

# MASTERARBEIT

im Masterstudiengang Politikwissenschaft  
gemäß der Studien- und Prüfungsordnung des Fachbereichs Politik- und Sozialwissenschaften  
der Freien Universität Berlin vom 8. Mai 2013

**Evidenzbasierte Politik in postfaktischen Zeiten –**  
Eine empirische Potentialanalyse wissenschaftlicher Evaluationen zur  
Steigerung der Evidenzbasierung in politischen Entscheidungsprozessen  
öffentlicher Institutionen

*Erstbetreuer: Prof. Dr. Thorsten Faas*

*Zweitbetreuer: Prof. Dr. Dieter Ohr*

**Vorgelegt von:**

Marisa Brecht

Berlin, am 21.09.2020

## **ABSTRACT**

Angesichts des omnipräsenten Drucks öffentlicher Institutionen, Regulationsqualität erhöhen sowie Regulationskosten senken zu müssen, sind Evaluationen als Methode evidenzbasierter Politik auf dem Vormarsch. Evidenz und Politik sollen demnach weniger in einem Spannungsverhältnis, sondern in einem komplementären Arbeitsverhältnis stehen, um Entscheidungsprozesse faktenbasierter und objektiver zu gestalten. Doch der Trend evidenzbasierter Evaluierungen steht der Debatte um eine zunehmende Postfaktizität in postmodernen Demokratien diametral entgegen. Vor dem Hintergrund dieses vermeintlichen Paradoxons wird im Rahmen der vorliegenden Forschungsarbeit untersucht, inwiefern evidenzbasierte Evaluationen in postfaktischen Zeiten das Potential haben, auf die politische Dynamik öffentlicher Entscheidungsprozesse Einfluss zu nehmen. Knapp 20 qualitative Leitfadeninterviews mit Expert\*innen der operativen Arbeitsebene liefern die empirische Basis einer Potentialanalyse und einer Systematisierung notwendiger Erfolgsbedingungen. Es wird aufgezeigt, dass Evaluationen eine evidenzbasierte Diskurs- und Entscheidungsgrundlage für politisch-normative Entscheidungen darstellen, wenngleich dies nicht zwangsläufig mit einer Rationalisierung oder De-Politisierung einhergeht. Die empirischen Ergebnisse stützen ein Plädoyer für die stärkere systemübergreifende Präsenz evaluativer Evidenz zugunsten integrierter und transparenter Entscheidungsprozesse.

# Inhaltsverzeichnis

<b>1. EINLEITUNG: EVALUATIVE EVIDENZ IN POSTFAKTISCHEN ZEITEN</b> .....	<b>1</b>
<b>2. THEORETISCHE HERLEITUNG UND STATE OF THE ART</b> .....	<b>6</b>
2.2 GRUNDLEGENDE KONZEPTE .....	7
2.2.1 <i>Evidenzbasierte Politik: Das epistemologische (Miss-)Verhältnis zwischen Evidenz und Politik</i> .....	7
2.2.2 <i>Postfaktizität: Buzzword oder postmoderner Zeitgeist?</i> .....	12
2.2 STAND DER FORSCHUNG.....	15
2.2.1 <i>Evaluierte Perspektive: Die Nachfrage des politischen Systems</i> .....	17
2.2.1.1 Entstehungs- und Entwicklungsgeschichte der veränderten Nachfrage.....	17
2.2.1.2 Institutionalisation, legislative Rahmenbedingungen und Umsetzungsdefizit.....	18
2.2.1.3 Ziele und Nutzungsfunktionen von Evaluationen in politischen Entscheidungsprozessen.....	20
2.2.1.4 Sollbruchstellen und hemmende Faktoren von und für Evaluierungen im politischen Kosmos.....	23
2.2.1.5 Hinreichende Bedingungen, begünstigende Faktoren und das reale Potential von Evaluierungen.....	26
2.2.2 <i>Evaluierende Perspektive: Das Angebot des wissenschaftlichen Systems</i> .....	29
2.2.2.1 Typologisierung, Politikfelder und Akteurslandschaft.....	29
2.2.2.2 Nicht realisiertes Potential und Änderungsbedarf in der Scientific Community .....	32
2.2.3 <i>Zivilgesellschaftliche Perspektive: Die Partizipation des sozialen Systems</i> .....	33
2.3 FORSCHUNGSFRAGEN UND HYPOTHESEN.....	36
<b>3. METHODIK UND FORSCHUNGSDESIGN</b> .....	<b>39</b>
3.1 SEMISTRUKTURIERTE LEITFADENINTERVIEWS.....	39
3.2 SAMPLING: AUSWAHL DER STICHPROBE.....	40
3.3 AUFBAU UND INHALT DES LEITFADENS .....	42
3.4 SETTING: VORBEREITUNG UND DURCHFÜHRUNG DER INTERVIEWS .....	43
3.5 DATENAUFBEREITUNG UND ANALYSEMETHODE.....	44
<b>4. FORSCHUNGSERGEBNISSE</b> .....	<b>47</b>
4.1 EVALUATION ALS EVIDENZBASIERTE DELIBERATIONSGRUNDLAGE FÜR POLITISCHE ENTSCHEIDUNGEN .....	47
4.1.1 <i>H1: Die verstärkte Evidenzbasierung öffentlicher Entscheidungsprozesse</i> .....	47
4.1.2 <i>H2: Die vermeintliche Rationalisierung und De-Politisierung öffentlicher Entscheidungsprozesse</i> .....	49
4.2 ERFOLGSBEDINGUNGEN FÜR EVIDENZBASIERTE ENTSCHEIDUNGEN IN POLITIK, WISSENSCHAFT UND GESELLSCHAFT .....	53
4.2.1 <i>Das politische System: Evaluations- und Fehlerkultur</i> .....	53
4.2.2 <i>Das wissenschaftliche System: Knowledge Brokering</i> .....	55
4.2.3 <i>Das soziokulturelle System: Bürger*innennahe Vermittlung</i> .....	57
4.3 DISKUSSION UND PRAKTISCHE IMPLIKATIONEN DER FORSCHUNGSERGEBNISSE .....	59
<b>5. FAZIT: KÄMPFE, DIE EVALUATIONEN (NICHT) KÄMPFEN MÜSSEN</b> .....	<b>63</b>

<b>6.</b>	<b>LITERATURVERZEICHNIS.....</b>	<b>67</b>
<b>7.</b>	<b>ANHANG .....</b>	<b>86</b>
7.1	ANFRAGE EXPERT*INNENINTERVIEWS .....	86
7.2	LEITFADEN: EVALUIERENDE PERSPEKTIVE.....	87
7.3	LEITFADEN: EVALUIERTE UND AUFTRAGGEBENDE PERSPEKTIVE.....	88
7.4	TRANSKRIPTE EXPERT*INNENINTERVIEWS .....	89
7.4.1	<i>Expertin 1</i> .....	89
7.4.2	<i>Expertin 2</i> .....	95
7.4.3	<i>Expertin 3</i> .....	101
7.4.4	<i>Experte 4</i> .....	107
7.4.5	<i>Experte 5</i> .....	113
7.4.6	<i>Experte 6</i> .....	129
7.4.7	<i>Expertin 7</i> .....	139
7.4.8	<i>Experte 8</i> .....	148
7.4.9	<i>Expertin 9</i> .....	160
7.4.10	<i>Experte 10</i> .....	170
7.4.11	<i>Experte 11</i> .....	180
7.4.12	<i>Experte 12</i> .....	195
7.4.13	<i>Experte 13</i> .....	204
7.4.14	<i>Experte 14</i> .....	221
7.4.15	<i>Experte 15</i> .....	230
7.4.16	<i>Expertin 16</i> .....	244
7.4.17	<i>Experte 17</i> .....	256
7.6	MAXQDA CODEBUCH .....	267
<b>8.</b>	<b>EIDESSTAATLICHE ERKLÄRUNG .....</b>	<b>281</b>

## **Abbildungsverzeichnis**

Abbildung 1: Übersicht der befragten Expert*innen.....	40
Abbildung 2: Theoretisches Idealmodell eines Evaluierungsprozesses .....	59

# 1. Einleitung: Evaluative Evidenz in postfaktischen Zeiten

*“(...) governments are turning to evidence of performance for legitimacy since it is no longer guaranteed solely by democratic political processes”*  
(Sanderson 2002: 2)

Nicht erst seit der Corona-Pandemie ist eine Debatte über die politische Nutzung und Verlässlichkeit wissenschaftlicher Erkenntnisse und Evidenzen in einer sich rasant ändernden Welt entfacht (Polisphere 2020, Zeit 2020). Die Sehnsucht nach einfachen Erklärungen ist groß, die Illusion einer Objektivität ebenso. Betrachtet man postfaktische Anekdoten wie die US-amerikanische Präsidentschaft, die Kampagne zum Brexit-Referendum, Klimawandel-Leugnungen oder Covid19-Verschwörungstheorien, so ist der politisch Beobachtende versucht, die Frage, ob wissenschaftliche Evidenz und Tagespolitik sonderlich gut zusammenpassen, mit einem nüchternen Nein zu beantworten.

Auf dem Ziel, Evidenz und Politik zu vereinen, basiert der Ansatz evidenzbasierter Politikgestaltung (u.a. Davies 2004; 2012, Sanderson 2002, Weiss 1999). Regierungen stehen angesichts Globalisierung, Digitalisierung, Klimakrise und demographischem Wandel vor enormen Herausforderungen bei gleichzeitig knappen öffentlichen Ressourcen. *Evidence-based policymaking* oder auch *evidence-informed policymaking* strebt an, politische Entscheidungsprozesse durch den Einsatz wissenschaftlicher und empirischer Evidenzen faktenbasierter, objektiver und transparenter zu gestalten. Steuergelder sollen effizienter eingesetzt, Bundesgesetze effektiver umgesetzt und staatliche Leistungen optimiert werden. Obwohl die Anfänge jenes Vorhabens lange zurückliegen, setzte sich die Idee und das Mantra *„what matters is what works“* zuerst 1997 unter der britischen New Labour Regierung durch – mit dem gezielten Bruch *„with past administrations, in the sense that policy making should be pragmatic, based on what works best, rather than ideological position“* (Campbell 2002: 89). Die Bewegung ist der anhaltende Optimismus bezüglich des Potentials, sozialen Fortschritt durch die Anwendung rationaler Vernunft zu erreichen (Sanderson 2002). Evidenzbasierte Politik soll versachlichen und rationalisieren sowie die gesellschaftliche Entwicklung auf Basis von geprüftem und belastbarem Wissen gestalten und vorantreiben (Head 2010). Die Prämisse des modernistischen Projekts ist demnach ein durch wissenschaftliche Erkenntnis angetriebener soziopolitischer Fortschritt (Bronk 1999).

Der Einsatz systematischer Evaluationen stellt neben Impact Assessments, Politikfolgenabschätzungen und Spending Reviews nur eine von vielen Methoden einer evidenzbasierten Politik dar. Evidenzbasierte Evaluation ist ein evidenz-, indikator- und datenbasiertes Bewertungsinstrument zur Überprüfung und Analyse politischen Handelns und dessen Wirksamkeit; welches sowohl arbeitspolitische Maßnahmen auf Bundesebene als auch kommunale Förderprojekte, entwicklungspolitische Programme oder Landesgesetze betreffen kann. Idealistische Positionen zeigen Evaluation als Instrument für rationalere, nachhaltigere und gemeinwohlorientiertere Entscheidungen. Mit dem Ziel, Regulationsqualität zu erhöhen sowie Regulationskosten zu senken, haben sich seit der Jahrtausendwende in allen Ländern der OECD Evaluationssysteme institutionalisiert (Weiland 2013). Doch jener Trend evidenzbasierter Evaluierungen (Leeuw/Furubo 2008) scheint der Wahrnehmung einer zunehmenden Postfaktizität in Politik und Gesellschaft diametral gegenüber zu stehen (vgl. Buch et al. 2019, Dobes 2017; 2018, Harsin 2015, Picciotto 2019). Diesem vermeintlichen Paradox der beiden Phänomene soll im Rahmen dieser Forschungsarbeit ausreichend Raum gegeben werden. Es wird die zentrale Forschungsfrage untersucht, inwiefern evidenzbasierte Evaluationen in postfaktischen Zeiten das Potential haben, auf die politische Dynamik öffentlicher Entscheidungsprozesse Einfluss zu nehmen.

Im Rahmen der vorliegenden Arbeit wird sich insbesondere auf summative bzw. ex post Evaluierungen in der Abschlussphase eines Policy Cycles konzentriert. Diese bilanzieren, ob eine politische Maßnahme die jeweilige beabsichtigte Wirkung hatte und dabei Einflüsse anderer Faktoren weitgehend ausschließen sowie nichtbeabsichtigte Wirkungen identifizieren können. Von evaluierender Evidenz wird erwartet, dass sie Politik kohärenter macht, die Effizienz ihrer Umsetzung erhöht sowie ihre Wirksamkeit steigert. Kernfragen in der Evaluationsforschung sind demnach: Was ist evaluative Evidenz und wann, wie, wo und von wem sollte sie in politische Entscheidungsprozesse einbezogen werden? Während sich Literatur zur Evaluationsdurchführung oder -implementierung oft auf eine abhängige Nutzungsvariable konzentriert und eine analytische Haltung einnimmt, ist die evidenzbasierte Politikliteratur regelrecht Teil einer politischen Bewegung und geht damit mit einem durchaus normativen Anspruch einher, d.h. der Evidenzbasierung oder gar Rationalisierung von politischen Entscheidungsprozessen. Während eine zu untersuchende Evidenzbasierung politischer Prozesse das Ziel hat, im Wettkampf mit verschiedensten Einfluss- und Entscheidungsfaktoren vermehrt Fakten und Evidenzen Bedeutungskraft zu verschaffen, geht der angestrebte Prozess einer Rationalisierung oder gar De-Politisierung einen Schritt weiter: Es wird intentioniert,

evidenz- und indikatorgeleitete sowie wertneutrale Entscheidungsprozesse von politischen Entscheidungsträger\*innen zu generieren.

Die Forschung hinkt in diesem aufsteigenden Feld allgemein hinterher (Stephenson 2019); einen Mangel an aktuellen Studien gibt es insbesondere in Bezug auf die demokratische Nutzung von Evaluationsergebnissen in politischen Prozessen (Sager 2018) sowie den de facto Einfluss von Evaluationen auf Politikgestaltungsprozesse und deren Ergebnisse (Hertin et al. 2009). Auch die Debatte über das Verhältnis von Demokratie, Postfaktizität und Evaluation ist längst nicht abgeschlossen (Widmer 2009). Zur Kontextualisierung der vorliegenden Arbeit wird dementsprechend sowohl auf die klassische erkenntnistheoretische Differenzierung zwischen positivistischen sowie konstruktivistischen Ansätzen eingegangen als auch auf Theorien sozialer Differenzierung in Bezug auf die Systeme Wissenschaft und Politik sowie auf Evaluationsforschung zu u.a. Funktionslogiken, Nutzungsmotiven und politischer Einflussnahme.

Zur Überprüfung der Forschungsfrage wurden neben einer umfangreichen Literatur- und Dokumentenrecherche in einer zweimonatigen Datenerhebungsphase knapp 20 semistrukturierte und qualitative Leitfadeninterviews mit Expert\*innen und Praktiker\*innen der direkt betroffenen, operativen Arbeitsebene durchgeführt. Die befragten Evaluierenden mit teilweise jahrzehntelanger Evaluationserfahrung waren sowohl bei öffentlichen Forschungsinstituten, Internationalen Organisationen als auch internationalen „Big 5“-Managementberatungen beschäftigt. Befragte mit einer evaluierten oder auftraggebenden Perspektive arbeiteten hingegen für staatliche Stellen wie Ministerien oder öffentlich finanzierte Träger wie Stiftungen, Vereine oder Förderbanken. Die praxisnahen Einblicke der Expert\*innen lieferten die Basis für die vorliegende Analyse des Evaluationspotentials und die Systematisierung notwendiger Erfolgsvoraussetzungen. Sowohl die theoretische Herleitung als auch die empirische Analyse berücksichtigen die für die Forschungsfrage wichtige Differenzierung der drei Systeme: die evaluierte bzw. auftraggebende Perspektive, die die Nachfrage des politischen Systems symbolisiert; die evaluierende bzw. auftragsdurchführende Perspektive, die das Angebot des wissenschaftlichen Systems umfasst und die zivilgesellschaftliche Perspektive, d.h. die Partizipation des sozialen Systems repräsentiert.

Auf dem Forschungsstand und den Expert\*inneninterviews aufbauend, soll mit der vorliegenden Arbeit im Rahmen eines eklektizistischen Ansatzes und einer Fokussierung auf die abhängige

Variablen dreierlei geleistet werden: *Erstens*, eine empirische Potentialanalyse evidenzbasierter Evaluationen, indem bisherige empirische Erkenntnisse überprüft und mit aktuellen, eigenständig erhobenen Daten abgeglichen und neu systematisiert werden; *zweitens*, die Selektion, Strukturierung und Auswertung von Gemeinsamkeiten der Expert\*innenaussagen, um konkrete (Miss-)Erfolgsfaktoren zu identifizieren und darauf aufbauend praktischen Handlungsbedarf aufzuzeigen; *drittens*, eine Analyse der bislang vernachlässigten Schnittstelle zwischen Evaluation und Postfaktizität.

In der Forschungsarbeit wird gezeigt, dass Evaluationen eine evidenzbasierte Diskurs- und Entscheidungsgrundlage für politisch-normative Entscheidungen darstellen. Eine klare Mehrheit der Expert\*innen bestätigte das Potential von Evaluationen für eine zunehmende Evidenzbasierung (H1), dies geht jedoch nicht zwangsläufig mit einer Rationalisierung oder De-Politisierung und einer Ablöse dogmatischer oder ideologischer Kämpfe einher (H2). Es wird deutlich, dass die Befragten überzeugt davon sind, dass der Politik das Politische zu nehmen, kein Kampf ist, der durch Evaluierungen gefochten werden kann und sollte. Doch indem sie zu einer verstärkten Arbeitsteilung zwischen Politik und Wissenschaft und zu einer gesteigerten Kohärenz der beiden Systeme beitragen, können Evaluationen als vorgelagerte evidenzbasierte Voraussetzung für darauf aufbauende meinungsbasierte Diskurse angesehen werden und als neutrale Evidenzbasis für politische Entscheidungen dienen. Die Forschungsergebnisse zeigen auf, inwiefern Evaluationserzeugnisse die politische Dynamik bereichern, ideologische und interessenbasierte Ansichten in Frage stellen und so die Qualität des politischen Diskurses in demokratischen Prozessen verbessern (vgl. Pollitt 2006, Sager 2018, Schlaifer 2017). Zudem wurden Erfolgsbedingungen zur langfristigen Nutzung von Evaluationsergebnissen erarbeitet und analysiert, wie sie über die kurzfristige und instrumentelle Nutzung hinaus ihren Weg in den politischen Diskurs finden können. Praktische Implikationen waren in diesem Zusammenhang drei zentrale Voraussetzungen für die erfolgreiche Beauftragung, Durchführung und Nutzung von Evaluierungen: Der Aufbau einer Evaluations- und Fehlerkultur für progressive institutionelle Lernprozesse, Knowledge Brokering für eine effektivere und beidseitig profitable Zusammenarbeit des wissenschaftlichen und politischen Systems sowie eine bürger\*innennahe Vermittlung durch mediale Kanäle und öffentliche Partizipationsformate. Auf den Forschungsergebnissen aufbauend zeichnet sich ein Plädoyer für eine verstärkt zu nutzende Komplementarität der Systeme Wissenschaft und Politik und für eine stärkere Präsenz von Evaluation in Demokratien über die Nutzung durch unmittelbar Involvierte hinaus.

Die Ausarbeitung der Forschungsfrage ist wie folgt gegliedert und strukturiert: Zu Beginn wird die Forschungsthematik theoretisch beleuchtet. Dies umfasst die Vorstellung der grundlegenden Konzepte, d.h. evidenzbasierte Politik und das epistemologische (Miss-)Verhältnis zwischen Politik und Gesellschaft sowie das Phänomen Postfaktizität. Es folgt die Analyse des Forschungsstands im Rahmen dessen die bisherigen theoretischen und empirischen Erkenntnisse zur evaluierten, evaluierenden und zivilgesellschaftlichen Perspektive gesichtet, selektiert, kategorisiert und bewertet werden sowie die Herleitung der Forschungsfragen und Hypothesen. Im Anschluss wird das Untersuchungsdesign mitsamt der Datengrundlage und der angewandten Methodik vorgestellt, d.h. das Sample und Setting der Leitfadeninterviews sowie die Datenaufbereitung und Analysemethode. Es folgt die Beschreibung, Analyse und Interpretation der Forschungsergebnisse sowie die Überprüfung der Hypothesen, die Analyse der Erfolgsbedingungen und die Diskussion praktischer Implikationen. Das Fazit inklusive einer kritischen Reflexion des Erkenntnisgewinns und bestehender Limitationen werden am Ende der vorliegenden Arbeit dargestellt. Im Anhang finden sich aus Transparenz- und Nachvollziehbarkeitsgründen die Interviewanfrage, die beiden Interviewleitfäden, die Interviewtranskripte sowie das Codebuch mitsamt aller Code-Definitionen und -Häufigkeiten.

## 2. Theoretische Herleitung und State of the Art

Die Relevanz von Evaluierungen hat als Instrument der Erfolgs- und Wirkungskontrolle sowohl national als auch international kontinuierlich zugenommen. Internationale Meilensteine wie der rasante Anstieg an Evaluationen innerhalb der EU in den 90ern (Pollitt 1998, Stern 2009), die Pariser Erklärung (OECD 2005), die Accra Agenda for Action (OECD 2008) oder die EU Better Regulation Guidelines (Europäische Kommission 2017) haben Evaluierungsvorhaben auch auf deutscher Ebene Vorschub geleistet (Leeuw 2006, Schwab 2009). Doch trotz aller Evaluationsanstrengungen wurden bisher kaum systematische Evaluationen durchgeführt, die die Frage, welche politischen Maßnahmen unter welchen Bedingungen wie funktionieren, präzise beantworten konnten (Hertin et al. 2009, Knill/Tosun 2015, Stephenson 2019). Das BMZ (2008) spricht sogar von einem regelrechten quantitativen und qualitativen „Evaluation Gap“: Aus quantitativer Perspektive werden zu wenige Evaluationen mit Fokus auf Outcomes oder gar Impacts durchgeführt und aus qualitativer Perspektive sind sie häufig methodisch unzureichend oder werden nicht systematisch durchgeführt.

Angesichts der Tatsache, dass Evaluierungen auf dem Vormarsch sind und deren Systematisierung, Professionalisierung und Institutionalisierung stetig steigt, hängt die Forschung in diesem aufsteigenden Feld zurück (Stephenson 2019). Hertin et al. (2009) konstatiert, dass während der Aufbau und die Verbreitung evaluativer Verfahren im Mittelpunkt einer Reihe von vergleichender Studien stand (z.B. OECD 1997; 2004, Radaelli 2005), die vorhandene Literatur wenig darüber aussagt, welchen Einfluss sie auf die tatsächlichen Prozesse und Ergebnisse der Politikgestaltung haben sowie was ihr genuiner Sinn und Zweck ist. Auch Knill und Tosun (2015) bestätigen den Mangel an aktueller Forschung zum Effekt von Evaluationsaktivitäten auf die Gestaltung und Umsetzung von Policies, also deren genuine Funktions- und Leistungsfähigkeit. Es wird betont, dass nicht ausreichend erforscht wurde, wie genau Evaluationsergebnisse von den unterschiedlichen Akteuren aufgegriffen werden und ob sich dadurch die Dynamik politischer Entscheidungen ändert. Auch wurde noch nicht genügend Forschung betrieben, um systematische Vergleiche zu ermöglichen oder die kontemporäre Dynamik durch aktuelle Studien zu erfassen.

## 2.2 Grundlegende Konzepte

Evidenzbasierte Politikgestaltung und das ausgerufenen postfaktische Zeitalter sind zwei zentrale Konzepte für die Ausarbeitung der vorliegenden Forschungsarbeit – konträr und gleichzeitig komplementär. Im Folgenden wird zunächst auf theoretische Ansätze zum epistemologischen Verhältnis zwischen Evidenz und Politik und danach auf das Buzzword Postfaktizität eingegangen, um die Forschungsthematik zu kontextualisieren und einen Bogen zu spannen für die nachfolgende Sichtung und Analyse der bisherigen theoretischen und empirischen Erkenntnisse über evidenzbasierte Evaluierungen.

### 2.2.1 *Evidenzbasierte Politik: Das epistemologische (Miss-)Verhältnis zwischen Evidenz und Politik*

*“an approach that helps people make well-informed decisions about policies, programs, and projects by putting the best available evidence from research at the heart of policy development and implementation”*  
Davies (2004: 3)

Der moderne, demokratische Rechtsstaat ist mit dem Anspruch konfrontiert, nicht nur Gesetze im Sinne des Volkes zu erlassen, sondern diese auch wirksam umzusetzen. Zunehmend komplexe gesellschaftliche Probleme und ein steigender Grad politischer Unsicherheit stehen knappen öffentlichen Mitteln und dennoch umfangreichen Staats- und Wohlfahrtstätigkeiten gegenüber. Die Bedeutung von Wirksamkeitskontrollen öffentlichen Handelns wächst.

Eine mögliche Antwort auf diese Herausforderungen ist evidenzbasierte Politikgestaltung. Sie basiert auf der Idee, das verfügbare, meist wissenschaftlich generierte Wissen verstärkt in den politischen Prozess einzubringen, diesen dadurch zu rationalisieren und auf diese Weise bessere oder zumindest besser fundierte politische Entscheidungen zu ermöglichen (Jun/Grabow 2008, Howlett 2009). Die „politische Rhetorik“ biete eine Alternative zu ideologischer bzw. glaubensbasierter Politikgestaltung und sei gefragter denn je, so Head (2010: 77). Sie baue auf einem stark ökonomisch geprägten Diskurs über eine effizientere und effektivere Regulierung im öffentlichen Raum auf und basiert auf der Erwartungshaltung, dass die Qualität von politischen Entscheidungen durch Evidenzbasierung zunehme (Weiland 2013). Evidenzbasierte Politik sei ein Instrument, um Politik zu versachlichen und zu rationalisieren. Ziel sei es, die gesellschaftliche Entwicklung auf Basis von geprüftem und belastbarem Wissen zu gestalten und voranzutreiben (Head 2010). Dieser Argumentation folgend soll Evidenz

systematisch generiert werden, um die Folgen von Regulation vorauszusagen, dann zu bewerten und dadurch Handlungsbedarf bzw. Optimierungspotential zu identifizieren. Politikfolgenabschätzung und -evaluation stellen den methodischen Rahmen für eine solche evidenzbasierte Politik dar.

Die Bewegung des *evidence-based policymaking* (EBP) sei der anhaltende Optimismus bezüglich des Potentials, sozialen Fortschritt durch die Anwendung rationaler Vernunft zu erreichen (Sanderson 2002). Das modernistische Projekt basiere auf der Annahme eines durch wissenschaftlichen Fortschritt angetriebenen soziopolitischen Fortschritts (Bronk 1999). Es sei letztlich ein Versprechen für einen post-ideologischen Regierungsansatz, ein Ansatz, der Evidenz in das Zentrum politischer Entscheidungsprozesse stellt (Davies et al. 1999: 3). Mit der Debatte über evidenzbasierte Politik wird auch die traditionelle erkenntnistheoretische Debatte neu entfacht, ob und inwieweit politische Bewertungen überhaupt auf objektiven Einschätzungen basieren können oder durch subjektive Wahrnehmungen der involvierten Akteure geprägt sind:

Rationalistisch-positivistische Positionen beziehen sich unter anderem auf die Quantifizierbarkeit und Objektivierbarkeit von Ereigniszusammenhängen, die Möglichkeit allgemeingültiger Gesetze sowie rational agierende Individualakteure mit einer klaren Hierarchie an Präferenzen (vgl. Buchanan/Tullock 1962, Downs 1957). Dieser klassische positivistische Ansatz vertrete ein lineares Modell der politischen Problemlösung (Weiland 2013): Wissenschaftliche Analysen wie Folgeabschätzungen oder Evaluationen liefern evidenzbasierte Inputs mit darauf aufbauenden, wertfreien Handlungsempfehlungen, welche wiederum von rationalen politischen Entscheider\*innen in konkrete Policy-Outputs umgesetzt werden.. Eine bessere Evidenzbasis führe – dieser rationalistisch-optimistischen Erwartungshaltung des EBP folgend – zu fundierteren politischen Analysen und dies wiederum zu einer Erweiterung der Problemlösungskapazitäten von Regierungsorganen (Head 2010). Dies suggeriert, dass Politik darauf ausgerichtet sei, klar identifizierte Probleme anzugehen, es einen zentralen Entscheidungstragenden gäbe, der eine Politikoption auf der Grundlage des erwarteten Nettonutzens auswählt, und dass „bessere“ Informationen notwendigerweise zu einer „rationaleren“ Politik führen (Hertin et al. 2009).

