

3. Methodisches Vorgehen und Fallauswahl

Ein zentrales Forschungsproblem liegt häufig in der kausalen Zuordnung von Veränderungen, in diesem Fall zu den Digitalisierungsfolgen (Wollmann 2006: 211). Wie können die empirisch beobachtbaren Veränderungen entlang der vier Transformationsdimensionen in einen kausalen Zusammenhang mit der unabhängigen Variable, namentlich dem Digitalisierungsprozess, gesetzt werden? Forschungsstrategien wie die Erhebung eines Vorher- und eines Nachher-Stands konnten hierfür nicht durchgeführt werden, genauso wenig wie Quasi-Experimente, bei denen zwei sehr ähnliche Organisationen verglichen werden und nur bei einer der Stimulus ‚Digitalisierung‘ wirkt. Hier kommt zum Tragen, dass Digitalisierung in den Verwaltungen allgegenwärtig ist und sich keineswegs auf die Einführung des digitalen Bauantrags, von i-Kfz usw. reduzieren lässt. Vor allem die Betrachtung verwaltungsinterner Variablen wie die Veränderungen von Organisation und Arbeit kann nicht ohne die Betrachtung von E-Mail-Kommunikation, Fachverfahrenssoftware, Onlineterminvereinbarung und E-Akte etc. erfolgen. Neue E-Government-Bürgerdienste sind insofern nur die gut sichtbare Spitze des Eisbergs der Verwaltungsdigitalisierung (Lenk 2011: 319). Hinzu kommt, dass Verfahren wie i-Kfz und ELSTER aufgrund gesetzlicher Anforderungen in allen Verwaltungen eingeführt werden.

Kausalitätsproblem

Damit verbunden ist auch die Unbestimmtheit von Digitalisierung, die als Konzept und Kategorie kaum trennscharf ist. Zum einen ist Digitalisierung faktisch nicht zu trennen von wesentlich früheren Formen von „Verwaltungsautomation“ (Brinkmann et al. 1974; Brinkmann/Kuhlmann 1990; Grimmer 1986). Zum anderen steht die Verwaltungsdigitalisierung auch in unverkennbaren Sinnzusammenhängen mit Verwaltungsreformprogrammen wie Open Government, New Public Management, der Neuordnung der Verwaltung an Bürger-Lebenslagen und dem Neuen Steuerungsmodell (vgl. Wewer 2020; Pollitt/Bouckaert 2017; Dunleavy et al. 2006; Bogumil et al. 2007; Bogumil et al. 2019; Liedtke/Tepper 1989). Folglich ist Digitalisierung in der einen oder anderen Form bereits in jeder deutschen Verwaltung anzutreffen. Dies erschwert eine wirkliche Kontrolle für diesen Stimulus.

Für Institutionenevaluationen empfiehlt Wollmann (2000: 207–209) daher ein fallstudienbasiertes Vorgehen. Aus methodologischer Sicht bieten Fallstudien eine hohe Aussagekraft gegenüber dem untersuchten Fall (interne Validität). Auch viele Variablen können so intensiv beforscht werden (Prittwitz 2007: 10–11), was insbesondere die Analyse von Kausalbeziehungen zwischen Variablen im Wege des *Process Tracing* ermöglicht (Siewert 2017). Wesentlich für die angestrebte Feststellung einer Kausalbeziehung ist dabei die Untersu-

Fallstudiendesign

chung der zugrundeliegenden Beziehungen von intervenierenden Variablen auf die Abhängige. Auf deutsch: Um nachzuweisen, dass Digitalisierung ursächlich für eine bestimmte Wirkung ist, ist der zugrundeliegende Mechanismus offenzulegen (z.B. Digitalisierung geht einher mit mehr Computerarbeit → mehr Computerarbeit führt zu mehr Sitzen am Schreibtisch → führt ggf. zu gesundheitlichen Muskel-Skelett-Problemen) (Siewert 2017: 243–244). Die reine Feststellung, dass sich nach Start eines Digitalisierungsprojekts eine Variable geändert hat, reicht dagegen nicht für einen Kausalbeleg aus (ebd.). Vor allem in den Sozialwissenschaften ist es allerdings meist nicht möglich, Kausalitätsmechanismen vollumfänglich zu erklären, für die Inferenz von Kausalität (*causal inference*) sollten mögliche Mechanismen aber zumindest in minimaler Art und Weise skizziert werden (ebd.: 245). Geht es darum, Empfehlungen für eine bessere Gestaltung der Implementierungsprozesse zu formulieren, ist dieses detaillierte Verständnis von Mechanismen zudem sehr wertvoll.

