errechnete Konfidenzintervall umfasst den Wertebereich $[-.062; .003]$ und lässt damit keine Ablehnung der Nullhypothese zu.

Modell 6: Einwerben von DFG-Drittmitteln

Abschließend wird nun der Stellenwert des Einwerbens von Drittmitteln in der Forschungsarbeit der Professorinnen und Professoren betrachtet. Dies geschieht in Modell 6 zunächst mit dem Fokus auf DFG-Drittmittel, während in der nachfolgenden Ergebnisdarstellung zu Modell 7 die sonstigen Drittmittel näher analysiert werden.

Der in Hypothese *H2.1* formulierte Interaktionseffekt konnte auch für den Indikator *Höhe der eingeworbenen Drittmittel der DFG* nicht nachgewiesen werden ($B = .06$, $p = .688$). Berechnet wurde entsprechend das in Tabelle 34 dargestellte Regressionsmodell. Für dieses Modell zeigten sich bei der Überprüfung der Modellvoraussetzungen keine gravierenden Verstöße gegen diese. Auch die mittels des Breusch-Pagan-Tests

⁷¹ Auf die Möglichkeit, den Test als einseitiges Testproblem aufzufassen und den p -Wert entsprechend zu halbieren, wurde bereits weiter oben hingewiesen. Bei dem vorliegenden theoretischen Modell erscheint dieses Vorgehen aber wie erwähnt nicht gerechtfertigt. Auch durch den Ausschluss der Ausreißerfälle hätte ein p -Wert unter dem 5%-Niveau ‚erkaufft‘ werden können. Darauf wurde aufgrund fehlender inhaltlicher Begründung ebenfalls verzichtet.

überprüfte Varianzhomogenität der Residuen ist gegeben ($\chi^2 = 2.95$, $p = .086$). Zu berücksichtigen sind allerdings die sehr unterschiedlich stark besetzten Gruppen bei der unabhängigen Variablen *Einsatz Indikator „Höhe der eingeworbenen Drittmittel der DFG“* (vgl. die deskriptiven Analysen in Kapitel 6.1.2). Bei dieser gaben in der Regressionsstichprobe nur 16.3 % ($n = 78$) der Befragten an, dass der Indikator nicht eingesetzt wird. Dem gegenüber stehen 83.7 % ($n = 400$) der Befragten, die angaben, dass der Indikator genutzt wird. Die Aussagekraft der Ergebnisse – zumindest bei dieser Variablen – ist damit sicherlich eingeschränkt.

Für Modell 6 wurden insgesamt 12 Ausreißerfälle hinsichtlich der Regressorwerte ermittelt. Diese wurden nach inhaltlicher Überprüfung der Datensätze für die weiteren Analysen beibehalten.

Tabelle 34: *Multiple Regression mit der abhängigen Variablen „Stellenwert des Einwerbens von DFG-Drittmitteln“*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
<i>Konstante</i>	.409	.709	0.58	.564
Einsatz Indikator „Höhe der eingeworbenen Drittmittel der DFG“ (0 kein Einsatz / 1 Einsatz)	.113	.147	0.77	.443
Bedeutung „Höhe der eingeworbenen Drittmittel der DFG“ für Reputationszuwachs	.101	.065	1.56	.120
Beitrag „Durchführung von Drittmittelprojekten der DFG“ für wissenschaftlichen Erkenntnisgewinn	.414	.060	6.92	.000
Index Intrinsische Motivation	.237	.101	2.35	.019
Alter	-.015	.012	-1.27	.205
Beschäftigungsdauer	-.005	.014	-0.36	.722
Geschlecht (0 männlich / 1 weiblich)	.096	.122	0.79	.431
Dummy Erziehungswissenschaft	.269	.159	1.69	.092
Dummy Politikwissenschaft	.643	.209	3.07	.002
Dummy Psychologie	.660	.161	4.11	.000
Dummy Soziologie	.830	.185	4.49	.000
<i>R</i>²		.285		
Korrigiertes <i>R</i>²		.268		
<i>F</i>(11, 466)		16.89 (<i>p</i> < .001)		
<i>n</i>		478		

Anmerkung. *B*: Unstandardisierter Regressionskoeffizient; *SE*: Standardfehler von *B*; *t*: t-Statistik von *B*; *R*²: Bestimmtheitsmaß; *F*: F-Statistik mit Freiheitsgraden; Referenzkategorie für Fachdummies: Wirtschaftswissenschaften

Quelle: eigene Berechnungen

Der F-Test des Gesamtmodells zeigt mit einem Wert von $F(11, 466) = 16.89$ eine höchst signifikante ($p < .001$) Abweichung von der Nullhypothese, dass alle Steigungsparameter den Wert Null annehmen.

Für die aus dem theoretischen Modell abgeleiteten Einflussvariablen ergeben sich lediglich für den eingeschätzten Beitrag des Durchführens von DFG-Drittmittelprojekten für den wissenschaftlichen Erkenntnisgewinn ($B = .41, p < .001$) sowie für die intrinsische Forschungsmotivation ($B = .24, p = .019$) signifikante positive Steigungsparameter. Zudem zeigen sich signifikante Fachunterschiede, die die Ergebnisse des deskriptiven Fächervergleichs aus Kapitel 6.1.6 widerspiegeln. In den Fächern Politikwissenschaft ($B = .64, p = .002$), Psychologie ($B = .66, p < .001$) und Soziologie ($B = .83, p < .001$) fällt der eingeschätzte Stellenwert der Durchführung von Drittmittelprojekten der DFG in der Forschungsarbeit der Professorinnen und Professoren demnach jeweils signifikant höher aus als in der Referenzkategorie der Wirtschaftswissenschaften. Für alle anderen Variablen lassen sich keine signifikanten Effekte nachweisen. Im Hinblick auf die Variable *Einsatz Indikator „Höhe der eingeworbenen Drittmittel der DFG“* ist das Ergebnis aber aufgrund der beschriebenen stark ungleich verteilten Gruppengrößen nur mit Vorsicht zu interpretieren.

Das korrigierte Bestimmtheitsmaß deutet mit einem Wert von $R^2 = .27$ auf eine eher schlechte Anpassung des Modells an die vorliegenden Daten hin. Wie die Ergebnisse zu den einzelnen Regressoren bereits vermuten ließen, fällt die Varianzaufklärung entsprechend niedrig aus. Entscheidende unabhängige Einflussvariablen scheinen in diesem Modell also zu fehlen. Dies könnte z. B. die generelle Zufriedenheit mit der finanziellen Grundausstattung sein sowie die individuelle Einschätzung der Notwendigkeit, diese durch zusätzliche Drittmittel aufzustocken. Jemand, der seine Forschungsarbeit auch ohne Drittmittelfinanzierung als ausreichend befriedigend ansieht, wird dem Einwerben von Drittmitteln beispielsweise einen geringeren Stellenwert in der eigenen Forschungsarbeit einräumen als jemand, für den das Betreiben von Forschung ohne unterstützende Drittmittel undenkbar ist.

Aufgrund der ungleich verteilten Gruppengrößen bei der unabhängigen Variablen *Einsatz Indikator „Höhe der eingeworbenen Drittmittel der DFG“* ist auch die Aussagekraft des Tests des Mediations- bzw. des vermuteten Verdrängungseffektes als problematisch anzusehen. Daher sei hier nur der Vollständigkeit halber darauf hingewiesen, dass sich für die Forschungstätigkeit *Einwerben von Drittmitteln der DFG* keine signifikanten Ergebnisse ergaben, welche die Verdrängungsthese stützen bzw. einen über die intrinsische Motivation medierten Zusammenhang zwischen der unabhängigen und der abhängigen Variablen vermuten lassen.

Modell 7: Einwerben von sonstigen Drittmitteln

Für die sonstigen Drittmittel gelten hinsichtlich der Variablen *Einsatz Indikator „Höhe der sonstigen eingeworbenen Drittmittel“* dieselben methodischen Einschränkungen aufgrund der ungleichen Gruppengrößen wie bei den DFG-Drittmitteln. 15.3 % ($n = 73$) der Befragten in der Regressionsstichprobe gaben an, dass ein solcher Indikator in Steuerungs- und Anreizverfahren an ihrer Fakultät nicht zum Einsatz kommt, während 84.7 % ($n = 404$) der Befragten angaben, dass der Indikator eingesetzt wird. Die folgenden Ergebnisse sind in Bezug auf die genannte Variable daher nur sehr eingeschränkt aussagekräftig.

Ein Interaktionseffekt zwischen dem Einsatz des Forschungsindikators *Höhe der eingeworbenen sonstigen Drittmittel* und der eingeschätzten Bedeutung der eingeworbenen Drittmittel für den Reputationszuwachs in der Scientific Community kann nicht nachgewiesen werden ($B = -.13, p = .404$). Auch in Modell 7 wird daher kein Produktterm der beiden Variablen integriert.

Im Gegensatz zum vorangegangenen Modell ergeben sich bei Modell 7 nach Durchführung des Breusch-Pagan-Tests Probleme hinsichtlich der geforderten Varianzhomogenität der Residuen ($\chi^2 = 5.47, p = .019$). Bei der Berechnung des Regressionsmodells wurde daher wieder auf die robusten Standardfehler zurückgegriffen. Die übrigen Modellvoraussetzungen können überwiegend als erfüllt angesehen werden. Die Ausreißerdiagnostik ergab 21 Fälle, die hinsichtlich der Hebelwerte bzw. *Cooks D* über den jeweiligen kritischen Schwellenwerten lagen, allerdings inhaltlich keine Auffälligkeiten zeigten. Diese Fälle wurden daher in der Analyse belassen.

Auch beim vorliegenden Modell kann die Nullhypothese des globalen Overall-F-Tests mit einem Wert von $F(11, 465) = 20.10$ ($p < .001$) abgelehnt werden. Die Parameterwerte und Ergebnisse der Signifikanztests zu den einzelnen Regressoren sind in Tabelle 35 aufgeführt.

Wie auch bei den DFG-Drittmitteln zeigen sich in Modell 7 nur wenig signifikante Beziehungen zwischen der abhängigen Variablen und den Regressoren. Für die theoretisch hergeleiteten Variablen ergibt sich allein für den eingeschätzten Beitrag der Durchführung von sonstigen Drittmittelprojekten für den wissenschaftlichen Erkenntnisgewinn ein signifikant positiver Regressionskoeffizient ($B = .50, p < .001$). Der bei den DFG-Drittmitteln noch signifikante Zusammenhang zwischen der abhängigen Variablen und der intrinsischen Motivation lässt sich für die sonstigen Drittmittel nicht mehr nachweisen ($B = .05, p = .664$). Den fächervergleichenden deskriptiven Analysen entsprechende (vgl. Kapitel 6.1.6) signifikante Fachunterschiede ergeben sich hinsichtlich der Referenzkategorie der Wirtschaftswissenschaften für die Erziehungswissenschaften ($B = .51, p = .003$), die Politikwissenschaften ($B = .55, p = .007$) und die Soziologie ($B = .40,$

$p = .035$). Die Modellgüte ist mit einem korrigierten Bestimmtheitsmaß von $R^2 = .26$ insgesamt als eher schlecht zu bewerten (vgl. zu möglichen fehlenden erklärenden Variablen die Ergebnisdarstellung zu Modell 6).

Tabelle 35: *Multiple Regression mit der abhängigen Variablen „Stellenwert des Einwerbens von sonstigen Drittmitteln“*

	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
<i>Konstante</i>	.623	.693	0.90	.369
Einsatz Indikator „Höhe der sonstigen eingeworbenen Drittmittel“ (0 kein Einsatz / 1 Einsatz)	.107	.178	0.61	.545
Bedeutung „Höhe der sonstigen eingeworbenen Drittmittel“ für Reputationszuwachs	.101	.072	1.41	.160
Beitrag „Durchführung von sonstigen Drittmittelprojekten“ für wissenschaftlichen Erkenntnisgewinn	.496	.064	7.80	.000
Index Intrinsische Motivation	.045	.104	.43	.664
Alter	.007	.012	.53	.594
Beschäftigungsdauer	-.008	.014	-0.61	.544
Geschlecht (0 männlich / 1 weiblich)	-.032	.126	-0.26	.799
Dummy Erziehungswissenschaft	.508	.173	2.94	.003
Dummy Politikwissenschaft	.548	.204	2.69	.007
Dummy Psychologie	-.089	.193	-.46	.645
Dummy Soziologie	.401	.190	2.11	.035
<i>R</i>²		.280		
Korrigiertes <i>R</i>²		.263		
<i>F</i>(11, 465)		20.10 (<i>p</i> < .001)		
<i>n</i>		477		

Anmerkung. *B*: Unstandardisierter Regressionskoeffizient; *SE*: Standardfehler von *B*; *t*: t-Statistik von *B*; *R*²: Bestimmtheitsmaß; *F*: F-Statistik mit Freiheitsgraden; Referenzkategorie für Fachdummies: Wirtschaftswissenschaften

Quelle: eigene Berechnungen

Ein signifikanter Mediationseffekt bzw. eine signifikanter negativer Effekt des Einsatzes des Indikators *Höhe der eingeworbenen sonstigen Drittmittel* auf die intrinsische Forschungsmotivation konnte auch für das hier betrachtete Modell nicht nachgewiesen werden. Die entsprechenden Testergebnisse werden wie bereits bei Modell 6 aufgrund der Verteilungsproblematik der stark ungleichen Gruppengrößen bei der unabhängigen Variablen nicht weiter im Detail berichtet.

Nachdem nun das Vorgehen bei der Berechnung der sieben Regressionsmodelle und die dazugehörigen Resultate dargestellt wurden, folgt im nächsten Kapitel eine zusammenfassende Darstellung der Ergebnisse im Hinblick auf die in Kapitel 4.2 formulierten Forschungshypothesen sowie die inhaltliche Interpretation der gewonnenen Daten. Davon werden nochmals knapp die Ergebnisse der deskriptiven Analysen in Erinnerung gerufen.

6.3 Zusammenfassung und Interpretation der wesentlichen Ergebnisse

6.3.1 Diskussion der Ergebnisse im Hinblick auf die Fragestellungen der Arbeit und die Forschungshypothesen

Eine wesentliche Zielsetzung der Governance-Reformen im Hochschulbereich unter dem Zeichen des New Public Managements ist die Intensivierung des Governancemechanismus des Wettbewerbs. Während sich der Mechanismus der akademischen Selbstorganisation von den Reformprozessen noch weniger stark beeinflusst zeigt, ist – neben ebenfalls sichtbaren Verschiebungen der ‚Regler‘ der Mechanismen der staatlichen Regulierung, der externen Steuerung und der hierarchischen Selbststeuerung – eine deutliche Stärkung des Wettbewerbselements zwischen und innerhalb der Hochschulen zu beobachten (vgl. Kapitel 2.1.2). Als Instrumente dieses Wettbewerbs lassen sich die auf Landes- und hochschulinterner Ebene eingeführten indikatorenbasierten Steuerungs- und Anreizverfahren verstehen, die in der deutschen Hochschullandschaft mittlerweile teils starke Verbreitung gefunden haben (vgl. Kapitel 2.2 und 2.4).

In dieser Studie wurde nun mit dem Fokus auf die Sozialwissenschaften untersucht, inwieweit Professorinnen und Professoren in Deutschland mit dem Einsatz der verschiedenen Verfahren konfrontiert sind. Dabei zeigte sich, dass eine flächendeckende Implementierung von hochschulinternen Steuerungs- und Anreizverfahren in den Sozialwissenschaften bislang nicht beobachtet werden kann. Am häufigsten wurde angegeben, dass eine *indikatorenbasierte Mittelzuweisung zwischen Hochschulleitung und Fakultät* zum Einsatz kommt. Mit 58.1 % sind aber selbst bei diesem meistgenannten Verfahren nur etwas über die Hälfte der befragten Professorinnen und Professoren von dem Einsatz des Verfahrens betroffen. Es folgen mit ähnlich hohen Antworthäufigkeiten die Verfahren *Zielvereinbarungen zwischen Hochschulleitung und Fakultät* (55.3 %) sowie die *indikatorenbasierte Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl* (54.0 %). *Zielvereinbarungen auf der Ebene Fakultät-Institut/Lehrstuhl* betreffen nur etwa ein Viertel (25.5 %) der befragten Professorinnen und Professoren. *Benchmarkingverfahren*

spielen außer in den Wirtschaftswissenschaften (hier geben ebenfalls etwa ein Viertel der Befragten an, dass diese eingesetzt werden) so gut wie keine Rolle.

Im Vergleich zu den Ergebnissen der Studie von Bogumil et al. (2013), in der für die indikatorenbasierte Mittelzuweisung eine Implementierungsquote an deutschen Universitäten von über 90 % (sowohl auf fakultätsinterner als auch universitätsinterner Ebene) und für universitätsinterne Zielvereinbarungen (Ebene Hochschulleitung – Fakultät) eine Quote von 75 % ermittelt wurde (vgl. Kapitel 2.2), machen die Ergebnisse der vorliegenden Studie deutlich, dass die Verfahren in den Sozialwissenschaften dagegen vergleichsweise selten zum Einsatz kommen. In anderen Fächern wie der Medizin sind einzelne Verfahren dagegen wesentlich stärker verbreitet (vgl. Karmann 2014, S. 462).

In den Steuerungs- und Anreizverfahren finden eine Reihe verschiedener Forschungsindikatoren Verwendung, wobei – wie in anderen forschungsevaluativen Verfahren und Hochschulrankings auch (vgl. Kapitel 2.3.2) – insbesondere die Drittmittelhöhe (sowohl DFG- als auch sonstige Drittmittel), publikationsbasierte Indikatoren (hier vor allem über alle Fächer hinweg die Publikationen mit Peer-Review) sowie die Anzahl der Promotionen von zentraler Bedeutung sind (vgl. Kapitel 6.1.2). Zwischen den einzelnen Verfahren zeigen sich dabei Unterschiede hinsichtlich der verwendeten Indikatoren: Während die Anzahl der Promotionen vor allem in der indikatorenbasierten Mittelzuweisung zum Einsatz kommt (auf beiden Ebenen), werden publikationsbasierte Indikatoren vor allem auf fakultätsinterner Ebene in der indikatorenbasierten Mittelzuweisung und in Zielvereinbarungen eingesetzt.

In fächervergleichender Perspektive machen sich hinsichtlich der verwendeten Indikatoren die unterschiedlichen Fach- bzw. Publikationskulturen bemerkbar (vgl. hierzu auch die Analysen zur Bedeutung verschiedener Aspekte der Forschungsarbeit für den Reputationszuwachs in der Scientific Community in Kapitel 6.1.4). So wird in der Psychologie und den Wirtschaftswissenschaften beispielsweise die Anzahl der Publikationen ohne Peer-Review-Verfahren gegenüber den anderen Fächern signifikant seltener als Indikator eingesetzt, während die Anzahl von Publikationen in high-ranked journals in diesen beiden Fächern deutlich öfter als Indikator zum Einsatz kommt. Auch in Bezug auf die Verwendung von Drittmitteln (in der Psychologie kommen diese fast immer zum Einsatz, in den Wirtschaftswissenschaften dagegen signifikant seltener als in den anderen Fächern) und der Anzahl von Promotionen (in den Wirtschaftswissenschaften ebenfalls signifikant seltener) als Indikator, zeigen sich fachspezifische Besonderheiten (vgl. im Detail Kapitel 6.1.2).