Prämisse hierfür seien rationale und unitäre Entscheidungsträger\*innen, die eher als Manager\*innen als Politiker\*innen fungieren und letztlich jene Policy-Option wählen, die den

objektiv größten Nettonutzen für die Gesellschaft verspricht (Weiland 2013). Diese positivistische Sichtweise fasst wissenschaftliches Wissen als wertfrei, empirisch verifizierbar und als Ergebnis objektiver, quantitativer sowie wissenschaftlicher Methoden zusammen. Die Sichtweise zielt auf einen kumulativen Wissensprozess und verstärkt iterativen Politiklernzyklus ab, der durch die enge Berücksichtigung von Forschungs- und Evaluierungsergebnissen ermöglicht wird (Lerner/Lasswell 1951). Dieser technokratische und vermeintlich apolitische Ansatz ist jedoch starker Kritik ausgesetzt und hat zu intensiven Auseinandersetzungen zwischen Theoretiker\*innen geführt: Sabatier und Jenkins-Smith (1993) betonen, eine rationale und lineare Konzeption des Politikprozesses sei kein empirisch belastbares Modell. Die politische Entscheidungsfindung wäre viel mehr durch Diskontinuitäten, dynamische Veränderungen und eine lose Kopplung zwischen Problemen und Politiken gekennzeichnet. Die Wahl aus einer Bandbreite von verfügbaren politischen Optionen sei Ergebnis institutioneller Pfadabhängigkeiten und komplexer Akteurs- und Interessenkonstellationen. Clarence (2002) kritisiert des Weiteren, dass durch die Priorisierung quantitativer Präzision und technischem Fachwissen stillschweigend Inklusion und Partizipation als Grundlagen für Reformen abgelehnt werden würden.

Viele Autor\*innen fordern auf der Kritik aufbauend neue partizipatorische, diskursive oder kritische Formen der Politikanalyse und -bewertung (deLeon 1992, Dryzek 1990, Fischer 2009, Fischer/Forester 1993, Majone 1989, Saretzki 2012). In einem epistemologischen Kontrast zur rationalistischen Perspektive stehen sozialkonstruktivistische und post-positivistische Positionen (Berger/Luckmann 1966, Crotty 1998, Guba/Lincoln 1989, Stucki 2018). Vertreter\*innen jener Ansätze betonen im Gegensatz dazu die Relativität von Wissen und die diskursive Natur von Politikformulierung. Politik sei ein primär deliberativer Prozess: Bei der Interpretation und Evaluation von Politikoptionen gehe es um das komplexe Zusammenspiel von Werten, Prinzipien und Interessen in unterschiedlichsten Machtkontexten und nicht um faktische Evidenzen (Fischer/Forester 1993, Majone 1989, Radaelli 1995). Evidenz sei nur ein Mittel zum Zweck: Sie werde genutzt, um Entscheidungen öffentlichkeitswirksam vermeintlich zu rationalisieren, die eigentlich auf der Basis anderer Größen getroffen wurden. Politikfolgenabschätzungen und -evaluationen würden durch politische Akteure strategisch genutzt, um ihre jeweiligen Positionen zu stärken und ihre politischen Entscheidungen als vermeintliche neutrale wissenschaftliche Ergebnisse zu legitimieren. Diese postmoderne Kritik verwirft die Vorstellung einer objektiven, messbaren Welt. Selbst empirisches Wissen sei sozial konstruiert, kulturell kontingent und abhängig von einer "radikalisierten Subjektivität":

*“We find ourselves abroad in a world in which social theory and analysis is no longer able, with any credibility, to provide a warrant for political practice and ethical decision making”*

Smart (1999: 63)

Der quantitativen Herangehensweise wird von Vertreter\*innen konstruktivistisch-erkenntnistheoretischer Positionen ein begrenztes Potential zugesprochen. Wissenschaftliches Wissen könne keinen Anspruch auf Objektivität erheben, sondern viel mehr eine breitere Aufklärung fördern. Politikentwicklung wird als ein deliberativer Prozess gesehen, der multiple Überzeugungen, Prinzipien und Handlungen unter multiplen Bedingungen abwägt, die jeweils mehrere Rahmen für die Interpretation und Bewertung der Welt bieten (Dryzek 1990, van der Knaap 1995). Politische Lernprozesse seien sozial bedingt, diskursiv oder argumentativ und würden die Entscheidungsträger\*innen lediglich in unsystematischen und diffusen Formen erreichen.

Owens et al. (2004) lehnt die Polarisierung zwischen dem technisch-rationalen Modell und einer post-positivistischen Kritik sowohl aus empirischen als auch aus theoretischen Gründen ab. Realismus biete einen Mittelweg zwischen den überoptimistischen Ansprüchen des rationalistisch-objektivistischen Ansatzes und dem überpessimistischen Nihilismus der Sozialkonstruktivisten (Sanderson 2002). Die realistische Tradition erlebt ein gewisses Revival im Kontext des Bestrebens, Politik evidenzbasierter zu gestalten (Archer 1995, Archer et al. 1998, Kirk 1999, Pawson/Tilley 1997, Searle 1995, Trigg 2001). Der „neue“ Realismus legt nahe, dass involvierte Akteure trotz eines evidenzbasierten Wandels, der sowohl wünschenswert als auch möglich ist, nicht erwarten können, ein politisches System aufzubauen, das sich in erster Linie auf objektive Forschungsergebnisse stützt (Head 2010). Politische Entscheidungen in der realen Welt werden danach nicht aus empirisch-analytischen Modellen abgeleitet, sondern aus Politik und praktischem Urteilsvermögen sowie aus dem Zusammenspiel von Fakten, Normen und Handlungsoptionen. Was als Evidenz gilt, sei nicht universell, sondern vielfältig und anfechtbar. Dieser Argumentation folgend bediene sich der politische Entscheidungsprozess in demokratisch-westlichen Ländern der Rhetorik der rationalen Problemlösung und der Effektivität des Politikmanagements, aber der politische Prozess selbst sei und bleibe unscharf und konflikthaft – schlichtweg politisch.

Sanderson (2002) kommt zu einer alle drei Ansätze einschließenden Schlussfolgerung: Zwar untergrabe die zunehmende Komplexität der sozialen Systeme allmählich jegliche Gewissheit in Bezug auf soziales Wissen, doch jene Charakteristiken der postmodernen Welt, die das

„*bonfire of the certainties*“ erzeugt haben, würden nicht bedeuten, dass analytische oder kognitive Grundlagen für politisches Entscheiden und Handeln nicht mehr nötig seien (Sanderson 2002: 7). Es bedeute auch nicht, dass eine systematische Kenntnis des menschlichen Handelns nicht möglich sei, sondern viel mehr, dass keine fundamentalistischen Garantien gemacht werden könnten. Mit einem realistischen Ansatz in der Politik- und Programmevaluierung in Verbindung mit Governance-Prozessen, die eine ernsthafte Verpflichtung verkörpern, aus einer solchen Evaluierung lernen zu wollen, bestehe die Chance, eine rationale Grundlage für die Lenkung sozialen Wandels zu kollektiv gewünschten Zielen zu schaffen.

Auch Weiland (2013) stellt im Vergleich der erkenntnistheoretischen Politikmodelle die Persistenz des rationalistischen Ansatzes und den Optimismus bezüglich der Rolle rationalen Wissens fest. Trotzdem appelliert sie für eine Abkehr von einem rationalistischen Governance-Modell und einem Streben nach rationaler Politik zugunsten eines diskursiven Modells und einer Entwicklung hin zu offeneren Governance-Strukturen. Es müsse ein Mittelweg zwischen beiden Modellen der Wissensbasierung getroffen werden (Fischer 1995). In Abhängigkeit zum jeweiligen Kontext und den jeweiligen Anforderungen sei entweder eindeutiges Faktenwissen oder reflexives Wissen vonnöten. Der technokratische Ansatz, der auf eine Rationalisierung von politischen Entscheidungen durch den Gebrauch von wissenschaftlichem Wissen abziele, würde einen Beitrag zu einer gesteigerten prozeduralen Legitimität leisten. Legitimität von Regulierung hänge schließlich wesentlich von glaubwürdigen und transparenten Verfahren ab. Evidenzbasierte Methoden würden Gründe für Regulations- und Politikentscheidungen angeben sowie Abwägungsprozesse offenlegen. Eine Evidenzbasierung solle aber vielmehr als Möglichkeitsstruktur verstanden werden, welches das Potential biete, neo-korporative und elitäre Politiknetzwerke aufzubrechen. Evidenzbasierte Methoden und Verfahren können demzufolge Vehikel sein, um Wissen aus unterschiedlichsten Quellen zusammenzutragen und konfligierende Werte dadurch sichtbar zu machen. Dieser neo-pluralistische Ansatz hat zum Ziel, einer Vielzahl von unterschiedlich starken Akteuren die Möglichkeit zu verschaffen, Einfluss auf Entscheidungen zu nehmen und dadurch den politischen Prozess ergebnisoffener und zurechenbarer zu gestalten. Instrumente wie Politikfolgenabschätzungen oder Evaluationen seien in diesem Rahmen Mittel der Deliberation. Bei diesem Modell der Wissensbasierung von Politik läge der Fokus nicht auf De-Politisierung oder einer wirtschaftlichen Effizienzsteigerung, sondern auf Good Governance und inklusiver Deliberation. Evidenz- und

Wissensbasierung sei weniger ein Analyseinstrument, sondern viel mehr die Möglichkeit zu reflexiven und sozialen Lernprozessen.

### 2.2.2 Postfaktizität: Buzzword oder postmoderner Zeitgeist?

Es steckt zugleich eine gewisse Faszination und Abgedroschenheit hinter dieser Terminologie. Die vor einigen Jahren losgetretene Debatte über ein postfaktisches Zeitalter<sup>1</sup> basiert auf der Wahrnehmung, dass vor allem populistische Parteien sich unbeabsichtigt oder beabsichtigt zunehmend über vermeintlich gesellschaftlich akzeptierte Fakten und konventionell etablierte wissenschaftliche Erkenntnisse hinwegzusetzen scheinen (u.a. Dobes 2017; 2018, Harsin 2015, McIntyre 2018, Picciotto 2019, Sismondo 2017, Zeit 2016a). Begriffe wie *post-truth* und *post-facts*, die vor 5 Jahren noch weitgehend unbekannt waren, werden nun in den Medien, in (populär-)wissenschaftlicher Literatur und im öffentlichen Diskurs regelrecht inflationär verwendet (u.a. D’Ancona 2017, Keyes 2004, Lewandowsky et al. 2017). Angeführte Gründe sind wachsende wirtschaftliche Ungleichheit, zunehmende politische Polarisierung, sinkendes Vertrauen in Medien, Wissenschaft und Politik. Die Spannweite reicht von beinahe alltäglichen Falschdarstellungen von Tatsachen (z.B. über zwanzigtausend nachgewiesene Lügen von Donald Trump: Washington Post 2020) über die Erfindung abstruser Erklärungsmodelle in noch unerforschten Gebieten (z.B. das Wuchern von Verschwörungstheorien durch Covid-19: ZDF 2020) bis hin zum Bestreiten offensichtlicher, vielfach bewiesener Fakten (z.B. die Leugnung des Holocausts (Tagesschau 2020) oder des Klimawandels (Tagesspiegel 2020)).

In der Wissensgesellschaft ist „wahres“ bzw. auf intersubjektiv nachvollziehbaren Methoden basierendes, verlässliches Wissen die namensgebende und bedeutendste Kraft für gesellschaftlichen Fortschritt. Das aufklärerische Ideal und das Erfordernis von Faktizität ergibt sich aus den Geltungsansprüchen der Theorie des kommunikativen Handelns und der Diskursethik sowie aus dem damit verbundenen Prinzip der Validierungsfähigkeit, d.h. die Fähigkeit seine Argumente durch belegbare Fakten zu untermauern (Habermas 1973; 1981; 1992). Etliche weitere Klassiker der Philosophie beschäftigen sich mit dem Konglomerat aus

---

<sup>1</sup> Das Oxford English Dictionary wählte „post-truth“ zum internationalen Wort des Jahres 2016 mit der Begründung, dass das Konzept durch den Aufstieg von Social Media sowie durch das wachsende Misstrauen gegenüber dem „Establishment“ und dessen Angebot an Fakten Verbreitung gefunden habe (Oxford Languages 2016). Postfaktisch wird definiert wie folgt: *“relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief”*. Auch in Deutschland benannte die Gesellschaft für deutsche Sprache (GfdS) im gleichen Jahr „postfaktisch“ einstimmig zum Wort des Jahres (Gesellschaft für deutsche Sprache 2016). Entscheidend waren die Signifikanz und Popularität des Wortes.

vermeintlicher Wahrheit, Wissenschaft und Politik (u.a. Arendt 1987, Kant 1784). Doch im wissenschaftlichen System spricht man in den seltensten Fällen von „Wahrheit“ oder dergleichen, da dieser Begriff eine Endgültigkeit und Unwiderruflichkeit suggeriert, der der Komplexität der Wissensproduktion, seiner inhärenten Unabgeschlossenheit und Anfälligkeit für Fehler und Limitationen nicht gerecht werden würde. Dennoch stößt die Wahrnehmung des Gegenkonzepts, d.h. der Postfaktizität, und die damit einhergehende, leichtfertige Ablehnung eines hoch zu schätzenden und vermeintlich selbstverständlichen Gutes, d.h. Faktizität, auf großen Unmut und Desorientierung. Mehr als 83% der Befragten im Eurobarometer sehen Fake News und fehlerhafte Informationen in Online-Medien als Gefahr für demokratische Strukturen (Europäische Kommission 2018). Politik wird demnach von Teilen der Bevölkerung nicht nur aufgrund ihrer Werte und Überzeugungen, sondern auch aufgrund ihrer Rationalität und angewandten empirischen Evidenz bewertet und entsprechend sanktioniert. Etliche Initiativen, Kundgebungen, Berichterstattungen und Demonstrationen beziehen sich in ihrem Für und Wider auf das Konstrukt alternativer Fakten. Unter anderem die globale Bewegung March for Science ging gegen die Etablierung einer postfaktischen Ära und für die Unverhandelbarkeit wissenschaftlicher Fakten als Grundlage gesellschaftlicher Diskurse auf die Straße (Spektrum 2017, Zeit 2017). Die moderne Gesellschaft hat erst durch die jüngsten Zuwiderhandlungen eine Vorstellung darüber erlangt, wie wichtig Wissen und jene Institutionen, die dieses Wissen herstellen und sichern, sind. Die teilgesellschaftliche Beunruhigung und erregten Reaktionen gegen den omnipräsenten Populismus und politischen Antiintellektualismus zeigen, dass die moderne Wissenschaft und die damit einhergehende Rationalität seit der Aufklärung eine enorme Autorität als feste Institution der modernen Gesellschaft erlangt hat. Wissenschaft gilt laut der repräsentativen Bevölkerungsumfrage „Wissenschaftsbarometer 2019“ im Gegensatz zu anderen Systemen wie Wirtschaft, Medien und Politik nach wie vor und mit Abstand als in hohem Maße vertrauenswürdig (Wissenschaft im Dialog 2019).

Obwohl oder gerade weil der Begriff „postfaktisch“ inflationär verwendet wird, wird er aufgrund multipler Gründe kritisiert. Neben dem polemischen und populärwissenschaftlichen Gebrauch steht die dem Begriff zugrundeliegende Annahme im Zentrum der Kritik, dass zu einem bestimmten Zeitpunkt der Weltgeschichte Faktizität als akzeptierter Standard des politischen Lebens nicht nur gepredigt, sondern auch gelebt wurde (Telepolis 2016, Zeit 2016b). Es wird argumentiert, dass nicht nur in der Gegenwart, sondern auch in der Vergangenheit gesellschaftliche und politische Akteure versucht haben, den öffentlichen Diskurs bewusst mit unwahren Behauptungen für sich zu gewinnen. Der öffentliche Diskurs

sei niemals reinen Fakten unterlegen. Subjektive Wahrnehmungen und Werte würden unvermeidlich miteinließen.

Trotz der teilweise berechtigten wird der sich etablierte Begriff des postfaktischen Zeitalters aufgrund seiner bereits stattgefundenen Kontextualisierung und Etablierung im öffentlichen Diskurs zur Ausarbeitung des vorliegenden Forschungsthemas verwendet, da er das zu beschreibende Phänomen einer Postfaktizität und die sinnverwandten Assoziationen im breiteren Sinne treffend erfasst. Desinformation ist im politischen Kosmos sicher nichts Neues. Was sich hingegen tatsächlich geändert und den Ausruf eines postfaktischen Zeitalters provoziert hat, ist, dass es für Entscheidungsträger\*innen nicht mehr in gleichem Maße notwendig erscheint, falsche Behauptungen als Fakten zu verschleiern oder Widersprüche mit überprüfbaren Tatsachen auszugleichen, wie das in der Vergangenheit der Fall war (Hendricks/Vestergaard 2017, Marschall 2017, Sarasin 2017). Wissenschaftliche Fakten scheinen Autorität eingebüßt zu haben. Ein AfD-Kommunalpolitiker verteidigte beispielsweise die Kritik an den von seiner Partei kommunizierten Zahlen zu geflüchteten Menschen wie folgt (Deutschlandfunk 2016): *„Es geht nicht nur um die reine Statistik, sondern es geht darum, wie das der Bürger empfindet. (...) Das heißt: Das, was man fühlt, ist auch Realität.“* Wenn Tatsachen mit „alternativen Fakten“ koexistieren und als Grundlage der öffentlichen Meinungsbildung dienen und wenn überprüfbare Fakten selbst Gegenstand politischer Auseinandersetzungen werden, um der politischen Logik der eigenen Argumentation zu folgen, gibt es für die politische Debatte keinen einheitlichen Maßstab mehr. Jener Wandel der politischen Kommunikation definiert den postfaktischen Zustand, eine neue Phase in der Auseinandersetzung mit der politischen Lüge – mit langfristigen Konsequenzen für die demokratische Kultur.

## 2.2 Stand der Forschung

Eine zunehmende Evidenzbasierung der Politik wird insbesondere in einer Zeit populistischer Strömungen und postfaktischer Diskurse, die ihre politische Strategie auf fehlerhaften und falschen Informationen aufbauen, von Wissenschaftler\*innen als unabdingbar gesehen (Buch et al. 2019). Wissenschaftlich fundierte Studien, Analysen und Evaluationen allein werden das Problem postfaktischer Politikstrategien zwar nicht lösen können, aber sie können eine evidenzbasierte Basis für besser informierte Entscheidungen liefern – für jene Akteure, die sich dieser annehmen wollen bzw. dafür empfänglich sind (Buch et al. 2019). Wählerinnen und Wähler können durch externe, wissenschaftliche Erfolgskontrollen in die Lage versetzt werden, beurteilen zu können, ob öffentliche Gelder, Maßnahmen und Regulierungen effektiv eingesetzt werden, um priorisierte gesellschaftliche Ziele de facto zu erreichen (Bundeszentrale für politische Bildung 2017: 1): *„Meinungen können (...) nicht gegen Forschungsergebnisse ausgespielt werden, sondern finden an diesen ihre Grenze.“*

Die Theorie sozialer Differenzierung stellt das Kernstück von evidenzbasierten Modernisierungsansätzen dar (Luhmann 1983; 1988; 1997, Schimank 1996). Der Systemtheorie zufolge besteht die moderne Gesellschaft aus unterschiedlich spezialisierten Teilsystemen, die nach ihrer eigenen Logik und Struktur sowie eigenen Prinzipien funktionieren, eigenen Kommunikationsregeln folgen und jeweils eine Eigendynamik entwickeln. Jene funktionale Spezialisierung dient einer gesellschaftlichen Arbeitsteilung in z.B. Politik, Recht, Wissenschaft, Wirtschaft. Modernisierungstheoretiker\*innen sind sich einig, dass die einzelnen Dimensionen bzw. Subsysteme zwar über eine relative strukturelle und kulturelle Selbstständigkeit verfügen, aber nicht autonom agieren, sondern in einem Austausch und Verflechtungszusammenhang zu den Systemen stehen (Pollack 2016). In diesen Teilsysteminterdependenzen sieht Schimank (2001) den Schlüssel zum Verständnis moderner Gesellschaft: Sie steigern zwar die gesellschaftliche Leistungsfähigkeit, verursachen aber unweigerlich auch Konkurrenzsituationen und Kompetenzstreitigkeiten, insbesondere im Falle der für diese Forschungsarbeit zentralen Systeme Politik und Wissenschaft. Zusätzlich bedingt die Globalisierung eine zunehmend technische und räumliche Entgrenzung sowie einen stetig steigenden Grad an Komplexität und sozialer Differenzierung (Alexander 2001, Beck 1997). Das hat zur Folge, dass einzelne Akteure, aber auch ganze Subsysteme immer weniger an nationale Systemgrenzen gebunden sind, der Veränderungsdruck auf bestehende Teilsysteme

verstärkt wird und Differenzierungs- und Anpassungsprozesse auf der Ebene bestehender nationaler Institutionen ausgelöst werden.

Der weltweite Aufstieg der Evaluation ist als eine Reaktion auf diese neuen Herausforderungen zu verstehen (Meyer 2015, Stockmann/Meyer 2017). Die systemtheoretische Vorstellung einer ausdifferenzierten Gesellschaft, deren Teilsysteme einem Entwicklungsweg folgen, bedeutet nämlich auch, dass dieser Prozess aktiv gestaltet und gesteuert werden kann. Stockmann und Meyer (2017) betonen in diesem Zusammenhang die Notwendigkeit als auch die Möglichkeit, Evaluationen als rationales Instrument zur Unterstützung jener gesellschaftlichen Steuerung und somit als notwendiges Produkt der Modernisierung zu verstehen. Evaluationen können in modernen und ausdifferenzierten Gesellschaften der Rationalisierung von politischen Debatten dienen, als politisches Steuerungsinstrument, aber auch als Aufklärungsinstrument und damit die empirische Basis für eine umfassende Selbstreflexion bilden (Stockmann/Meyer 2017). Evaluative Kontrollen seien somit nicht nur ein technokratisches Instrument, sondern würden sich auch für einen reflexiven Einsatz und Kritik an der Moderne selbst eignen. Diese Überzeugung basiert auf einem Komplementaritätskonzept zwischen Politik und Verwaltung (Svaras 2001). Bereits 1986 entwickelte Habermas die Idee, dass das politische System als Reaktion auf die wachsende Komplexität der modernen Gesellschaft drei Formen annehmen könne: Entweder würde eine wenig informierte Politik eine professionalisierte Verwaltung dominieren (*Dezisionismus*), eine wissenschaftlich arbeitende Verwaltung eine wenig spezialisierte Politik dominieren (*Technokratie*) oder beide arbeiten auf eine nichthierarchische und produktive Art und Weise zugunsten gemeinsamer gesellschaftlicher Ziele zusammen (*Pragmatismus*).

Das politische Subsystem ist neben dem sozialen und kulturellen eines der drei grundlegenden Systeme und bildet die Nachfrageperspektive der Evaluationen. Im nachfolgenden Kapitel wird darauf eingegangen, wie und warum Evaluationen in parlamentarischen und verwaltungstechnischen Strukturen verankert sind und von jenen genutzt werden. Das kulturelle System hingegen ist der Angebotsseite und dem Wissenschaftssystem zuzuordnen. In diesem Sinne wird wie folgt darauf eingegangen, wie Evaluationen durchgeführt werden und welche Besonderheiten sich aus der wissenschaftlichen Arbeitsweise ergeben. Das soziale System, im engeren Sinne die Zivilgesellschaft, umfasst die Verbreitung und gesellschaftliche Akzeptanz von Evaluationen sowie zivile Forderungen an das politische System, Evaluationen zu Kontroll- und Legitimierungszwecken für eine rationale Entscheidungsfindung zu nutzen.

Demzufolge wird dahingehend auch auf den öffentlichen Diskurs über Evaluationsergebnisse und die Einbindung zivilgesellschaftlicher Interessen eingegangen. Der Theorie der sozialen Differenzierung folgend, wird bei jedem Subsystem auf das jeweilige theoretische Ideal bzw. den Sollzustand sowie auf die bisherigen empirischen Erkenntnisse bzw. den Ist-Zustand sowie bedingende Faktoren eingegangen.

### *2.2.1 Evaluierter Perspektive: Die Nachfrage des politischen Systems*

Die Entstehung und Entwicklung der EBP-Bewegung basiert auf sowohl Angebots- als auch Nachfragefaktoren. Letztere beziehen sich auf die Nachfrage nach sozial- und wirtschaftswissenschaftlichen Studien durch Regierungs-, Gesetzgebungs- oder Behördenorganen und das damit einhergehende Commitment der Geldgeber\*innen zugunsten staatlich finanzierter Forschung (Head 2010). Zur späteren Analyse der nachfragenden bzw. auftraggebenden Perspektive des politischen Systems wird im Folgenden der Forschungsstand zur Entstehungsgeschichte von Evaluierungsmethoden, zur derzeitigen Institutionalisierung und den legislativen Rahmenbedingungen, zu den Zielfunktionen sowie zu den Sollbruchstellen und hemmenden Faktoren gesichtet und bewertet.

#### *2.2.1.1 Entstehungs- und Entwicklungsgeschichte der veränderten Nachfrage*

Evaluationen können als Erfindung der Moderne verstanden und deren kurze Entstehungsgeschichte in drei Wellen zusammengefasst werden (Stockmann/Meyer 2017). Treibende Kräfte waren das politische System und die öffentliche Administration, die die Nachfrage repräsentierten, Evaluationsziele und -aufgaben formulierten sowie notwendige Ressourcen bereitstellten:

Der erste Boom des Evaluationsvorhabens nahm im Deutschland der 1960er Jahre zuerst die Form einer von der Beratungswirtschaft beherrschten Implementationsforschung an, im Rahmen derer Reformen mit sozialen Experimenten gleichgesetzt wurden (Campbell 1969, Wollmann 1997). Die durch die Ölpreiskrise bedingten ökonomischen Einschränkungen in den 70er und 80er Jahren erstickten hingegen tiefergehende Reformanstrengungen in der zweiten Welle ab. Die öffentlichen Budgetprobleme forderten jedoch auch eine stärkere Priorisierung und Selektion, wodurch der Fokus klar auf Effektivität und Effizienz von politischen Maßnahmen gesetzt wurde, um damit politische Entscheidungen gegenüber der Öffentlichkeit

zu legitimieren. In jener Welle verkam Evaluation zu einem taktischen Instrument, um bereits getroffene Policy-Entscheidungen im Nachhinein mit Evaluationsergebnissen zu rechtfertigen (Wollmann 2017, Roth 2004). Mit den 1990er Jahren begann die heutige dritte Phase der Evaluation, die primär vom Aufstieg des New Public Management (Behn 1998) und der Public Choice Theorie (u.a. Buchanan/Tullock 1962, Dunleavy 2014) angetrieben wurde. Im Kontext jener neuen Managementmodelle erfuhr das Instrument der Evaluation eine enorme Aufwertung und zunehmende Beachtung (Wollmann 1999; 2017, Stockmann 2006). Evaluationen wurden als angewandte Sozialforschung anerkannt und analysierten fortan nicht nur Planungsprozesse und Leistungserbringung (Input), sondern auch die de facto erbrachten Leistungen (Output), die Zielerreichung (Outcome) und die insgesamt ausgelösten intendierten sowie nicht intendierten Wirkungen (Impact). Erstmals wurden Forderungen laut, die Leistungsfähigkeit einer Administration systematisch an ihren Erfolgen zu messen sowie alle politisch intendierten Ziele wirkungsorientiert zu formulieren und mit Indikatoren überprüfbar zu machen.

