

4. Operationalisierte sozialethische Bewertung

4.1 Die moralische Situation mit Blick auf Fluggastkontrollen

Die Entwicklung der drei Typiken im letzten Kapitel habe ich mit der Absicht unternommen, diejenigen sozialethischen Wertungskonflikte zu identifizieren und zu strukturieren, mit denen bei der Gewährleistung öffentlicher Sicherheit durch Techniken der Fluggastkontrollen typischerweise gerechnet werden muss. Zusammen mit der Skizze der habitualisierten Denk- und Handlungsweisen in diesem Bereich, die ich in Kapitel 2 ausgearbeitet habe, ergibt sich dabei eine ‚moralische Landkarte‘, welche eine grobe Orientierung zu derartigen (nicht-ökonomischen) Folgeproblemen leisten kann. Grundannahme hinter dieser Ausarbeitung war, dass die Entwicklung und der Einsatz von diesen Techniken als moralisch relevantes Handlungsfeld gelten kann. Wie ich in Abschnitt 1.1.1 ausgeführt habe, impliziert dies die Annahme, dass über Entscheidungsprozesse in Politik und Gesellschaft Techniken der Fluggastkontrolle in wesentlichen Aspekten vor einem Wertungshorizont und mit Blick auf die Zukunft *gestaltet werden können*. Politische Debatten zur Ausgestaltung der Kontrollprozesse (sei es in Fachausschüssen oder auch in einer breiteren Öffentlichkeit), wie auch Entscheidungen vor Ort zur konkreten Implementierung sind demzufolge zumindest grundsätzlich einer moralphilosophisch argumentierenden Beratung zugänglich. In diesem Sinne stand hinter den letzten beiden Kapiteln auch die Absicht, Fluggastkontrollen als ein gesellschaftspolitisches Problem technischen Sicherheitshandelns zu begreifen und die für eine moralpragmatisch argumentierende, innovative Technikbewertung zentralen sozialethischen Problemkonstellationen zu beschreiben.

Über meinen moralpragmatischen Ansatz bin ich dabei, übereinstimmend mit neueren TA-Ansätzen,¹³⁹ von sogenannten Betroffenheitsöffentlichkeiten ausgegangen, nämlich der Gruppe aller Passagiere (*flying public*). Wie in Kapitel 2 deutlich wurde, sind dies aber keineswegs die einzigen relevanten Stakeholder; und neben den in Kapitel 3 entwickelten sozialethischen Risikotypen sind bei Entscheidungen zur Ausgestaltung der Kontrollprozesse durchaus auch noch andere Wertungen relevant. Insbesondere der durch die Kontrollen beabsichtigte Schutz vor Angriffen sowie ökonomische Interessen müssen hier genannt werden. Doch kann und soll eine über Kapitel 2 hinausgehende Betrachtung dieser Wertungen hier

¹³⁹ Dies gilt zwar grundsätzlich auch für die Ansätze der partizipativen TA, soll an dieser Stelle aber stärker im Sinne der in Abschnitt 1.1.2 formulierten Forderung nach einem alternativen Ansatz in der Technikfolgenforschung verstanden werden: „Weg von der Abschätzung und Bewertung einer Technik als solche hin zu einer Gesellschaftsanalyse – ‘the proper point of departure is not to assess technology but to assess society’“ (PETERMANN 1999: 26).

nicht geleistet werden. Folglich muss die moralische Landkarte im Kontext einer noch weiter gefassten moralischen Entscheidungssituation gesehen werden.¹⁴⁰

Dabei ist es wichtig nochmals hervorzuheben, dass sich weitere relevante Wertungen in ähnlicher Weise um die drei eigenständigen Problemtypiken herum *gruppieren*, aber nicht zu einer einzigen, umfassenden Wertungstypik *hierarchisch integrieren* lassen. Die Integration der drei Typiken sozialethischer Probleme, wie auch eine noch weitergehende Integration zu einer allgemeinen Typik, die auch ökonomische und schutzbezogene Wertungen umfassen würde, ist nicht möglich, weil die verschiedenen relevanten Wertungen zueinander inkommensurabel sind: In der konkreten Entscheidungssituation treten diese Wertungen in je unterschiedliche Sinnzusammenhänge und werden zudem in ähnlichen Entscheidungssituationen auch nicht unter einem gemeinsamen normativen Konzept diskutiert.