Inwieweit der Einsatz dieser Indikatoren in hochschulinternen Steuerungs- und Anreizverfahren in einem Zusammenhang mit dem Forschungsverhalten der Professorinnen und Professoren steht, wurde mit Hilfe der im vorangegangenen Kapitel dargestell-

ten multiplen Regressionsmodelle statistisch überprüft. Vor dem Hintergrund der organisations- und professionstheoretischen Überlegungen aus Kapitel 3 und dem motivationstheoretischen Analyserahmen aus Kapitel 4.1 werden die Ergebnisse dieser Berechnungen nun im Hinblick auf die abgeleiteten Forschungshypothesen (vgl. Kapitel 4.2) inhaltlich diskutiert. Wo es sinnvoll oder notwendig erscheint, wird zudem auf die Ergebnisse der bivariaten Analysen aus Kapitel 6.2.1 Bezug genommen.

Zunächst werden der Reihe nach die Forschungshypothesen aus Kapitel 4.2 zu den Ergebnissen der einzelnen Modelle in Bezug gesetzt. Im Anschluss folgen kurze Anmerkungen zu weiteren Erkenntnissen und in einem eigenen Teilkapitel schließlich die zusammenfassende Ergebnisdarstellung.

H1: Der Einsatz eines Forschungsindikators in hochschulinternen Steuerungs- und Anreizverfahren korreliert positiv mit dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

Die theoretische Annahme, dass indikatorenbasierte Steuerungsversuche als external regulierte Form der Motivation bei den sozialwissenschaftlichen Professorinnen und Professoren wirksam werden und so in einem (schwachen) Zusammenhang mit dem Forschungsverhalten stehen, lässt sich anhand der tätigkeitsspezifischen Regressionsmodelle nur vereinzelt bestätigen. Während sich in den bivariaten Analysen außer für die Indikatoren *Anzahl betreuter Promotionen* sowie *Höhe der eingeworbenen sonstigen Drittmittel* durchgehend signifikant schwache bis mittelstarke positive Zusammenhänge mit dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit zeigten, verschwinden diese unter Kontrolle der Fachzugehörigkeit und der anderen herangezogenen Drittvariablen größtenteils. Ein signifikanter schwacher bis mittelstarker positiver Zusammenhang zeigt sich in der multivariaten Betrachtung lediglich bei den publikationsbasierten Indikatoren *Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* und *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals*. Im Gegensatz zu den anderen untersuchten Indikatoren werden diese beiden Indikatoren – neben den Drittmittelindikatoren – beim Einsatz in den Steuerungs- und Anreizverfahren besonders stark gewichtet (vgl. Kapitel 6.1.2 und 6.1.3). Dies könnte eine mögliche Erklärung dafür sein, warum nur der Einsatz dieser beiden Indikatoren in einem signifikanten Zusammenhang mit dem Forschungsverhalten steht, während sich bei den übrigen Indikatoren keine signifikanten Zusammenhänge nachweisen

lassen.⁷² Hypothese *H1* kann damit nur mit Einschränkung auf die beiden genannten Indikatoren bestätigt werden und ist ansonsten abzulehnen.

In Anbetracht der bereits ausführlich diskutierten Steuerungsproblematik aufgrund der Spezifika der Organisation Universität und der akademischen Profession selbst (vgl. Kapitel 3) erscheint dieses Ergebnis nicht gänzlich überraschend. Ein leichter bis moderater Zusammenhang zwischen dem Einsatz von Forschungsindikatoren in indikatorenbasierten Steuerungsverfahren und dem Forschungsverhalten zeigt sich demnach nur bei entsprechend starker Gewichtung der Indikatoren in den Steuerungs- und Anreizverfahren. Daneben spielen sicherlich auch Spezifika der jeweiligen Forschungstätigkeit bzw. der eingesetzten Forschungsindikatoren selbst eine Rolle. Denkbar ist, dass den Anforderungen publikationsbasierter Forschungsindikatoren leichter entsprochen werden kann als Bewertungskriterien wie der Drittmittelhöhe oder der Anzahl betreuter Promotionen: Während letztere auch stark von externen Faktoren beeinflusst sein können (z. B. Passung zwischen eigenem Forschungsschwerpunkt und Förderschwerpunkten bei der Einwerbung von Drittmittelprojekten, passende Promovierende), lässt sich die Erhöhung des eigenen Publikationsoutputs sicherlich deutlich unabhängiger von solchen äußeren Faktoren und damit selbstbestimmter forcieren.⁷³

H2.1: Je höher die eingeschätzte Bedeutung einer Forschungstätigkeit für den Reputationszuwachs in der Scientific Community ist, desto höher ist der eingeschätzte Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

Hypothese *H2.1* lässt sich für alle publikationsbasierten Forschungstätigkeiten (Modell 1 – 4) bestätigen. Je höher die Bedeutung der jeweiligen Publikationsform für den Reputationszuwachs in der Scientific Community eingeschätzt wird, desto höher fällt auch der Stellenwert des Verfassens entsprechender Publikationen in der Forschungsarbeit der Professorinnen und Professoren aus. Die theoretisch vermutete hohe Relevanz der Scientific Community als handlungsleitender Orientierungspunkt (vgl. Kapitel 3.2) bestätigt sich also zumindest für die Publikationsarbeit. Da das Publizieren von wissenschaft-

⁷² Die Drittmittelindikatoren werden zwar wie erwähnt in den Verfahren ebenfalls stark gewichtet, allerdings ergeben sich bei den entsprechenden Regressionsmodellen (Modell 6 und Modell 7) Probleme hinsichtlich der stark unterschiedlich besetzten Gruppengrößen bei der unabhängigen Variablen (Einsatz Indikator ja/nein; vgl. Kapitel 6.2.2.2). Der nicht signifikante Effekt auf die abhängige Variable ist in diesen beiden Modellen möglicherweise auf diese Verteilungsproblematik bei der unabhängigen Variablen zurückzuführen.

⁷³ Wobei natürlich auch hier externe Faktoren wie zeitliche Ressourcen, thematische Schwerpunktsetzungen von Zeitschriften oder das Gutachterverfahren für den Publikationserfolg hinderlich sein können. Insgesamt dürfte bei publikationsbasierten Tätigkeiten für die Professorinnen und Professoren aber eher die Möglichkeit gegeben sein, selbst an den entsprechenden ‚Stellschrauben zu drehen‘, um den Publikationsoutput zu steigern.

lichen Erkenntnissen in Zeitschriften und anderen Formaten das primäre Austauschmedium zwischen den Mitgliedern der Scientific Community darstellt und die Sichtbarkeit und Reputation eines Wissenschaftlers oder einer Wissenschaftlerin durch Publikationen wesentlich (mit)bestimmt wird, ist dieses Ergebnis durchaus plausibel.

Die in den bivariaten Analysen noch signifikanten Zusammenhänge für die Forschungstätigkeiten *Betreuung von Promotionen* sowie dem *Einwerben von Drittmitteln* (DFG und *sonstige*) lassen sich im Rahmen der multivariaten Analysen nicht mehr bestätigen. Bei den Promotionen könnte eine Erklärung für den fehlenden Zusammenhang in der geringen Sichtbarkeit von diesen in der Scientific Community zu finden sein. Die wenigsten Professorinnen und Professoren dürften umfassende Kenntnis über den aktuellen Stand der betreuten Doktorarbeiten ihrer Kolleginnen und Kollegen haben. Ein entsprechender Reputationszuwachs durch eine Erhöhung der Anzahl der Promovierenden bliebe demnach aus. Zwar attestierten in der Befragung durchaus einige Professorinnen und Professoren ($n = 94$; 14.7 %) der Anzahl betreuter Promotionen eine (sehr) hohe Bedeutung für den Reputationszuwachs in der Scientific Community. Entsprechend der Annahmen von Hypothese *H2.1* sollte sich das in einem höheren Stellenwert der Betreuung von Promotionen in der Forschungsarbeit niederschlagen. Möglich ist allerdings hier, dass die Befragten bei der Beantwortung der Frage nach der Bedeutung für den Reputationszuwachs mehr auf qualitative denn auf quantitative Aspekte abzielten: Aufmerksamkeit und damit verbundene Reputation wird aus dieser qualitativen Perspektive dann nicht über die Anzahl der Promotionen, sondern durch besonders qualitativ hochwertige Arbeiten mit Relevanz für die Scientific Community erzielt. Ein Reputationsgewinn in diesem Sinne lässt sich dann nicht durch eine einfache Erhöhung der Anzahl der Promovierenden erzielen bzw. durch eine intensivere Betreuung der Promovierenden forcieren, was den fehlenden signifikanten Zusammenhang zwischen den beiden Variablen in Modell 5 erklären könnte.⁷⁴

Im Hinblick auf den nicht signifikanten Zusammenhang bei der Forschungstätigkeit *Einwerben von Drittmitteln* könnte eine Rolle spielen, dass die Möglichkeiten und die Erfolgsaussichten zur Einwerbung von Drittmitteln je nach Forschungsgebiet und Forschungsausrichtung (Grundlagenforschung vs. anwendungsbezogener Forschung) variieren können. Beispielsweise nehmen für die Grundlagenforschung die Fördermöglichkeiten der DFG einen hohen Stellenwert in der Forschungsarbeit ein, während für die anwendungsbezogene Forschung auch privatwirtschaftliche Fördergelder von hoher Bedeutung sein können (vgl. hierzu auch die bivariaten Analysen in Kapitel 6.2.1 zum Stel-

⁷⁴ Wobei die Art und Intensität der Betreuung der Promovierenden natürlich Auswirkungen auf die Qualität der Arbeiten haben kann. Allerdings lässt sich durch eine intensive Betreuung allein nicht garantieren, dass die spätere Arbeit in der Scientific Community in der gewünschten Weise rezipiert wird.

lenwert von Grundlagenforschung und anwendungsbezogener Forschung in der Forschungsarbeit der Professorinnen und Professoren). Ebenso kann es für Professorinnen und Professoren bestimmter populärer Forschungsgebiete einfacher sein, Fördergelder einzuwerben, als für Wissenschaftlerinnen und Wissenschaftler mit ‚exotischen‘ Forschungsthemen. Die Bedeutung für den Reputationszuwachs könnte dann von beiden Gruppen jeweils hoch eingeschätzt werden, aber während sich dies bei den einen durch verstärktes Einwerben von Drittmitteln bemerkbar macht, steht diese Einschätzung bei den anderen mangels Förderangeboten oder Erfolgsaussichten in keinem Zusammenhang mit dem Forschungsverhalten. Wenn also dem Einwerben von Drittmitteln grundsätzlich eine hohe Bedeutung für den Reputationszuwachs zugesprochen wird, muss sich dies demnach nicht zwangsläufig im konkreten Forschungsverhalten niederschlagen.

H2.2: Die Höhe der eingeschätzten Bedeutung einer Forschungstätigkeit für den Reputationszuwachs in der Scientific Community moderiert den positiven Zusammenhang zwischen dem Einsatz eines Forschungsindikators in hochschulinternen Steuerungs- und Anreizverfahren und dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren in folgender Weise: Die Stärke des Zusammenhangs ist bei einer hoch eingeschätzten Bedeutung für den Reputationszuwachs geringer als bei einer niedrig eingeschätzten Bedeutung für den Reputationszuwachs.

Der in Hypothese H2.2 formulierte Interaktionseffekt kann nur in Modell 1 für die Forschungstätigkeit *Publizieren in Zeitschriften mit Peer-Review-Verfahren* statistisch nachgewiesen werden. Da sich mit Ausnahme von Modell 1 in keinem anderen Modell ein signifikanter Effekt nachweisen lässt – insbesondere nicht in Modell 4 mit der inhaltlich recht ähnlichen Forschungstätigkeit *Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals* – kann Hypothese H2.2 in dieser Form nicht bestätigt werden.

Die Annahme einer starken Orientierung im Forschungsverhalten an den Anforderungen der Scientific Community ist unter dem Blickwinkel der Ergebnisse zu Hypothese H2.1 empirisch zu bestätigen. Nur scheint auch bei einer niedrigen eingeschätzten Relevanz einer Forschungstätigkeit für den Reputationszuwachs in der Scientific Community die externe Anreizwirkung durch indikatorenbasierte Steuerungsversuche zu gering auszufallen, um alternativ motivational wirksam zu werden. Statt auf einen möglichen Reputationsgewinn zu verzichten und stattdessen den Kriterien der Steuerungsversuche gerecht zu werden, werden die begrenzten zeitlichen Ressourcen möglicherweise lieber in andere Forschungstätigkeiten investiert, die für die Scientific Community eine höhere Relevanz besitzen.

Dass sich bei der Forschungstätigkeit *Publizieren in Zeitschriften mit Peer-Review-Verfahren* dennoch ein signifikanter Interaktionseffekt ergibt, könnte damit zusammenhängen, dass eben hier die Gewichtung des Indikators in den Steuerungsverfahren (meist) höher angesetzt wird und damit die Anreizwirkung durch die Steuerungsversuche höher ausfällt. Der Interaktionseffekt stellt sich hinsichtlich dieser Forschungstätigkeit dann wie folgt dar: Während sich bei einer hoch eingeschätzten Bedeutung des Publizierens in Zeitschriften mit Peer-Review-Verfahren für den Reputationszuwachs der hohe Stellenwert dieser Forschungstätigkeit in der eigenen Forschungsarbeit überwiegend über die Orientierung an der Scientific Community erklären lässt, kann bei einer gering eingeschätzten Relevanz der Forschungstätigkeit für den Reputationszuwachs die indikatorenbasierte Steuerung alternativ motivational wirksam werden.

Allerdings müsste sich nach diesem Erklärungsmuster ein ähnlicher Zusammenhang bzw. ein signifikanter Interaktionseffekt auch bei der Forschungstätigkeit *Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals* zeigen, bei welcher der korrespondierende Forschungsindikator ebenfalls meist hoch in den Steuerungsverfahren gewichtet wird. Da dies nicht der Fall ist, könnte es sich bei dem signifikanten Interaktionseffekt in Modell 1 auch schlicht um ein statistisches Artefakt handeln. Aufgrund der wenig konsistenten Ergebnisse in den verschiedenen Modellen zu Hypothese *H2.2* wird diese daher abgelehnt. Der Zusammenhang zwischen den drei betrachteten Variablen lässt sich damit statistisch wohl eher als Addition zweier Haupteffekte (Einsatz Forschungsindikator und Bedeutung für Reputationszuwachs) verstehen, die gemeinsam – mehr oder weniger stark – auf das Forschungsverhalten wirken.

H3: Je höher der eingeschätzte Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn ist, desto höher ist der eingeschätzte Stellenwert der korrespondierenden Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

Hypothese *H3* kann durchgehend für alle hier berechneten Regressionsmodelle bestätigt werden. Unabhängig von der untersuchten Forschungstätigkeit wird also – den Hypothese *H3* zugrundeliegenden theoretischen Annahmen nach – das Streben nach wissenschaftlichem Erkenntnisgewinn als internalisierte Form der Motivation (PSM) wirksam und steht so in einem direkten Zusammenhang mit dem Forschungsverhalten.

Auch wenn ein exakter Vergleich der Effektstärken anhand der unstandardisierten Regressionskoeffizienten nicht möglich ist, lassen die *B*-Werte fast durchgehend einen moderaten bis mittelstarken Zusammenhang mit der abhängigen Variablen erkennen. Aufgrund der durchgehend signifikanten Zusammenhänge und den jeweiligen Effektstärken kann damit die Variable *Eingeschätzter Beitrag für den wissenschaftlichen Erkennt-*

nisgewinn als einflussreichster Prädiktor für den eingeschätzten Stellenwert der verschiedenen Forschungstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren identifiziert werden.

H4.1: Je höher die intrinsische Forschungsmotivation ist, desto höher ist der eingeschätzte Stellenwert einer Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren.

Im Hinblick auf den Zusammenhang zwischen der intrinsischen Forschungsmotivation und dem Forschungsverhalten sind die modellspezifischen Ergebnisse wieder weniger eindeutig. Während die intrinsische Forschungsmotivation in einem signifikant positiven Zusammenhang mit den Forschungstätigkeiten *Publizieren in Zeitschriften mit Peer-Review-Verfahren*, *Betreuung von Promotionen* sowie dem *Einwerben von Drittmitteln der DFG* steht, ergeben sich für die übrigen Forschungstätigkeiten keine signifikanten Effekte. Auffällig an diesem Ergebnis ist, dass mit Ausnahme der Betreuung von Promotionen jene Forschungstätigkeiten positiv mit der intrinsischen Forschungsmotivation zusammenhängen, denen auch eine hohe Bedeutung für den Reputationszuwachs zugesprochen wird (vgl. Kapitel 6.1.4). Einzig für die Forschungstätigkeit *Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals*, welche hinsichtlich der Bedeutung für den Reputationszuwachs ebenfalls hoch eingeschätzt wird, ergibt sich kein signifikanter Zusammenhang. Hier zeigt sich aber zumindest in der bivariaten Analyse eine signifikante positive Korrelation zwischen den beiden Variablen.

Die erfolgreiche Durchführung dieser, mit einem hohen Reputationsgewinn verbundenen Tätigkeiten (Veröffentlichung eines Aufsatzes in einer peer-reviewed Zeitschrift, Einwerben eines DFG-Drittmittelprojektes) ist nun oft auch mit einem höheren zeitlichen Aufwand bzw. höheren Hürden (z. B. (strengere) Gutachterverfahren, höher angesetzte qualitative Standards) verbunden, als dies bei den anderen Tätigkeiten der Fall ist.⁷⁵ So muss ein eingereicherter Artikel zur Veröffentlichung in peer-reviewed Zeitschriften den qualitativen Standards und der Ausrichtung der Zeitschrift entsprechen und gegebenenfalls auf Aufforderung der Gutachter überarbeitet werden. Ebenso ist das Verfassen von Drittmittelanträgen zum Einwerben von DFG-Fördergeldern in der Regel zeit- und arbeitsintensiver als beispielsweise das Einwerben manch privatwirtschaftlicher Gelder. Zudem ist auch die Gefahr höher, dass der investierte Aufwand trotz aller Bemühungen nicht zum Erfolg führt und der ausgearbeitete Artikel oder Projektantrag an den Urteilen

⁷⁵ Dies soll keinesfalls bedeuten, dass z. B. in nicht begutachteten Zeitschriften keine qualitativ hochwertigen Artikel erscheinen. Insgesamt kann aber davon ausgegangen werden, dass die Ausarbeitung eines Artikels zur Einreichung in einer peer-reviewed Zeitschrift mit einem höheren Aufwand verbunden ist als die Publikation eines nicht-begutachteten Artikels.

der Gutachter scheitert. Berücksichtigt man die genannten Charakteristika dieser Forschungstätigkeiten, erscheint es plausibel, dass insbesondere Professorinnen und Professoren mit einer hohen intrinsischen Forschungsmotivation diese erschwerenden Hindernisse in Kauf nehmen und den Forschungstätigkeiten einen entsprechend hohen Stellenwert in ihrer Forschungsarbeit einräumen.

Für die Forschungstätigkeit *Betreuung von Promotionen*, deren eingeschätzter Stellenwert ebenfalls in einem signifikant positiven Zusammenhang mit der intrinsischen Forschungsmotivation steht, treffen obige Ausführungen nicht zu. Allerdings ist auch hier die Annahme plausibel, dass diejenigen Professorinnen und Professoren mit einer höheren intrinsischen Forschungsmotivation der Betreuung von Promotionen als Teil ihrer Forschungsarbeit einen höheren Stellenwert einräumen, als dies weniger intrinsisch motivierte Kolleginnen und Kollegen tun.

Insgesamt lässt sich Hypothese *H4.1* damit mit Einschränkungen bestätigen. Es zeigen sich die theoretisch erwarteten positiven Zusammenhänge mit der abhängigen Variablen. Diese sind allerdings – neben der Betreuung von Promotionen – auf diejenigen Forschungstätigkeiten beschränkt, deren Bedeutung für den Reputationszuwachs hoch eingeschätzt wird und deren Durchführung zudem mit einem entsprechenden Mehraufwand verbunden ist.