Generalisierbarkeit
der Ergebnisse

Eine Schwäche von Fallstudien besteht darin, dass Ergebnisse von wenigen untersuchten Einzelfällen nicht verallgemeinerungsfähig sind, d.h. keine externe Validität aufweisen (vgl. Wollmann 2000: 207–209; Prittwitz 2007: 11; Gerring 2004): Die einzelne Untersuchung von digitaler Transformation in einem Amt ist schwerlich aussagekräftig für andere Ämter oder gar die gesamte deutsche Verwaltungslandschaft. Diese Untersuchung erweitert ihre Reichweite daher anhand einer Inferenzkette und Methodentriangulation (vgl. Wollmann 2000: 209). Der Hergang der Untersuchung ist dreistufig aufgebaut. Untersucht werden (1) einzelne Behörden (z.B. Finanzamt XY), die (2) einem konkreten, bei der Digitalisierung weiter fortgeschrittenem Verwaltungsbereich zugehören (z.B. Steuer- oder Bauverwaltung). Durch die Verbindung von qualitativen Fallstudien in mehreren Behörden des gleichen Verwaltungsbereichs mit quantitativer Erhebung sollen valide Aussagen für den jeweiligen Verwaltungsbereich ermöglicht werden. Ein Inferenzschluss auf (3) die größere Verwaltungslandschaft in Deutschland kann daraus methodologisch nicht erfolgen, die Auswahl und der Vergleich der untersuchten Verwaltungsbereiche liefert aber wichtige Trendergebnisse.

Auswahl der Unter-
suchungsbereiche

Bei der Auswahl der Untersuchungsbereiche bietet es sich an, auf fortgeschrittene digitale Verwaltungsbereiche zurückzugreifen, in denen schon langjährige Erfahrungen mit Verwaltungsdigitalisierung und E-Government-Leistungen vorhanden sind und die eine Relevanz für eine größere Anzahl von Bürgern haben. Hierzu werden die kommunale Bauaufsicht, die Kfz-Zulassung, die Steuerverwaltung und Elternleistungen rund um die Geburt eines Kindes betrachtet. Diese Verwaltungsbereiche sind durch eine, für Deutschland sehr fortgeschrittene digitale Reife gekennzeichnet, um Veränderungen durch Digitalisierung überhaupt feststellen zu können. Es handelt

sich also methodologisch um eine konzeptionell reflektierte Fallauswahl (*purposeful selection*) (Wollmann 2000: 208). Auch wenn dementsprechend kein echtes quasiexperimentelles Verfahren möglich ist, wird die Überprüfbarkeit des Zusammenhangs zwischen den untersuchten Variablen durch den Vergleich der Verwaltungsbereiche gestärkt. Während die vier Verwaltungsbereiche mit Blick auf die abhängige Variable ähnlich gehalten sind (vergleichsweise hohe digitale Reife), sind sie so ausgewählt, dass sich eine Varianz im Hinblick auf die Unabhängige ergibt. Wie in Tabelle 4 aufgeführt, unterscheiden sich die Implementierungsprozesse im Hinblick auf Steuerung und Umsetzung der Digitalisierungsprogramme.