Eine informierte, abwägende Entscheidung zu einer zukünftigen Ausgestaltung von Fluggastkontrollen wird, so die Hypothese, neben Fragen des Schutzes und neben wirtschaftlicher Überlegungen daher auch die oben identifizierten sozialethischen Risiken einbeziehen müssen. Geschieht dies nicht, so werden sich Konfliktkonstellationen manifestieren, die bei der Entscheidung nicht berücksichtigt wurden, die sich aber über die ausgearbeiteten Kategorien ‚Eingriff in die Privatheit‘, ‚diskriminierende Effekte‘ sowie ‚Restriktion gesellschaftlicher Offenheit‘ hermeneutisch erschließen lassen. Über diese moralische Landkarte lassen sich demnach „begründete Zukunftserwartungen“ formulieren, die für „jedermann nachvollziehbar sind“ (GRUNWALD 1999: 75). Manifestierten sich diese erwarteten Konfliktkonstellationen hingegen wiederkehrend nicht oder traten typischerweise weitere, nicht berücksichtigte sozialethische Wertungskonflikte auf, die direkt mit der Gewährleistung öffentlicher Sicherheit durch Fluggastkontrollen zusammenhängen, so erwiese sich die moralische Landkarte als verbessерungsbedürftig. Hierin konkretisiert sich demnach ein Kriterium der moralpragmatischen Tauglichkeit für die drei oben entwickelten Typiken.

Damit diese Typiken aber tatsächlich als Planungsgrößen für kollektives zukunftsbezogenes Handeln fungieren können, wie Grundwald (1999: 75) es für eine gelungene philosophische TA fordert, bedarf es noch eines weiteren methodischen Schritts. Denn damit die Wertungskonflikte nicht nur in ihrem Bedeutungsgehalt erschlossen, sondern tatsächlich in einer abwägenden Entscheidung sinnvoll berücksichtigt werden können, muss zudem auch eingeschätzt werden können, wie stark sich diese Wertungskonflikte für verschiedene Kontroll-Optionen jeweils im Vergleich manifestieren. Ganz im Sinne Grunwalds (1999: 77) müssen also an die Entwicklung der drei Typiken außerdem explizite Über-

¹⁴⁰ Wie bereits erwähnt ist das in diesem Kapitel vorgestellte Framework in einer früheren Form in dem deutlich breiter aufgestellten EU-Forschungsprojekt XP-DITE entwickelt worden, bei dem diese beiden Wertungsdimensionen über eigenständige Bewertungsmodelle ebenfalls berücksichtigt werden.

legungen zur Operationalisierung¹⁴¹ angeschlossen werden, wenn diese in tatsächlichen Entscheidungssituationen eine praxisnahe Hilfestellung bieten sollen. Dabei geht es, wie in Abschnitt 1.1.2 ausgeführt, nicht um eine Vorhersage der Zukunft – etwa in Form der Frage, wie heftig öffentliche oder mediale Proteste zu diskriminierenden Effekten ausfallen werden –, sondern um eine vorausgreifende Bearbeitung gesellschaftspolitischer Wertungskonflikte und gegebenenfalls um die vorausschauende Entwicklung von Kompromisslösungen (BAYERTZ 2008: 167; MESSNER 2012).

Angesichts der zum Teil reflexartigen Antworten auf neue Bedrohungen in der zivilen Luftfahrt läge hierin ein konkreter gesellschaftspolitischer Gewinn: Weder müsste eine von vornherein ablehnende Haltung hinsichtlich der staatlichen Gewährleistung von Sicherheit eingenommen werden, noch müssten Sicherheitsversprechen unkritisch übernommen werden. Vielmehr könnten verschiedene Optionen hinsichtlich absehbarer wertungsrelevanter Folgen miteinander verglichen werden, wobei auch die in Abschnitt 1.1.2 aufgeworfene Frage berücksichtigt werden könnte, „welchen Preis – in Form von Geld, Freiheit, Gerechtigkeit oder Privatheit ... wir bereit [sind] für den Wert ‚Sicherheit‘ zu bezahlen“ (AMMICH QUINN & RAMPP 2009: 143). Erst in diesem weitergehenden Kontext kann von einer gelingenden situationsbezogenen Selbstverständigung die Rede sein – und erst zusammen mit situativen Operationalisierungen kommt diese hermeneutische Aufgabe letztlich zur vollen, produktiven Entfaltung. Denn soll die hier geleistete moralphilosophische Expertise tatsächlich im oben ausgeführten Sinne *angewandte* Ethik sein, so darf sich diese Figur der hermeneutischen Selbstverständigung nicht nur über die relevanten Absichten und Wertungen erstrecken; stattdessen müssen die angebotenen „intellektuellen Werkzeuge“ (DEWEY 1984a: 347) auch die konkreten gesellschaftspolitischen *Konsequenzen* umfassen.