#### *2.2.1.2 Institutionalisierung, legislative Rahmenbedingungen und Umsetzungsdefizit*

Im internationalen Vergleich sind Evaluationen und das Konzept evidenzbasierter Politik mittlerweile etabliert und institutionalisiert, insbesondere in großen westlichen Demokratien, den OECD-Ländern und den Pionierstaaten USA und Großbritannien (Wollmann 2004). Auch die Verankerung des Evaluierungsvorhabens im Gesetzgebungsprozess ist breiter Standard geworden (Davies et al. 2012, Normenkontrollrat 2013). In der Schweiz oder in Schweden haben Evaluationen sogar Verfassungsrang. Doch ein bundesweites Gesetz zur Regelung und Vorschrift von Evaluationen gibt es in Deutschland bis heute nicht. Im Gegensatz zu beispielsweise der EU ist auch die Durchführung von Programmevaluationen nicht verpflichtend vorgeschrieben. Doch obwohl Gesetzesbewertungen in Deutschland keine lange Tradition haben und noch immer rar sind (Gusy 2015), hat die Anzahl der Gesetze und Verordnungen, in denen Evaluations- und Berichtspflichten explizit vorgeschrieben sind, in den letzten drei Legislaturperioden stark zugenommen (Konzendorf 2009). Vor allem in den Sozialgesetzbüchern wurden Evaluationen, Erfolgskontrollen, Wirkungsprüfungen oder Berichtspflichten fest verankert. Eine klare Statistik, in wie vielen und welchen Bundes- und Landesgesetzen solche Klauseln aufgenommen wurden, ist dennoch nicht verfügbar.

Ein Meilenstein der deutschen Institutionalisierung war der im Jahr 2000 hinzugefügte Passus in der Gemeinsame Geschäftsordnung der Bundesministerien (GGO), der vorschreibt, dass der Gesetzgeber beim Erlass jeder Rechtsform eine systematische Erfassung und Bewertung von voraussichtlichen und tatsächlichen Folgen vornehmen soll (Bundesregierung 2020). Dabei muss erstens explizit auf die beabsichtigten sowie nicht beabsichtigten Wirkungen eines Gesetzes eingegangen werden. Zweitens muss erläutert werden, ob die Wirkungen einer nachhaltigen Entwicklung entsprechen und drittens, welche langfristigen Wirkungen erwartet werden (§44 Gesetzesfolgen). Ein weiterer Meilenstein war der 2013 in Kraft getretene Staatssekretärsausschuss „Bürokratieabbau und besondere Rechtsetzung“ mit dem die Bundesregierung erstmals einen verbindlichen Rahmen zur systematischen Überprüfung von Gesetzen und Verordnungen geschaffen hat (Bundestag 2014). Der Beschluss besagt, dass grundsätzlich alle bundesrechtlichen Regelungsvorhaben ab einem Schwellenwert von einer Million Euro Erfüllungsaufwand drei bis fünf Jahre nach ihrem Inkrafttreten im Rahmen einer ex-post-Wirkungsanalyse evaluiert und mit dem Ziel überprüft werden sollen, ob der Zweck des Regelungsvorhabens erreicht worden ist (Normenkontrollrat 2015; 2020). Das gängige rechtfertigende Argument für die mangelnde Beauftragung von Evaluationen, d.h. die fehlenden zeitlichen, finanziellen und personellen Kapazitäten, wird im Beschluss vorweggenommen, da festgehalten wird, dass auch mit überschaubarem Aufwand solide Evaluierungsberichte produzierbar seien.

Evaluation in Deutschland hat sich in den letzten Jahren und Jahrzehnten insbesondere durch Trends aus dem angelsächsischen Raum, wo Evaluationen als wissenschaftliche Verfahren zur Bewertung politischer Projekte bereits deutlich früher an Bedeutung gewannen hatten, dynamisch entwickelt. Trotz dieser de facto stattfindenden Institutionalisierung und der steigenden Nachfrage von Entscheidungsträger\*innen (Speer et al. 2015) ist Evaluation in Deutschland eine noch recht unterentwickelte Disziplin (Coryn et al. 2017). In internationalen Vergleichsstudien rangiert Deutschland in der Bewertung der Evaluationskultur nur im Mittelfeld (Furubo et al. 2002, Jacob et al. 2015, Normenkontrollrat 2013). Evaluationen haben hier nicht den gleichen Stellenwert wie in den USA oder in Großbritannien, da die deutschen Auftraggeber meist nicht gleichzeitig auch die Entscheidungskompetenz besitzen, die jeweiligen Evaluationsergebnisse auch umzusetzen (Schöffski 2008). Außerdem ist die Evaluationspraxis der einzelnen Ministerien und Behörden bisher höchst unterschiedlich und einheitliche Strategien, Pläne oder gar eine Evaluations-Policy der Bundesregierung fehlen.

Zwar sind diese fehlenden gesetzlichen Rahmenbedingungen zu kritisieren, schwerwiegender ist jedoch das Vollzugsdefizit der rudimentär bestehenden Vorschriften (Konzendorf 2009). Unter anderem der Bundesrechnungshof (BRH) und der Normenkontrollrat (NKR) sind wichtige treibende Kräfte für die fortschreitende Institutionalisierung von Evaluationen in Deutschland und problematisieren daher auch das eklatante Anwendungsdefizit (Stockmann/Meyer 2017): Der BRH appelliert bereits seit 1989 für die Ausweitung der öffentlichen Evaluationsaktivitäten und die Schaffung organisatorischer und methodischer Voraussetzungen für deren Durchführung, um die Wirksamkeit von Subventionen überprüfen und die Bewilligung von Haushaltsmitteln an die Vorlage solcher Evaluationen binden zu können (Bundesrechnungshof 1989, Müller 2011). Doch ein BRH-Gutachten stellt fest, dass Behörden es häufig unterlassen, Evaluationen bzw. Erfolgs- und Wirkungskontrollen durchzuführen (Bundesrechnungshof 2013). Dies wird von den Verantwortlichen damit gerechtfertigt, dass der Einsatz enormer personeller und finanzieller Ressourcen keinem praktischen Nutzen gegenüberstehe. Evaluationen würden somit laut BRH nicht als notwendige kontinuierliche Aufgabe verstanden und daher trotz der Verpflichtung nach §44 der Bundeshaushaltsordnung nur in wenigen Fällen durchgeführt (Bundesrechnungshof 2013). Mitarbeitenden sei nicht bekannt, wie diese überhaupt durchzuführen seien und in vielen Fällen würden nur die beabsichtigten Wirkungen untersucht. Der BRH empfiehlt in diesem Zusammenhang in Querschnittseinheiten Evaluierungskompetenz aufzubauen, Mindeststandards und praxistaugliche Leitfäden zu erarbeiten, ex post und ex ante Evaluationen verstärkt zu verknüpfen, ex-post-Evaluationen zum Bestandteil der politischen Vorhabenplanung zu machen sowie eine externe und systematische Qualitätssicherung verpflichtend festzuschreiben.

### *2.2.1.3 Ziele und Nutzungsfunktionen von Evaluationen in politischen Entscheidungsprozessen*

Aufgrund der Vielzahl an unterschiedlichsten Kategorisierungen und Typologisierungen der Evaluations- und Wissensnutzung (u.a. Feldman 1989, Radaelli 1995, Rich 1979, Weiss 1978a; 1978b) wurde die folgende eigene Übersicht mit einer hinreichenden, aber nicht erschöpfenden Auswahl grundlegender Funktionen von Evaluierungen zusammengestellt, die das politische System betreffen:

*Funktion 1: Generierung von empirischen Daten und wissenschaftlichen Informationen (u.a. Cronbach 1982, Furubo 1994, Hertin et al. 2009, Romsdahl 2005)*

Die grundlegendste instrumentelle Funktion von Evaluationen besteht darin, wissenschaftliche Informationen und Erkenntnisse über die Wirkungsweise von staatlichen Maßnahmen zu generieren und damit die Entscheidungsgrundlage für eine evidenzbasierte Ausgestaltung von Politik bilden zu können. Beispielsweise können generierte Längsschnittdaten Trends und Probleme aufzeigen, Warnungen signalisieren und darauf aufbauend politische Reaktionen auslösen. Es ist zudem möglich, dass Entscheidungsträger\*innen zu einem politischen Thema nicht genügend Informationen vorliegen oder den vorliegenden Informationen misstrauen, weshalb sie sich an sozialwissenschaftliche Produkte und Evaluationen als Korrektiv wenden.

*Funktion 2: Ausübung von Kontrolle (u.a. Pollitt 2006, Stockmann 2017)*

Durch die Notwendigkeit zur Dokumentation und Informationsoffenlegung geht jede Evaluationsform auch zwangsläufig mit einer Kontrollsituation für die öffentliche Hand einher. Dies scheint nicht nur aus einer gesellschaftlichen Perspektive wünschenswert, sondern auch aus einer überparteilichen: Da nicht alle politischen Parteien und Interessensgruppen ihre Meinung zur evaluierten Thematik gleich stark artikulieren können, können Erfolgskontrollen die bestehenden Einflussasymmetrien in einer pluralistischen Demokratie abfedern.

*Funktion 3: Steuerung und Optimierung (u.a. Janett 2003, Weiss 1987; 1999)*

Evaluationen sind in einem immer komplexer werdenden Wirkungsgefüge staatlicher Institutionen unverzichtbar, um neben tagespolitischem Feedback und administrativen, routinierten Kontrollen den Bedarf an tiefergehenden Analysen des öffentlichen Sektors zu decken. Sie sind eine wichtige Steuerungsressource und Orientierungshilfe im Projekt- oder Programmmanagement, um bestehende Politiken zu verbessern, getroffene Annahmen und Erfolgskriterien zu revidieren und Ansätze zu rekonzeptualisieren. Der Steuerungsnutzen besteht somit in der Leistung, die Steuerungsfähigkeit jener Projekte und Programme durch Evaluierung zu erhöhen.

*Funktion 4: Institutionelles Lernen (u.a. Janett 2003, Schöffski 2008)*

Die durch aktive Steuerung entstehende Policy-Entwicklung und -Veränderung ist eng mit dem Konzept des institutionellen und organisationalen Lernens verknüpft. Die Generierung von neuen Informationen und die darauf aufbauende Revision von Programm- oder Vollzugsstrukturen kann zu Lernprozessen führen und aktiv zum Qualitätsmanagement einer lernenden Organisation beitragen.

*Funktion 5: Strategie und Instrumentalisierung (u.a. Davies 2012, Kogan 1999, Marston/ Watts 2003)*

Selbstredend können Evaluationen und wissenschaftliche Erzeugnisse auch für nicht originäre Zwecke, d.h. politisch oder symbolisch eingesetzt werden, und dazu dienen, strategisches Framing zu betreiben, eine gewünschte Botschaft in die Öffentlichkeit zu transportieren, eine bereits getroffene Entscheidung zu rechtfertigen, Unterstützung zu mobilisieren, Koalitionen aufzubauen, Unentschlossene zu überzeugen, gegnerische Argumente zu verwerfen und Skeptiker zu überzeugen. Evaluationen können als politische Zertifizierung genutzt werden, um Politiker\*innen vermeintliche Integrität und Rationalität zu verleihen. Nicht immer ist ein bewusster bzw. strategischer Missbrauch von Evidenz der Fall, dennoch besteht die Gefahr, dass Evaluationen als politisches Instrument und nicht als wissenschaftliches Bewertungsinstrument verwendet werden. Des Weiteren zeigen die vielfachen strategischen Nutzungsmöglichkeiten, wie machtvoll Wissen im Rahmen politischer Profilierung geworden ist.

Weiland (2013) fasst diese Funktionen in zwei übergeordnete und zentrale Argumente für die Nutzung von Evaluationen zusammen, die sich auf die beiden unterschiedlichen epistemischen Modelle – rationalistisch und offen-diskursiv – rückbeziehen: Das Argument der Versachlichung und Eindeutigkeit bezieht sich auf das Ziel, Politikkohärenz zu fördern sowie Politik über die interessen- und wertbasierten Behauptungen politischer Akteure erhaben zu machen (gelistete Funktionen 1 und 3). Das Argument der Reflexivität besagt hingegen, dass evidenzbasierte Verfahren dazu beitragen können, die Anzahl der in der Politikgestaltung berücksichtigten Aspekte zu erhöhen, insbesondere die häufig unbeachteten oder übersehenen Nebenfolgen von Politik (gelistete Funktionen 2 und 4).

#### *2.2.1.4 Sollbruchstellen und hemmende Faktoren von und für Evaluierungen im politischen Kosmos*

In der Verwendungsforschung herrscht ein breiter Konsens darüber, dass systematische und fest in den politischen Entscheidungsprozess eingebettete und wissenschaftliche Evaluierungen bisher nur punktuell und peripher im deutschen Politikkosmos eingesetzt werden (Weiss 1987). Und selbst wenn diese höchsten Standards entsprechen, deren Wirkung auf und Nutzung durch das politische System äußerst gering oder gar problematisch ist (Amara et al. 2004, Landry et al. 2003, Nelson et al. 1987, Nutley et al. 2007, Rich/Oh 2000, Shulock 1999). Der Effekt von Policy oder Impact Assessments auf politische Entscheidungen ist häufig marginal (Owens et al. 2004, Weiss 1982; 1999); eine direkte oder gar unmittelbare Nutzung von Evaluierungsergebnissen durch Entscheidungsträger\*innen zeichnet sich nicht ab (Baklien 1983, Weiss 1999). Die Erwartungshaltung von evaluierenden Akteuren, dass die generierten Daten in den Entscheidungsprozess einfließen und Aktionen beeinflussen werden, wird oft enttäuscht. Selten setzen politische Akteure die von der Forschung gelieferten Antworten zur Lösung eines politischen Problems ein (Weiss 1982).

Die Umwandlung wissenschaftlicher Erkenntnisse in politisches Handeln ist und bleibt problematisch (Janett 2003). Es gibt umfangreiche Forschung über die Barrieren, die der Umsetzung von Forschungs- und Evaluierungsergebnissen in der politischen Praxis im Wege stehen und darüber, wie diese Barrieren überwunden werden können. Im Folgenden werden jene hemmenden und hindernden Faktoren, die primär in der Verantwortung des politischen Systems liegen, analysiert:

*Erstens*, die Mehrheit an politischen Entscheidungen kommt nicht durch rationale und formale Entscheidungsprozesse zustande. Es gibt etliche Faktoren außer sozialwissenschaftliche Evidenz, die die politische Dynamik beeinflussen (Davies 2004, Nelson et al. 1987, Walker 2000): Werte, Ideologie, Fachwissen, Erfahrung, Urteilsvermögen, Zweckmäßigkeit, Ressourcen, politische und bürokratische Kultur, Lobbyist\*innen, Interessengruppen, Wähler\*innen und die Notwendigkeit, schnell auf Eventualitäten zu reagieren. Dass durch Evaluierende generierte Wissen konkurriert mit anderen im Umlauf befindlichen Kenntnissen und Faktoren um eine Anhörung (Weiss 1999). Programm- und Politikentscheidungen sind das Ergebnis einer Vielzahl von Aktionen und Reaktionen durch eine Vielzahl von Akteuren ohne einen unitären Entscheidungstragenden (Weiss 1987).

*Zweitens*, die bloße Verfügbarkeit verlässlicher Forschung sichert nicht deren späteren Einfluss und Wirkung (Head 2010). Das traditionelle Problem liegt in der Frage, wie man politische Entscheidungsträger\*innen dazu bringen kann, externes, von Wissenschaftler\*innen entwickeltes Wissen handlungsleitend zu verwenden (Nutley et al. 2007). Öffentliche Amts- und Entscheidungsträger\*innen handeln auf der Grundlage dessen, was sie bereits wissen, d.h. ihres aktuellen Wissensbestands. Sie suchen oft nicht aktiv nach neuen Informationen, haben häufig ein geringes Bewusstsein für Forschungs- und Evaluierungsergebnisse und greifen auf sozial akzeptable Antworten zurück bzw. begnügen sich mit dem Status quo (Weiss 1987). Wenn Entscheidungsträger\*innen gar bereits über umfangreiches Wissen über ein Thema verfügen oder wenn sie starke ideologische Verpflichtungen und Eigeninteressen haben, werden Evaluierungen noch weniger genutzt.

*Drittens*, der zeitliche Faktor ist problematisch. In Bezug auf medizinisches Wissen schätzen Mold und Peterson (2005: 14), dass es durchschnittlich 17 Jahre dauert bis 14% der originären Forschungserkenntnisse in Form konkreter Veränderungen für die Patient\*innenpflege umgesetzt werden. Davies (2012) argumentiert in Hinblick auf die wesentlich kürzere Geschichte und Tradition der EBP-Bewegung, dass die Zeitspanne für die konkrete Umsetzung von politischer Evidenz im Vergleich somit noch wesentlich länger dauern müsste. Des Weiteren erfolgt politische Entscheidungsfindung in der Regel innerhalb von Tagen, Wochen und Monaten, wohingegen die Sammlung von qualitativ hochwertigem Forschungsmaterial viele Monate und Jahre erfordert (Davies 2012, Head 2010). Das bedeutet zwangsläufig, dass lediglich das beste zum jeweiligen Zeitpunkt verfügbare Material in den gegenwärtigen politischen Entscheidungsprozess einfließt, während gleichzeitig eine solidere Faktenbasis für die künftige mittel- bis langfristige Politikgestaltung entwickelt werden muss. Durch die kurzen Evaluationsfristen (oft bereits nach wenigen Jahren) und die kurzen Durchführungszeiträume (typischerweise wenige Monate bis zu einem Jahr) können Policies, die ihre Wirkungen erst nach längerer Zeit entwickeln oder gar konzeptionelle politische Lernprozesse, oft nicht aufgedeckt werden (Hertin et al. 2009).

*Viertens*, Hertin et al. (2009) bestätigen, dass Evaluationsverpflichtungen in vielen Fällen eine geläufige „Tick the box“-Haltung provozieren und dazu führen, dass die Verfahren nicht als Instrument zur Entscheidungsunterstützung, sondern als bürokratische Hürde betrachtet werden. Das National Audit Office (NAO) von Großbritannien konstatiert kritisch, das Analyseniveau von Evaluierungen sei bislang schwach und die Anzahl der Policy-Optionen

begrenzt (National Audit Office 2007). Evaluierungen verkämen dadurch zu einer leidigen Compliance-Aufgabe statt zu einem inhärenten Bestandteil der Politikgestaltung zu werden (National Audit Office 2009). Insgesamt wären bei 42% der vom NAO überprüften Folgenabschätzungen nur eine einzige Option in Betracht gezogen worden – zusätzlich zur Option, nichts zu tun (National Audit Office 2010: 14).

*Fünftens*, der hohe zeitliche, personelle und finanzielle Aufwand für evaluierende Organisationen symbolisiert gerade für den öffentlichen Sektor mit vielfach eingeschränkten Ressourcen eine enorme Hürde und scheint zunächst nachteilig für die Leistungsfähigkeit und Bereitstellung staatlicher Leistungen (Kuhlmann et al. 2004).

*Sechstens*, wie alle wissenschaftlichen Beweise sind Evaluationen oft probabilistisch und mit einem gewissen Grad an Unsicherheit behaftet (Davies 2012). Mängel in der Forschung selbst sind nicht auszuschließen (Romsdahl 2005). In der Praxis fallen wissenschaftlich-methodische Standards von Evaluationsstudien angesichts deren Ziels, praktisches Handlungswissen zu generieren, hinter ihr Potential zurück (Hertin et al. 2009).

*Siebtens*, die Analyse beschränkt sich in Deutschland oft auf die Analyse der direkten und messbaren wirtschaftlichen und administrativen Kosten, wodurch wenig Raum bleibt für weitsichtige Reflexionen zum tiefer liegenden Problem, zu den kausalen Effekten oder alternativen Maßnahmen (Hertin et al. 2009).

*Achtens*, das Problem der externen und internen Einflussnahme besteht bei der Vertragsforschung (Morris/Clark 2013, Wollmann 2007). Politisch-institutionelle Auftraggeber tragen die Kosten der Evaluation und definieren daher Evaluationsgegenstand, Forschungsfragen und zuweilen Methodik. Dass die Auftraggebenden grundlegend an einem für sie günstigen Evaluationsergebnis interessiert sind und externe Evaluierende an Folgeaufträgen ist oft kontraproduktiv. Auch involvierte Stakeholder sind in der machtvollen Lage, die Durchführung und die Ergebnisse von Evaluationen anzufechten und politisch zu beeinflussen (Eckhard/Jankauskas 2018, Orr 2010).

*Neuntens*, oft werden politische Ziele und Policy-Ziele vermischt. Der Anstoß für Evaluationen ist häufig politisch motiviert (Bovens et al. 2008, McConnell 2010a; 2010b, Rossi et al. 2004). Evidenz wird dabei vielfach symbolisch verwendet, um bereits feststehende politische

Positionen zu untermauern und strategisch für politische Interessen zu nutzen (Hertin et al. 2009). Diese technokratische Herangehensweise kann dem demokratischen Prozess der Politikgestaltung zuwiderlaufen (Marston/Watts 2003). Evaluationen, die versuchen, eine eindeutige Handlungsempfehlung oder schlichtweg die beste Antwort zu geben, werden von der einen Seite als Munition im politischen Wettbewerb verwendet und von der anderen geleugnet (Weiss 1987).

Evidenz allein kann Nutzer\*innen nicht sagen, was sie tun oder wie sie handeln sollen (Davies 2012). Wenn, wie es oft der Fall ist, multiple Interessen und Ideologien in einem kontroversen Zusammenhang stehen, wird Evaluierung bzw. Evidenz nicht das entscheidende Element sein (Weiss 1999). Politik und nicht Evidenz sorgt in Demokratien dafür, divergierende Interessen zu versöhnen und Kompromisse zu erreichen. Grundlegende Problemdefinitionen, auf denen politisches Handeln beruht, sind durch Evaluierungen kaum zu beeinflussen (Lampinen 1992: 244). Die Hoffnung, effizienzsteigernde Methoden und Tools der Privatwirtschaft auf den öffentlichen Sektor auszuweiten, kann somit nur in gewissem Maße erfüllt werden. Entscheidungen im privaten Sektor haben im Vergleich eine höhere Wirtschaftlichkeitsorientierung, die Ziele sind leichter zu quantifizieren und es sind weniger involvierte Gruppen zufrieden zu stellen (Nelson et al. 1987). Im Gegensatz dazu legt der politische Kosmos mit seinen Eigenheiten und Charakteristiken dem Evaluierungsvorhaben weitaus mehr Steine in den Weg.

#### *2.2.1.5 Hinreichende Bedingungen, begünstigende Faktoren und das reale Potential von Evaluierungen*

*“What is needed [from evaluation] is information that supports negotiation rather than information calculated to point out the “correct” decision”*  
(Cronbach 1980: 4)

Trotz der zuvor dargestellten Limitationen können Evaluierungen Vielfaches leisten. Die Tatsache, dass Evidenz nicht immer gezielt und direkt auf politische Entscheidungen wirkt, bedeutet nicht, dass sie wenig oder gar keinen Einfluss hat (Davies 2012). Forschungsergebnisse werden in öffentliche politische Debatten und interne Prozesse des öffentlichen Sektors einbezogen und zeigen so, welche Policies und Programme unter welchen Bedingungen funktionieren und tragen dadurch zur Verbesserung dieser bei (Adams 2004, Boaz et al. 2008, Maynard 2006). Sozialwissenschaftliche Informationen werden in einer

entscheidungsvorbereitenden Rolle verwendet und liefern Informationen, die als Grundlage für fundierte Urteile und die tatsächlichen politischen Entscheidungen dienen (Rich 1979, Caplan 1975, Davies 2012). Sie stellen alte Ideen und sichere Annahmen in Frage, zeigen neue Perspektiven, Konzepte und Handlungsmöglichkeiten auf, helfen bei der Neuordnung der politischen Agenda, schaffen ein geteiltes Themenverständnis zwischen Entscheidungsträger\*innen als Grundlage für weitere Policy-Arbeit und generieren wichtige Informationen, die zur Gestaltung und Neuausrichtung der öffentlichen Politik beitragen (Weiss 1999, Lampinen 1992). Evaluierungsergebnisse, die in der größeren Policy Community bekannt werden, haben die Chance, die Bedingungen der politischen Debatte, die Sprache, in der sie geführt wird sowie die Ideen, die als relevant erachtet werden, zu beeinflussen (Weiss 1987). Ideen und Konzepte, die aus Evaluationen stammen, können u.a. durch Expert\*innenkommissionen, Panels, Berater\*innen, Interessengruppen, berufliche und soziale Netzwerke sowie Medien verbreitet werden: *“as evaluation moves into currency and becomes accepted by informed publics, it can change the premises that are taken for granted”* (Weiss 1987: 11).

Viele Forscher\*innen und evaluierende Organisationen konstatieren zudem, dass Evaluationen zur Aufklärung des Policy-Diskurses beitragen (Rivlin 1971, Bulmer 1986, OECD 1980, Kallen 1982, Weiss 1980). Verallgemeinerungen und Implikationen aus Evaluierungen sickern in das Bewusstsein und den Wissenstand der Entscheidungsträger\*innen und verändern die Art und Weise, wie Themen formuliert und Alternativen entworfen werden (Weiss 1987). Jedoch ist ein langfristiger Einfluss auf politische Aufklärung prinzipiell kaum messbar und schwer als konkretes Produkt der Evaluierung identifizierbar (Weiss 1999). Es wurde nachgewiesen, dass sich vermeintlich in der Bedeutungslosigkeit verschwundene Evaluationsstudien langfristig als relevant erwiesen und kontinuierlichen Einfluss hatten – wenngleich nicht zwangsläufig auf jene politischen Entscheidungen für die sie ursprünglich antizipiert wurden (Anderson/Biddle 1991, Bulmer 1987, Weiss/Bucuvalas 1980). Weiss (1999) argumentiert, die Sozialwissenschaftler\*innen würden nach der falschen Art von Einflussnahme am falschen Ort und zur falschen Zeit suchen. Es sei wichtig, einzugestehen, dass Politik nicht zu einem festen Zeitpunkt und in einer linearen, transparenten Weise mit einer begrenzten Anzahl von Multiplikatoren Gestalt annehme.

Evaluationen schaffen neben ihrer eigentlichen Bewertungsleistung Transparenz für staatliches und behördliches Handeln. Das kann sich wiederum positiv auf die Legitimität von politischen

Entscheidungen auswirken. Ressourcen können effizienter eingesetzt und Reformen angestoßen werden. Evaluationsergebnisse haben durch ihren unabhängigen, wissenschaftlichen und empirischen Charakter das Potential, politischem Handeln größere Glaubwürdigkeit beizumessen. Die Prüfung und Kontrolle von politischen Entscheidungen und die damit verbundenen Feedbackprozesse können des Weiteren langfristig zu politischen und gesellschaftlichen Lernprozessen beitragen (Owens et al. 2004, Weiss 1999).

Es gibt eine Vielzahl an Faktoren und Bedingungen im politisierten Kontext von Evaluierungen, die den Einsatz und die Nutzung derer begünstigen und fördern:

*Erstens*, eine politische Kultur, die geprägt ist von Transparenz und Rationalismus und die offen ist für Beweise und wissenschaftliche Argumente (Head 2010, Weiss 1999). Evidenzbasierte Politik und Evaluierungserkenntnisse sind nicht selten unerwünscht und ein politisches Risiko. Politische Akteure jeglicher Art wollen sich nicht freiwillig starker öffentlicher Kritik aussetzen, wenn die Programmsergebnisse möglicherweise enttäuschend ausfallen könnten. Der Aufbau einer Kultur der Evaluierung und des Lernens ist entscheidend – idealerweise mit partiübergreifender Unterstützung über die politische Kluft hinweg. Es muss ein Klima der Erwartung geschaffen werden, dass Entscheidungsträger\*innen willens sind, Studien in Auftrag geben, Experten hinzuziehen, Evaluationen zu zitieren und in politikrelevante Forschung und Leistungsinformationssysteme zu investieren. Wenn Evaluationen einen Bildungs- und Lernzweck haben und nicht nur als Tool der Wirtschaftlichkeitsprüfung verwendet werden, wird deren Beitrag wesentlich größer sein (Edelenbos/van Buuren 2005).