H4.2: Der Einsatz eines Forschungsindikators in hochschulinternen Steuerungs- und Anreizverfahren korreliert negativ mit der intrinsischen Forschungsmotivation.

Die auf der *Motivation Crowding Theory* beruhende Verdrängungshypothese *H4.2* kann nur für den Einsatz der Forschungsindikatoren *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* sowie *Anzahl anderer Publikationen* bestätigt werden. Bei diesen ergibt sich jeweils ein signifikanter negativer, allerdings eher schwacher Zusammenhang mit der intrinsischen Forschungsmotivation. Ein Abgleich mit den Ergebnissen der bivariaten Analysen (vgl. Kapitel 6.2.1) zeigt, dass dieser negative Zusammenhang insbesondere auf die Substichprobe der Psychologie zurückzuführen ist. Nach den Annahmen der *Motivation Crowding Theory* kann ein Verdrängungseffekt der intrinsischen Motivation immer dann auftreten, wenn der Versuch der externen Einflussnahme als kontrollierend wahrgenommen und damit das Kompetenz- und Autonomieerleben eingeschränkt wird (vgl. Kapitel 4.1.2). Nicht begutachtete Publikationen und andere Publikationsformen als Zeitschriftenartikel sind nun in der Publikationskultur der Psychologie von geringerer Relevanz und deren Bedeutung für den Reputationszuwachs und Beitrag für den wissenschaftlichen Erkenntnisgewinn werden eher niedrig eingeschätzt (vgl. Kapitel 6.1.4 und 6.1.5). Angenommen werden kann daher, dass der externe Versuch, den Publikationsoutput bei diesen Publikationsformen mittels indikatorenbasierter Steuerung zu erhöhen, bei Professorinnen und Professoren der Psychologie aufgrund der geringen

Relevanz dieser Publikationsformate in ihrem Forschungsalltag vermehrt als kontrollierend im obigen Sinne wahrgenommen wird. Vor diesem Hintergrund wäre es daher nachvollziehbar, dass sich bei den genannten Forschungstätigkeiten speziell in der Psychologie ein negativer Effekt auf die intrinsische Forschungsmotivation durch den Einsatz indikatorenbasierter Steuerungsverfahren zeigt.

Gegen diese Annahme spricht allerdings, dass sich in den Wirtschaftswissenschaften, bei denen die Publikationskultur und die Einschätzungen der Befragten zur Bedeutung der genannten Publikationsformate für den Reputationszuwachs und deren Beitrag für den wissenschaftlichen Erkenntnisgewinn ähnlich gelagert sind, in den bivariaten Analysen keine signifikante negative Korrelation zeigt. Ob hier weitere fachspezifische Unterschiede (z. B. höhere Akzeptanz von Steuerungsversuchen in den Wirtschaftswissenschaften) zu berücksichtigen sind, kann an dieser Stelle nicht überprüft werden. Generell stellen die Einstellungsmuster der Befragten und die Wahrnehmung der Steuerungsinstrumente einen blinden Fleck in den vorliegenden Analysen dar. Um mögliche Zusammenhänge zwischen indikatorenbasierten Steuerungsversuchen und der intrinsischen Forschungsmotivation näher analysieren zu können, wäre es daher zukünftig hilfreich, die Einstellungsmuster und Wahrnehmung der Befragten im Hinblick auf den Einsatz und die Ausgestaltung der Steuerungsverfahren zu erheben (vgl. Kapitel 7.2). Diese weiterführenden Analysen könnten möglicherweise einen Beitrag zur Erklärung der auf den ersten Blick wenig konsistenten Ergebnisse zu Hypothese *H4.2* leisten.

In Bezug auf die vorliegenden Ergebnisse bleibt daher festzuhalten, dass sich durchaus Hinweise auf einen möglichen Verdrängungseffekt der intrinsischen Forschungsmotivation durch den Einsatz indikatorenbasierter Steuerungsverfahren zeigen. Allerdings erlauben die teils inkonsistenten Ergebnisse zu Hypothese *H4.2* keine eindeutige Annahme oder Ablehnung der Forschungshypothese, so dass hier auf weiteren Forschungsbedarf verwiesen werden muss.

Weitere Erkenntnisse aus den multivariaten Analysen

Nach der Darstellung der Ergebnisse zu den Hypothesentests werden im Folgenden nun knapp weitere Ergebnisse aus den multivariaten Analysen erläutert, bevor die zentralen Erkenntnisse im Anschluss nochmals gebündelt dargestellt werden.

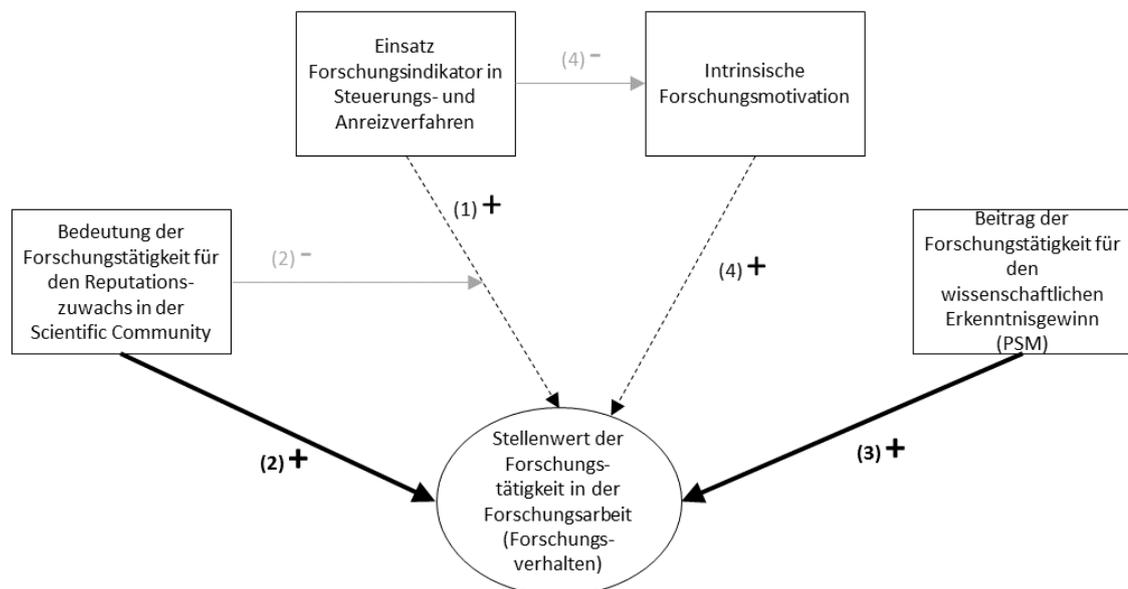
Auch in den multivariaten Analysen können fachspezifische Unterschiede beobachtet werden, die größtenteils die Erkenntnisse aus den fächervergleichenden deskriptiven und bivariaten Analysen widerspiegeln. Beispielsweise zeigt sich, dass der Stellenwert des *Publizierens in Zeitschriften ohne Peer-Review-Verfahren* und des *Verfassens anderer Publikationen* in den Erziehungswissenschaften signifikant höher eingeschätzt wird als in den Wirtschaftswissenschaften.

Zudem ergibt sich bei den Forschungstätigkeiten *Publizieren in Zeitschriften mit Peer-Review-Verfahren* sowie *Publizieren in Zeitschriften mit hohem Impact Factor / in high ranked journals* ein signifikanter negativer Zusammenhang zwischen dem Alter der Befragten und dem eingeschätzten Stellenwert dieser Forschungstätigkeiten. Dies deckt sich mit der These unterschiedlicher altersabhängiger wissenschaftlicher Sozialisationsprozesse mit Auswirkungen auf das jeweilige Publikationsverhalten (vgl. Kapitel 6.2.2.1). Die in den bivariaten Analysen teilweise noch signifikanten Zusammenhänge zwischen der Beschäftigungsdauer bzw. dem Geschlecht und dem Forschungsverhalten (vgl. Kapitel 6.2.1) zeigen sich im multivariaten Fall dagegen nicht mehr. Diese Korrelationen sind damit auf andere Variablen wie die Fachzugehörigkeit (für Geschlecht) zurückzuführen.

6.3.2 Zusammenfassende Ergebnisdarstellung

Im Folgenden werden die Erkenntnisse aus den durchgeführten Analysen nochmals zusammengefasst. Eine grafische Darstellung der zentralen Ergebnisse in abgewandelter Form des bereits aus Kapitel 4.2 bekannten theoretischen Zusammenhangsmodells findet sich in Abbildung 11.

Abbildung 11: *Theoretisches Modell - Ergebnisdarstellung*



Anmerkung. Die fett formatierten Pfeile zeigen Zusammenhänge an, die im Rahmen dieser Studie empirisch bestätigt werden konnten. Die gestrichelten Pfeile beziehen sich auf Zusammenhänge, die nur mit Einschränkungen (je nach Modell bzw. Forschungstätigkeit) bestätigt werden konnten. Die grau eingefärbten Pfeile zeigen nicht-signifikante Zusammenhänge an, bei denen die zugrundeliegenden Forschungshypothesen (vorläufig) abzulehnen waren.

Quelle: eigene Darstellung

Der Einsatz von Forschungsindikatoren in indikatorenbasierten Steuerungs- und Anreizverfahren steht nur in einem schwachen Zusammenhang mit dem Forschungsverhalten der sozialwissenschaftlichen Professorinnen und Professoren. Signifikante schwache positive Zusammenhänge lassen sich bei den jeweiligen Modellen nur für den Einsatz von zwei Forschungsindikatoren (*Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren* sowie *Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals*) empirisch nachweisen, welche beide in den Verfahren auch besonders stark gewichtet werden.

Von deutlich höherer Bedeutung und für das Forschungsverhalten handlungsleitend sind vielmehr der mit den jeweiligen Forschungstätigkeiten verbundene Reputationszuwachs in der Scientific Community sowie der Beitrag für den wissenschaftlichen Erkenntnisgewinn. Motivationstheoretisch gesprochen sind hier also zwei (stark) internalisierte Formen der Motivation (vgl. Kapitel 4.2) besonders prägend für das Forschungsverhalten. Vor allem die durchgehend signifikanten positiven Zusammenhänge zwischen dem eingeschätzten Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn und dem Stellenwert in der eigenen Forschungsarbeit deuten darauf hin, dass sozialwissenschaftliche Professorinnen und Professoren auch vor dem Hintergrund eines zunehmenden Ökonomisierungsdrucks durch die Reformbemühungen des New Public Managements immer noch stark durch das (innere) Streben nach wissenschaftlichem Erkenntnisgewinn motiviert werden. Die professionstheoretische Annahme der ‚Zentralwertorientierung‘ (vgl. Kapitel 3.2) – im Falle der akademischen Profession in Form eben dieses wissenschaftlichen Erkenntnisgewinns (vgl. Kapitel 4.2) – kann damit mit den vorliegenden Analysen bestätigt werden.

Auch die der akademischen Profession nachgesagte hohe intrinsische Motivation konnte im Rahmen dieser Arbeit nachgewiesen werden (vgl. Kapitel 6.1.7). Dabei zeigte sich auch, dass eine hohe intrinsische Forschungsmotivation mit einem höheren Stellenwert von solchen Forschungstätigkeiten einhergeht, welche (oft) mit höherem Arbeitsaufwand und einer größeren Gefahr des Scheiterns verbunden sind (Publizieren in Zeitschriften mit Peer-Review-Verfahren, Einwerben von DFG-Drittmitteln).

Neben der starken handlungsleitenden Orientierung an den Kriterien der Scientific Community und dem Beitrag der Forschungstätigkeiten für den wissenschaftlichen Erkenntnisgewinn ist ein weiterer Grund für die eher schwache Steuerungswirkung indikatorenbasierter Steuerungs- und Anreizverfahren vermutlich auch in der eher generell ablehnenden Haltung sozialwissenschaftlicher Professorinnen und Professoren gegenüber diesen zu finden. Die Einstellung gegenüber den Steuerungsversuchen bzw. den damit verbundenen Verfahren wurde zwar nicht direkt erhoben, die Vielzahl an kritischen Anmerkungen im Fragebogen und Rückmeldungen per E-Mail, die den Steuerungsverfahren zudem teils eine generelle Wirkungslosigkeit attestieren (vgl. Kapitel 6.1.8), legen

aber einen entsprechenden Schluss nahe. Es ist daher anzunehmen, dass aufgrund dieser oftmals kritischen Einstellung bei einzelnen Professorinnen und Professoren auch eine bewusste Blockade- oder Verweigerungshaltung vorherrscht, die die Steuerungsversuche ins Leere laufen lässt.

Hinsichtlich eines möglichen negativen Effektes hochschulinterner Steuerungsversuche in Form des Verdrängungseffektes der intrinsischen Motivation erlauben die Ergebnisse wie bereits angemerkt keine eindeutige Aussage. Erste Hinweise auf einen solchen Effekt, die sich in den vorliegenden Analysen andeuten, müssen daher erst in ergänzenden Forschungsarbeiten empirisch validiert werden.

Diese und weitere Folgerungen für die Forschung sowie die Hochschulpraxis werden nun noch abschließend diskutiert. Davor folgt zunächst eine kurze Reflexion im Hinblick auf methodische Aspekte der Arbeit.

7. Abschließende Diskussion

7.1 Methodische Reflexion

Im Folgenden werden einige methodische Aspekte der vorliegenden Studie näher beleuchtet und mögliche mit diesen verbundene Probleme sowie Lösungsansätze diskutiert.

Zunächst ist auf die teilweise vorhandene Verteilungsproblematik durch stark ungleich besetzte Gruppengrößen (insbesondere bei der unabhängigen Variablen *Einsatz Indikator* in den Drittmittelmodellen) bzw. durch schiefe Verteilungen bei den quasi-metrischen Variablen hinzuweisen. Nicht auszuschließen ist, dass die teils stark schiefen Verteilungen in Form von Deckeneffekten (Häufung der Antworten zu einer Variablen im oberen Teil der Skala; vgl. Urban und Mayerl 2011, S. 319) die Varianzaufklärung negativ beeinflusst haben und so einige in der Grundgesamtheit signifikante Zusammenhänge in der Stichprobe unentdeckt bleiben mussten. Diese Problematik könnte bei ähnlichen zukünftigen Erhebungen durch die Verwendung breiterer Skalen entschärft werden (vgl. ebd.). Vor allem für die Variable *Intrinsische Forschungsmotivation*, die in der vorliegenden Studie eine (mit den theoretischen Annahmen zwar konsistente) starke Linksschiefe aufweist, würde die Verwendung einer breiteren Skala möglicherweise eine differenziertere Erfassung des Grades der intrinsischen Motivation erlauben.

Hinsichtlich der Fragenformulierungen und dem Antwortverhalten sind ebenfalls einige mögliche Einschränkungen in der Aussagekraft der Ergebnisse zu diskutieren. Bereits erwähnt wurde das Problem der Mehrdimensionalität beim Item *Andere Publikationen (Monografien, Beiträge in Sammelbänden etc.)* (vgl. Fußnote 40 in Kapitel 6.1.2). Dieser mehrdimensionalen Formulierung wäre eine jeweils separate Abfrage für die einzelnen Publikationsformen vorzuziehen gewesen, um Verzerrungen im Antwortverhalten zu vermeiden und eine eindeutige Interpretation der Ergebnisse zu diesem Item zu ermöglichen.

Des Weiteren können auch für diese Arbeit Verzerrungen aufgrund des in der sozialwissenschaftlichen Forschung häufig anzutreffenden Problems der sozialen Erwünschtheit (vgl. Diekmann 2004, S. 382–386) nicht ausgeschlossen werden. Insbesondere betrifft dies die Abfrage der intrinsischen Forschungsmotivation und die Einschätzungen zum Stellenwert der verschiedenen Forschungstätigkeiten in der Forschungsarbeit der Professorinnen und Professoren. So ist es denkbar, dass die der akademischen Profession generell zugesprochene hohe intrinsische Motivation (vgl. Kapitel 3.2) dazu führt, dass sich einzelne Befragte selbst als hoch motiviert einstufen, ohne dass dies der tatsächlichen Motivationslage entspricht. Ebenso könnten sich die Befragten ‚verpflichtet‘ fühlen, sich entsprechend ihrer Profession durch ein hohes Forschungsengagement auszuzeichnen und in der Folge den Stellenwert der einzelnen Forschungstätigkeiten in

ihrer Forschungsarbeit jeweils besonders hoch einstufen. Inwieweit solche Verzerrungen aufgrund sozialer Erwünschtheit hier von Relevanz für die statistischen Analysen sind, kann mit den vorliegenden Daten allerdings nicht überprüft werden.

Auch Fragereiheneffekte (vgl. Diekmann 2004, S. 398) können einen Einfluss auf das Antwortverhalten der Professorinnen und Professoren ausgeübt haben. Möglich ist beispielsweise, dass die Positionierung der Abfrage der eingeschätzten Bedeutung einer Forschungstätigkeit für den Reputationszuwachs vor der Abfrage des eingeschätzten Beitrags einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn dazu führt, dass Forschungstätigkeiten, denen eine hohe Bedeutung für den Reputationszuwachs zugesprochen wird, auch ein hoher Beitrag für den Erkenntnisgewinn attestiert wird. Solche Fragereiheneffekte könnten bei Folgeuntersuchungen durch Variationen der Fragenpositionierungen diagnostiziert und die Reihenfolge der Fragen gegebenenfalls angepasst werden.

Im Zuge der methodischen Reflexion dieser Arbeit soll auch nochmals kurz das bereits angesprochene Problem der Selbstselektion (vgl. Kapitel 6.1.8) erwähnt werden. In Anbetracht der teils sehr kritischen Haltung von sozialwissenschaftlichen Professorinnen und Professoren gegenüber indikatorenbasierten Steuerungsversuchen in der Wissenschaft sind zwei Szenarien hinsichtlich möglicher Stichprobenverzerrungen denkbar: Zum einen kann die ablehnende Haltung dazu führen, dass die stärksten Kritiker die Teilnahme an der Befragung von vornherein verweigern und damit in der Stichprobe unterrepräsentiert sind. Zum anderen kann eben diese ablehnende Haltung gerade als Antrieb dienen, an der Befragung teilzunehmen, um ‚seinem Ärger Luft zu machen‘. In letzterem Fall wären die Kritiker in der Stichprobe eventuell sogar überrepräsentiert. Auch diese möglichen Stichprobenverzerrungen aufgrund von Selbstselektion können mit den in dieser Studie erhobenen Daten nicht überprüft werden. In nachfolgenden thematisch ähnlich gelagerten Erhebungen könnten durch eine Abfrage der Einstellungen der Befragten zu Aspekten der indikatorenbasierten Steuerung solche Einstellungsmuster erfasst werden. Dies würde eine statistische Überprüfung des Verhältnisses von Kritikern, eher neutral eingestellten Befragten und Befürwortern ermöglichen. Bei einer solchen Vorgehensweise sollten aber zwingend die oben genannten Fragereiheneffekte berücksichtigt werden, um wiederum Verzerrungen aufgrund bzw. bei der Erhebung der Einstellungsmuster zu vermeiden.

