

birgt drei Problemlagen: Erstens verstärkt sie die Problematik der *vielen Variablen bei wenigen Fällen* (Scharpf 2000: 765f.). Indem lediglich ein Akteur pro Behörde befragt wird, ist aufgrund der geringen Fallzahl eine über einfache deskriptive Statistik hinausgehende quantitative Datenauswertung selten möglich. Zweitens ist die *Validität* der Aussagen von Führungskräften (oder mit dem Ausfüllen des Fragebogens beauftragter Subalternen) zu hinterfragen. So arbeiten diese Akteure nicht selbst innerhalb der interessierenden Strukturen und Prozesse, sondern nehmen eine herausgehobene, durch Informationsasymmetrien gekennzeichnete Position ein. Gerade in größeren, multifunktional organisierten Behörden wie Mittelinstanzen, Kommunalverwaltungen und Sammelbehörden sind sie sehr weit von der praktischen Vollzugstätigkeit einzelner Bereiche entfernt und verfügen im Regelfall auch nicht über die zur *fachlichen* Einschätzung der Vollzugsleistung notwendige Expertise. In der Folge können ihre Wahrnehmungen stark von jenen nachgeordneter Verwaltungsmitarbeiter abweichen. Abgesehen davon kann ihnen ein starkes Interesse an der Kontrolle der nach außen gelangenden Informationen und einer möglichst positiven Darstellung der Behördentätigkeit unterstellt werden (vgl. Kumar 1993; Enticott et al. 2009: 231), da sie zu Recht oder zu Unrecht für die erreichte Vollzugsleistung verantwortlich gemacht werden. Vor diesem Hintergrund erscheint die Praxis, aus den individuellen Aussagen dieser speziellen Akteurebene vorbehaltlos eine Zustandsbeschreibung der Behörden auf einer Makro-Ebene herauszulesen, äußerst fragwürdig. Die Folgerung von Walker und Enticott (2004: 432), dass über derartige Elitesurveys gewonnene Daten mit Vorsicht zu genießen sind, wird hier geteilt.

5.1.2. Föderalismus als valides most similar systems design

Aus den genannten Problemen der häufig in der Organisationsforschung gewählter Ansätze lassen sich Anforderungen an ein geeignetes Untersuchungsdesign für die hier bearbeiteten Forschungsfragen ableiten:

- Größtmögliche Varianz auf den als theoretisch relevant erachteten strukturellen Erklärungsfaktoren,
- größtmögliche Homogenität oder Kontrolle relevanter Kontextvariablen,
- eine ausreichend große Fallzahl zur Anwendung multivariater Analysemethoden und
- Erfassung nicht nur der Performanz-, sondern auch von Prozessvariablen der internen Verwaltungsorganisation.

Um diese Anforderungen zu erfüllen, greift das gewählte Forschungsdesign auf eine simple, aber erfolgversprechende Strategie zurück: Den *subnationalen Institutionenvergleich*. Als Form des Intrasystem-Vergleichs wird hierbei eine Grup-

pe von Subsystemen innerhalb eines Staates analysiert (vgl. Lijphart 1971: 687-690; Smelser 1976: 198-200; Snyder 2001: 95-97; Sager 2009). Um nicht *Unvergleichbares zu vergleichen* (vgl. Derlien 1992: 285-287) und die Problematik *zu vieler Variablen bei zu wenigen Fällen* zu mildern, eignet sich ein solches *most-similar systems Design* (Przeworski/Teune 1970; Meckstroth 1975; Faure 1994) am besten. Die untersuchten Einheiten unterliegen weitgehend identischen Rahmenbedingungen, weisen jedoch spezifische Besonderheiten auf, so dass sie nicht als eine homogene Einheit gelten können (Freitag/Vatter 2010: 181).⁹⁸ Snyder (2001: 94f.; vgl. Denk 2010: 30) benennt drei Vorteile dieser Vorgehensweise und diskutiert sie im Detail: Erstens kann damit das grundlegende Problem von Fallstudien oder kleinzahliger Vergleiche – das ungünstige Zahlenverhältnisse zwischen Variablen und Fallzahlen – durch eine Ausweitung der Vergleichsgrundlage deutlich abgemildert werden. Dies ermöglicht die Analyse auch komplexer Wirkungszusammenhänge. Zweitens können mittels einer restriktiven Fallauswahl solche Kontextfaktoren konstant gehalten werden, die potenziell das Verwaltungshandeln beeinflussen, aber nur begrenzt oder überhaupt nicht zu erfassen und deren individuelle Effekte deshalb nicht zu differenzieren sind.⁹⁹ Drittens sind Intrasystemvergleiche sensibel gegenüber systeminterner Varianz (vgl. Snyder 2001: 100; Dogan 1994), so dass unterschiedliche Wirkmechanismen etwa in verschiedenen Regionen eines Staates nicht verzerrend unter eine Kennzahl subsumiert werden.