Tabelle 4: Einordnung der Fallstudien

Umsetzungsdimension	Steuerungsdimension	
	zentral	dezentral
zentral	ELSTER	/
dezentral	i-Kfz	Digitaler Bauantrag, Elternleistungen

(Quelle: eigene Darstellung)

Insgesamt stützt sich diese Studie auf einen Methodenmix von qualitativen Fallstudien und einem quantitativen Fallvergleich im Wege einer Verwaltungsumfrage. In den Fallstudien geht es darum, Einsichten in die Natur jeder der digitalen Umwelten und Informationssysteme zu erhalten und die Auswirkungen auf Organisation, Arbeit und Personal, Leistung sowie Mitbestimmung zu erheben. Für die Datenerhebung wurden dazu umfangreich Dokumente und bisherige Forschung ausgewertet. Vor allem stützen sich die Fallstudien jedoch auf die Erfahrungen von einer Vielzahl an Experten aus den Verwaltungsbereichen, welche in qualitativen Experteninterviews (Meuser/Nagel 2009; Bogner et al. 2009) erhoben wurden.

Insgesamt wurden 70 Experteninterviews zwischen Januar 2021 und Juni 2022 mit Leitungen und Beschäftigten aus den entsprechenden Behörden sowie mit Ministerialverantwortlichen und Stakeholdern geführt (ausführlich siehe Tabelle 5). Als Experten der Implementationsebene wurden Sachbearbeitende, Amtsleitungen sowie Personalräte und IT- bzw. Digitalisierungsverantwortliche befragt, während steuerungspolitische Kontexte auch mit Ministerialvertretern erörtert wurden. Aus methodologischer Sicht ist diese Erhebung vielfachiger Perspektiven sowohl in mehreren Behörden als auch durch mehrere Personen in diesen Behörden relevant, um die Repräsentativität der Experteninterviews für die Fallstudien zu gewährleisten (Enticott

Datenerhebung
mittels Experten-
interviews

2004). Die Auswahl der Ämter erfolgte dabei nach theoretischen Überlegungen und der erreichten Digitalisierungsreife, gleichzeitig war eine gewisse Forschungspragmatik notwendig, um Feldzugang zu erreichen.

Die Interviews waren halb-standardisiert und wurden mithilfe von Interviewleitfäden durchgeführt. Diese setzten sich aus acht Blöcken von Fragen zusammen, die auf Basis der theoretischen Vorannahmen formuliert wurden. Die Leitfäden wurden dabei auch entsprechend der jeweiligen Position der/des Interviewten angepasst, indem z.B. der Block zu partizipationsbezogenen Fragen bei Personalräten umfangreicher behandelt wurde. Des Weiteren sollten die Leitfäden im Sinne der qualitativen Gegenstandsoffenheit auch Platz für narrative Passagen bieten, um ggf. neue, für die Fragestellung relevante Themen und Aspekte mitaufzunehmen (vgl. Meuser/Nagel 2009). Solche Aspekte wurden dann teilweise in weiteren Interviews übernommen und abgefragt. Der zeitliche Umfang der Interviews betrug zwischen 30 und 90 Minuten. Sie wurden überwiegend telefonisch oder per Videokonferenz geführt. In allen Untersuchungsfeldern wurden jedoch auch Interviews vor Ort in den Behörden geführt, um einen vertiefteren Eindruck zu erhalten.

Die Interviews wurden in der Mehrzahl aufgezeichnet, außer Interviewpartner waren mit einer Aufzeichnung nicht einverstanden oder wenn die Interviewsituation im Umfeld des Kundengeschäfts stattfand. Die Interviewtranskripte bzw. -protokolle wurden anschließend im Wege einer qualitativen Inhaltsanalyse kodiert und ausgewertet.

Tabelle 5: Überblick über durchgeführte Interviews

	Zahl untersuchter- Behörden	Sachbearbeitende	Amtsleitungen	Personalvertretende	IT und Digitalisierung sowie Datenschutz	Ministerien	Ausbildungs- management	Externe Stakeholder
Kfz-Zulassung (20 Interviews mit 23 Befragten)	7	7	7	3	4	2	0	0
Bauaufsicht (19 Interviews mit 19 Befragten)	9	6	6	2	3	0	0	1
Steuerverwaltung (17 Interviews mit 22 Befragten)	9	7	3	3	2	6	1	0
Elternleistungen (14 Interviews mit 15 Befragten)	8	1	8	3	1	1	0	0

(Quelle: eigene Darstellung)