Ein mehr grundlegendes Problem der vorliegenden Studie betrifft die – ebenfalls bereits an anderer Stelle angesprochene (vgl. Kapitel 6.1.2) – mögliche hierarchische Datenstruktur der Stichprobe. Diese hierarchische Datenstruktur ergibt sich daraus, dass jeweils Gruppen von Befragten derselben Organisationseinheit (Fakultät oder Universität) angehören und sich damit im Hinblick auf bestimmte Merkmale (z. B. Forschungs-

schwerpunkt, Erfahrungen mit Steuerungsinstrumenten etc.) ähnlicher sind als die übrigen Elemente der Stichprobe. Das Paradebeispiel solcher hierarchisch strukturierter Daten sind Schülerinnen und Schüler (Individualebene), die in verschiedenen Schulklassen (Gruppenebene) mit unterschiedlichen Lernumgebungen konfrontiert sind (jeweils andere Mitschülerinnen und Mitschüler, anderes Lehrpersonal, andere Räumlichkeiten etc.). Mögliche Effekte und Interaktionen zwischen den Gruppenmerkmalen und Individualmerkmalen der Befragten können vorzugsweise über Mehrebenenanalysen kontrolliert werden (vgl. Nezlek et al. 2006). Die Nichtbeachtung der Mehrebenenstruktur der Daten kann zu Verzerrungen bei der Schätzung von Effektstärken und zu fehlerhaften Signifikanztests führen (ebd., S. 213). Aus Gründen der Anonymität war in dieser Studie die Erfassung der Universität, an der die Befragten beschäftigt sind, allerdings nicht möglich. Auf die Durchführung einer Mehrebenenanalyse musste daher verzichtet werden. Da aber Professorinnen und Professoren derselben Universität, z. B. im Vergleich zu Schulklassen, deutlich weniger klar definierte Gruppen mit einheitlichen Gruppenmerkmalen bilden⁷⁶, dürfte mit eher geringen Verzerrungen durch Gruppenmerkmale und deren Interaktionen mit Individualmerkmalen zu rechnen sein. Gänzlich auszuschließen sind diese aber nicht. Eine mehrebenenanalytische Überprüfung der Effekte anhand von jeweils ausreichend großen Fachbereichen bzw. ausreichend großen Universitäten, die aufgrund der jeweiligen Stichprobengröße keine Rückschlüsse auf einzelne Professuren zulassen, könnte hierzu möglicherweise hilfreiche Erkenntnisse liefern.

Abschließend soll nochmals daran erinnert werden, dass das hier vorliegende querschnittliche Studiendesign streng genommen keine Aussagen über die Kausalität der betrachteten Zusammenhänge erlaubt. Die Annahmen über die Wirkrichtungen beruhen rein auf theoretischen Überlegungen und können ohne ein längsschnittliches Studiendesign statistisch nicht überprüft werden. Wie bereits ausgeführt (vgl. Kapitel 5.1), wurden in dieser Studie teils dennoch implizite oder explizite Kausalitätsaussagen formuliert, sofern die zugrundeliegenden theoretischen Annahmen entsprechende Wirkrichtungen nahelegten. Die zu präferierende Alternative in Form einer Längsschnittanalyse war aufgrund der zum Stand der Erhebung bereits erfolgten Einführung indikatorenbasierter Steuerungs- und Anreizverfahren nicht zu realisieren.

In diesem Teilkapitel wurden methodische Aspekte dieser Arbeit reflektiert und methodische Schwächen sowie diesbezügliche Handlungsempfehlungen für die weitere Forschung aufgezeigt. Im nun folgenden abschließenden Kapitel werden die Ergebnisse

⁷⁶ Allein der organisationale Aufbau von Universitäten nach dem Lehrstuhl-Prinzip zeigt die relative Unabhängigkeit der einzelnen Professur. Natürlich ‚teilen‘ sich die Professorinnen und Professoren eine gemeinschaftliche Umwelt in Form der jeweiligen Universität. Sie sind aber mitnichten denselben strikten gruppenbildenden Grenzziehungen unterworfen, wie das in Schulklassen der Fall ist. Die primäre Orientierung an der Scientific Community anstatt an der Organisation Universität (vgl. Kapitel 3.2) dürfte den Einfluss möglicher universitärer Gruppenmerkmale zudem weiter einschränken.

der Studie im Hinblick auf generelle Folgerungen für die (Hochschul-)Forschung und die Hochschulpraxis diskutiert.

7.2 Folgerungen für die Forschung und Hochschulpraxis

Zunächst muss darauf hingewiesen werden, dass es sich bei den vorliegenden Analysen nur um eine Momentaufnahme zur Situation in der sozialwissenschaftlichen Hochschullandschaft handeln kann. Mit den indikatorenbasierten Steuerungs- und Anreizverfahren wurde ein Untersuchungsgegenstand in den Blick genommen, der in der Art seines Einsatzes (Häufigkeit, Zielsetzung) und seiner Ausgestaltung (verwendete Indikatoren, Gewichtung) einem stetigen Wandel unterworfen ist. Auch Veränderungen in den politischen Rahmenbedingungen (Stichwort Governance-Wandel) sowie Wandlungsprozesse in den Forschungskulturen der einzelnen Fächer selbst (z. B. vermehrte Impact-Factor-Orientierung) können dazu führen, dass sich der Stellenwert universitärer Steuerungsversuche in den Sozialwissenschaften zukünftig in die eine oder die andere Richtung verschiebt. Bei der Interpretation der Ergebnisse sind daher der Erhebungszeitpunkt der Studie und die damit verbundene momentane Situation in der sozialwissenschaftlichen Hochschullandschaft zu berücksichtigen.

In den Analysen wurde zudem deutlich, dass sich innerhalb der Sozialwissenschaften zwischen den einzelnen Fächern teils große Unterschiede zeigen. Dies betrifft neben den in den Steuerungs- und Anreizverfahren eingesetzten Forschungsindikatoren auch die Einschätzungen hinsichtlich der Bedeutung der verschiedenen Forschungstätigkeiten für den Reputationszuwachs sowie deren Relevanz für den wissenschaftlichen Erkenntnisgewinn. Entsprechend nehmen die verschiedenen Forschungstätigkeiten in den Forschungskulturen der einzelnen Fächer einen jeweils unterschiedlich hohen Stellenwert ein. Vor dem Hintergrund dieser fachspezifischen Differenzen wäre es sicherlich zielführend, für eine umfassende Detailanalyse die fächervergleichende Perspektive zu verlassen und einzelne Fächer isoliert zu betrachten. Hierfür würde sich neben einem quantitativen auch ein qualitativer Forschungsansatz als methodischer Zugang anbieten. In diesem Zusammenhang wären neben den Sozialwissenschaften natürlich auch gänzlich andere Fächergruppen wie z. B. die Naturwissenschaften von Interesse. Dass sich je nach analysiertem Fach andere Ergebnisse in Bezug auf die Steuerungswirkung indikatorenbasierter Verfahren ergeben können, zeigte sich beispielsweise bereits in der Studie zu den Effekten der leistungsorientierten Mittelverteilung im Fach Medizin von Krempkow et al. (2013) (vgl. Kapitel 2.4). Für die Interpretation und für mögliche Schlussfolgerungen aus den Ergebnissen der vorliegenden Untersuchung ist daher zu berücksichtigen, dass diese auf die Sozialwissenschaften begrenzt bleiben müssen und nicht auf andere Fächer übertragbar sind.

Im Hinblick auf den Einsatz indikatorenbasierter Steuerungs- und Anreizverfahren könnte eine detaillierte isolierte Betrachtung einzelner Verfahren zu einem besseren Verständnis der Wirkungsweise beitragen. In dieser Arbeit zeigten sich bereits im Rahmen der bivariaten Analysen bei einzelnen Verfahren teils signifikante Zusammenhänge zwischen dem Einsatz eines Forschungsindikators in diesen und dem eingeschätzten Stellenwert der korrespondierenden Forschungstätigkeit, welche sich bei der verfahrensübergreifenden Analyse nicht bestätigen ließen. So korreliert beispielsweise der Einsatz des Forschungsindikators *Anzahl betreuter Promotionen* in den Verfahren *Leistungszulagen W-Besoldung* sowie *Bleibe-/Berufungsverhandlungen* positiv mit dem eingeschätzten Stellenwert der Betreuung von Promotionen in der Forschungsarbeit der Professorinnen und Professoren (vgl. Kapitel 6.2.1). In der verfahrensübergreifenden Betrachtung mit der unabhängigen Variablen *Einsatz Forschungsindikator „Anzahl betreuter Promotionen“ in Steuerungs- und Anreizverfahren insgesamt* ließ sich dagegen weder im bivariaten noch im multivariaten Fall ein signifikanter Zusammenhang nachweisen. Um zu überprüfen, ob sich diese verfahrensspezifischen Effekte auch unter Kontrolle von Drittvariablen reproduzieren lassen, sollten entsprechende multivariate Analysen durchgeführt werden. Bei der vorliegenden Studie war dies aufgrund der Stichprobengröße und der damit für die verfahrensspezifischen Analysen überwiegend zu geringen Fallzahlen nicht möglich.

Im Zusammenhang mit einer vertiefenden Einzelbetrachtung verschiedener Steuerungs- und Anreizverfahren sollte auch deren Ausgestaltung noch detaillierter in den Blick genommen werden. In dieser Arbeit ergaben sich Hinweise, dass die Gewichtung der Forschungsindikatoren in den Verfahren ein entscheidendes Kriterium darstellen kann, ob eine Steuerungswirkung zu beobachten ist. Hier könnten die Analyse der genauen Gewichtungsformeln sowie die Berücksichtigung weiterer Merkmale wie die Höhe der über die Verfahren vergebenen Finanzmittel einen zusätzlichen Erklärungsbeitrag zur Wirkungsweise der Steuerungs- und Anreizverfahren leisten.

Wie im Zuge der Erarbeitung des theoretischen Modells (vgl. Kapitel 3.3 und Kapitel 4.2) bereits ausgeführt wurde, kann sich eine Anreizwirkung zum Entsprechen der Steuerungsversuche auch aus einem zunehmenden Legitimationsdruck (Legitimation über Leistung bzw. Leistungstransparenz) auf die akademische Profession ergeben. Dieser Legitimationsdruck resultiert den theoretischen Annahmen zufolge aus einem gesellschaftlichen Vertrauensverlust in die Arbeit der akademischen Profession. Als zusätzliche erklärende Variable könnte daher erhoben werden, ob diejenigen Professorinnen und Professoren, die einen solchen Vertrauensverlust wahrnehmen bzw. sich einem erhöhten Legitimationsdruck ausgesetzt sehen, auch eher den Anforderungen hochschulinterner Steuerungsversuche entsprechen.

Im Hinblick auf die Analyse der Wirkungsweise indikatorenbasierter Steuerungs- und Anreizverfahren ist zudem zu berücksichtigen, dass in dieser Arbeit nur der Bereich Forschung in den Blick genommen und der Bereich der Lehre als zweiter großer Tätigkeitschwerpunkt der akademischen Profession ausgeklammert wurde. Da die Lehre ebenso wie die Forschung Gegenstand hochschulinterner indikatorenbasierter Steuerungsversuche sein kann, wäre im Sinne einer ganzheitlichen Betrachtungsweise eine Ausweitung der Analyse auf beide Bereiche wünschenswert. Insbesondere hinsichtlich des Spannungsverhältnisses zwischen Forschung und Lehre, welches sich aus den begrenzten zeitlichen Ressourcen der Professorinnen und Professoren ergibt (vgl. Jacob und Teichler 2011; Wissenschaftsrat 2008a), könnte der Einbezug der Lehre bei der Untersuchung der Wirkungsweise indikatorenbasierter Steuerungsversuche zusätzliche wertvolle Erkenntnisse liefern.

Einen weiteren Aspekt, welcher mit dem Themenfeld der indikatorenbasierten Steuerung verbunden ist, betrifft die ausführliche Analyse möglicher nichtintendierter Effekte auf das Forschungsverhalten der Professorinnen und Professoren (vgl. hierzu auch die kritische Diskussion der Forschungsindikatoren in Kapitel 2.3.2 dieser Arbeit). Abgesehen von der Überprüfung eines möglichen Verdrängungseffektes der intrinsischen Forschungsmotivation im Rahmen der Hypothesentests (vgl. Kapitel 4.2 und 6.2.2) wurde die Analyse anderer möglicher nichtintendierter Effekte in dieser Arbeit nur angeschnitten. So wurden im Rahmen der bivariaten Analysen beispielsweise Zusammenhänge zwischen dem Einsatz von Forschungsindikatoren und dem Stellenwert des Publizierens in Ko- bzw. Mehrautorenschaft (als Möglichkeit den Publikationsoutput zu erhöhen) analysiert. Die Ergebnisse hierzu sind aber aufgrund der fehlenden Berücksichtigung von Drittvariablen nur von geringer Aussagekraft (vgl. Kapitel 6.2.1). Weitere Forschung könnte sich also vertiefend auf die Analyse nichtintendierter Effekte wie den Auswirkungen auf das Publikationsverhalten („Salamitaktik“, Zitationskartelle), einer möglichen Tendenz zur Mainstreamforschung oder der Vernachlässigung der Lehre beziehen.⁷⁷

Wie bereits angemerkt, wurden in dieser Studie mögliche negative Zusammenhänge zwischen indikatorenbasierten Steuerungsversuchen und der intrinsischen Forschungsmotivation analysiert. Allerdings waren die Ergebnisse der Analysen auch hier wenig eindeutig, so dass die – in anderen Forschungskontexten empirisch oft bestätigte (vgl. Kapitel 4.1.2) – Verdrängungshypothese (vorläufig) abzulehnen war. Neben den erwähnten methodischen Problemen bei der Erfassung der Forschungsmotivation (starke Linksschiefe) könnte ein weiterer Grund für die wenig konsistenten Ergebnisse in der fehlenden Berücksichtigung der Einstellungsmuster bzw. Wahrnehmung der Befragten

⁷⁷ Vgl. zur Untersuchung der nichtintendierten Effekte auch das BMBF-Projekt „Nichtintendierte Effekte der Leistungsbewertung in der Wissenschaft“: <https://www.uni-bamberg.de/soztheorie/forschung/projekte/bmbf-projekt/> (28.02.2017).

gegenüber den indikatorenbasierten Steuerungsversuchen zu finden sein. Wie bereits ausgeführt, kommt es entsprechend den Annahmen der *Motivation Crowding Theory* nur dann zu einem Verdrängungseffekt der intrinsischen Motivation, wenn die externe Einflussnahme als kontrollierend empfunden wird. In dieser Arbeit wurde theoretisch angenommen, dass aufgrund der eher kritischen Haltung der (sozialwissenschaftlichen) akademischen Profession gegenüber externen Steuerungsversuchen diese primär als kontrollierend wahrgenommen werden (vgl. Kapitel 4.2). Vermutlich folgt die Wahrnehmung der Professorinnen und Professoren aber keinem so einheitlichen Muster. Vielmehr sind generationenspezifische und fachspezifische Unterschiede hinsichtlich der Einstellungen gegenüber indikatorenbasierten Steuerungsversuchen denkbar (vgl. auch die Unterscheidung zwischen Ökonomen und Traditionalisten in der Studie von Minssen et al. 2003). In Folgeuntersuchungen sollte daher überprüft werden, ob sich unter Kontrolle der Wahrnehmung bzw. Einstellungsmuster der Professorinnen und Professoren Ergebnisse zeigen, die mit den Annahmen der *Motivation Crowding Theory* im Einklang stehen.

Eine weitere theoretische Annahme dieser Arbeit, die ebenfalls nicht empirisch überprüft wurde, betrifft das Konzept der *Public Service Motivation*. Angenommen wurde hierbei, dass das Streben nach wissenschaftlichem Erkenntnisgewinn als ein kollektiver Orientierungspunkt der akademischen Profession aufgefasst werden kann (Mesoebene), welches als stark internalisierte Form der Motivation (*Public Service Motivation*) auf der Mikroebene des Handelnden wirksam wird. Ein hoher eingeschätzter Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn sollte entsprechend dieser Annahme mit einer hohen *Public Service Motivation* einhergehen. Diese sollte wiederum positiv mit dem Engagement hinsichtlich der Forschungstätigkeit assoziiert sein. Empirisch überprüft wurde in dieser Studie aber lediglich, inwieweit sich ein signifikanter Zusammenhang zwischen dem eingeschätzten Beitrag einer Forschungstätigkeit für den wissenschaftlichen Erkenntnisgewinn und dem Stellenwert dieser Forschungstätigkeit in der Forschungsarbeit der Professorinnen und Professoren ergibt. Die *Public Service Motivation* wurde dabei nicht direkt erhoben. Von Interesse wäre es daher zu überprüfen, ob sich die theoretische Annahme, dass die individuelle *Public Service Motivation* positiv mit dem eingeschätzten Beitrag für den wissenschaftlichen Erkenntnisgewinn korreliert ist, bestätigen lässt. Hierzu könnte auf bereits entwickelte Skalen zur Erfassung der *Public Service Motivation* zurückgegriffen werden (vgl. Perry et al. 2010), die noch an die Untersuchungspopulation der akademischen Profession anzupassen wären. Damit könnte auch ein Beitrag zur Konkretisierung des immer noch etwas vagen Theoriekonzeptes der *Public Service Motivation* geleistet und dessen Anwendbarkeit für die Forschungspraxis erhöht werden.

Unabhängig von den theoretischen Annahmen zur *Public Service Motivation* wurde durch die vorliegenden Analysen aber deutlich, welche handlungsleitende Relevanz dem wissenschaftlichen Erkenntnisgewinn für das Forschungsverhalten der Professorinnen und Professoren zukommt. Daneben bestätigte sich in den Analysen der hohe Stellenwert der Scientific Community (Reputation) als weiterer zentraler Orientierungspunkt wissenschaftlichen Handelns. Trotz eines gestiegenen Konkurrenzdrucks in der Hochschullandschaft und sichtbaren Verschiebungen der anderen ‚Governance-Regler‘ (hierarchische Selbststeuerung, staatliche Regulierung und externe Steuerung; vgl. Kapitel 2.1.2) bewahrt sich die akademische Profession damit ihre starke Positionierung und die damit verbundene Handlungsautonomie. Sozialwissenschaftliche Professorinnen und Professoren werden in ihrem Forschungsverhalten also in Übereinstimmung mit den professionstheoretischen Überlegungen aus Kapitel 3.2 wesentlich von der Orientierung an den Anforderungen ihrer Scientific Community und der Relevanz ihrer Forschungstätigkeiten für den wissenschaftlichen Erkenntnisgewinn geleitet. Ein wenig idealisierend ließe sich damit schlussfolgern, dass sich das Forschungsverhalten der akademischen Profession in diesem Sinne (überwiegend) an den an sie gerichteten gesellschaftlichen Erwartungen orientiert – die Professorinnen und Professoren forschen, um den wissenschaftlichen Erkenntnisgewinn voranzutreiben und orientieren sich dabei an den eigenen professionellen Standards und Anforderungen.

Bei dem Versuch der hochschulinternen Steuerung durch indikatorenbasierte Steuerungs- und Anreizverfahren wäre entsprechend darauf zu achten, nicht durch eine ‚falsche‘ Ausgestaltung der Verfahren Fehlanreize zu setzen, die konträr zur ursprünglichen Handlungsorientierung der akademischen Profession laufen. Dies könnte neben einer mangelnden Akzeptanz bzw. Abwehrhaltung gegenüber den Verfahren gegebenenfalls auch negative Auswirkungen auf die Motivation zur Folge haben. Erste Hinweise für einen solchen negativen Effekt auf die intrinsische Motivation ergaben sich bei den multivariaten Analysen für den Einsatz der Forschungsindikatoren *Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren* sowie *Anzahl anderer Publikationen* im Fach Psychologie (vgl. die Ergebnisdiskussion zu Hypothese *H4.2* in Kapitel 6.3.1). Die Ausgestaltung der Steuerungs- und Anreizverfahren sollte sich also an den fachspezifischen Forschungs- und Publikationskulturen orientieren. Beispielsweise geht die Nutzung von Zeitschriftenrankings in indikatorenbasierten Steuerungsverfahren bei wirtschaftswissenschaftlichen Professorinnen und Professoren vermutlich mit einer höheren Akzeptanz als bei Professorinnen und Professoren der Soziologie einher. Während sich bei den einen möglicherweise eine positive Anreizwirkung ergibt, wirkt sich bei den anderen die Anwendung dieses – in Bezug auf die spezifische Fachkultur – ‚unpassenden‘ Leistungskriteriums dagegen gegebenenfalls kontraproduktiv aus. Da der im Rahmen dieser Studie vermutete Interaktionseffekt (Bedeutung für Reputationszuwachs moderiert die

Steuerungswirkung) überwiegend nicht bestätigt werden konnte (vgl. Ergebnisdiskussion zu Hypothese *H2.2* in Kapitel 6.3.1), könnte sich durch eine ausreichende Berücksichtigung der Fachkulturen der jeweiligen Scientific Communities so im besten Fall, bei entsprechender Gewichtung der Indikatoren, ein zusätzlicher positiver Anreiz und damit ein additiver positiver Effekt auf das Forschungsengagement zeigen. Die Gefahr nicht-intendierter negativer Effekte könnte dagegen aufgrund einer höheren Akzeptanz der Professorinnen und Professoren gegenüber den Verfahren verringert werden.