Wenn es richtig ist, dass intelligente politische Entscheidungen letztlich auf einer vergleichenden Deliberation von Handlungsoptionen beruhen (DEWEY 1984a: 304), so lässt sich hierüber die Forderung nach einer situationsadäquaten Operationalisierung noch genauer spezifizieren. Denn der Bedarf an einer Operationalisierung besteht dann nicht in einer bemessenden Zurechnung zu universell anwendbaren Skalen – so als ließe sich für den Eingriff in private Lebensbereiche ein für alle Entscheidungssituationen übertragbares Maß analog zur räumlichen Ausdehnung definieren, das eine Messung wie mit einem Zollstock erlauben würde. Vielmehr geht es um die Gewährleistung einer *vergleichenden Einschätzbarkeit* verschiedener Handlungsoptionen, das heißt: verschiedener Prozesse der Fluggastkontrollen. Verglichen werden sollen die Prozesse dabei hinsichtlich

¹⁴¹ Unter ‚Operationalisierung‘ wird gemeinhin eine ‚Messbar-Machung‘ verstanden. Mit Blick auf eine operationalisierte Bewertung ist für mich dabei zentral, dass die Bewertung intersubjektiv nachvollziehbar ist, die Bewertungsergebnisse durch andere replizierbar sind, sowie dass für die Bewertung mit Blick auf das Erkenntnisinteresse dieser Arbeit pragmatische Tauglichkeitskriterien formuliert werden können.

der in Kapitel 3 ausgearbeiteten Problemtypen, die bereits so formuliert wurden, dass sie im Sinne eines graduellen ‚Mehr-oder-weniger‘ aufgefasst werden können. Die moralphilosophische Expertise, die hierbei entwickelt werden soll, zielt also darauf ab, verschiedene Möglichkeiten der Ausgestaltung von Fluggastkontrollprozessen hinsichtlich der einzelnen sozialethischen Problemkonstellationen zu vergleichen – ob also der Entwurf eines spezifischen Kontrollprozesses im Vergleich zu einem anderen Kontrollprozess etwa den Eingriff in die Privatheit der Passagiere deutlich verschärft, ob er deren private Lebensbereiche deutlich besser respektiert, oder ob sich das Problem bei beiden Optionen in etwa gleicher Schärfe manifestieren wird.

Aus dieser Überlegung heraus lässt sich die Forderung nach einer situationsadäquaten Operationalisierung noch in zwei weiteren Hinsichten spezifizieren. Zum einen wird deutlich, dass eine Vergleichbarkeit über alle bloß *denkbaren* Handlungsoptionen hinweg keineswegs notwendig und vermutlich sogar kontraproduktiv ist. Vielmehr geht es um die Vergleichbarkeit *realistischer* Handlungsoptionen, also um den Vergleich von Kontrollprozessen, die in naher bis mittlerer Zukunft tatsächlich an Flughäfen eingesetzt werden können. In Abschnitt 1.1.2 hatte ich bereits die in diesem Kontext in der TA verbreitete Unterscheidung von absehbaren ‚*Best-case*‘- und ‚*Worst-case*‘-Alternativen erwähnt. Zwar ist eine solche Zweiteilung für einen aussagekräftigen Vergleich aller Voraussicht nach zu grob, doch lässt sich über diese das Spektrum der vergleichenden Abwägung recht gut einschränken.

