

Policies sind von Politik und Verwaltung entwickelte Programme, die politisch als bearbeitenswert empfundene Problemlagen adressieren sollen. Ihre inhaltliche Ausgestaltung wird von den Annahmen der entscheidenden Akteure über Ursachen der Problemlage und damit kausal verbundenen Zusammenhängen bestimmt. Aufbauend auf diesen Annahmen werden in einem zweiten Schritt einzelne Instrumente und Maßnahmen entwickelt, die diesen Zusammenhängen entgegenwirken sollen. Die *Performanz* der innerhalb einer *policy* ergriffenen Maßnahmen kann an unterschiedlichen Punkten des politisch-administrativen Produktionsprozesses gemessen werden. Die Wahl dieses Messzeitpunkts stellt wie die Frage der Perspektive der Betrachtung eine für den Gang der weiteren Untersuchung richtungsweisende Entscheidung dar.

Dubnick (2005) bietet die grundlegendste Differenzierung möglicher Ansatzpunkte einer Leistungsmessung. Er unterscheidet zwischen dem Fokus auf die Handlung selbst und jenem auf das Handlungsergebnis (vgl. Tabelle 2). Während Performanz im Feld (P1) wertfrei den bloßen Sachverhalt einer Aufgabenausführung (Produktion) beschreibt, ohne deren Qualität in Ausführung oder Ergebnis besonders zu beachten, beinhalten die drei Felder (P2, P3, P4) normative Bewertungen (vgl. Van Dooren 2006: 14). Steht die Handlungsqualität im Vordergrund und spielt die Zielerreichung eine untergeordnete Rolle, so entspricht Performanz der *Kompetenz* (P2) der Mitarbeiter. Dem liegt die Annahme zugrunde, dass Organisationen mit besser qualifizierten Mitarbeitern eine höhere Qualität in ihrem Handeln erreichen und ihnen eine größere Handlungskapazität zugeschrieben wird. Liegt hingegen der Fokus nicht auf der Qualität der Handlungen, sondern auf der Zielerreichung, so entspricht Performanz den gelieferten *Ergebnissen* (P3). Liegt der Fokus gleichzeitig auf der Handlungsqualität und der Qualität der Zielerreichung, so kann dies mit dem relationalen Maß der *Produktivität* (P4) gleichgesetzt werden (vgl. Dubnick 2005: 392; Van Dooren 2006: 13ff.).

		Fokus auf die Qualität der Zielerreichung	
		<i>hoch</i>	<i>nieder</i>
Fokus auf die Qualität des Handelns	<i>nieder</i>	Produktion (P1)	Ergebnis (P3)
	<i>hoch</i>	Kompetenz (P2)	Produktivität (P4)

Tabelle 2: Performanzdimensionen

Quelle: Dubnick 2005: 392.

In modernen öffentlichen Verwaltungen sollen komplexe Sachverhalte mit begrenzten Ressourcen zuverlässig bearbeitet werden. Die Dubnick'sche Differenzierungsübung macht deutlich, dass ein holistischer Performanzbegriff die Qualität des *Handelns* in einer Verwaltung als auch die Qualität der *Zielerreichung* umfasst. Dabei ist fraglich, ob das Konzept der *Produktivität* (P4) als komplexe, die Dimensionen Handeln und Zielerreichung ohne hohen Informationsverlust integrierende Maßzahl konzeptioniert werden kann. Auch muss noch genauer spezifiziert werden, welche Aspekte des Handelns tatsächlich als Performanz aufzufassen, und welche eher den vorgelagerten Prozessen zuzuordnen sind.

In der Literatur werden als Messpunkte betrachtet: (1) der Ressourceneinsatz, also die Kosten des Verwaltungsvollzugs (*Input*), (2) die behördeninternen Prozesse, deren Transparenz, Legitimität und Qualität (*Throughput*), (3) die unmittelbare Leistungsproduktion der Verwaltung (*Output*), (4) deren unmittelbare Wirkung (*Outcome*) sowie schließlich (5) auch die tatsächliche Zielerreichung, gemessen an der ursprünglichen Intention des einzelnen staatlichen Programms (*Impact*). Zwar ist die Abgrenzung zwischen diesen Schritten in der Praxis teils schwierig, allerdings ist sie analytisch nutzbar, um eine erste Strukturierung von unterschiedlichen Performanzperspektiven vorzunehmen (vgl. Patton 1997; Wollmann/Kuhlmann 2010: 563).