Unter Berücksichtigung der in den vorliegenden Analysen nur schwachen nachweisbaren Steuerungswirkung indikatorenbasierter Verfahren, scheint der Nutzen von diesen aber möglicherweise ohnehin in anderen Bereichen als in der Steuerung des Forschungsverhaltens zu liegen. Wie bereits in verschiedenen Studien angemerkt wurde (vgl. Jaeger 2006; Krempkow et al. 2013; Schröder 2004), kann eine Funktion indikatorenbasierter Steuerungsverfahren die Schaffung von Transparenz im Hinblick auf die erbrachte Forschungsleistung darstellen. Der Einsatz der Verfahren hätte dann – ähnlich einem Benchmarking (vgl. Kapitel 2.2.4) – mehr informierenden als steuernden Charakter. Hierfür müssten die Verfahren zum einen entsprechend ausgestaltet sein: Es sollten alle relevanten Forschungstätigkeiten und soweit möglich auch qualitative Aspekte der Forschungsarbeit berücksichtigt und nicht nur auf die einseitige Erfassung bzw. Förderung von Kriterien, die beispielsweise für Hochschulrankings relevant sind, gesetzt werden. Zum anderen darf keine finanzielle Abhängigkeit von eventuellen über die Verfahren verteilten zusätzlichen Finanzmitteln entstehen. Das bedeutet, dass unabhängige Forschung auch (weiterhin) ohne die indikatorenbasierte Zuweisung von Finanzmitteln möglich sein muss. Durch die Transparenzfunktion ließen sich dann möglicherweise dennoch teilweise die durch den Einsatz der Verfahren erhofften Effektivitäts- und Effizienzsteigerungen in der Forschungsarbeit erzielen (vgl. Kapitel 2.1.2). Denn die durch die Transparenzfunktion geschaffene Vergleichslogik und die Offenlegung der Forschungsleistung kann dazu beitragen, dass diejenigen Professorinnen und Professoren, die – entgegen den der akademischen Profession zugeschriebenen Attribute – weniger stark intrinsisch und durch das Streben nach wissenschaftlichem Erkenntnisgewinn motiviert sind, dazu angeregt werden, ihr Forschungsengagement zu steigern.

Literaturverzeichnis

- Alberts, B. (2013). Impact Factor Distortions. *Science* 340 (6134), 787. doi:10.1126/science.1240319
- Alexander von Humboldt Stiftung (Hrsg.). (2009). *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen. Beiträge zur Beurteilung von Forschungsleistungen* (Diskussionspapiere der Alexander von Humboldt-Stiftung 12/2009, 2. Aufl.). Bonn. https://www.humboldt-foundation.de/pls/web/docs/F13905/12_disk_papier_publicationsverhalten2_kompr.pdf. Zugegriffen 06.06.2017.
- Allerbeck, K. R. (1978). Meßniveau und Analyseverfahren - Das Problem „strittiger Intervallskalen“. *Zeitschrift für Soziologie* 7 (3), 199–214. <http://www.zfs-online.org/index.php/zfs/article/viewFile/2355/1892>. Zugegriffen 02.03.2016.
- Altrichter, H. (2015). Theory and Evidence on Governance: Conceptual and Empirical Strategies of Research on Governance in Education. In J. Schrader, J. Schmid, K. Amos & A. Thiel (Hrsg.), *Governance von Bildung im Wandel. Interdisziplinäre Zugänge* (Educational Governance, Bd. 28, S. 25–44). Wiesbaden: Springer VS.
- Altrichter, H. & Heinrich, M. (2007). Kategorien der Governance-Analyse und Transformationen der Systemsteuerung in Österreich. In H. Altrichter, T. Brüsemeister & J. Wissinger (Hrsg.), *Educational Governance. Handlungskoordination und Steuerung im Bildungssystem* (S. 55–103). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Amsler, S. S. & Bolsmann, C. (2012). University ranking as social exclusion. *British Journal of Sociology of Education* 33 (2), 283–301. doi:10.1080/01425692.2011.649835
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (11. Aufl.). Berlin: Springer.
- Baltes-Götz, B. (2014). *Lineare Regressionsanalyse mit SPSS* (Zentrum für Informations-, Medien- und Kommunikationstechnologie (ZIMK) an der Universität Trier, Hrsg.), Trier. <https://www.uni-trier.de/fileadmin/urt/doku/linreg/linreg.pdf>. Zugegriffen 25.04.2017.
- Baltes-Götz, B. (2015). *Mediator- und Moderatoranalyse per multipler Regression mit SPSS* (Zentrum für Informations-, Medien- und Kommunikationstechnologie (ZIMK) an der Universität Trier, Hrsg.), Trier. <https://www.uni-trier.de/fileadmin/urt/doku/medmodreg/medmodreg.pdf>. Zugegriffen 25.04.2017.
- Batson, C. D. (1991). *The altruism question. Toward a social psychological answer*. Hillsdale, N.J: Lawrence Erlbaum Associates.

- Baum, J. A. C. (2011). Free-Riding on Power Laws: questioning the validity of the Impact Factor as a measure of research quality in organization studies. *Organization* 18 (4), 449–466. doi:10.1177/1350508411403531
- Benz, A., Lütz, S., Schimank, U. & Simonis, G. (Hrsg.). (2007). *Handbuch Governance. Theoretische Grundlagen und empirische Anwendungsfelder*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Berghoff, S. & Hornbostel, S. (2003). Das CHE hinter den sieben Bergen. *Perspektiven der Wirtschaftspolitik* 4 (2), 191–195. http://www.ssoar.info/ssoar/bitstream/handle/document/1522/ssoar-perspwpol-2003-2-berghoff_et_al-das_che_hinter_den_sieben.pdf. Zugegriffen 12.11.2013.
- Berghoff, S., Giebisch, P., Hachmeister, C.-D., Hoffmann-Kobert, B., Hennings, M. & Ziegele, F. (2011). *Vielfältige Exzellenz 2011: Forschung - Anwendungsbezug - Internationalität - Studierendenorientierung im CHE Ranking* (Arbeitspapier Nr. 149). Gütersloh: CHE gemeinnütziges Centrum für Hochschulentwicklung. http://www.che-ranking.de/downloads/CHE_AP149_Vielfaeltige_Exzellenz_2011_2_Auflage.pdf. Zugegriffen 27.01.2016.
- Boer, H. de, Enders, J. & Schimank, U. (2007). On the Way towards New Public Management? The Governance of University Systems in England, the Netherlands, Austria, and Germany. In D. Jansen (Hrsg.), *New Forms of Governance in Research Organizations* (S. 137–152). Dordrecht: Springer Netherlands.
- Bogumil, J. & Heinze, R. G. (2009a). Einleitung. In J. Bogumil & R. G. Heinze (Hrsg.), *Neue Steuerung von Hochschulen. Eine Zwischenbilanz* (Modernisierung des öffentlichen Sektors. Sonderband, Bd. 34, S. 7–12). Berlin: Edition Sigma.
- Bogumil, J. & Heinze, R. G. (Hrsg.). (2009b). *Neue Steuerung von Hochschulen. Eine Zwischenbilanz* (Modernisierung des öffentlichen Sektors. Sonderband, Bd. 34). Berlin: Edition Sigma.
- Bogumil, J., Burgi, M., Heinze, R. G., Gerber, S., Gräf, I.-D., Jochheim, L., Schickentanz, M. & Wannöffel, M. (2013). *Modernisierung der Universitäten. Umsetzungsstand und Wirkungen neuer Steuerungsinstrumente* (Modernisierung des öffentlichen Sektors: Sonderband, Bd. 41). Berlin: Edition Sigma.
- Bogumil, J., Jochheim, L. & Gerber, S. (2015). Universitäten zwischen Detail- und Kontextsteuerung: Wirkungen von Zielvereinbarungen und Finanzierungsformeln im Zeitvergleich. In P. Bungarten & M. John-Ohnesorg (Hrsg.), *Hochschulgovernance in Deutschland* (Schriftenreihe Hochschulpolitik, Bd. 10, S. 55–78). Bonn.

- Bornmann, L. & Daniel, H.-D. (2003). Begutachtung durch Fachkollegen in der Wissenschaft. Stand der Forschung zur Reliabilität, Fairness und Validität des Peer-Review-Verfahrens. In S. Schwarz & U. Teichler (Hrsg.), *Universität auf dem Prüfstand. Konzepte und Befunde der Hochschulforschung* (S. 207–225). Frankfurt/Main: Campus Verlag.
- Bornmann, L., Mutz, R. & Daniel, H.-D. (2013). Multilevel-statistical reformulation of citation-based university rankings: The Leiden ranking 2011/2012. *Journal of the American Society for Information Science and Technology* 64 (8), 1649–1658. doi:10.1002/asi.22857
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. Aufl.). Berlin: Springer.
- Bortz, J. & Döring, N. (2003). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3. Aufl.). Heidelberg: Springer.
- Bozeman, B. & Su, X. (2015). Public Service Motivation Concepts and Theory: A Critique. *Public Administration Review* 75 (5), 700–710. doi:10.1111/puar.12248
- Braun, D. (1999). New Managerialism and the Governance of Universities in a Comparative Perspective. In D. Braun & F.-X. Merrien (Hrsg.), *Towards a new model of governance for universities? A comparative view* (Higher education policy series, Bd. 53, S. 239–261). London: J. Kingsley.
- Bright, L. (2007). Does Person-Organization Fit Mediate the Relationship Between Public Service Motivation and the Job Performance of Public Employees? *Review of Public Personnel Administration* 27 (4), 361–379. doi:10.1177/0734371X07307149
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (PS Psychologie, 3. Aufl.). München: Pearson Studium.
- Burckhart, H. (2012). “Rankings are here to stay”. Ein Plädoyer für mehr Sachlichkeit in der Rankingdiskussion. *Forschung & Lehre* 19 (10), 822–824.
- CHE Centrum für Hochschulentwicklung (2012). Methodische Genauigkeit und öffentlicher Nutzen des CHE Hochschulrankings. Erwiderungen des CHE auf die Stellungnahme der DGS zum CHE Hochschulranking. http://www.che-ranking.de/downloads/DGS_Stellungnahme_CHE_%28%29_1438.pdf. Zugegriffen 25.04.2017.
- Clark, B. R. (1986). *The higher education system. Academic organization in cross-national perspective*. Berkeley: University of California Press.
- Clark, B. R. (1998). *Creating entrepreneurial universities. Organizational pathways of transformation* (Issues in higher education). Oxford: IAU Press.

- Cogan, M. L. (1953). Toward a Definition of Profession. *Harvard Educational Review* 23 (1), 33–50.
- Cohen, M. D., March, J. G. & Olsen, J. P. (1972). A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly* 17 (1), 1–25. doi:10.2307/2392088
- Daniel, H.-D., Teichler, U. & Enders, J. (1998). Beiträge der empirischen Hochschulforschung zur Evaluierung von Forschung und Lehre: Hochschul-Ranking, Studentische Beurteilung von Lehrveranstaltungen, Selbststeuerung der Wissenschaft durch Peer-Review. In U. Teichler, H.-D. Daniel & J. Enders (Hrsg.), *Brennpunkt Hochschule. Neuere Analysen zu Hochschule, Beruf und Gesellschaft* (S. 11–53). Frankfurt/Main: Campus.
- Deci, E. L. (1972). The effects of contingent and noncontingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance* 8 (2), 217–229. doi:10.1016/0030-5073(72)90047-5
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik* 39 (2), 223–238.
- Deci, E. L., Koestner, R. & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125 (6), 627–668. doi:10.1037/0033-2909.125.6.627
- Deemer, E. D., Martens, M. P. & Buboltz, W. C. (2010). Toward a Tripartite Model of Research Motivation: Development and Initial Validation of the Research Motivation Scale. *Journal of Career Assessment* 18 (3), 292–309. doi:10.1177/1069072710364794
- Deemer, E. D., Mahoney, K. T. & Ball, J. H. (2012). Research Motives of Faculty in Academic STEM: Measurement Invariance of the Research Motivation Scale. *Journal of Career Assessment* 20 (2), 182–195. doi:10.1177/1069072711420856
- Deutsche Forschungsgemeinschaft (Hrsg.). (2012). Systematik der Fächer und Fachkollegien der DFG für die Amtsperiode 2012-2015. http://www.dfg.de/download/pdf/dfg_im_profil/gremien/fachkollegien/amtsperiode_2012_2015/fachsystematik_2012_2015_de_grafik.pdf. Zugegriffen 20.01.2016.
- Deutscher Hochschulverband (2013). *Hochschullehrer Verzeichnis: 2013. Band 1: Universitäten Deutschland* (21. Aufl.), Berlin. http://www.degruyter.com/search?f_0=isbnissn&q_0=9783110304893&searchTitles=true. Zugegriffen 25.04.2017.
- Diekmann, A. (2004). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen* (11. Aufl.). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.

- Dobbins, M. & Knill, C. (2016). Reformen der Hochschulsteuerung in Europa: Konzepte, Messung und empirische Befunde. In J. Schmid, K. Amos, J. Schrader & A. Thiel (Hrsg.), *Internationalisierte Welten der Bildung. Bildung und Bildungspolitik im globalen Vergleich* (Wirtschafts- und Sozialpolitik, Bd. 16, S. 33–78). Baden-Baden: Nomos Verlagsgesellschaft.
- Enders, J. (1998). Berufsbild der Hochschullehrer. In U. Teichler, H.-D. Daniel & J. Enders (Hrsg.), *Brennpunkt Hochschule. Neuere Analysen zu Hochschule, Beruf und Gesellschaft* (S. 55–78). Frankfurt/Main: Campus.
- Fahrmeir, L., Künstler, R., Pigeot, I. & Tutz, G. (2003). *Statistik: Der Weg zur Datenanalyse* (4. Aufl.). Berlin: Springer.
- Federkeil, G. (2012). Mehr Transparenz für Hochschulprofile. Genügen die vorhandenen Instrumente der Leistungsbewertung ihrer spezifischen Zielsetzung? *Wissenschaftsmanagement* 18 (5), 22–25.
- Federkeil, G. (2013). Internationale Hochschulrankings. Eine kritische Bestandsaufnahme. *Beiträge zur Hochschulforschung* 35 (2), 34–48. <http://www.bzh.bayern.de/uploads/media/2-2013-Federkeil.pdf>. Zugegriffen 25.04.2017.
- Federkeil, G. & Buch, F. (2007). *Fünf Jahre Juniorprofessur. Zweite CHE-Befragung zum Stand der Einführung* (Arbeitspapier / CHE Centrum für Hochschulentwicklung), Gütersloh. http://www.che.de/downloads/CHE_Juniorprofessur_Befragung_AP_90.pdf. Zugegriffen 24.02.2016.
- Fejes, A. & Nylander, E. (2014). The Anglophone International(e): A Bibliometric Analysis of Three Adult Education Journals, 2005-2012. *Adult Education Quarterly* 64 (3), 1–18. doi:10.1177/0741713614528025
- Fernet, C., Senecal, C., Guay, F., Marsh, H. & Dowson, M. (2008). The Work Tasks Motivation Scale for Teachers (WTMST). *Journal of Career Assessment* 16 (2), 256–279. doi:10.1177/1069072707305764
- Ferrara, E. & Romero, A. E. (2013). Scientific impact evaluation and the effect of self-citations: mitigating the bias by discounting h-index. *Journal of the American Society for Information Science and Technology* 64 (11), 2332–2339. doi:10.1002/asi.22976
- Fetchenhauer, D. & Bierhoff, H.-W. (2004). Altruismus aus evolutionstheoretischer Perspektive. *Zeitschrift für Sozialpsychologie* 35 (3), 131–141.
- Fidler, B. & Parsons, C. (2008). World University Ranking Methodologies: Stability and Variability. *Higher Education Review* 40 (3), 15–34.