*Zweitens*, EBP-Ansätze und Evaluierungen werden vor allem in Demokratien mit multi-partizipativen Verhandlungs- und Entscheidungsprozessen, dezentralen politischen Systemen mit mehreren Machtzentren und einem starken politischen Wettbewerb gefördert (Weiss 1987; 1999, Head 2010). Auch sind reformistische Regierungen mit umfangreicher Wohlfahrtsprogrammatik eher bereit sich auf Evaluationen zu stützen als konservative Regierungen (Tarschys 1983): *“where many groups vie, where power is roughly equally distributed, or where important actors have not yet taken a stand, evaluative evidence can be of a significant value”* (Weiss 1987: 9)

*Drittens*, in den frühen Phasen des Politikzyklus sind Entscheidungsträger\*innen offener für neue Informationen als in späteren (Tarschys 1983). Zudem sind Entscheidungsträger\*innen

mit einem höheren Bildungsgrad und einem analytischen beruflichen Hintergrund offener für Daten und Methoden, die eine kritische Überprüfung der Evidenz erwarten lassen (Davies 2012). Das Cabinet Office (1999) in Großbritannien ließ sogar verlauten, dass Politiker\*innen spezielle EBP-Fähigkeiten und Kapazitäten bräuchten, um als intelligente Empfänger\*innen für komplexe Evidenzen agieren zu können. Es müssen zudem politische und organisatorische Anreize für die Verwendung evidenzbasierter Analyse und Beratung geschaffen werden (Romsdahl 2005, Weiss 1987). Evaluationsergebnisse müssen weit verbreitet und aktiv an Entscheidungsträger\*innen herangetragen werden.

In Ländern und Systemen, in denen die vorstehend beschriebenen Bedingungen herrschen, wird laut bisherigen empirischen Erkenntnissen die Evaluierung präserter, einflussreicher und bedeutender sein. Wenn Evaluation ihr Potential als treibende Kraft für Policy-Learning ausschöpfen will, muss sie vollständig in den laufenden Diskurs integriert werden und in der Lage sein, aktiv für „the evidential voice“ einzutreten (Sanderson 2002).

### *2.2.2 Evaluierende Perspektive: Das Angebot des wissenschaftlichen Systems*

Die beiden Systeme Politik und Wissenschaft haben eine eigene Subkultur mit subjektiven Werten und Dispositionen sowie einen eigenen Rahmen, in dem dieses Wissen existiert. Die diesen Systemen inhärenten Funktionsweisen werden besonders offensichtlich bei der Nutzung von durch Sozialwissenschaftler\*innen bereitgestellten Informationen durch politische Entscheidungsträger\*innen (Snow 2013). Im Folgenden wird erläutert, inwiefern sich die Art des produzierten Wissens, die Anzahl der Informationsquellen, die Zeitperspektive, die Belohnungssysteme und die grundlegenden Funktionslogiken unterscheiden. Zuerst wird hierbei auf die Durchführungscharakteristiken im wissenschaftlichen System eingegangen und dann auf hemmende und begünstige Faktoren für die Evaluationsnutzung für die sich Forschende zu verantworten haben.

#### *2.2.2.1 Typologisierungen, Politikfelder und Akteurslandschaft*

Ein jeweiliger Evaluationsgegenstand muss mithilfe geeigneter Methodik und Untersuchungsdesign empirisch erfasst und in bewert- und messbare Größen umgewandelt werden. In verschiedenen Phasen des Policy- oder Programm-Zyklus können zusätzliche Rückkopplungsschleifen, kontinuierliches Monitoring oder selektive Evaluationen eingesetzt

werden. Grundsätzlich sind hierbei drei Typen zu unterscheiden (Scriven 1967, Rossi et al. 2004, Stockmann/Meyer 2014): Preformative oder ex ante Evaluierungen liefern bereits während der Entscheidungsfindung und im Prozess der Problem- und Politikformulierung die Grundlage für das Policy-Design. Die Wirkungsanalyse nimmt dabei Einfluss auf die konkrete inhaltliche Ausgestaltung der Policy. Im Gegensatz dazu haben formative Evaluationen das Ziel, bereits umgesetzte Policies während der Umsetzungsphase zu verbessern, indem im Vergleich von Soll-Vorgaben und dem Ist-Zustand Empfehlungen zu deren Optimierung abgegeben werden. Summative bzw. ex post Evaluierungen finden letztlich sowohl in der Implementationsphase als auch in der Abschlussphase statt. Sie bilanzieren den Erfolg der Policy und liefern die Grundlage für die Entscheidung, inwiefern die Maßnahme weitergeführt wird. Dieser Evaluationsform mangelt es zwar im Vergleich zu den beiden anderen Formen an Gestaltungsmöglichkeiten; sie kann aber am ehesten als tatsächliche Wirkungsanalyse bezeichnet werden, da untersucht werden kann, ob die Maßnahme die jeweilige beabsichtigte Wirkung hatte, Einflüsse anderer Faktoren weitgehend ausgeschlossen und nichtbeabsichtigte Wirkungen identifiziert werden können.

Zur Durchführung von Evaluationen eignen sich sowohl experimentelle bzw. quasi-experimentelle als auch qualitative Verfahren. Erstere legen den Fokus auf Objektivität, Präzision und Kausalzusammenhänge. Experimentelle Evaluationsstudien ermöglichen unverzerrte Schätzungen der Policy-Auswirkungen und bewertende Aussagen darüber, ob die Policy effektiv war bzw. zu gewünschten Veränderungen geführt hat. Qualitative Verfahren hingegen legen den Schwerpunkt auf die Komplexität des Evaluationskontextes sowie auf Daten, die sich nicht quantifizieren lassen. Sie eignen sich daher für explorative Forschung, um Beziehungen zwischen Outputs und Outcomes bzw. Impacts nachzuweisen. Eine Kombination beider Untersuchungsdesigns ist ein notwendiger ganzheitlicher Ansatz; Daten werden in vertiefenden Studien analysiert, die sich u.a. auf kontrafaktische Argumente stützen, um die tatsächlichen mit den erwarteten Ergebnissen abzugleichen.

Evaluationen können sowohl intern als auch extern durchgeführt werden. Interne Evaluierende sind Angehörige derselben Organisation, die für die Durchführung der Policies oder Maßnahmen verantwortlich sind. Wenn jene Expert\*innen sogar in der gleichen Abteilung der Organisation sitzen, die für Implementation zuständig ist, ist von der Subform Selbstevaluation die Rede. Externe Expert\*innen können hingegen beispielsweise für Forschungsinstitute, kommerzielle Managementberatungen oder Internationale Organisationen tätig sein

(Bogumil/Jann 2009). Selbstredend haben beide Formen spezifische Vor- und Nachteile: Während bei internen der geringere bürokratische Aufwand, die hohe Sachkenntnis, die schnelle Umsetzung der Ergebnisse und die Glaubwürdigkeit nach innen wertgeschätzt wird, stehen dem eine geringere Unabhängigkeit, Methodenkompetenz und Bereitschaft zur Kritik gegenüber (Caplan et al. 1975, Dunn 1980, Rich 1979). Wenngleich eine größere Unabhängigkeit bei extern Evaluierenden wiederum gegeben ist und damit auch eine höhere Glaubwürdigkeit und Legitimität nach außen, geht dies oft mit geringeren Sachkenntnissen und potentiell abwehrenden Reaktionen und Reaktanzen innerhalb der betroffenen Organisation einher.

Fortschritte in der Institutionalisierung von Evaluationen wurden meist durch die anbietende, die evaluierende Seite vorangetrieben. Es haben sich Institutionen und Interessenvertretungen wie der globale Dachverband International Organization for Cooperation in Evaluation (IOCE), die European Evaluation Society (EES), das Centrum für Evaluation (CEval) sowie die Deutsche Gesellschaft für Evaluation e.V. (DeGEval) gebildet, die laut Satzung die Akzeptanz und Nutzbarmachung von Evaluationen in der Gesellschaft fördern (DeGEval 2014). Doch das Fehlen einer übergreifenden Evaluationspolitik oder für Evaluationen zuständige Ministerien, Abteilungen oder Behörden sowie die unterschiedliche Nachfrage des politischen Systems machen Evaluationen stark kontextabhängig. Evaluation hat sich politikfeldabhängig weitgehend unabhängig institutionalisiert. Eine bereichsübergreifende Institutionalisierung ist nur rudimentär feststellbar. Entwicklungszusammenarbeit (EZ) ist dabei in Deutschland und darüber hinaus das Politikfeld, das am häufigsten evaluiert wird. Wichtige und große Player in der EZ wie das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) (u.a. 2008), die Gesellschaft für Internationale Zusammenarbeit (GIZ) (u.a. 2018), die Kreditanstalt für Wiederaufbau (KfW) (u.a. 2004; 2011) und die Organisation for Economic Co-operation and Development (OECD) (u.a. 2005; 2008; 2016) haben das Evaluierungsvorhaben zu einem unvergleichbaren Maße systematisiert und gestärkt. Beispielsweise wurden Evaluationskriterien etabliert (Relevanz, Kohärenz, Effektivität, Effizienz, Wirkung und Nachhaltigkeit: OECD 2020a) und professionelle Evaluationsstandards gesetzt (Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit: Deutsche Gesellschaft für Evaluation 2016). Doch selbst in diesem Pionier-Politikfeld wird gegenwärtig weniger als 0,5% der staatlichen Mittel Evaluationen zugesprochen; für unabhängige Evaluationen lediglich 0,2% (OECD 2016: 133).

### 2.2.2.2 Nicht realisiertes Potential und Änderungsbedarf in der Scientific Community

Entscheidungsträger\*innen und Forscher\*innen haben gänzlich unterschiedliche Charakteristiken und Arbeitsweisen. Demnach sind auch die hemmenden oder begünstigenden Faktoren, die vom wissenschaftlichen System verursacht oder verantwortet werden, gänzlich von politischen zu unterscheiden:

*Erstens*, externe wissenschaftliche Impulse wie Evaluierungen müssen zuerst in Regeln der internen Systemsteuerung, d.h. in den politischen Kosmos, übersetzt werden. Jene Übersetzungsmechanismen wissenschaftlicher Expertise in staatliches Handeln müssen jedoch optimiert werden (Janett 2003, Weiss 1987). Es mangle den Forschenden an Vertrautheit zum politischen Kosmos und auf der anderen Seite den politischen Entscheidungsträger\*innen an Vertrautheit zum Wissenschaftskosmos (Amara et al. 2004, Lavis et al. 2005, Nutley et al. 2007, Ouimet et al. 2009). Während Forschende Evidenz als etwas empirisch Erwiesenes ansehen, das wissenschaftlichen Standards der Objektivität, Reliabilität und Validität entspricht, haben politische Entscheidungsträger\*innen einen eher utilitaristischen und problemlösungsorientierten Ansatz in Bezug auf die Evidenz (Campbell et al. 2007). Wenn die Nutzung von Forschung verbessert werden soll, müssen demnach beide Seiten ihre Herangehensweisen anpassen. Ein gegenseitiges Verständnis zwischen politischen Fachleuten, Forschenden und Entscheidungstragenden ist unabdingbar für den Erfolg von Erfolgskontrollen (Head 2010).

*Zweitens*, dem Forschungssektor fehlt es auch an Verständnis für die pragmatischen Bedürfnisse von Politiker\*innen; gegenwärtige Formate von Forschungsergebnissen sind zu lang, zu methodisch, zu unzugänglich (Caplan et al. 1975, Head 2010, Lavis et al. 2005, Romsdahl 2005). Es brauche eine benutzerfreundliche und einfache Ergebnisdarstellung mit klaren und pragmatischen Handlungsimplicationen für Verbesserungen, die sowohl technisch als auch politisch und wirtschaftlich durchführbar sind (Sanderson 2002, Weiss 1987). Auch sollten die untersuchten Variablen solche sein, die sowohl von Interesse für politische Entscheidungsträger\*innen sind als auch von diesen tatsächlich beeinflusst werden können (Nelson et al. 1987).

*Drittens*, eine hohe Interaktion zwischen Forschenden und Entscheidenden sowie die konstante Einbeziehung der Entscheidenden bei der Konzeptualisierung und Durchführung einer Studie

sei der beste Prädiktor für ihre spätere Nutzung (Lavis et al. 2005). Nelson et al. (1987) appelliert für eine synthetisierte sozialwissenschaftliche Forschung, die politische Implikationen, Bedürfnisse und Besonderheiten miteinschließt. Elementar sei, dass Forschende und Evaluierende realistische Erwartungen haben, was den Einfluss ihrer Arbeit auf die Politikgestaltung betrifft (Hertin et al. 2009, Weiss 1999).

*Viertens*, es braucht ein begünstigendes und florierendes Universitätssystem und Forschendenkohorten mit Kompetenzen in der Datenanalyse und Politikbewertung, die daran interessiert sind, politikorientierte Forschung zu betreiben sowie Institutionen, die die Kluft zwischen wissenschaftlichen Inputs und politischen Zielen überbrücken können (Head 2010, Nelson et al. 1987, Weiss 1999). Evaluationen werden weniger von Entscheidungsträger\*innen oder deren Mitarbeitenden proaktiv gesucht, sondern sie müssen von Evaluierenden selbst oder vermittelnden Dritten, wie Wissenschaftsakademien, ministeriellen Beratungsgremien, Think Tanks oder Medien an sie herangetragen werden.

### 2.2.3 *Zivilgesellschaftliche Perspektive: Die Partizipation des sozialen Systems*

*“We should communicate what we know not only in the rooms and corridors of power but also in the shops and cafes where ordinary people gather, an informed public is an important support for informed policy”*

(Weiss 1999: 483 nach Lindblom 1990)

Die Aufgaben des modernen Staates als Versorger und Gestalter haben ebenso zugenommen wie die Ansprüche der Staatsbürger\*innen an die Rechenschaft, Integrität und Legitimität staatlichen Handelns. Regierende Instanzen müssen demnach neben der Selektion, Ausgestaltung und Umsetzung politischer Ziele ihr Handeln, ihre Macht und ihren Einsatz von Ressourcen gut zu begründen wissen (Müller 2015, Stockmann/Meyer 2017). Bürger\*innen wollen gezeigt bekommen, dass Politik deren Norm- und Wertvorstellungen entspricht und entlang der gesellschaftlichen Prioritäten effektiv und effizient umgesetzt wird. Wenn Evaluationsergebnisse dann transparent dem sozialen System zur Verfügung gestellt werden, kann dies wichtige Dialogmöglichkeiten eröffnen.

Wie geht die Zivilbevölkerung und die Wähler\*innenschaft nun mit Evaluationsergebnissen bzw. deren Interpretation durch politische Akteure um? Allmählich nutzen auch zivilgesellschaftliche Organisationen wie Verbände, Interessensorganisationen, Stiftungen und ehrenamtlich tätige Vereine Evaluationen. Gesellschaftliche Forderungen nach einer kritischen

Überprüfung politischen Handelns sind allgegenwärtig. Doch Forderungen nach mehr Evaluationen werden weder von Medienvertreter\*innen noch von Repräsentant\*innen der Zivilgesellschaft an die Politik gerichtet. Wenn überhaupt, übernimmt der Staat die Rolle des Initiators – durch Ausschreibungen im eigenen Auftrag oder durch das Einfordern von Evaluationen bei Maßnahmen, die von ihm kofinanziert werden. Entsprechende Bemühungen und Evaluationsvorhaben im politischen und wissenschaftlichen System scheinen im gesellschaftlichen nicht honoriert zu werden. Trotz der zunehmenden Anzahl an Veröffentlichungen von Evaluationen und Studien scheint das öffentliche Interesse daran sehr gering (Stockmann/Meyer 2017). Es gibt mit Ausnahme der DeGEval und einzelnen wissenschaftlichen Einrichtungen keine gesellschaftliche Gruppe, Vereinigung oder Partei, die aktiv mehr Evaluationen fordert. Ein öffentlicher Diskurs über Evaluationsergebnisse ist ebenso fast nicht existent. Druck seitens der Wählerschaft wird nicht ausgeübt (Weiss 1999). Medien berichten nicht oder zurückhaltend über Evaluationen. Ausnahmen sind Evaluationen mit hoher Flughöhe, die in einem stark politisierten Kontext entstanden sind oder ein extremes Fazit ziehen, wie z.B. zur Mietpreisbremse (taz 2020), der DSGVO (Heise 2020) oder dem Netzwerkdurchsetzungsgesetz (NetzDG) (Netzpolitik 2018). Das soziale System ist bisher weder als passiver Nachfrager noch als aktiver Treiber für Evaluationen in Erscheinung getreten.

Dabei gibt es neben den bereits aufgeführten Nutzungsfunktionen für das politische System wie politische Steuerung oder institutionelles Lernen zwei zentrale Funktionen von Evaluierungen, die sich direkt auf das gesellschaftliche System beziehen:

*Funktion 1: Legitimation & Rechenschaftslegung (u.a. Janett 2003, Rossi et al. 2004, Stockmann/Meyer 2014)*

Evaluationen dienen der Rechenschaftslegung staatlicher Organe gegenüber der Gesellschaft, indem sie im Rahmen von z.B. Kosten-Nutzen-Analysen die Effizienz bzw. Wirtschaftlichkeit staatlicher Maßnahmen offenlegen und prüfen, ob öffentliche Ressourcen nutzbringend und im Sinne der Gesellschaft eingesetzt werden. Der Legitimitätsnutzen bezieht sich somit auf die gesteigerte Glaubwürdigkeit politischer Akteure durch jenen Überprüfungsprozess. Indem sozialwissenschaftlich gezeigt wird, ob und inwiefern die anvisierten politischen Ziele erreicht wurden, welche Wirkungen ausgelöst wurden und wie Kosten im Verhältnis zum Nutzen

stehen, kann das zivile Verständnis und die Akzeptanz für politische Entscheidungen erhöht werden.

*Funktion 2: Gesellschaftliche Aufklärung & deliberative Partizipation (u.a. Romsdahl 2005, Hertin et al. 2009, Weiss 1999, Schlauffer 2017, Patton 1997; 2002, Sager 2018)*

Der Aufklärungsnutzen entsteht durch die Bereitstellung einer empirischen Basis für die politische und gesellschaftliche Selbstreflexion und konzeptionelles Lernen. Da nicht ausschließlich die intendierten Wirkungen der Maßnahme überprüft werden, sondern auch nicht-intendierte Folgen erfasst werden, werden bestehende Überzeugungen und Prozesse hinterfragt, Möglichkeiten für politische Veränderung eröffnet und neue Informationen und Perspektiven in das politische System eingebracht. Evaluation und die Einbeziehung von Evidenz kann darauf aufbauend zu informierten, fundierten und demokratischen Debatten sowie einer hohen Diskursqualität im Sinne der Habermas'schen Theorie der deliberativen Demokratie beitragen (Habermas 1973; 1981; 1992).

In Bezug auf den gesellschaftlichen Nutzen von Evidenz und Evaluationen entwerfen Campell und Russo (1999) die Vision einer „experimenting society“. Wenn die beiden aufgeführten Funktionen der Legitimation und Aufklärung aktiver genutzt werden würden, könne sich eine undogmatische, für Selbstkritik offene Gesellschaft etablieren, die sich für Innovation, progressives Lernen und die Überprüfung der vermeintlich sozialen Wirklichkeiten einsetzt (White 2019).

## 2.3 Forschungsfragen und Hypothesen

Ein stabiles und homogenes Gleichgewicht der verschiedenen systemischen Funktionslogiken und eine aktive Arbeitsteilung wird in jenem idealen Fall ermöglicht, wenn die Nachfrage von Evaluationen und die Partizipation des politischen und gesellschaftlichen Systems sowie die Institutionalisierung des wissenschaftlichen Systems ähnlich hoch ausgeprägt sind (Stockmann/Meyer 2017). Im Vergleich der drei zentralen Teilsysteme lässt sich jedoch – wie bereits zuvor aufgezeigt – ein starkes Ungleichgewicht feststellen, welches in ausreichendem Raum in dieser Forschungsarbeit berücksichtigt werden soll. Im Rahmen der Arbeit wird sich bei der empirischen Datenerhebung und Forschungsfrage demnach auf die Wahrnehmung der wissenschaftlich-kulturellen als auch auf die gesellschaftlich-sozialen Perspektive auf die Forschungsthematik konzentriert. Dieser akademische Fokus hat zweierlei Gründe: Erstens wird durch die bereits skizzierte umfangreiche Forschung zu den Befindlichkeiten des politischen Systems auf der einen Seite (Davies 2012, Head 2010, Hertin et al. 2009, Weiss 1978a; 1978b; 1989; 1999) und den Forschungslücken auf Seiten des wissenschaftlichen Systems insbesondere auf die Repräsentation von Vertreter\*innen der Wissenschaft wertgelegt. Zweitens verwiesen bisherige empirische und theoretische Arbeiten hinsichtlich der Verantwortung und Fähigkeit, aktiv zur Optimierung politischer Entscheidungsprozesse beitragen zu können, vor allem auf Wissenschaft und Gesellschaft, und nicht etwa politische Entscheidungsträger\*innen (Campbell/Russo 1999, Nelson et al. 1987, Stockmann/Meyer 2017).

Zwar gibt es bereits umfangreiche Analysen zu den Zielfunktionen, dem politischen Kontext, den Durchführungsmodalitäten und den bedingenden Faktoren; der Forschungsstand zu den konkreten Einflüssen auf die Gestaltung und Umsetzung von Politik und demnach zur genuinen Funktions- und Leistungsfähigkeit von Evaluierungen ist aber gering (u.a. Hertin et al. 2009, Knill/Tosun 2015). Aus diesem Grund soll untersucht werden, inwiefern evidenzbasierte Evaluationen aus Sicht von unmittelbar beteiligten Expert\*innen und Praktiker\*innen das Potential haben, ihre primäre Zielfunktion zu erfüllen und die Evidenzbasierung in politischen Entscheidungsprozessen zu erhöhen:

*Forschungsfrage: Inwiefern haben evidenzbasierte Evaluationen in postfaktischen Zeiten das Potential auf die politische Dynamik öffentlicher Entscheidungsprozesse Einfluss zu nehmen?*

Obwohl einige wenige Autor\*innen auch die Brücke zum postfaktischen Zeitalter schlagen und die Verbindung zu populistischen Charakteristiken suchen (u.a. Buch et al. 2019, Dobes 2017; 2018, Picciotto 2019), ist der Zusammenhang zwischen Evaluationen und einer gesellschaftlichen oder politischen Postfaktizität ebenfalls wenig bis gar nicht untersucht worden. Insofern es die empirischen Erkenntnisse ermöglichen, wird daher im Rahmen der Analyse und Interpretation der Ergebnisse insbesondere auf den Zusammenhang zwischen der Systematisierung und Etablierung evidenzbasierter Evaluationen und postfaktischen Diskursen in Politik und Gesellschaft eingegangen.

Zur Beantwortung der zentralen Forschungsfrage, inwiefern evidenzbasierte Evaluationen das Potential innehaben, auf die politische Dynamik in Entscheidungsprozessen öffentlicher Institutionen Einfluss zu nehmen, soll im ersten Schritt untersucht werden, inwiefern Evaluationen aus Sicht von Expert\*innen das Potential haben, ihre primäre Zielfunktion zu erfüllen und die Evidenzbasierung in politischen Entscheidungsprozessen zu erhöhen (vgl. Cronbach 1982, Furubo 1994, Hertin et al. 2009, Romsdahl 2005, Weiland 2013):

*H1: Der Einsatz evidenzbasierter Evaluationen verstärkt unter spezifischen Bedingungen die Evidenzbasierung politischer Entscheidungsprozessen öffentlicher Institutionen.*

Während eine zu untersuchende Evidenzbasierung politischer Prozesse das Ziel hat, im Wettkampf mit verschiedensten Entscheidungsfaktoren vermehrt relevante Fakten und Evidenzen Bedeutungskraft zu verschaffen, geht der Prozess einer Rationalisierung oder gar De-Politisierung einen Schritt weiter und intentioniert, evidenz- und indikatorgeleitete sowie wertneutrale Entscheidungsprozesse von politischen Entscheidungsträger\*innen zu generieren. Inwiefern Evaluationen diesen Funktionen nachkommen (können), wird im Rahmen der zweiten Hypothese untersucht (vgl. Fischer/Forester 1993, Howlett 2009, Jun/Grabow 2008, Lane 1966, Majone 1989, Radaelli 1995, Stockmann/Meyer 2017):

*H2: Der Einsatz evidenzbasierter Evaluationen befördert unter spezifischen Bedingungen die Rationalisierung und De-Politisierung politischer Entscheidungsprozessen öffentlicher Institutionen.*

Zentraler Bestandteil beider Hypothesen ist die Prämisse, dass spezifische Bedingungen für die Beauftragung, Durchführung, Implementierung und Vermittlung von Evaluationen in den

Systemen Politik (vgl. Head 2010, Tarschys 1983, Weiss 1987; 1999), Wissenschaft (vgl. Lavis et al. 2005, Nutley et al. 2007, Olejniczak 2016) und Zivilgesellschaft (vgl. Patton 2002, Sager 2018, Schlauffer 2017, White 2019) erfüllt sind. Jene zum Erfolg von Evaluationen beitragenden Voraussetzungen sollen im Rahmen dieser Arbeit ebenfalls aufgrund der Expert\*innenaussagen systematisiert und ergebnisoffen analysiert werden.

### 3. Methodik und Forschungsdesign

Der Beantwortung der Forschungsfragen dienten zwei qualitative Hauptdatenquellen. Erstens boten relevante Dokumente u.a. Gesetzestexte, Grundsatzdokumente, Evaluierungsleitfäden, exemplarische Evaluationen und einschlägige Veröffentlichungen involvierter Akteure eine breite Grundlage für die Analyse. Zweitens konnten die Nutzung und Einflussnahme von Evaluationen in einem multikausalen politischen Prozess durch umfangreiche Interviews mit den Akteuren, die im Zentrum der Prozesse stehen, nachvollzogen werden. Sowohl die Durchführung als auch die Auswertung der Expert\*inneninterviews basiert auf den Erkenntnissen, Definitionen und Leitlinien von Gläser und Laudel (2010).

#### 3.1 Semistrukturierte Leitfadeninterviews

Aufgrund der gewählten Thematik, den komplexen Forschungsfragen und der intendierten Diversität der Antworten erschienen semistrukturierte und qualitative Leitfadeninterviews mit Expert\*innen als geeignetste Methodik (wie auch bei thematisch ähnlichen Studien wie Hertin et al. 2009 und Eckhard/Jankauskas 2018). Die wesentliche Aufgabe und Funktion jener Expert\*inneninterviews besteht darin, „*dem Forscher das besondere Wissen der in die Situationen und Prozesse involvierten Menschen zugänglich zu machen.*“ (Gläser/Laudel (2010:13). Die Methode bewahrt die nötige Flexibilität hinsichtlich der Strukturierung des Interviews, der Formulierungen und des individuellen Antwortverhaltens der Expert\*innen. Die gewählte Befragungsmethode ermöglicht des Weiteren einen offenen und natürlichen Dialog in einer möglichst ungezwungenen Atmosphäre. Die vorher zugesicherte Nicht-Nennung von personenbezogenen Daten sowie die strikte Anonymisierung der Transkripte ermöglicht zudem einen höchstmöglichen Grad an Offenheit und Authentizität. Dies erlaubte den Befragten einerseits, ohne potentielle Sanktionsängste hinsichtlich des Arbeitgebers wahrheitsgemäß zu antworten und erlaubte andererseits der Interviewerin, das Gespräch zu steuern und bei Bedarf und in Abhängigkeit zum individuellen Antwortverhalten des Gesprächspartners tiefer in bestimmte Fragen einzusteigen. Durch die festgelegte Leitfadenchronologie und -logik ermöglicht die gewählte Befragungsmethodik trotz der damit einhergehenden Breite der Antworten eine interpersonal-vergleichbare Analyse.