Sucht man nach Alternativen, die eine weitergehende Differenzierung zulassen, so wird zum anderen deutlich, dass die ebenfalls bereits erwähnten und in der TA recht verbreiteten Ampelmodelle eventuell ein Stück weit irreführend sind. So lässt sich über die Bewertungsstufen grün, gelb, rot zwar ein ‚besser‘ und ‚schlechter‘ in drei Schritten ausdrücken (bzw. in vier über die häufig genutzte Zwischenstufe orange), doch schwingt hier als loses Analogon zur Ampel zu meist in etwa der Bedeutungsgehalt ‚problemlos‘, ‚problembehaftet‘, ‚(kritisch)‘ bzw. ‚unzulässig‘ mit.¹⁴² Grundsätzlich scheinen solche mehrfach abgestuften Vergleichsbewertungen, wie sich in den nächsten Abschnitten noch verdeutlichen wird, durchaus für die Absicht dieser Arbeit geeignet, doch sollte selbst ein *Best-case*-Szenario hinsichtlich des Eingriffs in die Privatheit nicht darüber hinwegtäuschen, dass aufgrund der Grundfunktion der Aufdeckung auch hier ein grundsätzlicher Wertungskonflikt besteht. Ein solches ‚ethisches Framework‘ zur operationalisierten sozialethischen Bewertung realistischer Prozesse der Fluggastkontrollen im Sinne der vergleichenden, abgestuften Einschätzung soll im Folgenden vorgestellt werden.

¹⁴² Ein Beispiel hierfür wäre etwa der im EU-Projekt SURVEILLE entwickelte Ansatz zur Unterstützung der Entscheidungsfindung, der über eine dreistufige Ampel die Bewertungsstufen ‚use‘, ‚revise and reassess‘ und ‚reject‘ nutzt (SURVEILLE 2015).

4.2 Ansatz zur Bewertung der sozialethischen Risiken

4.2.1 Der Operationalisierungsansatz der Wertbaumanalyse

In Kapitel 3 hatte ich darauf verwiesen, dass im Rahmen der Wertbaumanalyse (KEENEY u. a. 1984) bereits die Idee entwickelt wird, an ausgearbeitete Typiken von relevanten Wert(ung)en eine Operationalisierung anzuschließen. Das Thema, in dessen Rahmen die Wertbaumanalyse entwickelt wurde, ist dabei die Energiepolitik Deutschlands in den 80er Jahren – das heißt der damalige politische Konflikt darum, wie die Energieversorgung künftig gestaltet werden solle. In ihrem Ausblick skizzieren die Autoren leider nur sehr grob, wie eine operationalisierte Bewertung auf Grundlage der Wertbäume konkret aussehen könnte:

„Der zusammengefaßte Wertbaum stellt eine strukturierte Liste der Kriterien dar, mit denen Energiesysteme verglichen und bewertet werden können ... Eine Bewertung mit Hilfe der Kriterien kann grundsätzlich auf zwei verschiedene Arten durchgeführt werden. Im ersten Fall wird eine Skala gefunden oder konstruiert werden, um die Kriterien ... zu messen und dann die mögliche Auswirkung jedes Energiesystems auf den Skalen abzuschätzen. Die Abschätzungen können deterministisch oder probabilistisch sein. Sie können aufgrund technisch-wissenschaftlicher Modelle, aufgrund physischer Messungen oder aufgrund von Erfahrungswerten von professionellen Eliten zustandekommen. Im zweiten Fall werden Energiesysteme lediglich in eine Rangordnung entsprechend der Abschätzung, wie stark jedes Kriterium erfüllt ist, gebracht. Für diese Aufgabe können wiederum wissenschaftlich verfügbare Modelle, Daten oder Experten-Meinungen benutzt werden. Diese Methode erfordert keine explizite Meßskala und eine Quantifizierung jedes Indikators.“ (KEENEY u. a. 1984: 84)

Anders ausgedrückt schlagen die Autoren der Wertbaumanalyse also folgendes Vorgehen vor: Die unterste Ebene des ‚Wertbaums‘ für Energiesysteme kann als Kriterienkatalog zur Bewertung genutzt werden. Dieser Katalog würde zum Beispiel die somit als wertungsrelevant definierten Kriterien „Belastung der Flora und Fauna in der regionalen Umwelt“ sowie „Versorgungssicherheit durch technische Störanfälligkeit“ enthalten (KEENEY u. a. 1984: 68). In der ersten vorgeschlagenen Variante würde versucht, alle diese Kriterien anhand naturwissenschaftlicher Kriterien zu quantifizieren. Ein einfacher Vorschlag könnte hier für das Kriterium der Versorgungssicherheit etwa sein, auf einer Skala von 0-365 zu bestimmen, an wie vielen Tagen im Jahr durchschnittlich mit einem Stromausfall bei Systemen gerechnet werden muss, die auf Kernenergie zurückgreifen, und an wie vielen Tagen bei solchen, die auf Solarenergie setzen. In der zweiten vorgeschlagenen Variante würde das Vorkommen von Stromausfällen stattdessen einer zuvor genauer definierten Rangordnung zugeordnet, zum Beispiel auf der fünfstufigen Skala ‚fast nie‘, ‚selten‘, ‚gelegentlich‘, ‚häufig‘ und ‚ständig‘.