- Fiedler, K. (2009). Psychologie. In Alexander von Humboldt Stiftung (Hrsg.), *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen. Beiträge zur Beurteilung von Forschungsleistungen* (Diskussionspapiere der Alexander von Humboldt-Stiftung 12/2009, 2. Aufl., S. 42–45). Bonn. https://www.humboldt-foundation.de/pls/web/docs/F13905/12_disk_papier_publicationsverhalten2_kompr.pdf. Zugegriffen 06.06.2017.
- Fischer, K. (2011). Soziale und kognitive Aspekte des Peer Review-Verfahrens. In K. Fischer & H. Parthey (Hrsg.), *Evaluation wissenschaftlicher Institutionen* (Wissenschaftsforschung Jahrbuch 2003, 2. Auflage, S. 23–62). Berlin: Gesellschaft für Wissenschaftsforschung e.V.
- Forschergruppe „Governance der Forschung“ (2009). Forschungspolitische Thesen der Forschergruppe „Governance der Forschung“: Rahmenbedingungen für eine leistungsfähige öffentlich finanzierte Forschung. In D. Jansen (Hrsg.), *Neue Governance für die Forschung. Tagungsband anlässlich der wissenschaftspolitischen Tagung der Forschergruppe „Governance der Forschung“, Berlin, 14. - 15. März 2007* (Interdisziplinäre Schriften zur Wissenschaftsforschung, Bd. 5, S. 131–145). Baden-Baden: Nomos.
- Frey, B. S. (1997). *Not just for the money. An economic theory of personal motivation*. Cheltenham, UK: Edward Elgar Publishing Ltd.
- Frey, B. S. (2003). Publishing as prostitution? - Choosing between one's own ideas and academic success. *Public Choice* 116 (1), 205–223. doi:10.1023/A:1024208701874
- Frey, B. S. & Jegen, R. (2001). Motivation Crowding Theory. *Journal of Economic Surveys* 15 (5), 589–611. doi:10.1111/1467-6419.00150
- Frey, B. S. & Osterloh, M. (1997). Sanktionen oder Seelenmassage? Motivationale Grundlagen der Unternehmensführung. *Die Betriebswirtschaft* 57 (3), 307–321.
- Fritsch, M., Henning, T., Slavtchev, V. & Steigenberger, N. (2007). *Hochschulen, Innovation, Region: Wissenstransfer im räumlichen Kontext*. Berlin: Edition Sigma.
- Fromm, S. (2003). *Multiple lineare Regressionsanalyse: Eine Einführung für Sozialwissenschaftler mit SPSS für Windows* (Baur, N. & Schulze, G., Hrsg.). Bamberg: Otto-Friedrich-Universität Bamberg, Fak. Sozial- und Wirtschaftswissenschaften, Professur für Methoden der empirischen Sozialforschung und Wissenschaftstheorie.
- Frost, J. & Brockmann, J. (2014). When qualitative productivity is equated with quantitative productivity: scholars caught in a performance paradox. *Zeitschrift für Erziehungswissenschaft* 17 (S6), 25–45. doi:10.1007/s11618-014-0572-8

- Gagné, M. & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior* 26 (4), 331–362. doi:10.1002/job.322
- Gillet, R. (1989). Research Performance Indicators Based on Peer Review. A Critical Analysis. *Higher Education Quarterly* 43 (1), 20–38.
- Glaser, H. A. (2012). Neue Zahlenspiele. Internationale Rankings von Universitäten. *Forschung & Lehre* 19 (3), 210–213.
- Gordon, R. A. (2010). *Regression analysis for the social sciences*. New York, NY: Routledge.
- Graß, D. (2015). Legitimation neuer Steuerung: Eine neo-institutionalistische Erweiterung der Governance-Perspektive auf Schule und Bildungsarbeit. In J. Schrader, J. Schmid, K. Amos & A. Thiel (Hrsg.), *Governance von Bildung im Wandel. Interdisziplinäre Zugänge* (Educational Governance, Bd. 28, S. 65–93). Wiesbaden: Springer VS.
- Grözinger, G. & Fromm, N. (2014). Auch Breiten- statt nur Spitzenförderung. Warum EPSCoR ein Vorbild für die DFG sein könnte. In E. Grande, D. Jansen, O. Jarren, A. Rip, U. Schimank & P. Weingart (Hrsg.), *Neue Governance der Wissenschaft: Reorganisation - externe Anforderungen – Medialisierung* (S. 113–138). Bielefeld: Transcript.
- Grözinger, G. & Leusing, B. (2006). *Wissenschaftsindikatoren an Hochschulen* (Internationales Institut für Management, Universität Flensburg, Hrsg.) (Discussion Paper Nr. 12).
- Gustin, B. H. (1973). Charisma, Recognition, and the Motivation of Scientists. *American Journal of Sociology* 78 (5), 1119–1134. <http://www.jstor.org/stable/2776628>. Zugegriffen 17.11.2016.
- Haas, H. (2009). Übersetzungsprobleme in der interkulturellen Befragung. *interculture journal* 8 (10), 61–77. <http://www.interculture-journal.com/index.php/icj/article/view/93/149>. Zugegriffen 14.10.2016.
- Hamilton, L. C. (2006). *Statistics with Stata* (8. Aufl.). Boston, Mass.: Brooks/Cole-Cengage Learning.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis. A regression-based approach*. New York, NY: Guilford Press.
- Hayes, A. F. & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression. An introduction and software implementation. *Behavior Research Methods* 39 (4), 709–722. doi:10.3758/BF03192961

- Hazelkorn, E. (2013). Reflections on a Decade of Global Rankings. What we've learned and outstanding issues. *Beiträge zur Hochschulforschung* 35 (2), 8–33. <http://www.bzh.bayern.de/uploads/media/2-2013-Hazelkorn.pdf>. Zugegriffen 25.04.2017.
- Heinze, R. G., Bogumil, J. & Gerber, S. (2011). Vom Selbstverwaltungsmodell zum Managementmodell? Zur Empirie neuer Governance-Strukturen im deutschen Hochschulsystem. In J. Schmid, K. Amos, J. Schrader & A. Thiel (Hrsg.), *Welten der Bildung? Vergleichende Analysen von Bildungspolitik und Bildungssystemen* (Wirtschafts- und Sozialpolitik, Bd. 7, S. 121–147). Baden-Baden: Nomos.
- Henke, J. & Dohmen, D. (2012). Wettbewerb durch leistungsorientierte Mittelzuweisungen? Zur Wirksamkeit von Anreiz- und Steuerungssystemen der Bundesländer auf Leistungsparameter der Hochschulen. *Die Hochschule* 21 (2), 100–120.
- Honecker, P. (2011). Image- und Reputationsmanagement als strategischer Prozess. Überlegungen zur Organisation von Hochschulkommunikation. *Wissenschaftsmanagement* 17 (5), 17–22.
- Hornbostel, S. (1997). *Wissenschaftsindikatoren. Bewertungen in der Wissenschaft*. Opladen: Westdeutscher Verlag.
- Hornbostel, S. (2001). Die Hochschulen auf dem Weg in die Audit Society: Über Forschung, Drittmittel, Wettbewerb und Transparenz. In E. Stöltzing & U. Schimank (Hrsg.), *Die Krise der Universitäten (Leviathan Sonderheft 20, S. 139–158)*. Wiesbaden: Westdeutscher Verlag.
- In der Smitten, S. & Jaeger, M. (2012). *Ziel- und Leistungsvereinbarungen als Instrument der Hochschulfinanzierung. Ausgestaltung und Anwendung* (Forum Hochschule Nr. 16). Hannover: Hochschul-Informations-System GmbH. http://www.dzhw.eu/pdf/pub_fh/fh-201216.pdf. Zugegriffen 26.10.2016.
- Jacob, A. K. & Teichler, U. (2011). *Der Wandel des Hochschullehrerberufs im internationalen Vergleich: Ergebnisse einer Befragung in den Jahren 2007/08* (Bundesministerium für Bildung und Forschung (BMBF), Hrsg.), Bonn. http://www.uni-kassel.de/wz1/pdf/BMBF_Hochschullehrerstudie2011_Druck.pdf. Zugegriffen 12.11.2013.
- Jacobsen, C. (2012). *Management interventions and motivation crowding effects in public service provision*. Aarhus: Forlaget Politica.
- Jacobsen, C., Hvitved, J. & Andersen, L. (2014). Command and motivation: How the perception of external interventions relates to intrinsic motivation and public service motivation. *Public Administration* 92 (4), 790–806. doi:10.1111/padm.12024

- Jacobsen, C. B. & Andersen, L. B. (2014). Performance Management for Academic Researchers: How Publication Command Systems Affect Individual Behavior. *Review of Public Personnel Administration* 34 (2), 1–24. doi:10.1177/0734371X13510850
- Jaeger, M. (2006). *Leistungsorientierte Budgetierung: Analyse der Umsetzung an ausgewählten Universitäten und Fakultäten/Fachbereichen* (HIS Kurzinformation A1/2006). Hannover: Hochschul-Informationssystem GmbH. http://www.dzhw.eu/pdf/pub_kia/kia200601.pdf. Zugegriffen 25.04.2017.
- Jaeger, M. (2009). Steuerung durch Anreizsysteme an Hochschulen. Wie wirken formelgebundene Mittelverteilung und Zielvereinbarungen? In J. Bogumil & R. G. Heinze (Hrsg.), *Neue Steuerung von Hochschulen. Eine Zwischenbilanz* (Modernisierung des öffentlichen Sektors. Sonderband, Bd. 34, S. 45–65). Berlin: Edition Sigma.
- Jaeger, M., Leszczensky, M., Orr, D. & Schwarzenberger, A. (2006). *Formelgebundene Mittelvergabe und Zielvereinbarungen als Instrumente der Budgetierung an deutschen Universitäten: Ergebnisse einer bundesweiten Befragung* (HIS Kurzinformation A13/2005). Hannover: Hochschul-Informationssystem GmbH. http://www.dzhw.eu/pdf/pub_kia/kia200513.pdf. Zugegriffen 25.04.2017.
- Jaeger, M. & In der Smitten, S. (2009). *Evaluation der leistungsbezogenen Mittelvergabe an die Berliner Hochschulen. Gutachten im Auftrag der Berliner Senatsverwaltung für Bildung, Wissenschaft und Forschung* (Forum Hochschule Nr. 1). Hannover: Hochschul-Informationssystem GmbH. http://www.dzhw.eu/pdf/pub_fh/fh-200901.pdf. Zugegriffen 26.10.2016.
- Jaeger, M. & In der Smitten, S. (2010). *Evaluation der leistungsbezogenen Mittelvergabe an die Hochschulen in Mecklenburg-Vorpommern. Gutachten im Auftrag des Ministeriums für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern* (Forum Hochschule Nr. 4). Hannover: Hochschul-Informationssystem GmbH. http://www.dzhw.eu/pdf/pub_fh/fh-201004.pdf. Zugegriffen 26.10.2016.
- Jansen, D. (2009). Neue Governance im deutschen Forschungssystem. Umsetzung und Wirkungen auf der Arbeitsebene der Forschung. In D. Jansen (Hrsg.), *Neue Governance für die Forschung. Tagungsband anlässlich der wissenschaftspolitischen Tagung der Forschergruppe „Governance der Forschung“, Berlin, 14. - 15. März 2007* (Interdisziplinäre Schriften zur Wissenschaftsforschung, Bd. 5, S. 37–59). Baden-Baden: Nomos.
- Jansen, D., Wald, A., Franke, K., Schmoch, U. & Schubert, T. (2007). Drittmittel als Performanzindikator der wissenschaftlichen Forschung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59 (1), 125–149. doi:10.1007/s11577-007-0006-1

- Jochheim, L. (2015). Zehn Jahre „W“. Effekte der leistungsorientierten Professorenvergütung. *Forschung & Lehre* 22 (4), 286–288.
- Karmann, A. (2014). Leistungsanreize an Fakultäten. LOM und ihre Fallstricke aus wirtschaftswissenschaftlicher Sicht. *Forschung & Lehre* 21 (6), 462–465.
- Kieser, A. (2010). Unternehmen Wissenschaft? *Leviathan* 38 (3), 347–367.
- Kieser, A. (2012). JOURQUAL – der Gebrauch, nicht der Missbrauch, ist das Problem. Oder: Warum Wirtschaftsinformatik die beste deutschsprachige betriebswirtschaftliche Zeitschrift ist. *DBW - Die Betriebswirtschaft* 72 (1), 93–110.
- Kieser, A. (2013, 14. August). So ertrickst man sich Reputation. Seriöse Forschung und die Jagd nach Ranking-Punkten sind nicht vereinbar. Die Grenze zum wissenschaftlichen Fehlverhalten ist bei Orientierung an Rangtabellen fließend. *Frankfurter Allgemeine Zeitung*.
- Kieser, A. & Kubicek, H. (1978). *Organisationstheorien I. Wissenschaftstheoretische Anforderungen und kritische Analyse klassischer Ansätze*. Stuttgart: W. Kohlhammer GmbH.
- Kieser, A. & Osterloh, M. (2012). „Den Unfug schnell beenden“. Professoren boykottieren das BWL-Ranking des Handelsblatts. *Forschung & Lehre* 19 (10), 820–821.
- Kleimann, B., In der Smitten, S. & Klawitter, M. (2015). Forschungserfahrung als Berufungskriterium. Anforderungen an zukünftige Professorinnen und Professoren. *Forschung & Lehre* 22 (8), 644–645.
- Kloke, K. (2014). *Qualitätsentwicklung an deutschen Hochschulen*. Wiesbaden: Springer Fachmedien.
- Kloke, K. & Krücken, G. (2012). Sind Universitäten noch lose gekoppelte Organisationen? Wahrnehmung und Umgang mit Zielkonflikten an deutschen Hochschulen aus der Organisationsperspektive unter besonderer Berücksichtigung der akademischen Lehre. In F. G. Becker, G. Krücken & E. Wild (Hrsg.), *Gute Lehre in der Hochschule. Wirkungen von Anreizen, Kontextbedingungen und Reformen* (S. 13–29). Gütersloh: Bertelsmann. http://www.pedocs.de/volltexte/2013/7815/pdf/Kloke_Kruecken_Sind_Universitaeten_2012.pdf. Zugegriffen 09.11.2016.
- Kock, S. (2013). Ärztliche Werbung im Wandel. Was darf ein Arzt wirklich? *Deutsches Ärzteblatt* 110 (46), 26–28. <http://www.aerzteblatt.de/epaperData/PRAXIS/2013/4/?startpage=26>. Zugegriffen 14.11.2016.

- Krempkow, R. & Schulz, P. (2012). Welche Effekte hat die leistungsorientierte Mittelvergabe? Das Beispiel der medizinischen Fakultäten Deutschlands. *Die Hochschule* 21 (2), 121–141.
- Krempkow, R., Landrock, U., Neufeld, J. & Schulz, P. (2013). *Intendierte und nicht-intendierte Effekte dezentraler Anreizsysteme am Beispiel der fakultätsintern leistungsorientierten Mittelvergabe in der Medizin. Abschlussbericht des Projektes GOMED-Governance Hochschulmedizin*. Berlin: IFQ - Institut für Forschungsinformation und Qualitätssicherung e.V. <http://www.forschungsinform.de/projekte/GOMED/GOMED-Abschlussbericht.pdf>. Zugegriffen 25.04.2017.
- Kroth, A. & Daniel, H.-D. (2008). Internationale Hochschulrankings. Ein methodenkritischer Vergleich. *Zeitschrift für Erziehungswissenschaft* 11 (4), 542–558.
- Krücken, G. (2006). Wandel - welcher Wandel? Überlegungen zum Strukturwandel der universitären Forschung in der Gegenwartsgesellschaft. *Die Hochschule* 15 (1), 7–18.
- Krücken, G. & Meier, F. (2006). Turning the university into an organizational actor. In G. S. Drori, J. W. Meyer & H. Hwang (Hrsg.), *Globalization and organization. World society and organizational change* (S. 241–257). Oxford: Oxford University Press.
- Kühl, S. (2015a). *Motivationsverdrängungseffekt. Die perverse Wirkung von Leistungsanreizen in der Wissenschaft* (Working Paper 2/2015). http://www.uni-bielefeld.de/soz/forschung/orgsoz/Stefan_Kuehl/pdf/Kuehl-Stefan-Working-Paper-2_2015-Motivationsverdrangungseffekt-17.09.15.pdf. Zugegriffen 25.04.2017.
- Kühl, S. (2015b). Reputation. Zur Funktion des Strebens nach Anerkennung in der Wissenschaft. *Forschung & Lehre* 22 (10), 804–806.
http://www.wissenschaftsmanagement-online.de/system/files/downloads-wimoarticle/1510_WIMO_Reputation_K%C3%9CHL.pdf. Zugegriffen 16.11.2016.
- Kurtz, T. (2002). *Berufssoziologie*. Bielefeld: Transcript.
- Lange, S. (2008). New Public Management und die Governance der Universitäten. *der moderne staat - Zeitschrift für Policy, Recht und Management* 1 (1), 235–248.
<http://www.budrich-journals.de/index.php/dms/article/view/2819/2354>. Zugegriffen 11.03.2014.
- Lau, M. Y., Cisco, H. C. & Delgado-Romero, E. A. (2008). Institutional and Individual Research Productivity in Five Nominated Multicultural Psychology Journals. *Journal of Multicultural Counseling & Development* 36 (4), 194–205.
- Leininger, W. (2009). Wirtschaftswissenschaften. In Alexander von Humboldt Stiftung (Hrsg.), *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen*.

- Beiträge zur Beurteilung von Forschungsleistungen* (Diskussionspapiere der Alexander von Humboldt-Stiftung 12/2009, 2. Aufl., S. 67–68). Bonn. https://www.humboldt-foundation.de/pls/web/docs/F13905/12_disk_papier_publicationsverhalten2_kompr.pdf. Zugegriffen 06.06.2017.
- Lenzen, D. (2012). Ranking, Rating - Steuerung und Motivation. Erfahrungen und Befunde zum Forschungsrating aus Sicht einer Universitätsleitung. *Erziehungswissenschaft* 23 (45), 13–25.
- Lippe, P. v. d. & Kladroba, A. (2002). Repräsentativität von Stichproben. *Marketing - Zeitschrift für Forschung und Praxis* 24 (2), 227–238.
- Long, J. S. & Ervin, L. H. (2000). Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician* 54 (3), 217–224. doi:10.2307/2685594
- Luhmann, N. (1964). *Funktionen und Folgen formaler Organisationen*. Berlin: Duncker & Humblot.
- Luhmann, N. (1975a). Interaktion, Organisation, Gesellschaft. In N. Luhmann (Hrsg.), *Soziologische Aufklärung 2* (S. 9–20). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Luhmann, N. (1975b). Komplexität. In N. Luhmann (Hrsg.), *Soziologische Aufklärung 2* (S. 204–220). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Luhmann, N. (1992). *Die Wissenschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.
- Maasen, S. & Weingart, P. (2006). Unternehmerische Universität und neue Wissenschaftskultur. *Die Hochschule* 15 (1), 19–45.
- Marginson, S. & van der Wende, M. (2007). To Rank or To Be Ranked: The Impact of Global Rankings in Higher Education. *Journal of Studies in International Education* 11 (3), 306–329. doi:10.1177/1028315307303544
- Meier, F. & Schimank, U. (2010). Organisationsforschung. In D. Simon, A. Knie & S. Hornbostel (Hrsg.), *Handbuch Wissenschaftspolitik* (S. 106–117). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Michael, S. O. (2005). The Cost of Excellence: The Financial Implications of Institutional Rankings. *International Journal of Educational Management* 19 (5), 365–382.
- Miller, C. C. (2006). Peer Review in the organizational and management sciences: Prevalence and effects of reviewer hostility, bias, and dissensus. *Academy of Management Journal* 49 (3), 425–431.

- Minssen, H. & Wilkesmann, U. (2003). Lassen Hochschulen sich steuern? *Soziale Welt* 54 (2), 123–144.
- Minssen, H., Molsich, B., Wilkesmann, U. & Andersen, U. (2003). *Kontextsteuerung von Hochschulen? Folgen der indikatorisierten Mittelzuweisung* (Abhandlungen zu Bildungsforschung und Bildungsrecht, Bd. 12). Berlin: Duncker & Humblot.
- Mintzberg, H. (1980). Structure in 5's: A Synthesis of the Research on Organization Design. *Management Science* 26 (3), 322–341.
- Mintzberg, H. (1992). *Die Mintzberg-Struktur. Organisation effektiver gestalten*. Landsberg/Lech: Verlag Moderne Industrie.
- Müller, H. (2012). Zitationen als Grundlage von Forschungsleistungsrankings – Konzeptionelle Überlegungen am Beispiel der Betriebswirtschaftslehre. *Beiträge zur Hochschulforschung* 34 (2), 68–92. <http://www.bzh.bayern.de/uploads/media/2-2012-Mueller.pdf>. Zugegriffen 10.12.2013.
- Nezlek, J. B., Schröder-Abé, M. & Schütz, A. (2006). Mehrebenenanalysen in der psychologischen Forschung. *Psychologische Rundschau* 57 (4), 213–223. doi:10.1026/0033-3042.57.4.213
- Oancea, A. (2014). Research assessment as governance technology in the United Kingdom: findings from a survey of RAE 2008 impacts. *Zeitschrift für Erziehungswissenschaft* 17 (S6), 83–110. doi:10.1007/s11618-014-0575-5
- Oberschelp, A. & Jaeger, M. (2015). Leistungsvergleiche als Instrument der Hochschulsteuerung: Ansätze, organisatorischer Kontext und Unterstützung des Steuerungshandelns. *Bibliotheksdienst* 49 (5), 475–494. doi:10.1515/bd-2015-0059
- Opp, K.-D. (2010). Kausalität als Gegenstand der Sozialwissenschaften und der multivariaten Statistik. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 9–38). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Parolo, P. D. B., Pan, R. K., Ghosh, R., Huberman, B. A., Kaski, K. & Fortunato, S. (2015). Attention decay in science. *Journal of Informetrics* 9 (4), 734–745. doi:10.1016/j.joi.2015.07.006
- Parsons, T. (1969). Research with Human Subjects and the „Professional Complex“. *Daedalus* 98 (2), 325–360. <http://www.jstor.org/stable/20023881>. Zugegriffen 15.11.2016.
- Parsons, T. & Platt, G. M. (1990). *Die amerikanische Universität. Ein Beitrag zur Soziologie der Erkenntnis*. Frankfurt am Main: Suhrkamp.