### 3.2 Sampling: Auswahl der Stichprobe

Zur Identifizierung potentieller Gesprächspartner\*innen wurden Expert\*innen gesucht, die nach der Definition von Gläser und Laudel (2010: 12) „als Quelle von Spezialwissen über die zu erforschenden (...) Sachverhalte“ dienen könnten. Die größte Herausforderung bei der Rekrutierung jener hochrangiger Expert\*innen und involvierten Akteuren bei einem politisch sensiblen Thema ist der Zugang. Für die Identifizierung möglicher Ansprechpartner\*innen wurde eine Liste der wichtigsten Evaluationsakteure und Arbeitgeber\*innen erstellt. Darauf aufbauend wurden jene Arbeitnehmer\*innen selektiert, die durch ihr Profil und ihre Evaluierungserfahrung am geeignetsten zur Beantwortung der Fragen schienen. Primäre Kriterien zur Kontaktaufnahme und Selektion waren erstens eine dezidierte, d.h. jahrelange oder meist jahrzehntelange, Evaluationserfahrung, zweitens ein etablierter Evaluationsakteur als Arbeitgeber, drittens eine höchstmöglich ausgewogene Soziodemographie zwischen den Befragten, um eine möglichst diverse Stichprobe sicherzustellen.

Bezüglich der Arbeitgeber\*innen war es essentiell, dass diese in Bezug auf ihre Formalisierung, ihre Politisierung, ihren institutionellen Aufbau und die allgemeine Schwerpunktsetzung variierten. Die Expert\*innen wurden folglich ungefähr hälftig der aktiv evaluierenden und der passiv evaluierten oder auftraggebenden Perspektive zugeordnet. Die jeweiligen Befragten mit einer evaluierenden Perspektive sind bei Arbeitgeber\*innen beschäftigt, die im Auftrag externer Institutionen in verschiedensten Kontexten Evaluationen durchführen – u.a. bei öffentlichen Forschungsinstituten, Internationalen Organisationen, kleineren Beratungsunternehmen oder internationalen Big 5-Managementberatungen. Befragte mit einer evaluierten oder auftraggebenden Perspektive arbeiten für staatliche Organe wie Ministerien oder öffentlich finanzierte Träger wie Stiftungen, Vereine oder Förderbanken.

Abbildung 1: Übersicht der befragten Expert\*innen

Name	Position	Arbeitgeber	Schwerpunkt	Datum
Expert*innen der evaluierenden Perspektive				
Expertin 1	Führungskraft	Internationale Organisation	Evaluierungen, Audits und Qualitätsmanagement	06.05.2020
Expertin 2	Evaluatorin	Internationale Organisation	Organisationsübergreifende Prozess- und Policy-Evaluierungen	07.05.2020

Expertin 3	Führungskraft und Gremiums-vorsitzende	Internationale Organisation	Managementevaluierungen und Inspektionen	08.05.2020
Experte 4	Evaluator	Internationale Organisation	Organisationsübergreifende Evaluierungen	19.05.2020
Experte 5	Führungskraft	Beratungs- und Prozessbegleitungsunternehmen	Evaluierungen und Wirkungsmonitoring in der Entwicklungszusammenarbeit	26.05.2020
Experte 6	Evaluator	Evaluierungsinstitut	Rigorese Wirkungsevaluierungen in der Entwicklungszusammenarbeit	28.05.2020
Expertin 7	Evaluatorin	Forschungs- und Beratungsunternehmen	Evaluationen von Gesetzen und Förderprogrammen auf bundesweiter sowie lokaler Ebene	29.05.2020
Experte 8	Führungskraft	Forschungszentrum	Evaluationen und Impact Assessments wirtschaftspolitischer Maßnahmen auf bundesweiter Ebene	29.05.2020
Expertin 9	Evaluatorin	Internationale Managementberatung	Ex-post Policy-Evaluationen und Impact Assessments im Auftrag der Europäischen Union	03.06.2020
Experte 10	Evaluator	Internationale Managementberatung	Evaluationen von Gesetzen auf nationaler und internationaler Ebene	04.06.2020
Experte 11	Evaluator, Forscher und Dozent	Forschungsinstitut	Gesetzesfolgenabschätzungen und Evaluationen im Auftrag von Bundes-, Landes- und Kommunalbehörden	08.06.2020
Expert*innen der evaluierten und auftraggebenden Perspektive				
Experte 12	Referatsleiter	Ministerium eines Bundeslandes	Evaluationen zu finanzpolitischen Maßnahmen und Fördermitteln	18.06.2020
Experte 13	Leiter der Evaluierungs-abteilung	Politische Stiftung	Evaluationen der internationalen Stiftungsarbeit sowie Qualitätsmanagement	18.06.2020
Experte 14	Referatsleiter und Dozent	Ministerium eines Bundeslandes	Wirtschafts- und industriepolitische Evaluierungen	18.06.2020
Experte 15	Leiter der Evaluierungs-abteilung	Politische Stiftung	Evaluationen, Wirkungsmonitoring und Qualitätssicherung der Stiftungsaktivitäten	19.06.2020
Expertin 16	Leiterin der Evaluierungs-abteilung	Staatlich finanziert Verein	Evaluationen kulturpolitischer und demokratiefördernder Maßnahmen	22.06.2020
Experte 17	Leiter der Evaluierungs-abteilung	Förderbank	Evaluationen von entwicklungspolitischen Förderprogrammen	24.06.2020

### 3.3 Aufbau und Inhalt des Leitfadens

Die Interviewleitfragen dienten zur Hinführung an die Beantwortung der Forschungsfragen sowie zur Überprüfung der Hypothesen. Des Weiteren sollten sie die nötige Ergebnisoffenheit und Tiefe sowie einen explorativen Charakter wahren, um neue Blickwinkel auf die Thematik und unerwartete Erkenntnisse nicht im Voraus zu verhindern (Gläser/Laudel 2010). Der Leitfaden diene darüber hinaus den Expert\*innen zur Vorbereitung auf und Orientierung während des Interviews sowie der Interviewerin zur Strukturierung und Durchführung des Gesprächs.

Die inhaltliche Struktur des Leitfadens wurde literaturgestützt entwickelt und in Anlehnung an die theoretischen und empirischen Erkenntnisse der zuvor vorgestellten Autor\*innen konzipiert. Auf Basis dieses Forschungsstandes und der zentralen Konzepte wurden zwei inhaltlich korrespondierende Leitfäden konzipiert; Nummer 1 für die evaluierende Perspektive mit 8 Leitfragen und Nummer 2 für die evaluierte Perspektive mit 9 Leitfragen, die zur Operationalisierung der Forschungsfrage dienten. Der Leitfaden beginnt mit einer kurzen Beschreibung des Forschungsvorhabens, des Auswertungsprozesses und einer Danksagung. Darauf folgen die Fragen in zwei thematischen Blöcken: Der erste Block beschäftigt sich deskriptiv mit dem Prozess und den Problematiken einer Evaluation und der zweite analysierend und interpretativ mit den Folgen und dem Potential derer. Der Vielfalt der beruflichen Hintergründe und dem Zeitrahmen der Interviews wurde Rechnung getragen. Alle Themen- und Fragenkomplexe waren offen konzipiert und sowohl sachlich-beschreibend als auch subjektiv-interpretativ:

#### *Prozess & Problematiken*

- Voraussetzungen für die Beauftragung, Durchführung, Verwertung und Implementierung von Evaluationen und Evaluationsergebnissen
- Herausforderungen und Problematiken in den verschiedenen Evaluationsphasen
- Einflussmöglichkeiten auf die Evaluationsmethoden, -akteure und -prozesse
- Das Verhältnis des finanziellen, personellen und zeitlichen Aufwands zur Umsetzung der Evaluationserkenntnisse
- Grundlegende Spannung zwischen und Vereinbarkeit von evidenz- und indikatorbasierten Evaluationen und subjektiver, meinungsbasierter Politik

### *Folgen & Potential*

- Einfluss auf eine Erhöhung der Effizienz bzw. Wirtschaftlichkeit, der Effektivität bzw. Wirksamkeit, der Nachhaltigkeit und Langfristigkeit, der Rechenschaftspflicht und Kontrolle, der Legitimation und Gemeinwohlverträglichkeit sowie eine Rationalisierung oder De-Politisierung
- Einfluss von systematischen Evaluationen auf ziviles Vertrauen in die Integrität, Effektivität und Gemeinnützigkeit öffentlicher Institutionen
- Einfluss eines quantitativen Anstiegs an öffentlichen Evaluationen auf den gesellschaftlichen Zuspruch für nicht-evidenzbasierte oder ideologische Politik

### 3.4 Setting: Vorbereitung und Durchführung der Interviews

Für die erste Kontaktaufnahme wurde die Business-Networking-Plattform LinkedIn gewählt, da zum einen im Gegensatz zu beruflichen Mail-Adressen eine gewisse wünschenswerte Distanz zum Arbeitgeber gewahrt werden konnte und zum anderen die Einsicht in das gegenseitige Profil, Vertrauen in die Integrität und Kompetenz für das gewählte Themengebiet der Interviewerin aufgebaut werden konnte.

Insgesamt wurden 26 Interviewanfragen mit individuellem Anschreiben versendet (Anhang 7.1). Das Schreiben enthielt eine individuelle Begründung für die Auswahl des jeweils Adressierten, die Vorstellung des akademischen und beruflichen Profils der Interviewerin, den Kontext der Befragung sowie Details zum Ablauf des Interviews. Von 26 Anfragen gab es 17 Rückmeldungen, womit eine Rücklaufquote von 65% zu verzeichnen ist. Durch den nachträglichen Kontakt zu Expert\*innen mit denen bereits Interviews durchgeführt wurden, kam es zur Weitervermittlung von Kontaktmöglichkeiten anderer potentiell interessierter Expert\*innen, die entsprechend wahrgenommen wurden. Mit allen sich rückmeldenden Expert\*innen wurden Interviews durchgeführt. Der überraschend hohe prozentuelle Anteil an Rückmeldungen spricht für die Aktualität und Relevanz der Thematik und das Interesse involvierter Akteure, sich in einer solchen Debatte aktiv einzubringen. Als weiterer Grund zur Teilnahme wurde die Zusicherung genannt, den finalen Forschungsbericht den zum Forschungsanliegen beitragenden Teilnehmenden zuzusenden.

Nach der Interviewzusage folgten in einem zweiten Schritt die Zusendung des Leitfadens sowie die Absprache von Termin und Details. Auf Wunsch wurde eine DSGVO-konforme

Datenschutzerklärung zur Verwendung der personenbezogenen Daten ausgehändigt und beidseitig unterschrieben. Die große Mehrheit der Interviews wurde mithilfe der Videokonferenz-Tools WebEx, Zoom, Skype und Microsoft Teams durchgeführt, um die Verständlichkeit und Verbindlichkeit zu erhöhen. Falls dies nicht erwünscht war, wurde das Interview telefonisch durchgeführt. Die Interviews erstreckten sich über einen Zeitrahmen von 40 bis 90 Minuten. Das jeweilige Einverständnis zur Gesprächsaufzeichnung und zur Verwendung des Transkripts wurde zu Beginn jedes Gesprächs eingeholt. Den Gesprächspartner\*innen wurde zugesichert, dass deren Identität gegenüber Dritten geschützt bleibt, die Transkripte anonymisiert und die Video- und Audio-Aufnahmen spätestens zum Zeitpunkt der Abgabe der vorliegenden Arbeit unwiderruflich gelöscht werden. Die Interviews begannen mit einer kurzen Einführung und Vorstellung beider Parteien und gingen dann in die konkrete Befragung anhand des Leitfadens über.

In der zweimonatigen Datenerhebungsphase in den Monaten Mai und Juni 2020 wurden somit die individuellen Einschätzungen von 17 Expertinnen und Experten eingeholt. Diese bilden die empirische Datengrundlage für die Ausarbeitung dieser Arbeit.

### 3.5 Datenaufbereitung und Analyseverfahren

Im Anschluss an alle Interviews wurden die Video- und Audio-Dateien detailliert und wortgetreu zur Dokumentation und späteren Auswertung transkribiert (Anhang 7.4). Small Talk und für die Forschungsfragen irrelevante Exkurse wurden ausgelassen.

Nach Prüfung und Abwägung verschiedener inhaltlich strukturierender und evaluativer Inhaltsanalyseverfahren (von u.a. der Grounded Theory nach Glaser/Strauss 2010, Mayring 2020, Kuckart 2016 und Glaser/Laudel 2010) wurde sich für eine empirisch begründete Theoriebildung durch Kombination von deduktivem und induktivem Codieren nach Glaser und Laudel (2010) entschieden, einer Modifizierung der von Mayring konzipierten Methode der qualitativen Inhaltsanalyse. Es wurde demzufolge im ersten Schritt eine deduktive Grobcodierung vorgenommen, indem zunächst thematische Hauptkategorien durch die Ableitung aus theoretisch hergeleiteten Forschungsfragen und Interviewleitfragen entwickelt wurden. Im zweiten Schritt folgte eine induktive Feincodierung und das Bestimmen von ausdifferenzierten Subkategorien am Interviewmaterial, um das vorliegende Material zu kontrastieren und abstrahieren.

Als technisches Tool wurde MAXQDA, eine Software zur Analyse qualitativer Daten, zur Auswertung und Codierung der Interviews verwendet. Entlang des beschriebenen Forschungsinteresses und dem Interviewleitfaden wurden zuerst Codes deduktiv erstellt und auf das Material angewandt. Wenn während der Kodierung ersichtlich wurde, dass die im ersten Schritt erstellten deduktiven Codes nicht erschöpfend waren, wurden zusätzliche Über- und Unter-codes induktiv erstellt, wodurch im Sinne von Gläser und Laudel eine kontinuierliche Anpassung und Ergänzung der Ausprägungen am Material stattfand.

Das Codesystem beinhaltet die folgenden 8 Codes und 17 Subcodes:

- Code: Voraussetzungen und Erfolgsfaktoren
  - Subcode: Phase 1: Beauftragung
  - Subcode: Phase 2: Durchführung
  - Subcode: Phase 3: Implementierung
- Code: Herausforderungen und Problematiken
- Code: Beeinflussung und strategische Instrumentalisierung
  - Subcode: Nicht-Evaluierende
  - Subcode: Evaluierende
- Code: Kosten-Nutzen-Verhältnis
- Code: Spannungsverhältnis Wissenschaft und Politik
- Code: Funktionen und Potential
  - Subcode: Verstärkung institutioneller und politischer Lernprozesse
  - Subcode: Rationalisierung und De-Politisierung
  - Subcode: Erhöhung der Legitimation und Gemeinwohlverträglichkeit
  - Subcode: Erhöhung der Rechenschaftspflicht und Erfolgskontrolle
  - Subcode: Erhöhung der Nachhaltigkeit und Langfristigkeit
  - Subcode: Erhöhung der Effektivität und Wirksamkeit
  - Subcode: Erhöhung der Effizienz und Wirtschaftlichkeit
- Code: Nutzung und Vermittlung
  - Subcode: Auftragnehmer bzw. Wissenschaft
  - Subcode: Zivilgesellschaft und Medien
  - Subcode: Auftraggeber bzw. Politik

- Code: Zurückdrängung nicht-evidenzbasierter Diskurse
  - Subcode: Gesellschaft
  - Subcode: Politik

Das erstellte Codebuch wird im Anhang 7.5 offengelegt. Die Befragten sind zwecks der Anonymisierung in chronologischer Reihenfolge durchnummeriert. Zitate oder Analysen, die sich auf konkrete Interviews beziehen, werden in Bezug auf die Nummer des Absatzes im jeweiligen Interview kenntlich gemacht, z.B. [Expertin 3: 12].

## 4. Forschungsergebnisse

Unter Rückbezug auf die Forschungsfrage, inwiefern evidenzbasierte Evaluationen das Potential haben, auf die politische Dynamik in Entscheidungsprozessen öffentlicher Institutionen Einfluss zu nehmen, und die damit verbundenen Hypothesen folgt die Darstellung, Beschreibung und Interpretation der zentralen Forschungsergebnisse der qualitativen Interviewauswertung und -analyse. Zunächst wird das Potential einer verstärkten Evidenzbasierung öffentlicher Entscheidungsprozesse (H1) und das Potential einer Rationalisierung und De-Politisierung jener Prozesse (H2) untersucht sowie eine vergleichende Schlussfolgerung beider Hypothesen gezogen. Im zweiten Schritt werden die wesentlichen Erfolgsbedingungen und Voraussetzungen für das in der Überprüfung der Hypothesen dargelegte Potential strukturiert und analysiert. Abschließend folgt eine Diskussion über die Forschungsergebnisse und deren praktische Implikationen.

### 4.1 Evaluation als evidenzbasierte Deliberationsgrundlage für politische Entscheidungen

Während eine zu untersuchende Evidenzbasierung politischer Prozesse das Ziel hat, im Wettkampf mit verschiedensten Einfluss- und Entscheidungsfaktoren vermehrt relevante Fakten und Evidenzen Bedeutungskraft zu verschaffen, geht der Prozess einer Rationalisierung oder gar De-Politisierung einen Schritt weiter und intentioniert, evidenz- und indikatorgeleitete sowie wertneutrale Entscheidungsprozesse von politischen Entscheidungsträger\*innen zu generieren. Inwiefern Evaluationen diesen Funktionen nachkommen (können), wird wie folgt empirisch untersucht.

#### 4.1.1 H1: Die verstärkte Evidenzbasierung öffentlicher Entscheidungsprozesse

Während sich die Expert\*innen bei den theoretisch-deduktiv erstellten Funktionskategorien und deren Gewichtung oft uneins waren (z.B. Experte 17: 12 vs. Experte 14: 25), so ergibt sich bei der Frage nach der grundlegendsten Kernfunktion von Evaluierungen – der Evidenzbasierung – ein unstrittiges Bild: Das Potential von evidenzbasierten Evaluationen sei für ein tragfähiges Checks-and-Balances-System immens, da sich Entscheidungsträger\*innen in ihren Entscheidungsprozessen vermehrt auf belastbare und messbare Informationen aus Evaluationen stützen, damit transparentere Entscheidungen erleichtert werden würden,

konstatieren etwa Experte 4 (21/25) und 15 (11). Evaluationen würden eine gemeinsame, sachliche und unideologische Faktenbasis schaffen, auf Basis derer man wiederum auf einer übergeordneten Ebene subjektiv-politisch diskutieren könne (u.a. Experte 5: 35, Experte 8: 17, Experte 6: 35). Evaluationen könnten somit als vorgelagerte Voraussetzung für evidenzbasiertes Handeln angesehen werden. Auch wenn die reine Existenz und Verfügbarkeit von Evaluierungen zwar keinerlei Garantie für deren Nutzung sei, biete die „Evaluierungsfunktion (...) die Möglichkeit, evidenzbasierte Entscheidungsfindung überhaupt erst zu ermöglichen.“ (Expertin 4: 19). Je evidenz- und faktenbasierter Evaluierungen seien, die Entscheidungstragenden zur Verfügung gestellt werden, desto mehr könne man von einem faktenbezogenen, wenngleich nicht faktengesteuerten Entscheidungsfindungsprozess ausgehen (Expertin 4: 3/19). Es sei „schwieriger, eine Maßnahme einzuführen, die rein politisch auf bestimmten Interessen begründet ist, wenn eine Evaluation durchgeführt worden ist (...)“ (Expertin 9: 26).

Emotionale Debatten über normative Zielvorstellungen würden Evaluationen weder ersetzen wollen noch können, so Experte 8 (17). Doch was Evaluation leisten könne, sei eine konkrete evidenzbasierte Entscheidungsgrundlage darzulegen und zu stärken in Situationen, in denen politische Ziele bereits festgelegt wurden oder wenn konträre Zielsysteme zur Debatte stehen und Unsicherheiten zu beantworten sind, z.B. mit welchen Maßnahmen bestimmte Ziele erreicht werden könnten oder ob die kalkulierbaren Folgen den politisch gesetzten Zielen entsprechen (Experte 5: 27). Eine Vielzahl der Expert\*innen betonte, dass Evaluationen eine wichtige evidenzbasierte Diskurs- und Deliberationsgrundlage für politische und wertorientierte Entscheidungen bilden (u.a. Experte 8: 17, Experte 5: 27, Expertin 4: 19).

Die auf einer Evidenzbasierung aufbauenden, weiterführenden Evaluationsfunktionen sind nicht minder relevant. Laut den Expert\*innen tragen Evaluationen dazu bei, dass vermeintliche Selbstverständlichkeiten im Rahmen von Lernprozessen hinterfragt werden (u.a. Experte 17: 12, Experte 5: 33), die individuellen Entscheidungen gegenüber Geldgebern und Partnern legitimiert werden (u.a. Expertin 16: 18, Experte 13: 26), Steuergelder durch Erfolgskontrollen gerechtfertigt eingesetzt werden (u.a. Experte 15: 9, Experte 13: 4), politische Maßnahmen langfristig und unabhängig von Legislaturperioden konzipiert werden (u.a. Experte 11: 19, Experte 10: 22) sowie in der Folge die Effektivität (u.a. Experte 11: 29, Experte 13: 16) sowie Effizienz (u.a. Experte 6: 33, Expertin 9: 24) öffentlicher Regulierungen erhöht werden. Experte 10 (18) betonte, hier müsse das bisher zu wenig genutzte Potential etwa zur Kontrolle von

parlamentarischem und staatlichem Handeln stärker genutzt werden. Evaluationen könnten ein „Korrektiv“ sein, sodass „sich Politik tatsächlich auf eine systematischere Art und Weise mit ihren eigenen Fehlern befasst“ (Experte 13: 32).

Die Analyse der Expert\*innenaussagen offenbart wie im Folgenden dargestellt umfangreiche Einschränkungen des Potentials, Evidenzen aus Evaluationen in Entscheidungsprozesse politischer Entscheidungsträger\*innen einfließen zu lassen. Dennoch bestätigt eine klare Mehrheit der Expert\*innen das Potential zu einer verstärkten Evidenzbasierung bei Erfüllung der nötigen, aber derzeit noch nicht geschaffenen Voraussetzungen (H1):

*H1: Der Einsatz evidenzbasierte Evaluationen befördert unter spezifischen Bedingungen die Evidenzbasierung politischer Entscheidungsprozessen öffentlicher Institutionen.*

#### *4.1.2 H2: Die vermeintliche Rationalisierung und De-Politisierung öffentlicher Entscheidungsprozesse*

Neben den anderen bereits erwähnten Funktionskategorien wie etwa die Steigerung der Evidenzbasierung oder die Erhöhung der Effizienz, Effektivität und Langfristigkeit gingen die Expert\*innen am häufigsten auf die Funktion der Rationalisierung und De-Politisierung ein. Die Kritik und Ernüchterung unter den Expert\*innen hinsichtlich eines potentiellen Rationalisierungspotentials bzw. De-Politisierungspotentials war jedoch groß (u.a. Expertin 3: 26, Experte 6: 19, Expertin 4: 23, Experte 15: 15). Zwar würde durch die reine Existenz des Evaluationsprozesses eine Entpolitisierung suggeriert werden, am Ende würden jedoch durch die klare Hierarchie zwischen den politischen Auftraggebenden und der zuarbeitenden Wissenschaft, die politischen Interessen dominieren, so Experte 11 (29).

*„(...) meistens ist die Erfahrung so, dass die Evaluatoren durchaus geneigt sind, in gewissem Umfang den politischen Erwartungen des Auftraggebers entgegen zu kommen.“  
(Experte 14: 9)*

Der evaluative Kontext ist und bleibt hochpolitisiert (Experte 10: 22, Expertin 9: 28, Expertin 3: 2). Im Gegensatz zu naturwissenschaftlichen Gesetzmäßigkeiten seien die prinzipiell nicht verifizierbaren Aussagen in Evaluationen jederzeit anfechtbar, instrumentalisierbar und korrumpierbar (Experte 11: 29). Expertin und Evaluatorin 3 (2) kritisierte dieses Spannungsverhältnis wie folgt:

*„When I talk about politics getting in the way, I mean that in a way that is negative. Absolutely negative. (...) political agendas are put ahead of evidence and efficiency and effectiveness.“*

Eine Vielzahl der Expert\*innen bestätigte, dass Fachpolitiker\*innen zwar oft Evaluationen begrüßen und auch aufgreifen, politische Entscheidungen jedoch in der letzten Instanz nicht aufgrund von Sachlichkeit und Evidenz getroffen werden würden (u.a. Experte 6: 19, Experte 12: 12/22). Dies sei jedoch nicht zwangsläufig illegitim, da politische Entscheidungsträger\*innen diese letztlich auch verantworten müssten (Experte 12: 22, Expertin 9: 28) und Evidenzen nicht alle anderen Aspekte überlagern müssten (Experte 15: 11, Experte 5: 27). Ziel sei keine „komplett objektivierbare Politik“ (Experte 13: 34). Die vorherrschende Entscheidungssteuerung durch subjektive Meinungen und Werte sei zwar unumstößlich, es sei allerdings auch nicht zielführend, dies als etwas gänzlich Abträgliches zu sehen (Expertin 9: 28). Denn die „Gefahr, dass die Evidenzlogik der Evaluation sich im Rahmen der zunehmenden Neoliberalisierung bestimmter Politikbereiche (...) verselbstständigt“ sei ebenfalls groß, so Experte 5 (27).

Letztlich sei es auch elementar, dass die Abnehmer von Evaluationen nicht tatsächliche Politiker\*innen sein sollten, sondern die operative Arbeitsebene, die die Kompetenz und Kapazitäten hat, sich den Evaluationsergebnissen gerecht anzunehmen (Experte 15: 5/13). Denn wenngleich die Zielfestlegung den Politiker\*innen obliege, werde die konkrete Ausgestaltung der Gesetze und Politiken bzw. detaillierte steuerungsrelevante Fragen in den Ministerien und deren Referate beschlossen (Experte 10: 18). Evaluierungen könnten und sollten demnach keine großen wertegeleiteten Entscheidungen steuern und versuchen, den individuellen Wert oder die Priorisierung von politischen Zielen festzulegen. Aus Interessenpolitik sollte sich Evaluationen bestmöglich raushalten, um nicht als „Totschlagargument“ benutzt zu werden, bestimmte Politiken zu rechtfertigen (Experte 5: 27). Experte 15 (15) betonte, dass Konflikte zu kanalisieren und die eigene Machtposition durchzusetzen, ein „seit Äonen gleicher“ Prozess sei. Große strategische oder ideelle Fragen würden nicht durch Evaluierungen abgeleitet, seien nicht evidenzbasiert und würden es auch nie sein. Man könne so viele Evidenzen haben wie man wolle, wenn diese nicht zu einer Abwandlung der eigenen strategischen Position passen, würden sie durch die Gefahr vor Stimmenverlust auch nicht berücksichtigt:

*„wenn's positiv evaluiert ist, dann ist gut und wenn es negativ evaluiert ist, wird's trotzdem gemacht, weil es ja politisch getrieben war“  
(Experte 12: 32)*

Auch in Bezug auf postfaktische Diskurse in der politischen Arena wird ein kleiner Beitrag von Evaluationen zwar eingeräumt (u.a. Experte 6: 33, Experte 13: 32), aber ein signifikanter Einfluss negiert (u.a. Experte 10: 26, Experte 5: 39). Als primärer Grund dafür wird die strategische Instrumentalisierung bzw. Politisierung von Evaluationen aufgeführt (Expertin 1: 25, Expertin 3: 14/26/28, Experte 11: 29, Experte 10: 10/18, Experte 13: 4):

*„When people don't like your results, they attack you and they attack the organization you represent. (...) It's not pretty. Whenever you conduct controversial evaluations or reviews, you can expect it. (...) You better be an organization with impeccable integrity and reputation and you individually better be same.“*

(Expertin 3: 24)

Von vielfachen Situationen der strategischen Einflussnahme war die Rede; Experte 15 (17) etwa berichtete, dass die AfD zu Beginn ihres Bundestagsmandats auffallend viele Evaluierungen in Debatten herangezogen habe oder sie in kleinen Anfragen angefordert habe, „um deren Begründungsmuster in Bundestagsdebatten zu unterfüttern“ und den „Anschein einer Wissenschaftlichkeit [zu] vermitteln.“ Beispielsweise forderte die AfD in Bundestagsanträgen eine Evaluierung des NetzDG (Bundestag 2020), des „deutschen politisch-militärisch-zivilen Engagements in Afghanistan“ (Bundestag 2018) oder der Gesamtkosten der „Willkommenskultur“ (AfD 2020). Expertin 3 (28) konstatiert: “you have to be careful what you are evaluating because it's so easily discredited by populist politics.” Auch Experte 6 (25-27) schilderte, dass Evaluationsberichte in Ausschüssen durch die politisch begründeten Motive der Vorsitz innehabenden Partei, in diesem Fall der CSU, zu spezifischen Themen nicht zugelassen wurden.