Die reine *Input*-Orientierung erweist sich als ungeeignete Mess- und Steuerungsgröße: *Inputs* spiegeln lediglich die Kosten staatlicher Programme ohne Bezug zur produzierten Leistung. Die Steuerung über *Inputs* – wie die auch heute noch bei Politikern beliebte „Rasenmähermethode“ (vgl. Banner 1987: 54; Röber 2005: 85; Holtkamp 2010: 52ff.) ist zwar höchst effektiv in der Kontrolle des Ausgabenniveaus. Da diese Steuerungsform jedoch sämtliche funktionalen Notwendigkeiten des Vollzugs ignoriert und fälschlicherweise von einem linearen Zusammenhang zwischen *Inputs* und produzierter Verwaltungsleistung auszugehen scheint (vgl. Bauer et al. 2007b: 244; Ebinger/Bogumil 2008: 25f.), sind *Input*-Maßzahlen keine validen Indikatoren für tatsächlich erbrachte Verwaltungsleistung (vgl. Klingebiel 1999: 41). Vielmehr ist Ressourceneinsatz eine unabhängige, erklärende Variable, die jedoch nur im Zusammenspiel mit weiteren Variablen einen Beitrag zur Erklärung der Performanz einer Verwaltung leisten kann.

Haus und Heinelt (2005: 15) führen eine weitere Performanzdimension ein, indem sie das ursprünglich auf den demokratischen Entscheidungsbildungsprozess bezogene Legitimationskonzept des *Throughput* (Zürn 1998: 240f., 2000) auf verwaltungsinterne *Prozesse* ummünzen. Eine hohe *Throughput*-Legitimation wird danach erzielt, wenn Institutionen, Verwaltungsvorgänge und das Zustandekommen von Entscheidungen transparent und nachvollziehbar sind. Diese Dimension verdient Berücksichtigung in der empirischen Analyse, da für Bürger und Unternehmen transparente und zurechenbare Prozesse von verwal-

tungsinternen Kriterien der Ablauforganisation abhängen, nämlich von klaren Kompetenzzuweisungen und Entscheidungsregeln. Inwiefern diese Voraussetzungen gegeben sind, scheint über eine Mitarbeiterbefragung erfassbar.

Der *Output* beschreibt das konkrete Produkt staatlicher Tätigkeit – bspw. die Zahl der Entleihungen in einer Stadtbibliothek oder die Bereitstellung und Nutzung eines Beratungsangebots zur Kindergesundheit (vgl. Patton 1997; Atkinson et al. 2005; Wollmann/Kuhlmann 2010). Die Mehrheit der Institutionenevaluierungen setzt auf dieser Ebene des Produktionsprozesses an, da sich hier sowohl ein direkter Abgleich mit explizit gesetzten Produktionszielen sowie der Vergleich mit dritten, dieselbe oder eine ähnliche Aufgabe vollziehenden Organisation anbietet. Auch können zahlreiche Probleme später ansetzender Betrachtungen vermieden werden. So zielen die *Outcome*- und *Impact*-Betrachtungen auf die tatsächliche Wirkung auf die Zielgruppe oder die Zielerreichung gemessen am gesetzten Globalziel ab. Im Regelfall ist die Zuordnung des einzelnen *Outputs* zu einem *Outcome* oder gar dem *Impact* problematisch.²¹ Die Zielerreichung erfordert oft eine Verhaltensänderung der anvisierten Zielgruppe (vgl. Van Dooren et al. 2006: 12; Ritz/Sager 2010: 127). Andere politisch-administrative Programme und sozio-ökonomische Zusammenhänge können intervenieren, Wirkungen können verzögert oder verzerrt auftreten und nicht-intendierte Effekte müssen berücksichtigt werden (vgl. Bouckaert/Balk 1991; Meyer/Gupta 1994; van Thiel/Leeuw 2002). Angesichts dieser Komplexität warnt Eisner (2006: 153)²² mit explizitem Bezug auf die Umweltverwaltung:

„We often judge the performance of environmental regulation in light of the concentrations of principal pollutants, assigning regulators the credit (or blame) for the results. Unfortunately, policy analysts often forget that “policy outcomes” are more correctly viewed as the product of a complicated set of relationships and decisions [...].”