- Perry, J. L. & Wise, L. R. (1990). The Motivational Bases of Public Service. *Public Administration Review* 50 (3), 367–373.
- Perry, J. L., Hondeghem, A. & Wise, L. R. (2010). Revisiting the Motivational Bases of Public Service: Twenty Years of Research and an Agenda for the Future. *Public Administration Review* 70 (5), 681–690. doi:10.1111/j.1540-6210.2010.02196.x
- Pfadenhauer, M. (2003). *Professionalität. Eine wissenssoziologische Rekonstruktion institutionalisierter Kompetenzdarstellungskompetenz*. Opladen: Leske + Budrich.
- Power, M. (1999). *The audit society. Rituals of verification*. Oxford: Oxford University Press.
- Rassenhövel, S. & Dyckhoff, H. (2006). Die Relevanz von Drittmittelindikatoren bei der Beurteilung der Forschungsleistung im Hochschulbereich. In S. v. Zelewski & N. Akca (Hrsg.), *Fortschritt in den Wirtschaftswissenschaften. Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen* (Gabler Edition Wissenschaft, S. 85–112). Wiesbaden: Deutscher Universitäts-Verlag.
- Rauhvargers, A. (2011). *Global university rankings and their impact* (European University Association (EUA), Hrsg.), Brüssel. http://www.eua.be/pubs/Global_University_Rankings_and_their_Impact.pdf. Zugegriffen 12.11.2013.
- Rauhvargers, A. (2013). *Global university rankings and their impact. Report II* (European University Association (EUA), Hrsg.), Brüssel. http://www.eua.be/Libraries/publications-homepage-list/EUA_Global_University_Rankings_and_Their_Impact_-_Report_II. Zugegriffen 26.04.2017.
- Rauhvargers, A. (2014). Where Are the Global Rankings Leading Us? An Analysis of Recent Methodological Changes and New Developments. *European Journal of Education* 49 (1), 29–44. doi:10.1111/ejed.12066
- Ridder, D. d. (2010). Reputationsmanagement an Hochschulen. Den guten Ruf strategisch planen. *Wissenschaftsmanagement* 16 (4), 53–54.
- Robinson-García, N. & Calero-Medina, C. (2014). What do university rankings by fields rank? Exploring discrepancies between the organizational structure of universities and bibliometric classifications. *Scientometrics* 98 (3), 1955–1970.
- Roessler, I. (2013). *Was war? Was bleibt? Was kommt? 15 Jahre Erfahrungen mit Rankings und Indikatoren im Hochschulbereich* (CHE Arbeitspapier Nr. 167). Gütersloh: Centrum für Hochschulentwicklung gGmbH. https://www.che.de/downloads/CHE_AP_167_Erfahrungen_mit_Rankings_und_Indikatoren.pdf. Zugegriffen 26.04.2017.

- Ryan, R. M. (1982). Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory. *Journal of Personality and Social Psychology* 43 (3), 450–461.
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55 (1), 68–78. doi:10.1037/0003-066X.55.1.68
- Schedler, K. (2003). Der Aspekt Benchmarking (an Universitäten). In J. Lange, W. Benz & K. Schedler (Hrsg.), *Wettbewerb, Kooperation, Benchmarking. Geeignete Instrumente für Hochschulentwicklung?* (Fortbildungsprogramm für die Wissenschaftsverwaltung, Bd. 80, S. 29–44). Weimar: Bauhaus-Universitätsverlag.
- Schenker-Wicki, A. (1996). *Evaluation von Hochschulleistungen. Leistungsindikatoren und Performance Measurements*. Wiesbaden: Deutscher Universitätsverlag.
- Schimank, U. (2005). Die akademische Profession und die Universitäten. „New Public Management“ und eine drohende Entprofessionalisierung. In T. Klatetzki & V. Tacke (Hrsg.), *Organisation und Profession* (Organisation und Gesellschaft, S. 143–164). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schimank, U. (2007). Die Governance-Perspektive: Analytisches Potenzial und anstehende konzeptionelle Fragen. In H. Altrichter, T. Brüsemeister & J. Wissinger (Hrsg.), *Educational Governance. Handlungskoordination und Steuerung im Bildungssystem* (S. 231–260). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schimank, U. (2008). Ökonomisierung der Hochschulen: eine Makro-Meso-Mikro-Perspektive. In K. S. Rehberg (Hrsg.), *Die Natur der Gesellschaft. Verhandlungen des 33. Kongresses der Deutschen Gesellschaft für Soziologie in Kassel 2006* (S. 622–635). Frankfurt am Main: Campus.
- Schimank, U. (2009). Governance-Reformen nationaler Hochschulsysteme – Deutschland in internationaler Perspektive. In J. Bogumil & R. G. Heinze (Hrsg.), *Neue Steuerung von Hochschulen. Eine Zwischenbilanz* (Modernisierung des öffentlichen Sektors. Sonderband, Bd. 34, S. 123–137). Berlin: Edition Sigma.
- Schimank, U. (2015). Identitätsbedrohungen und Identitätsbehauptung: Professoren in reformbewegten Universitäten. In V. v. Groddeck & S. M. Wilz (Hrsg.), *Formalität und Informalität in Organisationen* (S. 277–296). Wiesbaden: Springer VS.
- Schimank, U. & Volkmann, U. (2008). Ökonomisierung der Gesellschaft. In A. Maurer (Hrsg.), *Handbuch der Wirtschaftssoziologie* (S. 382–393). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Schleich, W. P. (2009). Quantenoptik. In Alexander von Humboldt Stiftung (Hrsg.), *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen. Beiträge zur Beurteilung von Forschungsleistungen* (Diskussionspapiere der Alexander von Humboldt-Stiftung 12/2009, 2. Aufl., S. 88–91). Bonn. https://www.humboldt-foundation.de/pls/web/docs/F13905/12_disk_papier_publicationsverhalten2_kompr.pdf. Zugegriffen 06.06.2017.
- Schluchter, W. (1974). Legitimationsprobleme der Medizin. *Zeitschrift für Soziologie* 3 (4), 375–396.
- Schmeiser, M. (2006). Soziologische Ansätze der Analyse von Professionen, der Professionalisierung und des professionellen Handelns. *Soziale Welt* 57 (3), 295–318. http://www.soziale-welt.nomos.de/fileadmin/soziale-welt/doc/SozWelt_06_03.pdf. Zugegriffen 17.11.2016.
- Schmidt, B. & Weishaupt, H. (2008). Forschung und wissenschaftlicher Nachwuchs. In K.-J. Tillmann, T. Rauschenbach, R. Tippelt & H. Weishaupt (Hrsg.), *Datenreport Erziehungswissenschaft 2008* (S. 113–138). Opladen: Budrich.
- Schmidt, B., Tippelt, R. & Gogolin, I. (2005). Lehrevaluation. In I. Gogolin, H.-H. Krüger, D. Lenzen & T. Rauschenbach (Hrsg.), *Standards und Standardisierungen in der Erziehungswissenschaft* (Zeitschrift für Erziehungswissenschaft. Beiheft 4, S. 227–241). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schmidt-Hertha, B. (2012). Forschung und Publikationskulturen. In W. Thole, H. Faulstich-Wieland, K.-P. Horn, H. Weishaupt & I. Züchner (Hrsg.), *Datenreport Erziehungswissenschaft 2012* (S. 159–180). Opladen: Budrich.
- Schmidt-Hertha, B., Thalhammer, V. & Müller, M. (2015). Studiengangs-Reputation als Indikator für Lehrqualität? In I. Langemeyer, M. Fischer & M. Pfadenhauer (Hrsg.), *Epistemic and learning cultures. Wohin sich Universitäten entwickeln* (S. 72–84). Weinheim: Beltz Juventa.
- Schmidt-Hertha, B., Tafertshofer, L. & Tippelt, R. (2016). Forschung und Publikationskulturen. In H.-C. Koller, H. Weishaupt & I. Züchner (Hrsg.), *Datenreport Erziehungswissenschaft 2016* (S. 159–178). Leverkusen: Budrich.
- Schmidt-Hertha, B., Tafertshofer, L. & Werner, E. (2016). *Qualität von sozialwissenschaftlichen Studiengängen und Reputation von Studienstandorten aus Sicht von Professorinnen und Professoren*. Unveröffentlichter Abschlussbericht zum Projekt QualRep.

- Schneider, W. L. (2009). *Grundlagen der Soziologischen Theorie. Band 2: Garfinkel - RC - Habermas - Luhmann* (3. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schrader, J., Schmid, J., Amos, K. & Thiel, A. (2015). Governance im Bildungsbereich – Erträge, Desiderate und Potenziale der Forschung. In J. Schrader, J. Schmid, K. Amos & A. Thiel (Hrsg.), *Governance von Bildung im Wandel. Interdisziplinäre Zugänge* (Educational Governance, Bd. 28, S. 3–21). Wiesbaden: Springer VS.
- Schröder, T. (2004). Der Einsatz leistungsorientierter Ressourcensteuerungsverfahren im deutschen Hochschulsystem. Eine empirische Untersuchung ihrer Ausgestaltung und Wirkungsweise. *Beiträge zur Hochschulforschung* 26 (2), 28–57. <http://www.bzh.bayern.de/uploads/media/2-2004-schroeder.pdf>. Zugegriffen 30.05.2014.
- Shih, T.-H. & Fan, X. (2008). Comparing Response Rates from Web and Mail Surveys. A Meta-Analysis. *Field Methods* 20 (3), 249–271. doi:10.1177/1525822X08317085
- Sieweke, J., Muck, J., Süß, S. & Haucap, J. (2014). Forschungsevaluation an Universitäten. Ergebnisse einer explorativen Studie rechts- und wirtschaftswissenschaftlicher Fakultäten. *Zeitschrift für betriebswirtschaftliche Forschung* 66 (4), 274–305.
- Singh, G., Haddad, K. M. & Chow, C. W. (2007). Are Articles in “Top” Management Journals Necessarily of Higher Quality? *Journal of Management Inquiry* 16 (4), 319–331. doi:10.1177/1056492607305894
- Starbuck, W. H. (2005). How Much Better Are the Most-Prestigious Journals? The Statistics of Academic Publication. *Organization Science* 16 (2), 180–200. doi:10.1287/orsc.1040.0107
- Statistisches Bundesamt (2013). *Bildung und Kultur. Personal an Hochschulen*. Fachserie 11, Reihe 4.4 (Statistisches Bundesamt, Hrsg.), Wiesbaden. <https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/PersonalHochschulen2110440137004.pdf>. Zugegriffen 04.01.2016.
- Statistisches Bundesamt (2015). *Bildung und Kultur. Personal an Hochschulen*. Fachserie 11, Reihe 4.4 (Statistisches Bundesamt, Hrsg.), Wiesbaden. <https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/PersonalHochschulen2110440147004.pdf>. Zugegriffen 08.01.2015.
- Statistisches Bundesamt (2016). *Bildung und Kultur. Personal an Hochschulen*. Fachserie 11, Reihe 4.4 (Statistisches Bundesamt, Hrsg.), Wiesbaden. <https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/PersonalHochschulen2110440157004.pdf>. Zugegriffen 21.10.2016.

- Steinhage, N. & Blossfeld, H.-P. (1999). *Zur Problematik von Querschnittsdaten: methodisch-statistische Beschränkungen von Querschnittsstudien bei der empirischen Überprüfung von Theorien* (Arbeitspapier / Sfb 186 62). Bremen: Universität Bremen, SFB 186 Statuspassagen und Risikolagen im Lebensverlauf. <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-57590>. Zugegriffen 15.01.2016.
- Stichweh, R. (2005). Wissen und die Professionen in einer Organisationsgesellschaft. In T. Klatetzki & V. Tacke (Hrsg.), *Organisation und Profession* (Organisation und Gesellschaft, S. 31–44). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Teichler, U. (2014). *Hochschulsysteme und quantitativ-strukturelle Hochschulpolitik. Differenzierung, Bologna-Prozess, Exzellenzinitiative und die Folgen* (Studienreihe Bildungs- und Wissenschaftsmanagement, Bd. 14). Münster: Waxmann.
- Tippelt, R., Schmidt, B. & Hippel, A. von (2010). Higher Education Evaluation in Germany. *Research in Comparative and International Education* 5 (1), 98–111. doi:10.2304/rcie.2010.5.1.98
- Trute, H.-H. & Pilniok, A. (2009). Von der Ordinarien- über die Gremien- zur Managementuniversität? Veränderte Governance-Strukturen der universitären Forschung und ihre normativen Konsequenzen. In D. Jansen (Hrsg.), *Neue Governance für die Forschung. Tagungsband anlässlich der wissenschaftspolitischen Tagung der Forschergruppe „Governance der Forschung“, Berlin, 14. - 15. März 2007* (Interdisziplinäre Schriften zur Wissenschaftsforschung, Bd. 5, S. 21–35). Baden-Baden: Nomos.
- Urban, D. & Mayerl, J. (2011). *Regressionsanalyse: Theorie, Technik und Anwendung*: VS Verlag für Sozialwissenschaften.
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C. & Vallieres, E. F. (1992). The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. *Educational and Psychological Measurement* 52 (4), 1003–1017. doi:10.1177/0013164492052004025
- Vandenabeele, W. (2007). Toward a public administration theory of public service motivation. *Public Management Review* 9 (4), 545–556. doi:10.1080/14719030701726697
- Vandenabeele, W. (2008). Government Calling: Public Service Motivation as an element in selecting government as an employer of choice. *Public Administration* 86 (4), 1089–1105. doi:10.1111/j.1467-9299.2008.00728.x

- Vandenabeele, W. (2009). The mediating effect of job satisfaction and organizational commitment on self-reported performance. More robust evidence of the PSM-performance relationship. *International Review of Administrative Sciences* 75 (1), 11–34. doi:10.1177/0020852308099504
- Weick, K. E. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly* 21 (1), 1–19. doi:10.2307/2391875
- Weingart, P. (2005). Impact of bibliometrics upon the science system. Inadvertent consequences? *Scientometrics* 62 (1), 117–131. doi:10.1007/s11192-005-0007-7
- Wenzel, H. (2005). Profession und Organisation. Dimensionen der Wissensgesellschaft bei Talcott Parsons. In T. Klatetzki & V. Tacke (Hrsg.), *Organisation und Profession* (Organisation und Gesellschaft, S. 45–71). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Wespel, J. & Jaeger, M. (2015). Leistungsorientierte Zuweisungsverfahren der Länder: Praktische Umsetzung und Entwicklungen. *Hochschulmanagement* 10 (3), 97–105.
- Wiesenthal, H. (2000). Markt, Organisation und Gemeinschaft als ‚zweitbeste‘ Verfahren sozialer Ordnungsbildung. In R. Werle, U. Schimank & R. Mayntz (Hrsg.), *Gesellschaftliche Komplexität und kollektive Handlungsfähigkeit* (Schriften des Max-Planck-Instituts für Gesellschaftsforschung, Köln, Bd. 39, S. 44–73). Frankfurt/Main: Campus.
- Wilhite, A. W. & Fong, E. A. (2012). Coercive Citation in Academic Publishing. *Science* 335 (6068), 542–543. doi:10.1126/science.1212540
- Wilkesmann, U. (2012). Auf dem Weg vom Gelehrten zum abhängig Beschäftigten? Zwei deutschlandweite Surveys zur Lehrmotivation von Professoren. In U. Wilkesmann & C. J. Schmid (Hrsg.), *Hochschule als Organisation* (S. 363–381). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Wilkesmann, U. & Schmid, C. J. (2012). The impacts of new governance on teaching at German universities. Findings from a national survey. *Higher Education* 63 (1), 33–52. doi:10.1007/s10734-011-9423-1
- Williams, R. (2008). Methodology, meaning and usefulness of rankings. *Australian Universities' Review* 2 (50), 51–58.
- Wissenschaftsrat (2008a). *Empfehlungen zur Qualitätsverbesserung von Lehre und Studium* (Drs. 8639-08). Berlin: Wissenschaftsrat. <http://www.wissenschaftsrat.de/download/archiv/8639-08.pdf>. Zugegriffen 12.11.2013.

- Wissenschaftsrat (2008b). *Leitfaden zu den Datenberichten für die Bewertungsgruppe Soziologie* (Drs. 8409-08). Köln: Wissenschaftsrat. http://www.wissenschaftsrat.de/download/Forschungsrating/Dokumente/Pilotstude_Forschungsrating_Soziologie/pilot-Leitf-Datbericht-Sozio.pdf. Zugegriffen 26.11.2013.
- Wolbring, T. (2013). *Fallstricke der Lehrevaluation. Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität* (Hochschule und Gesellschaft, Bd. 1). Frankfurt am Main: Campus Verlag.
- Wollersheim, J., Welppe, I. M. & Ringelhan, S. (2014). Sein und Sollen: Bewertung von Forschungsleistungen in den Wirtschaftswissenschaften. *Forschung & Lehre* 21 (4), 276–278.
- Würmseer, G. (2010). *Auf dem Weg zu neuen Hochschultypen. Eine organisationssoziologische Analyse vor dem Hintergrund hochschulpolitischer Reformen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Zerback, T., Schoen, H., Jakob, N. & Schlereth, S. (2009). Zehn Jahre Sozialforschung mit dem Internet — eine Analyse zur Nutzung von Online-Befragungen in den Sozialwissenschaften. In N. Jakob, H. Schoen & T. Zerback (Hrsg.), *Sozialforschung im Internet* (S. 15–31). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ziegele, F. & van Vught, F. (2013). „U-Multirank“ und „U-Map“ als Ansätze zur Schaffung von Transparenz im europäischen und globalen Hochschulsystem. Konzepte und Erfahrungen. *Beiträge zur Hochschulforschung* 35 (2), 50–74. <http://www.bzh.bayern.de/uploads/media/2-2013-Ziegele-vanVught.pdf>. Zugegriffen 26.04.2017.

Anhang

I. Fragebogen

Willkommen zur Befragung „Indikatorenbasierte Steuerungs- und Anreizverfahren und Forschungsverhalten von Professorinnen und Professoren“

Vielen Dank, dass Sie an der Befragung teilnehmen!

Der Fragebogen besteht überwiegend aus geschlossenen Fragen, bei denen Sie bitte die für Sie zutreffende Antwortoption ankreuzen.

Bei einigen Fragen ist eine 5-stufige Antwortskala vorgegeben (z.B. "1 sehr gering" bis "5 sehr hoch"), mit der Sie Ihre Einschätzung entsprechend der Werte der Skala abstimmen können.

Bitte beantworten Sie möglichst alle Fragen. Sollten Sie eine Frage dennoch nicht beantworten können oder wollen, lassen Sie die Frage einfach aus bzw. kreuzen Sie die „Weiß nicht“-Kategorie an.

Die Befragung ist natürlich freiwillig und erfolgt anonym. Die Ergebnisse werden nur in aggregierter Form ausgewertet, sodass keine Rückschlüsse auf Ihre Person möglich sind.

Sofern Ihr Browser Cookies akzeptiert, können Sie die Befragung jederzeit unterbrechen und durch erneutes Aufrufen des Links die Befragung an der Stelle fortsetzen, an der Sie abgebrochen haben.

Wenn Sie Fragen zur Thematik oder dem Fragebogen haben, melden Sie sich gerne bei

Dipl.-Soz. Lorenz Tafertshofer

lorenz.tafertshofer@uni-tuebingen.de oder 07071 29 74872.

Klicken Sie nun auf „Weiter“, um mit der Befragung zu beginnen.

Weiter >>

7%

In welchem Studienbereich bzw. Studienfach sind Sie tätig?

- Erziehungswissenschaften / Sonderpädagogik
- Politikwissenschaften
- Psychologie
- Sozialwesen / Soziale Arbeit / Sozialpädagogik
- Soziologie / Sozialwissenschaften
- Wirtschaftswissenschaften
- keiner der genannten, sondern:

Welche Position haben Sie aktuell an Ihrer Hochschule inne?