In einem zweiten Schritt soll dann für beide Varianten eine „Wertabwägung“ zwischen den einzelnen Kriterien eingebracht werden, die zum Ausdruck bringen soll, „wieviel jemand bereit ist, bei einem Kriterium aufzugeben, um im Austausch dafür ein ‚mehr‘ bei einem anderen Kriterium zu erreichen (Trade Offs).

Vereinfacht stellt sich die Frage, welche Kriterien wichtiger sind als andere und um wieviel.“ So ließe sich dann die „insgesamt ‚wertverträglichste‘ Option“ herausfiltern (KEENEY u. a. 1984: 84).

Ganz offenbar gibt es bei dieser sehr kurz gehaltenen Skizze einige konzeptuelle Unklarheiten. So scheint der erste Schritt der Zuordnung zu Messskalen noch vergleichsweise unproblematisch – zumindest solange unstrittig ist, wie sich das wert(ungs)relevante Kriterium konkret graduell bestimmen lässt. Denn dass in einer Situation, in der man eine möglichst hohe Ausfallsicherheit bei der Stromversorgung als wünschenswert identifizieren kann, ein Energiesystem als insgesamt bessere Option bewertet werden sollte, das – *ceteris paribus* – im Vergleich zu anderen Optionen seltener ausfällt, scheint eindeutig. Die Frage, wie das wertungsrelevante Kriterium tatsächlich jeweils epistemisch adäquat bestimmt werden kann, ist freilich problematischer, als es in der Skizze zunächst den Anschein hat: Wie bestimmt man die Ausfallsicherheit am besten, über Tage oder über Stunden? Berücksichtigt man, wie viele Haushalte betroffen sind? Wie sollen die Auswirkungen auf Wirtschaftsbetriebe adäquat einbezogen werden? Aus moralpragmatischer Sicht muss hierbei eingewandt werden, dass sich eine adäquate Skala sicher nicht einfach anhand objektiver Kriterien ‚finden‘ oder aus dem Wertbaum deduzieren lässt. Vielmehr stellen diese Skalen bewusst gewählte Konstrukte dar, die sich mit Blick auf das Erkenntnisinteresse dann als tauglich oder untauglich erweisen können.

Bei genauerer Betrachtung der in der Wertbaumanalyse vorgeschlagenen Wertabwägung wird aber überdeutlich, dass die Autoren in dieser Skizze einer möglichen Operationalisierung von höchst problematischen Annahmen ausgehen. Dies betrifft vor allem die implizite Voraussetzung, dass sich gesellschafts-politische Werte über punktuell erhobene Präferenzen abbilden lassen, weshalb sie dann auch davon ausgehen, dass die bestmögliche Entscheidung entsprechend rational vorweggenommen werden kann. Darauf, dass ein derartiger Wertbegriff mit meinem moralpragmatischen Ansatz inkompatibel ist, hatte ich schon in Abschnitt 3.1 hingewiesen.

Wie ich in Abschnitt 1.2.2 dargelegt habe, geht zwar auch Dewey (1988: 203–204) davon aus, dass sich individuelle, situationsgebundene Wertungen indirekt darüber beobachten lassen, welche konkreten Handlungen sie motivieren.¹⁴³ Zudem lassen sich nach Dewey (1988: 205–208) motivierende Wertungen und Interessen mit Blick auf erwartete Konsequenzen durch soziale Organisation auch auf eine kollektive Ebene heben.¹⁴⁴ Doch können motivierende Interessen nicht

¹⁴³ Wie ich in Abschnitt 1.2.2 ausgeführt habe, geht Dewey davon aus, dass eine persönliche Wertschätzung nicht einfach als eine passive Gefühls-Affizierung gefasst werden kann, sondern vielmehr immer auch auf affektiv-motivierende (*affective-motor*) Beweggründe verweist, also auf ein Verlangen oder Interesse, das ‚aktiv‘ zu bestimmten Handlungen motiviert.