Impact und *Outcome* sind offensichtlich problematische Indikatoren für die Performanz von Verwaltungen.²³ Deshalb soll in dieser Untersuchung die abhängige Variable nur auf der *Throughput*- und *Output*-Ebene erfasst werden.

21 Wilson (1989: 132) erläutert: „We cannot easily say whether the pupils were educated, the streets made safer, or some diseases prevented; but we can say whether every pupil got the same textbook, every citizen got the same police response, and every patient got the same vaccine.”

22 Vgl. ähnlich Van Dooren (2006: 13); Knill (2003: 170ff.).

23 Anders argumentieren Moynihan/Pandey (2005: 429f.): „Efforts to track performance tend to split between effectiveness measures that focus on operations or outputs and measures that focus on outcomes or results. As Richard Fenno (1966, 391) notes, “If, therefore, one attempts to measure bureau success by one of these measures [outputs or outcomes] to the exclusion of the other, he will be neglecting a discrete dimension of bureau performance.” Our measure of organizational effectiveness permits employees to rely on their organizational experience and understanding. Managers will be aware of the various goals of their agency, and they are in a good position to evaluate success in

2.3 Kategorisierung von Performanz

Der in der *Public Administration*- und Verwaltungsforschung allgegenwärtige Begriff der *organisationalen Performanz* beschreibt ein gesellschaftlich konstruiertes Phänomen mit variablem Inhalt (Anspach 1991; Reiter et al. 2009). Ob seiner Breite und unterschiedlichen Füllung ist das Konzept nicht unumstritten (vgl. Talbot 2000: 64; Reichard 2004: 341; Nullmeier 2004: 47). Schon 1974 beschrieb Derlien in seinem Aufsatz „Theoretische und methodische Probleme der Beurteilung organisatorischer Effizienz der öffentlichen Verwaltung“ eindrücklich den bereits damals langen Weg zur Entwicklung eines multidimensionalen Performanzbegriffs in der *Public Administration* Forschung (Derlien 1974: 1). Reding (1981: 18) konstatierte schon bald zynisch, dass die Verwendung des Effizienzbegriffs – hier nur eines von mehreren Unterkonstrukten unter dem Dach des Performanzbegriffs – mangels einer akzeptierten Konvention immer ineffizienter werde.

Auch die bundesdeutsche staatswissenschaftliche Organisationsforschung trieb in der Phase der sog. *Steuerungseuphorie* in den 1960er und 1970er Jahren (vgl. Scharpf 1973) die Ausdifferenzierung des Performanzbegriffs voran. Die systematische empirische Überprüfung der hier getroffenen Annahmen fand jedoch kaum statt. Dies änderte sich erst in den 1980er Jahren mit einer neuen Welle performanzorientierter Organisationsreformen. Dieses New Public Management (NPM) erfasste ausgehend von den konservativen Regierungen der angelsächsischen Staaten in diversen Transformationen die meisten OECD-Staaten (vgl. Christensen/Lægreid 2001a). Mit der Zielsetzung, den öffentlichen Sektor zu reformieren und leistungsfähiger zu machen, wuchs die Notwendigkeit präziser Performanzdefinitionen. Ein konsequent umgesetztes, von betriebswirtschaftlichen Konzepten abgeleitetes Steuerungsmodell erfordert ein *performance management* mit klar definierten politischen Zielen, Produkten der Verwaltungstätigkeit und Kennzahlen über Kosten und Zielerreichung: Was sind öffentliche Leistungen? In welche einzelnen Leistungsdimensionen lassen sich diese zerlegen und welche davon verdienen Berücksichtigung oder gar Priorisierung? Und schließlich, aber alles andere als unwichtig: Wie können die entsprechenden Variablen operationalisiert und gemessen werden? Die Übersetzung politischer Zielsetzungen in tatsächlich steuerungsrelevante Indikatoren, die Entwicklung

reaching these goals. Employee self-reports have the virtue of being global measures of performance, and it is therefore reasonable to expect that such evaluations will include an appropriate mix of output- and outcome-based criteria. A similar employee-based measure of performance has been employed by others (Molnar and Rogers 1976; Mott 1972), most recently by Brewer and Selden (2000), who point out that while the organizational psychology literature argues that organizations and individuals are interdependent, the employee perspective on performance is underemployed.“