Insbesondere die ersten beiden Punkte – Fallzahl und konstante Kontextfaktoren – sind für die Bearbeitung der hier gestellten deskriptiven und analytischen Forschungsfragen von Bedeutung. Durch den Rückgriff auf die deutschen Flächenländer stehen potenziell 13 Vergleichsfälle mit einer unüberschaubaren Zahl an Verwaltungsorganisationen zur Verfügung. Damit kann der in den theoretischen Modellen zu den Determinanten administrativer Performanz längst vollzogene Schritt von einem fallzentrierten zu einem variablenzentrierten Ansatz (vgl. Przeworski/Teune 1970) auf ein kausalanalytisch überprüfbares Modell übertragen werden.

Hierzu muss allerdings die in Snyders zweitem Punkt angemahnte restriktive Fallauswahl vollzogen werden, die aus dem theoretisch möglichen Universum an Landesbehörden eine beherrschbare Grundgesamtheit tatsächlich vergleichbarer *Verwaltungseinheiten* extrahiert. Mit *Verwaltungseinheit* ist eine um *einen zusammenhängenden Aufgabenkomplex* organisierte Behörde oder ein Teil einer

98 Beispiele bemerkenswerter internationaler Anwendungen dieses Ansatzes sind Putnam (1993), Vatter (2002), Gray/Hanson (2004), Dunn (2006).

99 Solche Faktoren sind vor allem in politischen Rahmenbedingungen der Makro-Ebene, sozioökonomischem Kontext, nationaler Verwaltungskultur und Rechtstradition sowie in Policy- und Aufgabenspezifika zu sehen (vgl. Przeworski/Teune 1970; Faure 1994).

Behörde gemeint. Dieser Begriff ist zentral für die weitere Untersuchung. Nur auf dieser Ebene kann tatsächliche Vergleichbarkeit erreicht werden, indem für die vollzogene *Aufgabe* und damit zusammenhängende Besonderheiten kontrolliert und diese so als erklärende Variable neutralisiert wird. Dies ist von herausragender Bedeutung, da der Aufgabenspezifität eine entscheidende Rolle vom Prozess des *policy-making* bis hin zur Implementation und ihrer Bewertung zugeschrieben wird (Lowi 1972; Windhoff-Héritier 1987: 48; Benz 1997; John/Cole 2000; Heinelt 2003; Lægreid et al. 2006). Aussagen über Einflussfaktoren administrativer Performanz werden kaum möglich sein, wenn in den betrachteten Verwaltungseinheiten verschiedene, sich in ihren Charakteristika fundamental unterscheidende *policies* vollzogenen würden; also bspw. sowohl regulativ als auch distributiv (vgl. Lowi 1964), politisch hoch salient und gänzlich irrelevant (vgl. RePass 1971) oder ausgeführt von hochqualifizierten Spezialisten und Hilfskräften (vgl. Freidson 2001). Da hier Aussagen über die *generellen* Determinanten von Verwaltungsleistung gesucht werden, müssen sich die Verwaltungseinheiten als Untersuchungseinheiten der Analyse hinsichtlich ihrer Aufgaben gleichen, die vielfältigen Merkmale des *Organisations-* oder *Verwaltungsmodells* müssen hingegen variieren.