¹⁴⁴ Dies wird etwa aus der folgenden oben bereits zitierten Textstelle deutlich: “The word ‚interest‘ suggests in a forcible way the active connection between personal activity and the

einfach als ‚Werte‘ im Sinne ablösbarer und verrechenbarer Entitäten verstanden werden. Die Abwägung von Werten besteht im Moralpragmatismus gerade nicht im Aufrechnen von punktuellen Präferenzen, sondern in einer reflektierenden *Evaluation*. Genauer gesagt erschöpft sich aufgrund der Unmöglichkeit, allen Wertungen gleichermaßen gerecht zu werden, die für eine Handlungsentscheidung nötige *Wertung über Wertungen* nicht in einfachen Gegenüberstellungen, sondern sie muss auf eine *Transformation des Werterlebens* hinauslaufen (DEWEY 1988: 208–209). Die in der moralischen Deliberation sich vollziehende Transformation von affektiv-motivierenden Interessen hin zu affektiv-reflektiert-motivierenden Interessen (DEWEY 1988: 237–238) zielt dabei in letzter Konsequenz auf eine Selbstverständigung darüber, welchen Charakter unsere Gesellschaft letztlich haben soll, und was dies für die vorliegende Entscheidungssituation konkret bedeutet.

Eine Überführung von Einzelwert(ung)en in abstrakt zu verrechnende Präferenzwerte umgeht hingegen genau diesen Selbstverständigungsprozess, der nicht nur für Deweys Moralpragmatismus, sondern auch aus hermeneutischer Sicht essentiell ist. Statt eine *Transformation* des Werterlebens anzustreben, löst eine solche Operationalisierung die je aktuell vorgängigen Wertungen aus den Prozessen des Erlebens heraus und übergibt sie einem abstrakten Kalkül, dessen Ergebnis es sein soll, die ‚wertverträglichste‘ Option als objektiv besten Handlungsplan auszuzeichnen.

Dabei sollen die eigentlichen Wertungskonflikte einerseits rational aufgelöst werden, müssten andererseits aber weiterhin *genau in gleicher Weise* gesellschaftlich wirksam sein. Anderweitig verlöre die rationale Auszeichnung ihre Aktualität – und damit auch ihre argumentative Grundlage, was bedeutet, dass sich das auf den Wertungskonflikten beruhende Ergebnis also sofort selbst revidieren müsste. Ein wirkliches *Verstehen* des zugrundeliegenden moralischen Konflikts kann hierbei also grundsätzlich nicht angestrebt werden, denn hierbei ist eine Veränderung des Werterlebens gerade die Voraussetzung.

In meiner Kritik an der Skizze zur Operationalisierung der Wertbaumanalyse spiegelt sich letztlich jene Kritik, die in der phänomenologischen und hermeneutischen Tradition allenthalben an operationalisierten Zugriffen auf normative Fragen geübt wird. Gerade in den Sozialwissenschaften hinterließ der in diesem Kontext zentrale Methodenstreit zwischen Vertretern qualitativ-normativer An-

conditions that must be taken into account in the theory of valuation ... It points to an activity which takes effect through the mediation of external conditions. When we think, for example, of the interest of any particular group, say the bankers' interest, the trade-union interest, or the interest of a political machine, we think not of mere states of mind but of the group as a pressure group having organized channels in which it directs action to obtain and make secure conditions that will produce specified consequences ... Whenever a person has an interest in something, he has a stake in the course of events and in their final issue – a stake which leads him to take action to bring into existence a particular result rather than some other one" (DEWEY 1988: 205–206).

sätze und quantitativ-empirischer Ansätze tiefe Gräben (JOHNSON & ON-WUEGBUZIE 2004: 14). Erst seit den 1990ern etablieren sich verstärkt sogenannte Multimethodenansätze (*mixed methods research*), die sich oft explizit auf den amerikanischen Pragmatismus als epistemologisches Fundament berufen: “Today, the primary philosophy of mixed methods research is that of pragmatism” (JOHNSON u. a. 2007: 113). Um einen besseren Einblick darin zu gewinnen, wie an die in Kapitel 3 ausgearbeitete Typik ein Ansatz zur Operationalisierung argumentativ angeschlossen werden kann, der die intendierte gesellschaftliche Selbstverständigung unterstützt und weiterführt, statt sie methodisch zu unterlaufen, soll im Folgenden zunächst auf eine reichhaltige Methodendebatte in den Politikwissenschaften verwiesen werden.