Die Ergebnisse der qualitativen Fallstudien in den in Tabelle 5 aufgeführten Behörden wurden mittels einer quantitativen Verwaltungsbefragung ergänzt. Ziel dieser zusätzlichen Untersuchung ist die Prüfung der Ergebnisse der Einzelfallstudien durch eine Vollerhebung aller Bauverwaltungen, Kfz-Zulassungen und Finanzämter in Deutschland. Durchgeführt wurde die Umfrage in den Untersuchungsbereichen Kfz-Zulassung, Bauverwaltung und Steuerverwaltung. Die Elternleistungen als vierte Fallstudie wurden bei der Verwaltungsbefragung ausgenommen, da es in Deutschland bislang nur sehr wenige Angebote von Kombi-Anträgen für Elternleistungen im Umfeld der Geburt eines Kindes gibt und daher keine Ergebnisse in der Fläche gesammelt werden können. Zielgruppen der Befragung waren dabei einerseits Beschäftigte dieser Verwaltungsbereiche, andererseits Personalratsmitglieder.

Die Verwaltungsbefragung wurde online mithilfe der Umfrage-Software „Unipark“ durchgeführt. Dazu wurde auf Basis der qualitativen Studienergebnisse ein umfassender Fragebogen entwickelt, der zugleich mithilfe von Filterfragen die Möglichkeit spezifischerer Nachfragen bot (z.B. wenn eine E-Akte vorhanden ist, ob zusätzlich eine Papierakte geführt wird). Auch konnte so ermöglicht werden, dass nur die jeweils für die Verwaltungsbereiche sowie Positionen der jeweiligen Umfrageteilnehmenden relevanten Fragen gestellt wurden. Die Umfrage wurde an Personen aus der Zielgruppe in einem Pilotversuch getestet, um die Verständlichkeit der Fragen sicherzustellen.

Um Zugang zur Umfrage zu erhalten, wurde ein Link bzw. QR-Code verschickt. Dieser Link wurde dazu an alle unteren Bauaufsichtsbehörden und alle Kfz-Zulassungsbehörden in Deutschland per E-Mail versendet. Dabei wurde eine Doppelstrategie angewandt, indem einerseits die Fachamtsleitung direkt angeschrieben wurde (falls denn eine Mailadresse oder ein Kontaktformular über die Webseite auffindbar war, ansonsten wurde die allgemeine Infomail der Kommunalverwaltung angeschrieben) und andererseits die Personalräte der Kommunalverwaltung, welche jeweils um Weiterleitung des Links an ihre Beschäftigten gebeten wurden.

In den Finanzämtern gestaltete sich die Ansprache von Beschäftigten ungleich schwieriger. Um die Herausforderungen beim Feldzugang zu der stark hierarchisch geprägten Steuerverwaltung zu bewältigen, wurde die Strategie verfolgt, mithilfe der Unterstützung der Gewerkschaft Ver.di und der Deutschen Steuergewerkschaft über E-Mail-Verteiler, soziale Medien und Printwerbung Beschäftigte und Personalvertreter aus den Finanzämtern aller Bundesländer zu erreichen. Allerdings führte dieses Verfahren auch zu einer deutlichen Überrepräsentation von Steuerbeschäftigten in Baden-Württemberg und Berlin. Um sicherzustellen, dass dies keine verzerrenden Auswirkungen auf die Ergebnisse der quantitativen Analyse hat, wurden statistische Tests durchgeführt, um die Beziehung zwischen dem Bundesland und relevanten

Variablen zumindest selektiv zu untersuchen. So wurden unter anderem für die Variablen „Digitalisierung führt zu Arbeiterleichterung“ und „Die Personalausstattung ist angemessen“, mit denen lineare Regressionen gerechnet wurden, der Exakte Fischer-Test angewandt. Die Ergebnisse zeigten keine signifikanten Zusammenhänge zwischen den untersuchten Variablen und dem Bundesland der Umfrageteilnehmenden. Davon ausgehend wurde auf die Anwendung von Gewichtungungsverfahren verzichtet.