*„Jeder, der irgendeine Meinung hat, findet auch irgendeine Studie dazu, die das unterstützen kann.“*  
(Expertin 7: 52)

Die Analyse der Aussagen zeigt deutlich, dass zwar eine klare Mehrheit der Expert\*innen das Potential für eine zunehmende Evidenzbasierung bestätigt (H1), dies gehe aber nicht zwangsläufig mit einer Rationalisierung oder De-Politisierung und einer Ablöse meinungsbasierter oder ideologischer Kämpfe einher, weswegen H2 nicht bestätigt werden kann. Es wird deutlich, dass die Befragten überzeugt davon sind, dass das Ziel „to counter the political“ (Expertin 3: 20) kein Kampf ist, der durch Evaluierungen gefochten werden sollte und auch nicht deren primären Zweck darstellt:

*H2: Der Einsatz evidenzbasierte Evaluationen befördert unter spezifischen Bedingungen die Rationalisierung und De-Politisierung politischer Entscheidungsprozessen öffentlicher Institutionen.*

Auf den vorherigen Analysen aufbauend, ergibt sich das Bild von Evaluation als einer Diskurs- und Entscheidungsgrundlage für evidenzbasierte – und nicht etwa evidenzbestimmte oder evidenzgeleitete – Entscheidungen. Evaluationen erscheinen zwar als vorgelagerte Voraussetzung für evidenzbasiertes Handeln (H1), die darauf aufbauenden normativen und subjektiven Diskurse sind jedoch genuin politisch und werden auf einer übergeordneten und aggregierten Ebene ausgetragen (H2).

Die klare Differenzierung zwischen einem evidenzbasierten Fundament und subjektiv politischen Entscheidungen in den Aussagen der Expert\*innen greift auf die vermeintliche Inkompatibilität der Systeme Wissenschaft und Politik zurück. Vielfach wird dabei das Spannungsverhältnis angesprochen, das durch die zwei systemischen Funktions- und Arbeitsweisen zu Stande kommt (u.a. Experte 5: 27, Experte 13: 20), wenngleich dies auch von bereichernder Natur sein kann (Experte 14: 10-11).

*„(...) der politische Prozess benötigt im Idealfall Ja-Nein-Antworten, schwarz-weiß: "Es ist gut, das machen wir weiter. Das ist schlecht, das funktioniert nicht". Aber Wissenschaft (...) ist (...) nicht schwarz und weiß, sondern ist immer mit Unsicherheit verbunden. (...) so ein Limitations-Kapitel wird in der Politik immer nicht so gerne gesehen, weil es die Ergebnisse schwammig macht und sie gerne mit harten Aussagen arbeiten.“*  
(Experte 11: 17)

Es sei die Natur der Wissenschaft, „dass sie lange dauert, dass sie grau ist, dass sie komplex ist“ (E15:13). Politiker aber müssten in einer Vielzahl an Themen sprechfähig sein, schnell Bewertungen abgeben und Entscheidungen treffen. Während Wissenschaft auf einer langfristigeren und übergeordneten Ebene ablaufe, sei Politik kurzfristiger und auf einer operativen Ebene anzusiedeln (E6:5). Das „daily business“ und die Erwartungen des politischen Kosmos passe oft nicht zur wissenschaftlichen Logik und Ethik. Experte 11 (17/19) sprach etwa von „weltfremden Anforderungen“ für Evaluierende:

*„(...) es ist absurd, da kriegen Sie Spiegelstrichlisten mit 40 Evaluationsfragen, die eigentlich gar nicht zu beantworten sind (...). Es ist ungefähr so: „Wir wollen jetzt eine Wirkungsanalyse, eine Effizienzanalyse, eine Prozessevaluation und in 5 Monaten hätten wir gerne einen Bericht“*  
(Experte 11: 17/19)

Eine akademische Elfenbeinturm-Haltung und die politische Gemengelage aus Partikularinteressen ergäben oft kein produktives Konglomerat (u.a. Experte 17: 6, Experte 8: 3). Das gegenseitige Verständnis und eine pragmatische Haltung beiderseits müsse daher ausgebaut bzw. bewahrt werden. Während auf der einen Seite das wissenschaftliche System aufhören solle, nur das zu belohnen, „was methodisch überzeugt, nicht inhaltlich“, müsse das politische System Ressourcen für unabhängige Forschung bereitstellen:

*„Ich sage da immer: „Wenn ich von McKinsey wäre, würdest du mich doch auch nicht fragen, ob ich es umsonst mache?““  
(Experte 17: 10)*

## 4.2 Erfolgsbedingungen für evidenzbasierte Entscheidungen in Politik, Wissenschaft und Gesellschaft

Im Folgenden werden die drei primären Kernbedingungen aufgeführt, die die Befragten in den Interviews als essentiell für die Verankerung von Evaluationen in Politik und Gesellschaft benannt und erörtert haben. Im folgenden Kapitel sollen die in den Hypothesen festgehaltenen spezifischen Bedingungen zur Realisierung des analysierten Potentials aufgeführt werden und praktische Implikationen aus den Interviews für die tatsächliche Steigerung der Evidenzbasierung in politischen Entscheidungsprozessen extrahiert werden. Deutlich wird, dass nicht nur die Beauftragung durch die Politik und die Verfügbarkeit durch die Wissenschaft eine große Rolle spielen, sondern auch die Vermittlung durch mediale Kanäle und damit die Zugänglichkeit für und Verarbeitung durch die Gesellschaft.

### 4.2.1 *Das politische System: Evaluations- und Fehlerkultur*

Die Mehrheit der Expert\*innen unterstreichen die Bedeutung einer institutionellen Evaluationskultur (u.a. Expertin 2: 16, Experte 12: 44). Statt „Naming and Shaming“ ist es die Aufgabe der evaluativen Erfolgskontrollen, eine konstruktive „Kultur des Scheiterns“ zu schaffen, in der Verbesserungsbedarf aufgezeigt wird, wovon die jeweilige Organisation auch langfristig profitiert (Experte 8: 9, Expertin 16: 18, Experte 12: 20/44). Das durch Evaluationen generierte Wissen müsse systematisch genutzt und an interne Prozesse rückgebunden werden (Experte 17: 4, Experte 15: 1, Experte 6: 5). Evaluationen seien ein klares „Kontrollinstrumentarium“ (Experte 15: 9) und es bestehe immer „die Gefahr, dass sich damit zeigt, dass etwas nicht wirkt oder nicht in der beabsichtigten Weise wirkt“ (Experte 8: 9). Doch

besonders den Evaluierten in der Organisation müsse der Mehrwert von Evaluationen besser aufgezeigt werden, um Reaktanzen vorzubeugen sowie Barrieren und Verharrungstendenzen abzubauen (u.a. Expertin 2: 2, Expertin 3: 8/10, Experte 8: 9, Experte 12: 14):

*“An interviewee once said it is like a dentist appointment when we go there.  
(Expertin 3: 8)*

Auch die Gefahr, dass Evaluationen innerhalb einer Organisation als lästige Compliance-Aufgabe wahrgenommen werden, sei noch immer stark verbreitet (u.a. Experte 10: 6, Experte 12: 4, Experte 14: 15, Experte 6: 9). Daher sei es essentiell, dass sich die jeweiligen Institutionen auch selbst Evaluationskompetenzen aneignen sowie in einer unabhängigen, aber nicht losgelösten Einheit Kapazitäten bündeln (Experte 12: 4/18, Experte 13: 12), z.B. in Form von eigenen Evaluierungsreferaten in allen Ministerien (Experte 12: 44). Dadurch könne die Einstellung „Man schreibt eine schöne Evaluierung und schickt es ans BMZ und danach macht man das, was man eh machen wollte“ sukzessive geändert werden (Experte 13: 20).

Ein weiterer elementarer Faktor sei klares Ownership bei den Auftragsverantwortlichen der Durchführungsorganisationen (u.a. Experte 6: 3/13, Experte 12: 16, Experte 11: 7), sodass nicht nur der Wille besteht, Evaluationen zu beauftragen und zu nutzen, sondern auch ausreichend Ressourcen hierfür bereitzustellen (Experte 17: 10). Es benötige insbesondere Einzelpersonen mit hoher Eigenmotivation und Erkenntnisinteresse, die sich aktiv für das Thema einsetzen und noch wenig systematisierte evaluative Einzelprozesse in Standardprozesse überführen (Expertin 2: 2, Experte 6: 3, Expertin 4: 9, Expertin 3: 4, Experte 8: 7, Experte 15: 1). Dass der Worst Case an operativen Herausforderungen „no time, no money, no data“ (Experte 17: 6) nicht eintritt, d.h. nicht ausreichend Zeit (u.a. Experte 6: 5, Experte 12: 20, Experte 15: 11), Daten (u.a. Experte 8: 9, Experte 17: 6) und Geld (u.a. Experte 12: 2, Experte 15: 3, Experte 11: 5) verfügbar sind, müsse letztlich der politische Auftraggebende sicherstellen. Wenngleich auch andere Stakeholder wichtige Voraussetzungen schaffen müssten, obliege der erste Schritt der Politik. Evaluationen könnten nur für politische Entscheidungen relevant sein, wenn Entscheidungsträger\*innen Bereitschaft und Interesse daran zeigen, dass Evidenz ausreichend generiert wird, diese auch aktiv anfordern, die Forschungsergebnisse respektieren, die Handlungsempfehlungen verwerten und in den Dialog einsteigen (Experte 15: 13, Experte 17: 10, Experte 13: 12).

Trotz eines vermeintlichen Ideals von genuinem, aus dem Inneren der jeweiligen Organisation kommenden Erkenntnisinteresse betonte eine Vielzahl an Expert\*innen die wichtige Rolle eines starken Rechtsfertigungs- und Legitimationsdrucks durch Dritte – wie etwa intermediäre Organisationen, Geldgeber oder Parlamente (Experte 11: 13/15, Experte 8: 13, Experte 17: 8, Experte 15: 7, Experte 15: 5). Im Rahmen eines konsultativen und partizipativen Ansatzes sei es essentiell, alle Beteiligte, d.h. Evaluierete, Stakeholder, Geldgeber und lokale Partner, in Feedbackschleifen und Austauschformate einzubinden, um die Chancen für die anschließende Nutzung der Ergebnisse durch alle Involvierte zu erhöhen (Expertin 4: 5/9, Experte 8: 1, Experte 15: 5, Experte 6: 5). Prozesse wie ein genaues Umsetzungsmonitoring oder Tracking-Systeme, in der sich evaluierte Akteure rechtfertigen müssen, welche Handlungsempfehlungen wie umgesetzt wurden, sind bislang Einzelfälle (Expertin 1: 9, Expertin 2: 6, Experte 6: 7). Dabei sei eindeutige Systematisierung der Nutzung sowie eine starke institutionelle Einbindung von Evaluationen elementar (Experte 6: 5, Experte 10: 6, Experte 13: 18, Experte 12: 2), um das Umsetzungsdefizit zu kompensieren (Experte 12: 4/44).

#### *4.2.2 Das wissenschaftliche System: Knowledge Brokering*

Mehrere Expert\*innen betonten die Wichtigkeit einer verständlichen Wissenschaftskommunikation und -vermittlung für die Realisierung des Evaluationspotentials (Experte 17: 10, Experte 14: 13, Experte 6: 5, Experte 12: 38-42). Auch wenn die Dienstleistung mit der Übergabe des Abschlussberichts beendet ist und in den meisten Fällen kein konkreter Follow-up-Prozess vorgesehen ist (Expertin 9: 7-8, Experte 11: 10-11), müsse die Scientific Community nicht nur die Fähigkeit optimieren, politisch relevante Evidenzen zu kreieren, sondern diese auch verkaufen zu können und aktives Knowledge Brokering zu betreiben (Experte 6: 5). Experte und Evaluator 15 (1) bezeichnete sich beispielsweise als „Salesman“, der „versuche (...) Evidenzen nach oben zu verkaufen“. Auch Evaluator 4 (9) betonte, man müsse Berichte regelrecht „vermarkten“, um einen Buy-In zu generieren. Experte 17 (10) führte an, dass das wissenschaftliche System bislang politiknahe Arbeit überhaupt nicht honoriere. Es solle aber nicht darum gehen „Experimente [zu] machen, um der Experimente Willen“, sondern um relevante Fragestellungen zu untersuchen getreu nach dem Motto „form follows function“ (Experte 17: 2):

*„(...) wenn Sie nach Uganda gehen, finden Sie hinter jedem Busch einen Doktoranden vom MIT, der irgendwas randomisiert. Das mag alles wissenschaftlich relevant sein, aber es ist komplett losgelöst von der Praxis (...)“ (Experte 17: 10)*

Das Bindeglied zwischen politischem und wissenschaftlichem System sei die Reduktion der zeitlichen Dauer und der methodischen Komplexität (Experte 6: 5). Evaluationen müssten sich vermehrt den Bedürfnissen der Politik anpassen sowie präzise und übersichtliche Reports mit klaren Handlungsempfehlungen entwickeln, die keinem Selbsterhaltungszweck dienen (Experte 12: 3):

*„Es sieht nach viel Papier aus und man kann ein schönes Foto machen, wenn der Geschäftsführer des (...) Instituts dann irgendeinem Staatssekretär oder der Abteilungsleitung dieses Gutachten in die Hand drückt, aber das hilft ja dann nichts. (...)“*  
(Experte 14: 13)

Nur realitätsnahe und pragmatische Handlungsempfehlungen, die unter den gegebenen politischen Rahmenbedingungen sowie mit begrenztem Budget und Personaleinsatz umsetzbar sind, seien von Entscheidungsträger\*innen auch de facto verwertbar (Experte 15: 5).

*“the recommendations (...) need to be credible, but also implementable and reasonable. There needs to be an anchor in reality for the evaluation. Evaluations that tend to go only at the ideal and not the pragmatic are worthless. They have to be pragmatic. You have to understand that resources are limited.”*  
(Expertin 3: 4)

Experte 8 (13) beschrieb in diesem Zusammenhang in aller Schärfe, dass Wissenschaft, die nicht in der Lage sei, ihre Ergebnisse in die Öffentlichkeit zu transferieren, ihren Zweck verfehle und lediglich dem Selbstzweck diene. Es bestehe eine „Bringschuld der Wissenschaft“, ihre Erkenntnisse so aufzubereiten, dass damit praktische Verbesserungen umgesetzt und öffentliche Gelder sinnvoll eingesetzt werden können. Evaluierungen sollten „als global public good nutzbar“ gemacht, d.h. durch Publikationen und Austausch in die Wissenschaftsgemeinschaft eingespeist werden (Experte 6: 5). Ein effizientes Wissensmanagement und Erfahrungsaustausch sowohl innerhalb der evaluierten Institution als auch zwischen Evaluierten und Evaluierenden und zwischen der evaluierten Institution und der Öffentlichkeit seien der Schlüssel für den langfristigen Einfluss von „lessons learned“ aus Evaluierungsberichten. Dies könne durch Abschlussworkshops mit den Geldgebern und Stakeholdern geschehen, in der nicht nur Ergebnisse, sondern auch Zukunftsszenarien präsentiert werden, oder im Rahmen organisierter Veranstaltungen für die interessierte Öffentlichkeit (Expertin 16: 16).

#### 4.2.3 *Das soziokulturelle System: Bürger\*innennahe Vermittlung*

Obwohl Expert\*innen betonten, dass eine Einbindung der Gesellschaft normativ wünschenswert sei, äußerten sie einstimmig Skepsis und Bedenken, dass die Zivilgesellschaft momentan kaum von Evaluationserzeugnissen erreicht werde (Experte 11: 33/37, Expertin 7: 56/58, Expertin 9: 28, Experte 6: 19), nicht an diesen interessiert sei (Experte 15: 17, Experte 8: 19, Expertin 9: 30, Expertin 16: 22, Experte 10: 24) oder nicht in der Lage sei, diese zu verwerten (Experte 11: 33/39, Experte 8: 9/19, Expertin 7: 54, Expertin 3: 26). Einige Expert\*innen sehen eine gesellschaftliche Nachfrage jedoch bei hochaktuellen, kontroversen oder besonders sensiblen Themen (Experte 12: 35-36, Experte 10: 24, Expertin 2: 18). Mehrere Expert\*innen äußerten, dass Bürger\*innen von der heutigen Komplexität politischer Prozesse überfordert seien und nicht genügend Kapazitäten für mehr Evidenz oder evidenzbasierte Evaluationen hätten (Experte 8: 21): „Ich glaube nicht, dass die Zivilgesellschaft in der Lage ist, die Komplexität und noch mehr Evidenzen zu absorbieren – auch wenn sie wüssten, dass es Evaluationen gibt.“ (Experte 15: 19). Ein quantitativer Anstieg an Evaluationen wurde sowohl von Auftragnehmenden als auch Auftraggebern mit großer Mehrheit als nicht zwangsläufig förderlich betrachtet (u.a. Experte 11: 37, Expertin 16: 8, Experte 5: 39, Experte 15: 17, Expertin 3: 26, Experte 10: 26).

Die Schraube, an der man drehen solle, sei Umsetzungsqualität und -schnelligkeit, nicht am quantitativen Anstieg (Expertin 7: 60/62, Experte 15: 17): „Ich glaube, dass ein „Mehr“ an Evaluation nicht ein Jota ändert. Ich glaube, dass eine andere Kommunikation Evaluation aufbricht.“ (Experte 15: 17) Die derzeitige Nicht-Erwähnung von Evaluationen in Bundestagsdebatten und der konventionellen Presse und die fehlende öffentliche Auseinandersetzung mit evaluativen Ergebnissen, um Politik einzuordnen, wird daher besonders kritisiert (u.a. Expertin 9: 28, Experte 15: 17). Der einzige Übertragungsweg, um Evaluationen einer breiteren Masse zugänglich zu machen, seien mediale Kanäle und eine etablierte Berichterstattung, die über eine Veröffentlichungspflicht hinaus gehe (Experte 11: 37, Expertin 16: 28, Experte 6: 29-31). Um die gesellschaftliche Partizipation und Nachfrage nach politischer Evidenz zu steigern, würden Medien in der Verantwortung stehen, Evidenzen aus Evaluationen transparent und bürger\*innenzugänglich zu präsentieren (Expertin 7: 56, Experte 5: 37, Expertin 3: 22, Expertin 2: 20). Dann bestehe auch das Potential, dass die Zivilgesellschaft mehr Vertrauen in die Integrität öffentlicher Institutionen und deren politische Maßnahmen schöpft und dass Politiker\*innen im Kampf um Wähler\*innenstimmen angehalten

werden, sich verstärkt auf Evidenzen zu berufen (u.a. Experte 5: 37, Experte 4: 21). Bürgerinnen und Bürger müssten vermehrt über die Wirksamkeit politischer Pläne aufgeklärt werden, um eine Kultur der Evidenz auch im gesellschaftlichen System zu verankern:

*„Mündige Bürger haben ein Recht darauf über das, was der Staat an Maßnahmen betreibt, aufgeklärt zu werden. Ich halte es für eine ureigene Aufgabe des Staates, diese Informationen bereit zu stellen. Das ist das idealtypische öffentliche Gut. Der mündige Bürger hat ein Anrecht darauf, Informationen darüber zu erlangen, was um ihn herum passiert.“*  
(Experte 8:19)

Jedoch entscheiden mediale Kanäle aufgrund einer begrenzten Aufnahmekapazität ihrer Rezipient\*innen, was ihnen transportwürdig erscheint (u.a. Experte 12: 38-42, Expertin 9: 28). Und genauso wie das im Falle der Politik gelte, so sei das Phänomen „data picking“ insbesondere auch in der Zivilgesellschaft und den Medien verankert (Expertin 3: 26). Expertin 16 (8) betonte, es reiche nicht, Evaluationsberichte zu publizieren, sondern man bräuchte Diskussionsformate mit der Öffentlichkeit und zugänglichere Begrifflichkeiten und Darstellungen: „Das einfach so zu subventionieren und anzunehmen, dass da ein Mehrwert entsteht, das würde ich als falschen Ansatz sehen, sondern da muss ich den Kontext schaffen, den Diskurs anpassen“. Um „Whitewashing“ (Experte 13: 28, Expertin 16: 20) bzw. „Window Dressing“ (Expertin 2: 10) bei Veröffentlichungen zu vermeiden, müssten zudem insbesondere unabhängig generierte Evaluationen von Medien und Zivilgesellschaft eingefordert werden. Außerdem müsse man zwangsläufig mit einer Verkürzung, Simplifizierung und Komprimierung der Evaluationsergebnisse rechnen, bemängelten Experte 6 (29) und Experte 11 (17):

*„Da habe ich damals geschrieben "unter bestimmten Bedingungen kann das Gesetz eine Wirkung haben" und die Medien machen daraus "Evaluation zeigt, Gesetz wirkt". (...) das ist das, was in der Öffentlichkeit dann ankommt“*  
(Experte 11: 17)

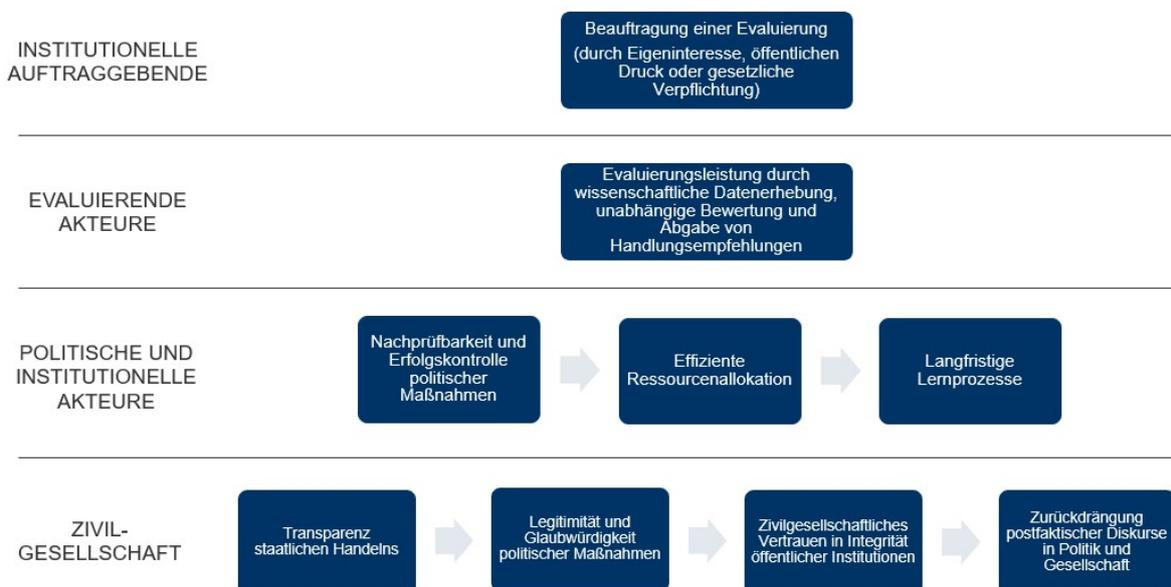
Evaluationen blieben im wissenschaftlichen und politischen System und würden deren Grenzen nicht überschreiten (Experte 11: 37, Expertin 9: 26). Die Verfügbarkeit von Evidenz in der Fachöffentlichkeit und in der allgemeinen Öffentlichkeit triffe immer mehr auseinander (Experte 5: 39). Diese Beschränkung auf fachöffentliche Kreise sah Experte 5 (37/39) jedoch nicht zwangsläufig problematisch an. Auch wenn dies normativ wünschenswert sei, sei eine Auseinandersetzung mit Evaluationen durch Massenmedien oder Tagespolitik nicht zwangsläufig notwendig. Wichtiger sei eine zunehmende Sensibilisierung in politischen Ausschüssen, Gremien oder Beiräten. Auch Experte 13 (28) relativierte, dass es anstelle einer

radikalen Partizipation fachfremder Bürger\*innen, verstärkte Partizipationsmöglichkeiten für zivilgesellschaftliche Akteure wie NGOs oder Gewerkschaften als „Mittlerinstanzen“ aufgrund ihrer „Scharnierfunktion“ weitaus zweckmäßiger seien. Die Relevanz der Partizipation und Nachfrage durch NGOs, Lobbygruppen oder nicht-staatliche und zivilgesellschaftliche Initiativen für die Verwertung und Nutzung von Evaluationen wurde mehrfach betont (Expertin 9: 8, Expertin 7: 20, Expertin 3: 22).

### 4.3 Diskussion und praktische Implikationen der Forschungsergebnisse

Die Erwartungshaltung an Evaluationen aus wissenschaftlicher Sicht sind hoch: Sie sollen als Erfolgskontrolle politischen Handelns zur Steigerung der Objektivität, Transparenz und Rechenschaftspflicht beitragen, das zivile Vertrauen in die Integrität und die Gemeinwohlorientierung politischer Institutionen erhöhen, deren Macht und Ressourcen legitimieren sowie nicht-evidenzbasierte, populistische Politik verdrängen (u.a. Adams 2004, Boaz et al. 2008, Kallen 1982, Lampinen 1992, Maynard 2006, Owens et al. 2004, Rivlin 1971, Weiss 1980; 1999).

Abbildung 2: Theoretisches Idealmodell eines Evaluierungsprozesses



Quelle: Eigene Darstellung

Doch in den Gesprächen mit Expert\*innen und Praktiker\*innen wurde äußerst deutlich, dass dieser theoretischen Idealvorstellung nicht nur ein einzelner Stein im Weg liegt. In der

Auswertung der Interviews wurde ersichtlich, was Evaluationen leisten können und was nicht: Evaluationen bieten eine wichtige Diskurs- und Deliberationsgrundlage für politische Entscheidungen und dienen damit als vorgelagerte Voraussetzung für evidenzbasiertes (und nicht evidenzbestimmtes oder evidenzgeleitetes) politisches Handeln. Doch eine solches Potential der Evidenzbasierung (H1) geht nicht zwangsläufig mit einer Rationalisierung oder gar De-Politisierung einher (H2). Die auf Evidenzen und Evaluationen aufbauenden normativen, wertorientierten und subjektiven Diskurse sind genuin politisch und finden nachgelagert statt. Evaluationen haben somit eine entscheidungsvorbereitenden Rolle inne, indem sie Informationen liefern, die als Grundlage für fundierte Urteile und die tatsächlichen politischen Entscheidungen dienen (vgl. Rich 1979, Caplan 1975, Davies 2012). Mit dem Einsatz von Evaluationen ergibt sich ein verstärkt faktenbezogener, wenngleich nicht faktengesteuerter Entscheidungsfindungsprozess.

Dass Evaluationen als Werkzeug evidenzbasierter Politik auch ein Instrument gegen die postfaktische Vertrauenskrise darstellen (vgl. Stephenson 2019), wurde unter den Expert\*innen gemeinhin abgelehnt und die Kritik und Ernüchterung bezüglich des hochpolitisierten evaluativen Kontextes war groß. Evidenzbasierte Politikgestaltung sei eine wichtige Errungenschaft, doch der Einfluss auf postfaktische bzw. nicht-evidenzbasierte Diskurse und deren vorantreibende Akteure ist nicht absehbar. Bislang besteht das Evidenzbasierungspotential lediglich in Evaluationen offen gegenüberstehenden, fachöffentlichen Kreisen. Wie schon Weiss (1999) und Hertin et al. (2009) betonten, ist es wichtig, pragmatisch zu berücksichtigen, was Evaluator\*innen leisten können und was nicht, sowie welche anderen Faktoren und verschiedenen Interessen das politische Konglomerat aus Fakten, Meinungen und Werten bedingen.