Nach der Festlegung der Untersuchungseinheiten muss nun auch die Ebene der Datenerhebung festgelegt und begründet werden. Zur Erfassung von Prozess- und Ergebnisdaten bietet sich der oben vorgestellte Rückgriff auf Perzeptionsdaten an. Allerdings muss die beschriebene Problematik einer fragwürdigen Überinterpretation der Aussagen von Behördenleitungen vermieden werden. Dies bedeutet jedoch nicht, dass diese Akteursebene gänzlich aus der Betrachtung ausgeklammert werden sollte. Auch wenn die Forschung zu *street level bureaucrats* Strukturen und Verwaltungsführung einen eher bescheidenen Einfluss einräumt (Maynard-Musheno/Moody 2003: 10), kann doch davon ausgegangen werden, dass diese Faktoren den Rahmen beeinflussen, innerhalb dessen das Vollzugspersonal seine Rolle ausfüllt (vgl. May/Winter 2007: 454-457). Auch erinnert Teodoro (2011: 7) mit Verweis auf renommierte Vertreter der organisationstheoretischen und *public administration* Forschung daran, dass Behördenmitarbeiter individuelle Entscheidungen auf der Basis ureigener Ziele und persönlicher Zwänge trafen, weshalb theoretische Erörterungen diese Ebene als Erklärungsfaktoren einbeziehen müssen.¹⁰⁰ Darüber hinaus muss die Wirkung von Umweltfaktoren und organisationalen Strukturen auf die Prozesse und Einstellungen von Mitarbeitern innerhalb der Behörden berücksichtigt werden. Diese Zusammenhänge werden in der verwaltungswissenschaftlichen Forschung bisher vernachlässigt (Pandey/Wright 2006: 513).

100 Teodoro bezieht sich auf die Entwicklung einer „theory of bureaucratic policy entrepreneurship“ (2011: 7), sein Argument kann jedoch als allgemeingültig angesehen werden.

Um verwaltungsinterne Prozesse erfassen zu können und die oben erläuterten Problemlagen auf Führungskräfte fokussierender Surveys zu umgehen, soll in dieser Untersuchung ein anderer Weg eingeschlagen werden. Für die Ebene der Prozesse und *Outputs* scheint nicht die Wahrnehmung des obersten Leitungspersonals, sondern jene der Mitarbeiter im Vollzug von besonderem Interesse zu sein. Dies erfordert zwangsläufig, dass die Datenerhebung diese für die empirische *public management* Forschung unüblich Zielgruppe adressiert. Es werden hier nicht einzelne Führungskräfte stellvertretend für ihre Behörde, sondern eine größere Zahl unmittelbar mit dem Aufgabenvollzug befassten Verwaltungsmitarbeiter befragt. Konkret sollen die mit Vollzugsaufgaben befassten Mitarbeiter in spezifischen Verwaltungseinheiten – der kleinsten mit der Bearbeitung eines zusammenhängenden Themenfeldes befassten Einheit – als Informanten dienen. Diese Akteure sollen Auskunft über Organisationsmerkmale, Prozessdaten auf der Meso-Ebene der Verwaltungsführung sowie zu Verhalten und Performanz auf der Mikro-Ebene der Verwaltungseinheiten geben. Die in diesem Abschnitt vorstellte Herangehensweise ermöglicht eine Vergleichbarkeit der untersuchten Fälle sowie die Generierung eines “large N“-Datensatzes zur hypothesengeleiteten Beantwortung der deskriptiven und analytischen Forschungsfragen. Um wie oben besprochen zum einen die Kontextvariablen möglichst konstant zu halten und zum anderen den *policy*-Effekt zu neutralisieren, ist ein Aufgabenfeld zu definieren, aus dem Organisationen als Untersuchungsfälle ausgewählt werden. Als solches Feld bietet sich die *Gewerbeaufsicht* an. Diese Auswahl und die vorgenommene Spezifizierung auf zwei Teilgebiete wird im Folgenden begründet.

5.2 Wahl des Untersuchungsfalls: Arbeitsschutz und Immissionsschutz

Wie oben ausgeführt, soll in Anlehnung an ein *most similar systems design* eine Reihe von schwer zu kontrollierenden Einflussfaktoren durch eine gezielte Fallauswahl konstant gehalten werden. Für die empirische Untersuchung wird ein Intrasystemvergleich angelegt, wozu innerhalb eines Föderalstaates ein zentralstaatlich reguliertes, jedoch auf subnationaler Ebene von Teilstaaten vollzogenes Politikfeld ausgewählt wird. Diese Auswahl wird in drei Schritten begründet: Als erstes ist ein Föderalstaat und seine Gliedstaaten festzulegen. Im Anschluss wird diskutiert, welche Art von *policy* – gemessen an ihrer Zielsetzung – für die Untersuchung besonders geeignet ist. Darauf aufbauend kann im dritten und letzten Schritt die Auswahl des konkreten Politikfeldes (Gewerbeaufsicht) mit seinen Spezialgebieten Arbeits- und Immissionsschutz begründet werden.

Aufgrund ihres ausgeprägten Verwaltungsföderalismus (Art. 30, 83 GG) ist die Bundesrepublik Deutschland prädestiniert für die empirische Untersuchung. So fallen dem Bund im Rahmen der *Ausschließlichen Gesetzgebung des Bundes*