Rücklauf

Insgesamt nahmen über den Umfragezeitraum von Januar bis März (Kommunen) sowie Mai bis Juli 2023 (Steuer) 1.752 Personen über den Online-Link teil.⁵ Hiervon waren 1.151 Beschäftigte der Steuerverwaltung, 271 in einer unteren Bauaufsicht beschäftigt und 155 Angehörige einer Kfz-Zulassung. Hinzukommen 174 sonstige kommunale Personalratsmitglieder. Die Teilnahme an der Umfrage erfolgte anonym, so dass auch die Angabe des Namens des zugehörigen Finanzamts bzw. der zugehörigen Kommune freiwillig war. Insoweit jedoch eine Identifizierung oder zweifelsfreie Unterscheidung der Ämter möglich war, lassen sich Aussagen über die Rücklaufquote schließen: Demnach haben mindestens 34 % der Finanzämter (160 von 473 Ämtern), 29 % der Kfz-Zulassungen (118 von ca. 394 Ämtern) und 21 % aller Bauaufsichtsbehörden (197 von 933 Ämtern) an der Befragung teilgenommen. Der Rücklauf wird damit als gut eingeschätzt.

Es ist jedoch wichtig zu betonen, dass die Umfrage keine statistische Repräsentativität beansprucht. Die Relevanz der Daten ergibt sich daraus, dass sie Einblicke in konkrete Kontexte und Erfahrungen bieten. In Kombination mit den qualitativen Interviewdaten ist es so möglich, spezifische Phänomene in Tiefe zu betrachten und zu kontextualisieren, sodass trotz der begrenzten Repräsentativität wertvolle Erkenntnisse gewonnen werden können. Im Rahmen der Verwaltungsbefragung stützen wir uns größtenteils auf subjektive Einschätzungen der Beschäftigten. Ziel der Umfrage ist es nicht, „objektive“ Messungen vorzunehmen, sondern die Ergebnisse unserer Analyse spiegeln dezidiert die Wahrnehmung der Befragten auf Vollzugebene wider.

5 Dies beinhaltet nur Teilnahmen, bei denen der Fragebogen bis zum Ende ausgefüllt wurde. Abgebrochene Umfragen wurden nicht verarbeitet; die Beendigung der Umfrage nach Unterbrechung war aber möglich. Gleichzeitig wurde über Browser-Cookies sichergestellt, dass Personen (über den gleichen Internetzugang) nicht mehrfach an der Befragung teilnehmen konnten.

Tabelle 6: Rücklauf nach Bundesländern

Bundesland (Teilnehmer N)	Teilgenommene Beschäftigte nach Verwaltungsbereich		
	Kfz (N = 155)	Bauaufsicht (N = 271)	Steuer (N = 1.151)
<i>Baden-Württemberg</i>	13 (8,4 %)	48 (18 %)	857 (74 %)
<i>Bayern</i>	28 (18 %)	39 (14 %)	18 (1,6 %)
<i>Berlin</i>	0	8 (3 %)	152 (13 %)
<i>Brandenburg</i>	4 (2,6 %)	20 (7,4 %)	27 (2,3 %)
<i>Bremen</i>	1 (0,6 %)	1 (0,4 %)	0
<i>Hamburg</i>	0	12 (4,5 %)	1 (<0,1 %)
<i>Hessen</i>	15 (9,7 %)	13 (4,8 %)	8 (0,7 %)
<i>Mecklenburg- Vorpommern</i>	0	2 (0,7 %)	0
<i>Niedersachsen</i>	17 (11 %)	27 (10 %)	23 (2,0 %)
<i>Nordrhein-Westfalen</i>	24 (15 %)	58 (21 %)	28 (2,4 %)
<i>Rheinland-Pfalz</i>	4 (2,6 %)	8 (3,0 %)	7 (0,6 %)
<i>Saarland</i>	7 (4,5 %)	6 (2,2 %)	11 (1,0 %)
<i>Sachsen</i>	24 (15 %)	9 (3,3 %)	7 (0,6 %)
<i>Sachsen-Anhalt</i>	8 (5,2 %)	4 (1,5 %)	3 (0,3 %)
<i>Schleswig-Holstein</i>	3 (1,9 %)	11 (4,1 %)	8 (0,7 %)
<i>Thüringen</i>	7 (4,5 %)	5 (1,8 %)	1 (<0,1 %)

(Quelle: eigene Zusammenstellung)