Die vorliegenden Forschungserkenntnisse machen deutlich, dass der technokratische Ansatz, der auf eine strenge Rationalisierung von politischen Entscheidungen durch den Gebrauch von wissenschaftlichem Wissen abzielt, nicht zielführend ist. Die qualitativen Analyseergebnisse untermauern viel mehr die Argumentation von Weiland (2013), demzufolge der Fokus von Instrumenten wie Evaluationen oder Politikfolgenabschätzungen nicht auf einer De-Politisierung liegt, die ohnehin nicht umsetzbar ist, sondern auf Good Governance. Evaluationen und evidenzbasierte Verfahren sind nicht nur ein Tool der Wirtschaftlichkeitsprüfung, sondern vielmehr Mittel der Deliberation, die die Möglichkeit bieten, reflexive und soziale Lernprozesse anzustoßen (vgl. Edelenbos/van Buuren 2005). Statt

einer rigorosen Rationalisierung von Politik können Evaluationen im Rahmen eines diskursiven Modells die Entwicklung hin zu offeneren Governance-Strukturen vorantreiben und dazu beitragen, den politischen Entscheidungsprozess ergebnisoffener, inklusiver und zurechenbarer zu gestalten. Weder eine Loslösung von jeglicher Faktenbasis noch die Abkehr von wertorientierter Meinungspolitik ist demokratietheoretisch wünschenswert; die Lösung liegt in einem Mittelweg zwischen beiden Modellen der Wissensbasierung (vgl. Fischer 1995), d.h. einem rationalistisch-positivistischen und einem sozialkonstruktivistischen bzw. post-positivistischen Ansatz.

Im Rahmen der vorliegenden empirischen Untersuchung wurde auf die Diskrepanz zwischen der stetig steigenden Systematisierung, Professionalisierung und Institutionalisierung von Evaluationen und der mangelnden Forschung auf diesem aufsteigenden Feld reagiert (vgl. Stephenson 2019). In Bezug auf jene Forschungslücke, die u.a. Hertin et al. (2009) sowie Knill und Tosun (2015) konstatierten, wurde mit der Forschungsarbeit ein Beitrag zur aktuellen Forschung über den Effekt von Evaluationsaktivitäten auf die Dynamik politischer Entscheidungen und Prozesse der Politikgestaltung sowie über die genuine Funktions- und Leistungsfähigkeit von Evaluationen geleistet. Es wurde veranschaulicht, wie genau Evaluationsergebnisse von den unterschiedlichen Akteuren aufgegriffen werden, welche Problematiken sich draus ergeben und was wesentliche Erfolgsbedingungen sind, um diesen vorzubeugen und das analysierte Potential von Evaluationen voll auszuschöpfen. Deutlich wurde, dass nicht nur die Beauftragung durch die Politik und die Verfügbarkeit durch die Wissenschaft eine große Rolle spielen, sondern auch die Vermittlung durch mediale Kanäle und damit die Zugänglichkeit für und die Verarbeitung durch die Gesellschaft. Praktische Implikationen waren in diesem Zusammenhang die drei zentralen Voraussetzungen für die erfolgreiche Beauftragung, Durchführung und Nutzung von Evaluierungen: Der Aufbau einer Evaluations- und Fehlerkultur für progressive institutionelle Lernprozesse, Knowledge Brokering für eine effektivere und beidseitig profitable Zusammenarbeit der Systeme Wissenschaft und Politik und eine bürger\*innennahe Vermittlung durch mediale Kanäle und öffentliche Partizipationsformate.

Eine weitere klare praktische Implikation, die sich aus den vorliegenden Ergebnissen ergibt, ist der Aufbau einer bundesweiten Open Data Plattform, die Evaluationsberichte jeglicher Art und Form in einer Online-Datenbank bündelt. Ein solches Tool würde Politik, Wissenschaft und Gesellschaft gleichermaßen ermöglichen, relevante Evaluationsstudien durch Filterkategorien

wie ‚Thema‘, ‚Gebiet‘ oder ‚Akteur‘ zu identifizieren und für politische, akademische oder institutionelle Lernprozesse weiterzuverwenden. Ähnliche Initiativen werden von Experte 6 (5) und 17 (2) anvisiert und existieren auch bereits in verschiedenen Subkontexten. Beispielsweise sind im OECD DAC Evaluation Resource Centre (DEReC) über 3.000 Evaluationsreports aus dem Bereich der Entwicklungszusammenarbeit frei verfügbar (OECD 2020b). Eine öffentliche, transparente und unabhängige Evaluationsplattform durch die nationale Akteure jeglicher Art emanzipiert werden, eigene Evaluationsberichte Anderen zur Verfügung zu stellen sowie Evaluationsberichte Anderer zu nutzen, würde systematische Lernprozesse und Synergien zwischen internen Evaluationen und externen Quellen ermöglichen (Experte 17: 2). Eine solche Austauschbasis könnte in einem großen Maße zu einem evidenzbasierten Kulturwandel und einem beidseitig profitablen Wissensaustausch beitragen. Eine solche auf Open Data, Open Access und Open Government Ansätzen (vgl. Janssen et al. 2012, Longo 2011, Meijer 2012) basierende Initiative würde öffentliche Teilhabe, intersektorale Zusammenarbeit sowie progressive Veränderungen fördern und damit dem Ideal von transparenten, partizipativen und kollaborativen öffentlichen Institutionen näher zu kommen (The White House 2009).

## 5. Fazit: Kämpfe, die Evaluationen (nicht) kämpfen müssen

*“Evaluators will never (...) take the politics out of policymaking”*  
(Weiss 1999: 483)

Inwiefern bedingt der Einsatz systematischer Evaluationen die Dynamik politischer Entscheidungen und öffentlicher Politikgestaltungsprozesse? Letztlich ergab sich auf Basis der Expert\*innenmeinungen ein durchaus ambivalentes Bild. Eine klare Mehrheit der Expert\*innen bejahte das Potential von Evaluationen, politische Entscheidungen evidenzbasierter zu gestalten (*H1*), verneinte jedoch eine damit einhergehende Rationalisierung bzw. De-Politisierung oder gar eine Ablöse meinungsbasierter, dogmatischer oder ideologischer Kämpfe (*H2*). Es wurde deutlich, dass die Befragten überzeugt davon sind, dass der Politik das Politische zu nehmen, kein Kampf ist, der durch Evaluierungen gefochten werden kann und sollte (vgl. Experte 13: 32). Im Zentrum der Forschungsarbeit steht vielmehr die Erkenntnis, dass Evaluationen eine evidenzbasierte Diskurs- und Entscheidungsgrundlage für politisch-normative Entscheidungen darstellen und darauf aufbauend die Diskursqualität in politischen Entscheidungsprozessen erhöhen können (vgl. Frey 2012, Howlett 2009, Pollitt 2006, Schlauffer 2017).

Des Weiteren wurde deutlich gemacht, dass evidenzbasierte Evaluationen eine konstruktive Arbeitsteilung vorantreiben – zwischen einem politischen System, welches Leitlinien, generalistische Ziele und subjektive Werte vorgibt und einem wissenschaftlichen System, welches optimale Policy-Designs entwickelt und deren Wirkungsweise evaluiert. Um sowohl politikrelevante als auch wertneutrale Informationen für die demokratische Entscheidungsfindung bereitzustellen, muss das Verhältnis zwischen Sender\*in und Empfänger\*in dieser Informationen jedoch neu kalibriert werden. Diese verstärkt zu nutzende Komplementarität zwischen Wissenschaft und Politik bringt nicht nur die Notwendigkeit eines sensiblen Umgangs mit verschiedenen Formen der Einflussnahme und des politischen Drucks auf Evaluierende mit sich, sondern eine Offenheit für und Konsolidierung von institutionellen und progressiven Lernprozessen (*Evaluations- und Fehlerkultur*), eine für Entscheidungsträger\*innen geeignete Aufbereitung evaluativer Ergebnisse (*Knowledge Brokering*) sowie das Angebot öffentlich-partizipativer Formate zur Diskussion evaluierten Themen (*Bürger\*innennahe Vermittlung*). Jene Faktoren sind notwendige Voraussetzungen dafür, dass Evaluation einen konstruktiven Beitrag zur demokratischen Entscheidungsfindung leisten kann.

Obwohl wie dargestellt nicht alle Thesen bestätigt werden konnten und eine Vielzahl an Einschränkungen und Bedingungen festgestellt wurden, liefert die vorliegende Forschungsarbeit einen wichtigen Beitrag zur Debatte rund um evidenzbasierte Politik in postfaktischen Zeiten in dreierlei Hinsicht:

*Erstens*, die empirische Potentialanalyse reagiert auf die mangelnde aktuelle Forschung im aufstrebenden Feld der evidenzbasierten Politik und Evaluationsforschung. Durch die aktuelle Aufarbeitung des Themas Evaluation wird ein Beitrag zur Bestandsforschung geleistet, indem bisherige empirische Erkenntnisse überprüft und mit aktuellen, eigenständig erhobenen Daten abgeglichen bzw. neu systematisiert wurden sowie der Effekt von Evaluationsaktivitäten auf die Dynamik politischer Entscheidungsprozesse und die Funktions- und Leistungsfähigkeit von Evaluationen untersucht wurde.

*Zweitens*, die Selektion, Strukturierung und Auswertung von Gemeinsamkeiten der Expert\*innenaussagen ermöglichte es, konkrete (Miss-)Erfolgsfaktoren zu identifizieren und darauf aufbauend praktischen Handlungsbedarf aufzuzeigen. Es wurde praxisnah veranschaulicht, welche Verhaltensweisen und -logiken unterschiedlichste Akteure im Evaluationsprozess übernehmen, welche Problematiken sich daraus ergeben und was wesentliche Erfolgsbedingungen sind, um diesen vorzubeugen und das analysierte Potential von Evaluationen voll auszuschöpfen. Die analysierten praktischen Implikationen sind wichtige Ansatzpunkte für eine progressive Evidenz- und Evaluationskultur.

*Drittens*, die bislang vernachlässigte Schnittstelle zwischen evidenzbasierter Politik, Evaluation und Postfaktizität wurde intensiver untersucht. Obwohl im Rahmen dieser Arbeit kein repräsentatives Bild geliefert werden konnte, wurden erste Implikationen für künftige Forschung geliefert.

Die vorliegende Methodik und Erkenntnisse sind kritisch zu reflektieren. Es gibt eine Reihe methodischer und praktischer Einschränkungen, die zur Kontextualisierung und Einordnung der vorliegenden Arbeit in Betracht gezogen werden müssen:

*Erstens*, die gewählte Methodik ist aufgrund der gewählten Thematik, den komplexen Forschungsfragen und der intendierten Ergebnisoffenheit als Argumentationsstütze für die

aufgestellten Thesen anzusehen und nicht als Kausalbeweis. Trotz der systematischen Selektion und der anvisierten bestmöglichen Diversität an Expert\*innen können aufgrund der geringen Fallanzahl keine kausalen Schlüsse gezogen werden. Repräsentativität, Generalisierbarkeit oder gar eine Kausalanalyse der untersuchten Phänomene wird im Rahmen dieser Arbeit nicht angestrebt. Der Fokus des Forschungsvorhabens beschränkte sich auf eine empirische Potentialanalyse. Hinsichtlich thematischer Kausalanalysen und systematisch-komparativer Vergleiche zwischen Staaten, Akteurstypen oder Politikfeldern verbleibt Bedarf für weitere Forschungsarbeiten.

*Zweitens*, im Rahmen der empirischen Erhebung wurden keine Politiker\*innen als direkte Entscheidungsträger\*innen und keine Bürger\*innen oder Dritte befragt. Es wurde sich auf die direkt betroffene, operative Arbeitsebene beschränkt. Weitere Forschung wird sich diesen fehlenden Perspektiven annehmen müssen. Erkenntnisreich wären unter anderem experimentelle Umfragen, die verdeutlichen, inwiefern nicht direkt beteiligte Bürger\*innen von Evaluationen beeinflusst werden.

*Drittens*, die vorliegenden subjektiven Meinungen der ausgewählten Expert\*innen ergeben sich aus deren Erfahrungen mit der Thematik in der Vergangenheit. Daraus prognostizistische Aussagen über das Potential von Evaluationen in der Zukunft abzuleiten, ist Ermessenssache. Auch gaben die befragten Expert\*innen in den Interviews ihre persönliche Wahrnehmung und subjektive Innensicht preis. Schlussfolgerungen auf externe Personen der Zivilgesellschaft oder andere Involvierte sind daher nur mit Vorsicht zu ziehen.

Dennoch leitet sich auf Basis der neu gewonnen empirischen Erkenntnisse ein klares Plädoyer für eine stärkere Präsenz evaluativer Evidenz für eine ambitionierte Zivilgesellschaft und Fachöffentlichkeit – jenseits der Nutzung durch direkt Involvierte – ab (vgl. Sager 2018). In einer Zeit, in der öffentliche Institutionen auf regionaler, nationaler und intergouvernementaler Ebene angesichts knapper öffentlicher Haushalte, zunehmendem Populismus, globaler Krisen und wachsenden politischen, ökologischen und gesellschaftlichen Anforderungen unter starken Druck geraten sind, ist eine stärkere Konzentration auf Evaluation von großer Bedeutung. Die Systematisierung von Evaluationen könnte der Schlüssel zu einer gesteigerten Rechenschaftspflicht und progressiven institutionellen Lernprozessen sein, die mittel- und langfristig kostengünstigere, wirksamere und nachhaltigere Politik ermöglichen. Evaluationen haben das Potential, die Komplementarität der beiden gesellschaftlichen Teilsysteme

Wissenschaft und Politik auszubauen, das in politischen Entscheidungsprozessen vertretene Maß an Inklusion, Legitimation und Reziprozität zu erhöhen und eine undogmatische, für Selbstkritik offene Gesellschaft zu fördern (vgl. Campbell/Russo 1999, Sager 2018, Schlaifer 2017, Stephenson 2019, White 2019).

Die Forschungsergebnisse zeichnen ein Bild von Evaluation als eine evidenzbasierte Deliberationsvoraussetzung für darauf aufbauende politisch-normative Entscheidungen. Evaluation ist demzufolge keine hinreichende Bedingung für evidenzbasierte Politik, aber ein notwendiges Element einer demokratischen und progressiven Regierungsführung. Eine verstärkte Evidenzbasierung verleiht dem politischen Kosmos mehr Integrität und Gewicht – in postfaktischen Zeiten wichtiger denn je.

## 6. Literaturverzeichnis

- Adams, D. (2004): Usable Knowledge in Public Policy, in: *Australian Journal of Public Administration* 63 (1), 29–42.
- Alexander, J. C. (2001): *Soziale Differenzierung und kultureller Wandel. Essays zur neofunktionalistischen Gesellschaftstheorie*. Frankfurt am Main/New York: Campus.
- Alternative für Deutschland (AfD) (2020): AfD im Bundestag fordert Offenlegung der Gesamtkosten der ‚Willkommenskultur‘. 15.01.2020 [online aufgerufen am 08.07.2020]: <https://afdkompakt.de/2020/01/15/afd-im-bundestag-fordert-offenlegung-der-gesamtkosten-der-willkommenskultur/>
- Amara, N.; Ouimet, M.; Landry, R. (2004): New Evidence on Instrumental, Conceptual, and Symbolic Utilization of University Research in Government Agencies, in: *Science Communication* 26 (1), 75–106.
- Anderson, D.; Biddle, B. J. (1991): *Knowledge for policy: Improving education through research*. Philadelphia: Falmer Press.
- Archer, M. S. (1995): *Realist social theory: the morphogenetic approach*. Cambridge: Cambridge University Press.
- Archer, M. S.; Bhaskar, R.; Collier, A. (1998): *Critical realism: essential readings*. London: Routledge.
- Arendt, H. (1987): *Wahrheit und Lüge in der Politik. Zwei Essays*. München.
- Baklien, B. (1983): The use of social science in a Norwegian Ministry: As a tool of policy or mode of thinking?, in: *Acta Sociologica* 26 (1), 33-47.
- Beck, U. (1997): *Was ist Globalisierung? Irrtümer des Globalismus – Antworten auf Globalisierung*. Frankfurt am Main: Suhrkamp.
- Behn, R. (1998): The new public management paradigm and the search for democratic accountability, in: *International Public Management Journal* 1 (2), 131–164.
- Berger, P. L.; Luckmann, T. (1966): *The Social Construction of Reality. A Treatise in the Sociology of Knowledge*. New York: Doubleday.

- Boaz, A.; Grayson, L.; Levitt, R.; Solesbury, W. (2008): Does evidence-based policy work? Learning from the UK experience, in: *Evidence & Policy: A Journal of Research, Debate and Practice* 4 (2), 233–253.
- Bogumil, J.; Jann, W. (2009): *Verwaltung und Verwaltungswissenschaft in Deutschland. Einführung in die Verwaltungswissenschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bovens, M.; Hart, P.; Kuipers, S. (2008): The politics of policy evaluation, in: Moran, M.; Rein, M.; Goodin, R.E. (Hrsg.): *The Oxford handbook of public policy*, Oxford: Oxford University Press, 319-335.
- Bronk, R. (1999): *Progress and the invisible hand. The philosophy and economics of human advance*. London: Warner.
- Buch, C. M.; Patzwaldt, K.; Riphahn, R. T.; Vogel, E. (2019): Verstehen — Entwickeln — Testen — Verbessern. Rahmenbedingungen für evidenzbasierte Politik, in: *Wirtschaftsdienst* 99 (2), 106-112.
- Buchanan, J.; Tullock, G. (1962): *The Calculus of Consent. Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press.
- Bulmer, M. (1986): *Social science and social policy*. London: Alleen & Unwin.
- Bulmer, M. (1987): *Social science research and government. Comparative essays on Britain and the United States*. Cambridge: Cambridge University Press.
- Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) (2008): *Wirkungsevaluierungen. Zum Stand der internationalen Diskussion und dessen Relevanz für die Evaluierung der deutschen Entwicklungszusammenarbeit*. BMZ Evaluation Division. Evaluation Working Paper [online aufgerufen am 02.05.2020]: [https://www.bmz.de/de/zentrales\\_downloadarchiv/erfolg\\_und\\_kontrolle/Evaluation\\_Working\\_Papers/BMZ\\_WP\\_Wirkungsevaluierung\\_2009.pdf](https://www.bmz.de/de/zentrales_downloadarchiv/erfolg_und_kontrolle/Evaluation_Working_Papers/BMZ_WP_Wirkungsevaluierung_2009.pdf)
- Bundesrechnungshof (1989): *Erfolgskontrolle finanzwirksamer Maßnahmen in der öffentlichen Verwaltung*. Stuttgart/Berlin/Köln: Kohlhammer.
- Bundesrechnungshof (2013): *Anforderungen an Wirtschaftlichkeitsuntersuchungen finanzwirksamer Maßnahmen nach § 7 Bundeshaushaltsordnung*. Stuttgart: Kohlhammer.

- Bundesregierung (2020): *Gemeinsame Geschäftsordnung der Bundesministerien (GGO)*. 30. Juli 2020 [online aufgerufen am 10.08.2020]: [https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/ministerium/ggo.pdf?\\_\\_blob=publicationFile&v=2](https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/ministerium/ggo.pdf?__blob=publicationFile&v=2)
- Bundestag (2014): *Unterrichtung durch die Bundesregierung. Bericht der Bundesregierung 2013 nach § 7 des Gesetzes zur Einsetzung eines Nationalen Normenkontrollrates. Bessere Rechtsetzung 2013: Erfolge dauerhaft sichern – zusätzlichen Aufwand vermeiden*. Drucksache 18/866. 18. Wahlperiode. 19.03.2014 [online aufgerufen am 21.06.2020]: <http://dipbt.bundestag.de/dip21/btd/18/008/1800866.pdf>
- Bundestag (2018): AfD fordert Bilanz der Afghanistan-Politik. Antrag Auswärtiges. 710/2018. 27.09.2018 [online aufgerufen am 08.07.2020]: <https://www.bundestag.de/presse/hib/570610-570610>
- Bundestag (2020): AfD fordert Evaluierung des NetzDG. Antrag Recht und Verbraucherschutz. 289/2020. 12.03.2020 [online aufgerufen am 028.07.2020]: <https://www.bundestag.de/presse/hib/687202-687202>
- Bundeszentrale für politische Bildung (2017): *"Wahres Wissen" und demokratisch verfasste Gesellschaft*. 24.03.2017 [online aufgerufen am 17.06.2020]: <https://www.bpb.de/apuz/245215/wahres-wissen-und-demokratisch-verfasste-gesellschaft>
- Cabinet Office (1999): *Professional policy making for the twenty first century*. Strategic Policy Team. London: Cabinet Office.
- Campbell, D. T. (1969): Reforms as experiments, in: *American Psychologist* 24 (4), 409–429.
- Campbell, D. T.; Russo, M. J. (1999): *Social experimentation*. Thousand Oaks: Sage Publications.
- Campbell, H. (2002): ‘Evidence-based policy’. The continuing search for effective policy processes, in: *Planning Theory and Practice* 3 (1), 89-90.
- Caplan, N.; Morrison, A.; Stambaugh, R. J. (1975): *The use of social science knowledge in policy decisions at the national level. A report to respondents*. Ann Arbor: Institute for Social Research University of Michigan.

- Clarence, E. (2002): Technocracy reinvented: The new evidence-based policy movement, in: *Public Policy and Administration* 17 (3), 1–11.
- Coryn, C. L. S.; Wilson, L. N.; Westine, C. D.; Hobson, K. A.; Ozeki, S.; Fiekowsky, E. L.; Greenman, G. D.; Schroter, D. C. (2017): A Decade of Research on Evaluation, in: *American Journal of Evaluation* 38 (3), 329–347.
- Cronbach, L. (1980): *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Cronbach, L. (1982): *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Crotty, M. (1998): *The foundations of social research. Meaning and perspective in the research process*. Thousand Oaks: Sage Publications.
- D'Ancona, M. (2017): *Post-truth. The new war on truth and how to fight back*. London: Ebury Press.
- Davies, H. T. O.; Nutley, S. M.; Smith, P. C. (1999): Editorial. What Works? The Role of Evidence in Public Sector Policy and Practice, in: *Public Money and Management* 19 (1), 3–5.
- Davies, P. (2004): *Is evidence-based government possible?* Jerry Lee Lecture. 4th Annual Campbell Collaboration Colloquium. Washington DC.
- Davies, P. (2012): The State of Evidence-Based Policy Evaluation and its Role in Policy Formation, in: *National Institute Economic Review*. 219 (1), R41-R52.
- DeLeon, P. (1992): The Democratization of the Policy Sciences, in: *Public Administration Review* 52 (2), 125-129.
- Deutsche Gesellschaft für Evaluation (2014): *Satzung vom 27.09.2006* [online aufgerufen am 04.08.2020]: <https://www.degeval.org/ueber-uns/satzung/>
- Deutsche Gesellschaft für Evaluation (2016): *Standards für Evaluation* [online aufgerufen am 04.08.2020]: [https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards\\_fuer\\_Evaluation.pdf](https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf)
- Deutschlandfunk (2016): *AfD-Wahlkampf in Berlin. Gefühlte Realität* [online aufgerufen am 08.08.2020]: [https://www.deutschlandfunk.de/afd-wahlkampf-in-berlin-gefuehlte-realitaet.1773.de.html?dram:article\\_id=365806](https://www.deutschlandfunk.de/afd-wahlkampf-in-berlin-gefuehlte-realitaet.1773.de.html?dram:article_id=365806)

- Dobes, L. (2017): *The Post-Truth Era in Government Evaluation of Major Projects and Policies* [online aufgerufen am 13.07.2020]: [https://crawford.anu.edu.au/sites/default/files/publication/crawford01\\_cap\\_anu\\_edu\\_a\\_u/2017-02/cswp1704.pdf](https://crawford.anu.edu.au/sites/default/files/publication/crawford01_cap_anu_edu_a_u/2017-02/cswp1704.pdf)
- Dobes, L. (2018): Making Facts and Telling Furphies with Economic Modelling - Cost–Benefit Analysis in the Post-Truth Era, in: Fabian, M.; Breunig, R. (Hrsg.): *Hybrid Public Policy Innovations - Contemporary Policy Beyond Ideology*, Routledge, 83–96.
- Downs, A. (1957): *An Economic Theory of Democracy*. New York: Harper & Brothers.
- Dryzek, J. S. (1990): *Discursive democracy. Politics, policy, and political science*. Cambridge: Cambridge University Press.
- Dunleavy, P. (2014): *Democracy, Bureaucracy and Public Choice. Economic Approaches in Political Science*. Hoboken: Taylor and Francis.
- Dunn, W. N. (1980): The two-communities metaphor and models of knowledge use: an exploratory case survey, in: *Knowledge* 1 (4), 515-536.
- Eckhard, S.; Jankauskas, V. (2018): The politics of evaluation in international organizations. A comparative study of stakeholder influence potential, in: *Evaluation* 25 (1), 62–79.
- Edelenbos, J.; van Buuren, A. (2005): The learning evaluation. A theoretical and empirical exploration, in: *Evaluation review* 29 (6), 591–612.
- Europäische Kommission (2017): *Better Regulation Guidelines. Evaluation and Fitness Checks* [online aufgerufen am 28.05.2020]: <https://ec.europa.eu/info/sites/info/files/better-regulation-guidelines-evaluation-fitness-checks.pdf>
- Europäische Kommission (2018): *Flash Eurobarometer 464. Report. Fake news and disinformation* [online aufgerufen am 20.07.2020]: <https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/survey/getsurveydetail/instruments/flash/surveyky/2183>
- Feldman, M. S. (1989): *Order without design. Information production and policy making*. Stanford: Stanford University Press.
- Fischer, F. (1995): *Evaluating Public Policy*. Chicago.

- Fischer, F. (2009): *Democracy and expertise. Reorienting policy inquiry*. Oxford: Oxford University Press.
- Fischer, F.; Forester, J. (1993): *The Argumentative turn in policy analysis and planning*. North Carolina: Duke University Press.
- Frey, K. (2012): *Evidenzbasierte Politikformulierung in der Schweiz. Gesetzesrevisionen im Vergleich*. Baden-Baden: Nomos.
- Furubo, J. E.; Leeuw, F. L.; Rist, R. C.; Sonnichsen, R. C. (1994): Learning from Evaluations: The Swedish Experience, in: Leeuw, F. L. (Hrsg): *Can governments learn? Comparative Perspectives on Evaluation and Organizational Learning*, New York: Routledge, 45-65.
- Furubo, J.-E.; Rist, R. C.; Sandahl, R. (2002): *International atlas of evaluation*. New Brunswick: Transaction Publications. Comparative policy analysis series.
- Gesellschaft für deutsche Sprache (2016): *GfdS wählt »postfaktisch« zum Wort des Jahres 2016*. 9.12.2016 [online aufgerufen am 17.06.2020]: <https://gfds.de/wort-des-jahres-2016/#>
- Gesellschaft für Internationale Zusammenarbeit (2018): *Wissen was wirkt – Das Evaluierungssystem der GIZ. Unternehmensstrategische Evaluierung* [online aufgerufen am 09.08.2020]: [https://www.giz.de/de/downloads/03\\_GIZ\\_EVAL\\_USE.pdf](https://www.giz.de/de/downloads/03_GIZ_EVAL_USE.pdf)
- Glaser, B. G.; Strauss, A. L. (2010): *Grounded theory. Strategien qualitativer Forschung*. Bern: Verlag Hans Huber.
- Gläser, J.; Laudel, G. (2010): *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Guba, E. G.; Lincoln, Y. S. (1989): *Fourth generation evaluation*. Newbury Park: Sage Publications.
- Gusy, C. (2015): *Von der Evaluation zur Evaluationsforschung. Das Beispiel der Sicherheitsgesetze*, in: Gusy, C. (Hrsg.): *Evaluation von Sicherheitsgesetzen*. Wiesbaden: Springer, 225-242.