- Juniorprofessor/in (Besoldungsgruppe W1)
- Professor/in (Besoldungsgruppe W2, W3)
- Professor/in (Besoldungsgruppe C2, C3, C4)
- apl. Professor/in
- Honorarprofessor/in
- Vertretung einer Professur
- emeritiert / im Ruhestand
- andere, und zwar:

<< Zurück

Weiter >>



Welche der folgenden Forschungsindikatoren werden an Ihrer Fakultät im Rahmen von Steuerungs- oder Anreizverfahren (z.B. formelgebundene Mittelzuweisung, Zielvereinbarungen, Leistungszulagen W-Besoldung, Benchmarking von Lehrstühlen) eingesetzt?

	Indikator wird eingesetzt	Indikator wird nicht eingesetzt	weiß nicht
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sonstige:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>



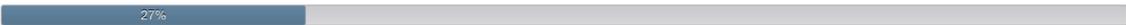
Bitte schätzen Sie ein, welche Gewichtung die folgenden Forschungsindikatoren in den Steuerungs- und Anreizverfahren an Ihrer Fakultät insgesamt haben.

Anmerkung: Zur Beantwortung der Frage ist es nicht erforderlich, dass Ihnen die exakten Gewichtungen (z.B. Gewichtungformel innerhalb formelgebundener Mittelzuweisung) bekannt sind. Geben Sie bitte nur an, wie stark die Indikatoren innerhalb der Steuerungs- und Anreizverfahren Ihrem subjektiven Empfinden nach insgesamt gewichtet werden.

	sehr geringe Gewichtung	1	2	3	4	sehr hohe Gewichtung
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>



Wie hoch schätzen Sie die Bedeutung der folgenden Aspekte der Forschungsarbeit für den Reputationszuwachs von Professorinnen und Professoren innerhalb Ihrer Scientific Community ein?

Anmerkung: Bitte beziehen Sie sich bei der Beantwortung der Frage auf diejenige fachspezifische Scientific Community, in der Sie sich am ehesten verorten würden / der Sie sich am ehesten zugehörig fühlen.

	sehr gering	1	2	3	4	sehr hoch
Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>					
Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>					
Andere Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>					
Publikationen in Zeitschriften mit hohem Impact-Factor / in high-ranked journals	<input type="radio"/>					
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>					
Anzahl betreuter Promotionen	<input type="radio"/>					
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>					
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>					
Sonstige:	<input type="radio"/>					

<< Zurück

Weiter >>

33%

Wie hoch schätzen Sie persönlich den Beitrag der folgenden Aspekte der Forschungsarbeit für den wissenschaftlichen Erkenntnisgewinn in Ihrer Disziplin insgesamt ein?

Anmerkung: Bitte beziehen Sie sich bei der Beantwortung der Frage darauf, welchen Beitrag (z.B. unter dem Gesichtspunkt der Wissensvermittlung oder der Sicherstellung qualitativer Standards) die verschiedenen Aspekte Ihrer Meinung nach für den wissenschaftlichen Erkenntnisgewinn leisten – unabhängig von externen Vorgaben oder Wertungen durch die Scientific Community.

	sehr gering					sehr hoch
	1	2	3	4	5	5
Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>					
Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>					
Andere Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>					
Publikationen in Zeitschriften mit hohem Impact-Factor / in high-ranked journals	<input type="radio"/>					
Zitationen	<input type="radio"/>					
Betreuung von Promotionen	<input type="radio"/>					
Durchführung von Drittmittelprojekten der Deutschen Forschungsgemeinschaft	<input type="radio"/>					
Durchführung von sonstigen Drittmittelprojekten (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>					
Sonstige:	<input type="radio"/>					

Die folgenden Aussagen beziehen sich auf Ihre Forschungstätigkeit. Bitte geben Sie an, inwieweit Sie den einzelnen Aussagen jeweils zustimmen.

	stimme überhaupt nicht zu				stimme voll und ganz zu
	1	2	3	4	5
Wenn ich forsche, fühle ich mich zufrieden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich forsche aus Freude an der Sache.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich rundum wohl, wenn ich in Forschungsarbeit involviert bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mir gefällt es sehr durch Forschung neue Erkenntnisse zu gewinnen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe das Bedürfnis wissenschaftliche Sachverhalte zu verstehen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich neue Erkenntnisse in meinem Forschungsfeld gewonnen habe, empfinde ich große Freude.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Forschungsarbeit an sich macht mir Spaß.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich betreibe Forschung um ihrer selbst willen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Zeit vergeht wie im Flug, wenn ich mit Forschungstätigkeiten beschäftigt bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>

40%

Wenn Sie an Ihre Forschungstätigkeit der letzten drei Jahre zurückdenken, welchen subjektiven Stellenwert hatten die folgenden Tätigkeiten für Sie in Ihrer Forschungsarbeit?

Anmerkung: Bitte beziehen Sie sich bei Ihrer Antwort nicht auf den absoluten aufgetragenen Zeitaufwand für die einzelnen Tätigkeiten, sondern darauf, wie wichtig es Ihnen war, sich in den jeweiligen Tätigkeiten zu engagieren.

	sehr geringer Stellenwert				sehr hoher Stellenwert
	1	2	3	4	5
Publizieren in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Publizieren in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verfassen anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Publizieren in Zeitschriften mit hohem Impact-Factor / in high-ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Publizieren als Koautor / in Mehrautorenschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Betreuung von Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Einwerben von Drittmitteln der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Einwerben von sonstigen Drittmitteln (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Betreiben von Grundlagenforschung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Betreiben von anwendungsbezogener Forschung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sonstiges:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>



Am Anfang des Fragebogens wurden Sie danach gefragt, welche Indikatoren an Ihrer Fakultät in Steuerungsverfahren und Anreizsystemen Verwendung finden. Bitte geben Sie nun an, welche Steuerungsverfahren und Anreizsysteme an Ihrer Fakultät eingesetzt werden.

Welche der folgenden Steuerungsverfahren und Anreizsysteme werden an Ihrer Fakultät eingesetzt?

	Verfahren wird eingesetzt	Verfahren wird nicht eingesetzt	weiß nicht
Formelgebundene/indikatorenbasierte Mittelzuweisung zwischen Hochschulleitung und Fakultät	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Formelgebundene/indikatorenbasierte Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zielvereinbarungen zwischen Hochschulleitung und Fakultät	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zielvereinbarungen zwischen Fakultät und Institut/Lehrstuhl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Benchmarking verschiedener Lehrstühle innerhalb der Fakultät	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sonstige:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>



Bitte geben Sie nun an, wie stark die Indikatoren innerhalb der folgenden Steuerungsverfahren und Anreizsysteme gewichtet werden. Wenn ein Indikator in dem genannten Verfahren nicht zum Einsatz kommt, kreuzen Sie bitte „Indikator wird nicht eingesetzt“ an.

Anmerkung: Zur Beantwortung der Fragen ist es nicht erforderlich, dass Ihnen die exakten Gewichtungen (z.B. Gewichtsformel innerhalb formelgebundener Mittelzuweisung) bekannt sind. Wenn Sie die genaue Gewichtung der Indikatoren innerhalb der Steuerungsverfahren nicht kennen, geben Sie bitte an, wie stark die Indikatoren innerhalb der Steuerungsverfahren Ihrem subjektiven Empfinden nach gewichtet werden.

Gewichtung der Indikatoren innerhalb der formelgebundenen/indikatorenbasierten Mittelzuweisung zwischen Hochschulleitung und Fakultät:

	sehr geringe Gewichtung					Indikator wird nicht eingesetzt	weiß nicht
	1	2	3	4	5		
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Gewichtung der Indikatoren innerhalb der formelgebundenen/indikatorenbasierten Mittelzuweisung zwischen Fakultät und Institut/Lehrstuhl:

	sehr geringe Gewichtung					Indikator wird nicht eingesetzt	weiß nicht
	1	2	3	4	5		
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>

60%

Gewichtung der Indikatoren innerhalb von Zielvereinbarungen zwischen Hochschulleitung und Fakultät:

	sehr geringe Gewichtung				sehr hohe Gewichtung	Indikator wird nicht eingesetzt	weiß nicht
	1	2	3	4			
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Gewichtung der Indikatoren innerhalb von Zielvereinbarungen zwischen Fakultät und Institut/Lehrstuhl:

	sehr geringe Gewichtung				sehr hohe Gewichtung	Indikator wird nicht eingesetzt	weiß nicht
	1	2	3	4			
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>

67%

Gewichtung der Indikatoren innerhalb von Benchmarking verschiedener Lehrstühle innerhalb der Fakultät:

	sehr geringe Gewichtung				sehr hohe Gewichtung	Indikator wird nicht eingesetzt	weiß nicht
	1	2	3	4			
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Gewichtung der Indikatoren innerhalb der Leistungszulagen im Rahmen der W-Besoldung:

	sehr geringe Gewichtung				sehr hohe Gewichtung	Indikator wird nicht eingesetzt	weiß nicht
	1	2	3	4			
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>

73%

Gewichtung der Indikatoren innerhalb von **Bleibe-/Berufungsverhandlungen:**

	sehr geringe Gewichtung				sehr hohe Gewichtung	Indikator wird nicht eingesetzt	weiß nicht
	1	2	3	4	5		
Anzahl Publikationen in Zeitschriften mit Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften ohne Peer-Review-Verfahren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl anderer Publikationen (Monografien, Beiträge in Sammelbänden, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl Publikationen in Zeitschriften mit hohem Impact Factor / in high ranked journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl der Zitationen / Zitationsindizes (z.B. h-Index)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzahl betreuter Promotionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der eingeworbenen Drittmittel der Deutschen Forschungsgemeinschaft	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Höhe der sonstigen eingeworbenen Drittmittel (z.B. EU, BMBF, Stiftungen, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<< Zurück

Weiter >>

80%

Zum Abschluss möchte ich Sie bitten, ein paar kurze Fragen zu Ihrer Person zu beantworten. Diese Angaben sind natürlich ebenfalls völlig freiwillig.

Welches Geschlecht haben Sie?

- weiblich
 männlich

In welchem Jahr sind Sie geboren?

Bitte wählen Sie Ihr Geburtsjahr aus der vorgegebenen Liste.

Auf welchem Weg haben Sie sich für die Professorenlaufbahn qualifiziert?

Mehrfachauswahl möglich

- Juniorprofessur
 Habilitation
 sonstiges:

<< Zurück

Weiter >>

87%

Wie lange sind Sie bereits als Professor/in tätig?

Bitte geben Sie eine Zahl oder Kommazahl ein.

ca. Jahre

In welchem Bundesland liegt die Universität, an der Sie beschäftigt sind?

Bitte wählen Sie das Bundesland aus der vorgegebenen Liste.

Befindet sich Ihre Universität in staatlicher, kirchlicher oder privater Trägerschaft?

- staatlich
 kirchlich
 privat
 sonstige:

<< Zurück

Weiter >>

93%

Wenn Sie noch weitere Anmerkungen zum Thema oder zur Befragung an sich machen möchten, haben Sie hier Gelegenheit dazu:

Klicken Sie nun auf "Weiter", um den Fragebogen abzusenden.

<< Zurück

Weiter >>

100%

Vielen Dank für Ihre Zeit und Mühe!

Sie können das Fenster nun schließen.

II. E-Mail Anschreiben

[PLATZHALTER ANREDE]

für eine Studie zum Thema „Indikatorenbasierte Steuerungs- und Anreizverfahren und Forschungsverhalten von Professorinnen und Professoren“, welche im Rahmen eines Dissertationsprojektes am Institut für Erziehungswissenschaften der Universität Tübingen durchgeführt wird, bitte ich Sie um Ihre Unterstützung durch die Teilnahme an dieser Online-Befragung.

Die Befragung thematisiert zum einen die Verbreitung und Ausgestaltung von indikatorenbasierten Steuerungs- und Anreizverfahren an Ihrer Fakultät (z. B. Zielvereinbarungen oder leistungsorientierte Mittelvergabe), zum anderen wird Ihre persönliche Einschätzung zu Forschungsindikatoren und deren Stellenwert innerhalb der Scientific Community sowie in Ihrer alltäglichen Forschungsarbeit erfragt.

Ziel der Studie ist es, mögliche Zusammenhänge zwischen der Intensität des Einsatzes fakultätsinterner Steuerungs- und Anreizsysteme und dem Forschungsverhalten von Professorinnen und Professoren zu ermitteln.

Die Befragung ist natürlich freiwillig und erfolgt anonym. Die Ergebnisse werden nur in aggregierter Form ausgewertet, sodass keine Rückschlüsse auf Ihre Person möglich sind.

Das Ausfüllen des Fragebogens wird, je nach Filterführung, etwa 10-15 Minuten Ihrer Zeit in Anspruch nehmen.

Bitte klicken Sie nun auf den folgenden Link oder kopieren Sie ihn in die Adresszeile Ihres Browsers, um mit der Befragung zu beginnen:

https://ww3.unipark.de/uc/umfrage_steuerung/

Die Befragung ist bis einschließlich 06.12.2015 für Sie freigeschaltet.

Wenn Sie Fragen zur Thematik oder zum Fragebogen an sich haben, melden Sie sich gerne per E-Mail (lorenz.tafertshofer@uni-tuebingen.de) oder Telefon (XXXXXX XX XXXXX).

Erste Ergebnisse der Studie werden voraussichtlich Ende 2016 verfügbar sein und können auf Wunsch gerne bei mir angefordert werden.

Vielen Dank für Ihre Unterstützung!

Mit freundlichen Grüßen

Dipl.-Soz. Lorenz Tafertshofer

III. Histogramme der Residuen (Multiple Regression)

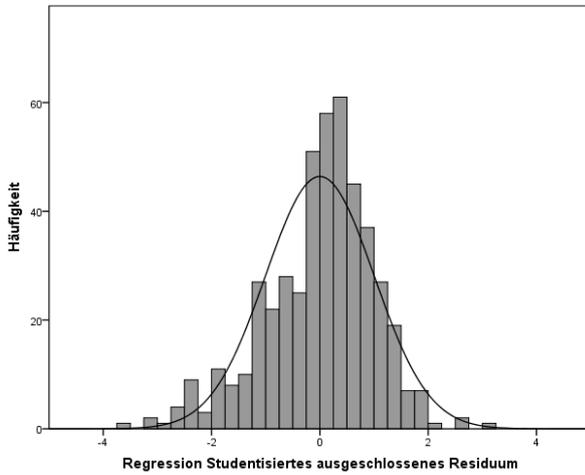
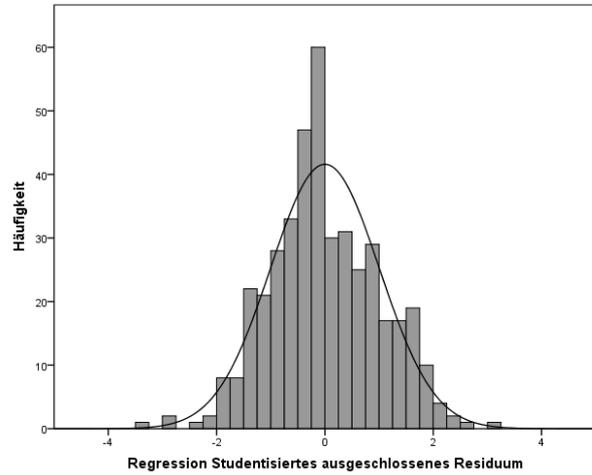
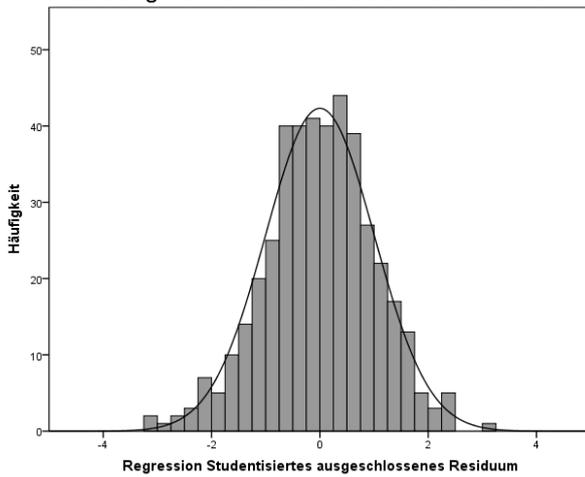
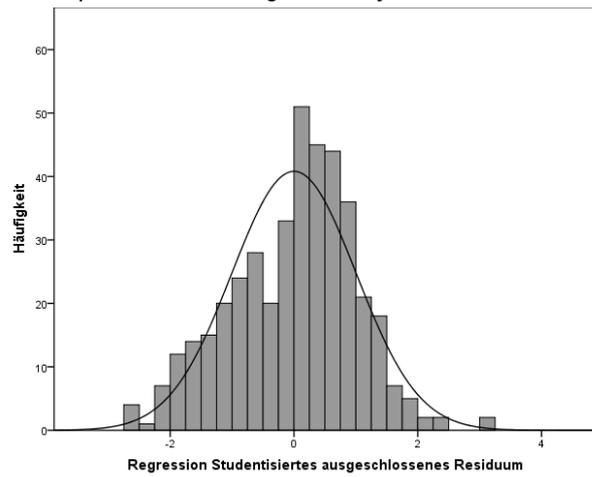
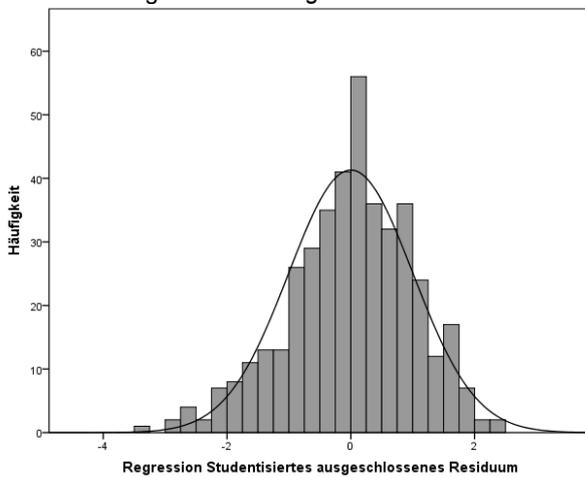
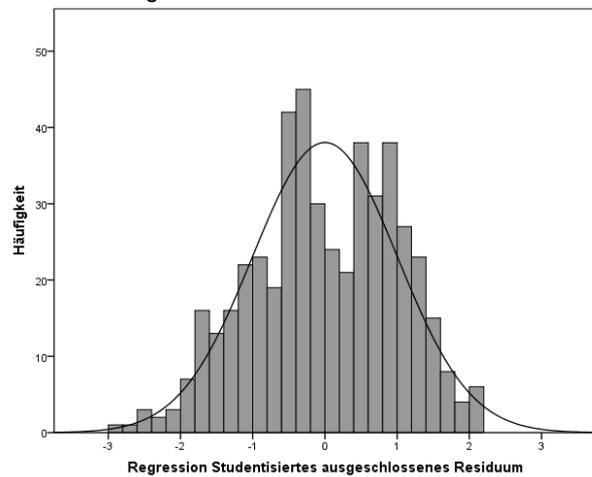
Abbildung A1: *Publizieren in Zeitschriften mit Peer-Review-Verfahren*Abbildung A2: *Publizieren in Zeitschriften ohne Peer-Review-Verfahren*Abbildung A3: *Verfassen anderer Publikationen*Abbildung A4: *Publizieren in Zeitschriften mit hohem Impact-Factor / in high ranked journals*Abbildung A5: *Betreuung von Promotionen*Abbildung A6: *Einwerben von Drittmitteln der DFG*

Abbildung A7: Einwerben von sonstigen Drittmitteln