- Habermas, J. (1968): *Technik und Wissenschaft als "Ideologie"*. Frankfurt am Main: Suhrkamp Verlag.
- Habermas, J. (1973): Wahrheitstheorien, in: Fahrenbach, H. (Hrsg.): *Wirklichkeit und Reflexion*. Pfullingen: Neske.
- Habermas, J. (1981): *Theorie kommunikativen Handelns*. Band 1: Handlungsrationalität und gesellschaftliche Rationalisierung. Band 2: Zur Kritik der funktionalistischen Vernunft. Frankfurt am Main: Suhrkamp.
- Habermas, J. (1992): *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt am Main: Suhrkamp.
- Harsin, J. (2015): Regimes of Posttruth, Postpolitics, and Attention, in: *Economies, Communication, Culture and Critique* 8 (2), 327–333.
- Head, B. W. (2010): Reconsidering evidence-based policy. Key issues and challenges, in: *Policy and Society* 29 (2), 77–94.
- Heise (2020): *DSGVO-Evaluation: Die Panik hat sich gelegt, die Welt dreht sich noch*. 24.06.2020 [online aufgerufen am 25.06.2020]: <https://www.heise.de/news/DSGVO-Evaluation-Die-Panik-hat-sich-gelegt-die-Welt-dreht-sich-noch-4794446.html>
- Hendricks, V. F.; Vestergaard, M. (2017): *Verlorene Wirklichkeit? An der Schwelle zur postfaktischen Demokratie*. 28.03.2017 [online aufgerufen am 01.08.2020]: <https://www.bpb.de/apuz/245212/verlorene-wirklichkeit-an-der-schwelle-zur-postfaktischen-demokratie?p=all>
- Hertin, J.; Turnpenny, J.; Jordan, A.; Nilsson, M.; Russel, D.; Nykvist, B. (2009): Rationalising the Policy Mess? Ex Ante Policy Assessment and the Utilisation of Knowledge in the Policy Process, in: *Environmental Planning A* 41 (5), 1185–1200.
- Howlett, M. (2009): Governance modes, policy regimes and operational plans. A multi-level nested model of policy instrument choice and policy design, in: *Policy Science* 42 (1), 73–89.
- Jacob, S.; Speer, S.; Furubo, J.-E. (2015): The institutionalization of evaluation matters. Updating the International Atlas of Evaluation 10 years later, in: *Evaluation* 21 (1), 6–31.

- Janett, D. (2003): Erfolgskontrolle staatlichen Handelns durch das Parlament, in: Schweizer, R. J.; Jeanrenaud, C.; Kux, S. (Hrsg.): *Verwaltung im 21. Jahrhundert*, Academic Press Fribourg, 97-109.
- Janssen; M.;Charalabidis, Y.; Zuiderwijk, A. (2012): Benefits, Adoption Barriers and Myths of Open Data and Open Government, in: *Information Systems Management (ISM)* 29 (4), 258-268.
- Jun, U.; Grabow, K. (2008): *Mehr Expertise in der deutschen Politik? Zur Übertragbarkeit des "Evidence-based policy approach"*. Zukunft Regieren. Beiträge für eine gestaltungsfähige Politik 1/2008. Bertelsmann Stiftung.
- Kallen, D. B. P. (1982): *Social science research and public policy-making. A reappraisal*. International SVO Workshop on Educational Research and Public Policy-Making, Council of Europe. NFER-Nelson.
- Kant, I. (1784): Beantwortung der Frage: Was ist Aufklärung?, in: *Berlinische Monatsschrift* (12), 481–494.
- Keyes, R. (2004): *The Post-Truth Era. Dishonesty and Deception in Contemporary Life*. New York: St. Martin's Press.
- Kirk, R. (1999): *Relativism and reality: a contemporary introduction*. London: Routledge.
- Knill, C.; Tosun, J. (2015): *Einführung in die Policy-Analyse*. Opladen, Stuttgart: Budrich.
- Kogan, M. (1999): The impact of research on policy, in: Coffield, F. (Hrsg.): *Speaking truth to power: Research and policy on lifelong learning*, Policy Press, 11-18.
- Konzendorf, G. (2009): Institutionelle Einbettung der Evaluationsfunktion in Politik und Verwaltung in Deutschland, in: Widmer, T.; Beywl, W.; Fabian, C. (Hrsg.): *Evaluation. Ein systematisches Handbuch*. Wiesbaden: Springer, 27-39.
- Kreditanstalt für Wiederaufbau (KfW) (2004): *Entwicklung finanzieren. Wirkungen messen*. Achter Evaluierungsbericht über die Projekte und Programme in Entwicklungsländern [online aufgerufen am 22.07.2020]: [https://www.kfw-entwicklungsbank.de/Download-Center/Dokumente-Evaluierung/8\\_Deutsch.pdf](https://www.kfw-entwicklungsbank.de/Download-Center/Dokumente-Evaluierung/8_Deutsch.pdf)
- Kreditanstalt für Wiederaufbau (KfW) (2011): *Wirkungen messen, Ergebnisse bewerten, für die Zukunft lernen*. 11. Bericht über die Evaluierung der Projekte und Programme in

- Entwicklungsländern 2009-2010 [online aufgerufen am 22.07.2020]: [https://www.kfw-entwicklungsbank.de/Download-Center/Dokumente-Evaluierung/11\\_Deutsch.pdf](https://www.kfw-entwicklungsbank.de/Download-Center/Dokumente-Evaluierung/11_Deutsch.pdf)
- Kuckartz, U. (2016): *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim Basel: Beltz Verlag.
- Kuhlmann, S.; Bogumil, J.; Wollmann, H. (2004): *Leistungsmessung und -vergleich in Politik und Verwaltung. Konzepte und Praxis*. Wiesbaden: VS Verlag für Sozialwissenschaften (Stadtforschung aktuell, 96).
- Lampinen, O. (1992): *The utilization of social science research in public policy*. Helsinki: VAPK Publishing.
- Landry, R.; Lamari, M.; Amara, N. (2003): The Extent and Determinants of the Utilization of University Research in Government Agencies, in: *Public Administration Review* 63 (2), 192–205.
- Lane, R. E. (1966): The Decline of Politics and Ideology in a Knowledgeable Society, in: *American Sociological Review* 31 (5), 649–662.
- Lavis, J.; Davies, H.; Oxman, A.; Denis, J.-L.; Golden-Biddle, K.; Ferlie, E. (2005): Towards systematic reviews that inform health care management and policy-making, in: *Journal of Health Services Research & Policy* 10 (1), 35–48.
- Leeuw, F. L. (2006): Evaluation in Europe, in: Stockmann, R. (Hrsg.): *Evaluationsforschung: Grundlagen und ausgewählte Forschungsfelder*, Münster: Waxmann, 57-76.
- Leeuw, F. L.; Furubo, J.-E. (2008): Evaluation Systems: What Are They and Why Study Them?, in: *Evaluation* 14 (2), 157–169.
- Lerner, D.; Lasswell, H. D. (1951): *The policy sciences*. Stanford: Stanford University Press.
- Lewandowsky, S.; Ecker, U. K.H.; Cook, J. (2017): Beyond Misinformation. Understanding and Coping with the “Post-Truth” Era, in: *Journal of Applied Research in Memory and Cognition* 6 (4), 353–369.
- Lindblom, C. E. (1990): *Inquiry and change. The troubled attempt to understand and shape society*. New Haven: Yale University Press.
- Longo, J. (2011): #Opendata: Digital-Era Governance Thoroughbred or New Public Management Trojan Horse?, in: *Public Policy & Governance Review* 2 (2), 38-52.

- Luhmann, N. (1983): *Legitimation durch Verfahren*. Frankfurt am Main: Suhrkamp.
- Luhmann, N. (1988): *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt am Main: Suhrkamp.
- Luhmann, N. (1997): *Die Gesellschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.
- Majone, G. (1989): *Evidence, argument, and persuasion in the policy process*. New Haven: Yale University Press.
- Marschall, S. (2017): *Lügen und Politik im "postfaktischen Zeitalter"*. 28.3.2017 [online aufgerufen am 01.08.2020]: <https://www.bpb.de/apuz/245217/luegen-und-politik-im-postfaktischen-zeitalter?p=all>
- Marston, G.; Watts, R. (2003): Tampering with the Evidence: A Critical Appraisal of Evidence-Based Policy-Making, in: *Australian Review of Public Affairs* 3(3), 143-163.
- Maynard, R. (2006): Presidential Address: Evidence-Based Decision Making: What Will It Take for the Decision Makers to Care?, in: *Journal of Policy Analysis and Management* 25(2), 249-265.
- Mayring, P. (2020): Qualitative Inhaltsanalyse, in: Mey, G.; Mruck, K. (Hrsg.): *Handbuch qualitative Forschung in der Psychologie*, Wiesbaden: Springer, 495-511.
- McConnell, A. (2010a): Policy Success, Policy Failure and Grey Areas In-Between, in: *Journal of Public Policy* 30 (3), 345–362.
- McConnell, A. (2010b): *Understanding policy success. Rethinking public policy*. New York: Palgrave Macmillan.
- McIntyre, L. (2018): *Post-Truth*. Cambridge: MIT Press.
- Meijer, A. J.; Curtin, D.; Hillebrandt, M. (2012): Open government: connecting vision and voice, in: *International Review of Administrative Sciences* 78 (1), 10–29.
- Meyer, W. (2015): Reflexive Modernisierung – oder doch: Evaluation als Ausrede?, in: Hennefeld, V.; Meyer, W.; Silvestrini, S. (Hrsg.): *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft*. Münster: Waxmann, 53-72.
- Mold, J. W.; Peterson, K. A. (2005): Primary care practice-based research networks. Working at the interface between research and quality improvement, in: *Annals of family medicine* 3 (1), 12-20.

- Morris, M.; Clark, B. (2013): You Want Me to Do What? Evaluators and the Pressure to Misrepresent Findings, in: *American Journal of Evaluation* 34 (1), 57–70.
- Müller, P. (2011): Performance Audit und Evaluation. Gemeinsamkeiten und Unterschiede zwischen den Prüfansätzen des Bundesrechnungshofes zur Aufgabenwahrnehmung der Bundesagentur für Arbeit und arbeitsmarktpolitischen Evaluationen, in: *Zeitschrift für Evaluation* 10 (1), 17-38.
- Müller, W. (2015): Von der angewandten Sozialforschung zur Sozialforschung der Verwendung, in: Hennefeld, V.; Meyer, W.; Silvestrini, S. (Hrsg.): *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft*. Münster: Waxmann, 15-40.
- National Audit Office (2007): *Evaluation of Regulatory Impact Assessments 2006-07*, London: National Audit Office.
- National Audit Office (2009): *Delivering High Quality Impact Assessments*, London: National Audit Office.
- National Audit Office (2010): *Assessing the Impact of Proposed New Policies*. Report by the Comptroller and Auditor General. HC 185 Session 2010–2011. London: National Audit Office.
- Nelson, C. E.; Roberts, J.; Maederer, C. M.; Wertheimer, B.; Johnson, B. (1987): The utilization of social science information by policymakers, in: *American Behavioral Scientist* 30(6), 569-577.
- Netzpolitik (2018): *Laut Angela Merkel wird das NetzDG gerade evaluiert. Aber das Justizministerium weiß von nichts*. 14.11.2018 [online aufgerufen am 01.09.2020]: <https://netzpolitik.org/2018/laut-angela-merkel-wird-das-netzdg-gerade-evaluiert-aber-das-justizministerium-weiss-von-nichts/>
- Normenkontrollrat (2013). *Gutachten zur Durchführung von Ex-post-Evaluierungen. Gute Praktiken und Erfahrungen in anderen Staaten* [online aufgerufen am 25.05.2020]: <https://www.normenkontrollrat.bund.de/resource/blob/267760/444152/a50b2b0987ab4865b514116498d73ba2/2014-02-11-evaluierungsstudie-data.pdf>
- Normenkontrollrat (2015): *Chancen für Kostenbegrenzung verbessert. Digitale Chancen tatsächlich nutzen!* [online aufgerufen am 25.05.2020]:

<https://www.normenkontrollrat.bund.de/resource/blob/267760/444056/e5c6b0557c4f393bad85bd888fe6a3ac/2015-10-012-nkr-jahresbericht-2015-data.pdf?download=1>

Normenkontrollrat (2020): *Gesamtkonzept Bürokratieabbau und bessere Rechtsetzung* [online aufgerufen am 04.06.2020]: <https://www.normenkontrollrat.bund.de/nkr-de/stellungnahmen/gesamtkonzept-buerokratieabbau-und-bessere-rechtsetzung-444134#tar-6>

Nutley, S. M.; Davies, H. T. O.; Walter, I. (2007): *Using evidence. How research can inform public services*. Bristol: Policy Press.

OECD (1980): *The utilisation of the social sciences in policy-making in the United States: case studies*. Washington D.C.: OECD Publications.

OECD (1997): *Regulatory Impact Analysis: Best Practices in OECD Countries*. Washington D.C.: OECD Publications.

OECD (2004): *Regulatory Impact Analysis (RIA) Inventory*. Paris: OECD Publications.

OECD (2005): *Paris Declaration on Aid Effectiveness* [online aufgerufen am 23.07.2020]: [https://read.oecd-ilibrary.org/development/paris-declaration-on-aid-effectiveness\\_9789264098084-en#page1](https://read.oecd-ilibrary.org/development/paris-declaration-on-aid-effectiveness_9789264098084-en#page1)

OECD (2008): *Accra Agenda for Action* [online aufgerufen am 23.07.2020] [https://read.oecd-ilibrary.org/development/accra-agenda-for-action\\_9789264098107-en#page1](https://read.oecd-ilibrary.org/development/accra-agenda-for-action_9789264098107-en#page1)

OECD (2016): *Evaluation Systems in Development Cooperation*. Paris: OECD Publications.

OECD (2020a): *Better Criteria for Better Evaluation. Revised and Updated Evaluation Criteria* [online aufgerufen am 12.09.2020]: <https://www.oecd.org/dac/evaluation/evaluation-criteria-flyer-2020.pdf>

OECD (2020b): *OECD DAC Evaluation Resource Centre* [online aufgerufen am 12.09.2020]: <http://www.oecd.org/derec/?hf=5&b=0&s=score>

Olejniczak K.; Raimondo E.; Kupiec, T. (2016): Evaluation units as knowledge brokers: Testing and calibrating an innovative framework, in: *Evaluation* 22(2), 168-189.

Orr, S. K. (2010): Exploring Stakeholder Values and Interests in Evaluation, in: *American Journal of Evaluation* 31 (4), 557–569.

- Ouimet, M.; Landry, R.; Ziam, S.; Bédard, P.-O. (2009): The absorption of research knowledge by public civil servants, in: *Evidence & Policy: A Journal of Research, Debate and Practice* 5 (4), 331–350.
- Owens, S.; Rayner, T.; Bina, O (2004): New agendas for appraisal: reflections on theory, practice, and research, in: *Environment and Planning A* 36 (1), 1943-1959.
- Oxford Languages (2016): *Word of the Year 2016* [online aufgerufen am 06.07.2020]: <https://languages.oup.com/word-of-the-year/2016/>
- Patton, M. Q. (1997): *Utilization-focused Evaluation: The New Century Text*. Thousand Oaks: Sage Publications.
- Patton, M. Q. (2002): A vision of evaluation that strengthens democracy, in: *Evaluation* 8 (1), 125-139.
- Pawson, R.; Tilley, N. (1997): *Realistic evaluation*. London: Sage Publications.
- Picciotto, R. (2019): Is evaluation obsolete in a post-truth world?, in: *Evaluation and program planning* 73, 88–96.
- Polisphere (2020): *Das Heinsberg-Protokoll und die Bedeutung von wissenschaftlicher Politikberatung* [online aufgerufen am 02.09.2020]: <https://www.polisphere.eu/de/politnews-das-heinsberg-protokoll-und-die-bedeutung-von-wissenschaftlicher-politikberatung/?fbclid=IwAR1WMfbiI9pRC2mKIOrNtIwWc4VxeAOckpEHVTUgjljR3dWabnrVogAz-Y#1550705845456-6c03b26d-a56e7ba6-8514788e-21751595-ca762121-7293>
- Pollack, D. (2016): Modernisierungstheorie – Revised: Entwurf einer Theorie moderner Gesellschaften, in: *Zeitschrift für Soziologie* 45 (4), 219-240.
- Pollitt, C. (1998): Evaluation in Europe: Boom or Bubble?, in: *Evaluation* 4 (2), 214–224.
- Pollitt, C. (2006): Performance Information for Democracy: The Missing Link?, in: *Evaluation* 12 (1), 38–55.
- Radaelli, C. M. (1995): The role of knowledge in the policy process, in: *Journal of European Public Policy* 2 (2), 159–183.

- Radaelli, C. M. (2005): Diffusion without convergence. How political context shapes the adoption of regulatory impact assessment, in: *Journal of European Public Policy* 12 (5), 924–943.
- Rich, R. F. (1979): *Translating evaluation into policy*. Beverly Hills: Sage Publications
- Rich, R.F.; Oh, C. H. (2000): Rationality and Use of Information in Policy Decisions: A Search for Alternatives, in: *Science Communication* 22 (2), 173–211.
- Rivlin, A. M. (1971): *Systematic Thinking for Social Action*. Washington: Brookings Institution Press.
- Romsdahl, R. J. (2005): When do environmental decision makers use social science, in: Brewer, G. D.; Stern, P. (Hrsg.): *Decision Making for the Environment: Social and Behavioral Science Research Priorities*. Washington D.C.: The National Academies Press.
- Rossi, P. H.; Lipsey, M. W.; Freeman, H.E. (2004): *Evaluation: A Systematic Approach*. Thousand Oaks: Sage Publications.
- Roth, R. (2004): „Reden Sie mit dem Pferd“. Bedingungen, Möglichkeiten und Grenzen Demokratischer Evaluationskultur in Deutschland, in: *Soziale Arbeit und Sozialpolitik. Sozial Extra* 28 (6), 6-9.
- Sabatier, P. A.; Jenkins-Smith, H. C. (1993): *Policy change and learning. An advocacy coalition approach*. Boulder: Westview Press.
- Sager, F. (2018): Policy evaluation and democracy: Do they fit?, in: *Evaluation and program planning* 69, 125–129.
- Sanderson, I. (2002): Evaluation, Policy Learning and Evidence-Based Policy Making, in: *Public Administration* 80 (1), 1–22.
- Sarasin, P. (2017): Fakten und Wissen in der Postmoderne. 28.3.2017 [online aufgerufen am 01.08.2020]:  
<https://www.bpb.de/politik/extremismus/rechtspopulismus/245449/fakten-und-wissen-in-der-postmoderne>
- Saretzki, T. (2012): The "argumentative turn" revisited: Demokratisierung von Policy-Analysen in partizipativen Projekten und diskursiven Designs?, in: Egner, B.; Haus, M.; Terizakis, G. (Hrsg.): *Regieren*, Wiesbaden: VS Verlag für Sozialwissenschaften, 57-74.

- Schimank, U. (1996): *Theorien gesellschaftlicher Differenzierung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schimank, U. (2001): Teilsysteminterdependenzen und Inklusionsverhältnisse. Ein differenzierungstheoretisches Forschungsprogramm zur System- und Sozialintegration moderner Gesellschaft, in: Barlösius, E.; Müller, H.-P.; Sigmund, S. (Hrsg.): *Gesellschaftsbilder im Umbruch. Soziologische Perspektiven in Deutschland*, Wiesbaden: VS Verlag für Sozialwissenschaften, 109-130.
- Schlauffer C. (2017). The contribution of evaluations to the discourse quality of newspaper content, *Evaluation and Program Planning* 69, 157-165.
- Schöffski, O. (2008): *Gesundheitsökonomische Evaluationen*. Berlin: Springer.
- Schwab, O. (2009): Evaluierung von Raumordnungspolitik in Deutschland, in: Widmer, T; Beywl, W.; Fabian, C. (Hrsg.): *Evaluation. Ein systematisches Handbuch*, Wiesbaden: Springer, 403-413.
- Scriven, M. (1967): The methodology of evaluation, in: Tyler, R.W; Gage, R.M.; Scriven, M. (Hrsg.): *Perspectives of Curriculum Evaluation*, Chicago: Rand McNally, 39-83.
- Searle, J.R. (1995): *The construction of social reality*. London: Penguin Books.
- Self, P. (1996): *Government by the market? The politics of public choice*. Basingstoke: Macmillan.
- Shulock, N. (1999): The paradox of policy analysis. If it is not used, why do we produce so much of it?, in: *Journal of Policy Analysis and Management* 18 (2), 226–244.
- Sismondo, S. (2017): Post-truth?, in: *Social studies of science* 47 (1), 3–6.
- Smart, B. (1999): *Facing modernity: ambivalence, reflexivity and morality*. London: Sage Publications.
- Snow, C. P. (2013): *The two cultures and the scientific revolution*. Mansfield Centre: Martino.
- Speer, S.; Pattyn, V.; De Peuter, B. (2015): The growing role of evaluation in parliaments: holding governments accountable?, in: *International Review of Administrative Sciences* 81(1), 37-57.

- Spektrum (2017): *March for Science: Ein Zeichen gegen postfaktisches Denken*. 23.04.2017  
[online aufgerufen am 16.07.2020]: <https://www.spektrum.de/news/march-for-science-in-deutschland/1453013>
- Stephenson, P. J.; Schoenefeld, J. J.; Leeuw, F. L. (2019): The Politicisation of Evaluation. Constructing and Contesting EU Policy Performance, in: *Politische Vierteljahresschrift* 60 (4), 663–679.
- Stern, E. (2009): Evaluation policy in the European Union and its institutions, in: *New Directions for Evaluation* 123, 67–85.
- Stockmann, R. (2006): *Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement*. Münster/München/Berlin: Waxmann (Sozialwissenschaftliche Evaluationsforschung, Band 5).
- Stockmann, R.; Meyer, W. (2014): *Evaluation. Eine Einführung*. Toronto: Opladen.
- Stockmann, R.; Meyer, W. (2017): Evaluation in Deutschland. Woher sie kommt, wo sie steht, wohin sie geht, in: *Zeitschrift für Evaluation* 2017 (2), 57-110.
- Stucki, I. (2018): Evidence-based arguments in direct democracy: The case of smoking bans in Switzerland, in: *Evaluation and Program Planning* 69, 148-156.
- Svara, J. (2001): The Myth of the Dichotomy: Complementarity of Politics and Administration in the Past and Future of Public Administration, in: *Public Administration Review* 61 (2), 176-183.
- Tagesschau (2020): *Holocaust-Leugnung. Fakes, falsche Gutachten und Lügen*. 27.01.2020  
[online aufgerufen am 10.08.2020]: <https://www.tagesschau.de/faktenfinder/holocaust-leugnung-101.html>
- Tagesspiegel (2020): *Fake News bei Youtube. 21 Millionen Klicks für Leugner des Klimawandels*. 16.01.2020 [online aufgerufen am 10.08.2020]:  
<https://www.tagesspiegel.de/politik/fake-news-bei-youtube-21-millionen-klicks-fuer-leugner-des-klimawandels/25440754.html>
- Tarschys, D. (1983): Fluctuations in the political demand for policy analysis. In: Spiro, S.E.; Yuchtman-Yaar, E. (Hrsg.): *Evaluating the welfare state*, New York: Academic Press, 375-385.

- Taz (2020): *Evaluation der Hamburger Mietpreisbremse: „Nahezu wirkungslos“*. 23.07.2020  
[online aufgerufen am 23.07.2020]: <https://taz.de/Evaluation-der-Hamburger-Mietpreisbremse/!5695545/>
- Telepolis (2016): *Wider die Rede vom "Postfaktischen"*. 10.12.2016 [online aufgerufen am 26.05.2020]: <https://www.heise.de/tp/features/Wider-die-Rede-vom-Postfaktischen-3562756.html>
- The White House (2009): *Transparency and Open Government. Memorandum for the Heads of Executive Departments and Agencies. Subject: Transparency and Open Government*. 21.01.2009 [online aufgerufen am 08.09.2020]: <https://obamawhitehouse.archives.gov/the-press-office/transparency-and-open-government>
- Trigg, R. (2001): *Understanding social science: a philosophical introduction to the social sciences*. Oxford: Blackwell.
- van der Knaap, P. (1995): Policy Evaluation and Learning, in: *Evaluation* 1 (2), 189–216.
- Walker, R. (2000): Welfare policy: tendering for evidence, in: Davies, H.T.O.; Nutley, S.M.; Smith, P. (Hrsg.): *What works?: Evidence-based policy and practice in public services*, Policy Press, 141-166.
- Washington Post (2020): *Trump has made more than 20,000 false or misleading claims* [online aufgerufen am 10.09.2020]: <https://www.washingtonpost.com/politics/2020/07/13/president-trump-has-made-more-than-20000-false-or-misleading-claims/>
- Weiland, S. (2013): Evidenzbasierte Politik zwischen Eindeutigkeit und Reflexivität, in: *Technikfolgenabschätzung – Theorie und Praxis* 22 (3), 9-15.
- Weiss, C. H. (1978a): Improving the linkage between social research and public policy, in: *Knowledge and policy: The uncertain connection*, 23-81.
- Weiss, C. H. (1978b): *Using social research in public policy making*. Lexington: Lexington Books (Policy Studies Organization Series, 11).
- Weiss, C. H. (1982): Policy research in the context of diffuse decision making, in: *Journal of Higher Education* 53 (6), 619–39.

- Weiss, C. H. (1987): *Evaluation for decisions: Is anybody there? Does anybody care?* Plenary address for meeting of the American Evaluation Association, Boston.
- Weiss, C. H. (1999): The Interface between Evaluation and Public Policy, in: *Evaluation* 5(4), 468–486.
- Weiss, C. H.; Bucuvalas, M. J. (1980): *Social science research and decision-making*. Columbia University Press.
- White, H. (2019): The twenty-first century experimenting society. The four waves of the evidence revolution, in: *Palgrave Communications* 5 (1), 1-7.
- Widmer, T. (2009): The contribution of evidence-based policy to the output-oriented legitimacy of the state, in: *Evidence & Policy: A Journal of Research, Debate and Practice* 5 (4), 351-372.
- Wissenschaft im Dialog (2019): *Wissenschaftsbarometer 2019* [online aufgerufen am 20.05.2020]: [https://www.wissenschaft-im-dialog.de/fileadmin/user\\_upload/Projekte/Wissenschaftsbarometer/Dokumente\\_19/Broschuere\\_Wissenschaftsbarometer2019.pdf](https://www.wissenschaft-im-dialog.de/fileadmin/user_upload/Projekte/Wissenschaftsbarometer/Dokumente_19/Broschuere_Wissenschaftsbarometer2019.pdf)
- Wollmann, H. (1997): Evaluation in Germany. Between early start, continuous performance, and the retarded professionalization, in: *EES Newsletter* 3, 4-5.
- Wollmann, H. (1999): Politik- und Verwaltungsmodernisierung in den Kommunen. Zwischen Managementlehre und Demokratiegebot, in: *Die Verwaltung* 32 (3), 345-376.
- Wollmann, H. (2004): Leistungsmessung („performance measurement“) in Politik und Verwaltung: Phasen, Typen und Ansätze im internationalen Überblick, in: Kuhlmann, S.; Bogumil, J.; Wollmann, H. (Hrsg.): *Leistungsmessung und -vergleich in Politik und Verwaltung*. Wiesbaden: VS Verlag für Sozialwissenschaften, Stadtforschung aktuell (96).
- Wollmann, H. (2007): Policy evaluation and evaluation research, in: Fischer, F.; Miller, G.J.; Sidney, M.S. (Hrsg.): *Handbook of policy analysis: theory, politics, and methods*, Boca Raton: CRC Press, 393-402.
- Wollmann, H. (2017): Entwicklungslinien von Evaluation und Evaluationsforschung in Deutschland. Zwischen Vorgeschichte, Aufbruch und Konsolidierung, in: *Zeitschrift für Evaluation* 16 (2), 32-52.

ZDF (2020): *Covid-19-Pandemie - Wie die rechte Szene die Corona-Krise benutzt*. 26.03.2020  
[online aufgerufen am 07.08.2020]:  
<https://www.zdf.de/nachrichten/politik/coronavirus-rechte-gruppen-covid-19-104.html>

Zeit (2016a): *Lügen: Die Erde ist eine Scheibe*. 28.08.2016 [online aufgerufen am 07.08.2020]:  
<https://www.zeit.de/2016/36/luegen-politik-donald-trump-rudy-giuliani-thomas-de-maiziere-gefuehlte-wahrheit/komplettansicht>

Zeit (2016b): *Wissenschaft: Die postfaktische Universität*. 15.12.2016 [online aufgerufen am  
03.07.2020]: <https://www.zeit.de/2016/52/wissenschaft-postfaktisch-rationalitaet-ohnmacht-universitaeten>

Zeit (2017): *March for Science: "Forschung basiert auf Fakten, nicht auf Meinungen"*.  
22.04.2017 [online aufgerufen am 07.08.2020]: <https://www.zeit.de/wissen/2017-04/march-for-science-demonstration-wissenschaftler-gruende>