4.2.2 Operationalisierungen in der Demokratiemessung¹⁴⁵

Die Zählung von Häufigkeiten und der Ausdruck qualitativer Unterschiede

Für die rationale Gestaltung von Politikinhalten spielt eine wie auch immer gearbeitete Bemessung von wertungsrelevanten Konzepten schon lange eine zentrale Rolle. So dienen beispielsweise Arbeitslosenzahlen in der Politik nicht einfach nur einer allgemeinen, neutralen Beschreibung der Sachlage, sondern sie ‚bemessen‘ ein gesellschaftliches *Problem*, das durch gestaltende Politik adressiert werden soll. Eine hohe Arbeitslosenquote ist *normativ schlechter* als eine niedrige und über sie lässt sich der Erfolg oder das Scheitern von politischen Programmen auch ein Stück weit normativ einschätzen.

Arbeitslosenquoten beruhen dabei in aller Regel auf einem bloß zählenden Vorgehen. Allerdings gibt es viele andere normative Konzepte, die sich ganz offensichtlich nicht adäquat durch die Bestimmung von reinen Häufigkeiten fassen lassen. Ein prominentes Beispiel hierfür ist das Problem der Armut: Viele Bemessungen dieses gesellschaftspolitischen Problems beruhen analog zur Arbeitslosenquote darauf, dass zunächst ein möglichst aussagekräftiges Kriterium als ‚Armutsgrenze‘ definiert wird und sodann für ein Land oder eine Region die jeweiligen Häufigkeiten bestimmt werden (relative Kopfzahl, *head count ratio*). Dabei wird allerdings schon lange kritisiert, dass dieses Kriterium verhältnismäßig wenig über die *Qualität* der Armutsteilung aussagt:

“This ratio, which we shall call the head-count ratio H , is obviously a very crude index ... The measure is ... completely insensitive to the distribution of income among the poor. A pure transfer of income from the poorest poor to those who are better off will either keep H unchanged, or make it go down – surely a perverse response.” (SEN 1976: 219)

¹⁴⁵ Dieser Abschnitt wurde im Rahmen des EU-Projekts XP-DITE in einer kürzeren und deutlich weniger ausgearbeiteten Fassung bereits an anderer Stelle veröffentlicht (VOLKMANN 2013b: 14–17).

Dementsprechend gibt es seit einiger Zeit die Bemühung, theoretische Ansätze zu entwickeln, die qualitative Unterschiede bei normativen Konzepten empirisch erheben und adäquat ausdrücken können. Parallel hat sich zudem auch eine reichhaltige Methodendebatte zu derartigen Ansätzen entwickelt, die oft als ‚normative Messungen‘ bezeichnet werden. Im Lichte dieser Forschungsdebatten zur normativen Messung muss letztlich auch die Skizze zur Operationalisierung der Wertbaumanalyse gesehen werden. Leider wird nicht immer in gebührender Weise auf die in dieser Methodendebatte ausgearbeiteten Qualitätsmerkmale zurückgegriffen (FOSTER 1994: 365), weshalb es sich lohnt, anhand solcher Methodendiskurse zunächst die *Grundidee der normativen Messung* zu entwickeln.

In den Politikwissenschaften hat sich diese Herausforderung einer ‚qualitativ-empirischen Bemessung‘ normativer Konzepte unter anderem hinsichtlich des Konzepts der Demokratie gestellt. Denn die vormals verbreitete, rein zählende Zuordnung zu den Kategorien ‚demokratischer‘ und ‚autoritärer Staat‘ stellte sich im Laufe der Zeit als inadäquat heraus, weil so die zum Teil profunden Unterschiede zwischen den einzelnen Demokratien überdeckt werden:

„In den vergangenen zwei Jahrzehnten entwickelte sich die empirische Demokratieforschung von einem politikwissenschaftlichen Randthema zu einem der wichtigsten Forschungsbereiche der vergleichenden Politikwissenschaft ... Als einer der Auslöser für dieses gestiegene Interesse kann hierbei die Tatsache gelten, dass bestehende Konzepte zwar eine größtenteils robuste Differenzierung von Systemtypen ermöglichen, jedoch nicht die Qualitätsunterschiede abbilden können, die sich bei einer genaueren, qualitativen Betrachtung zwangsläufig feststellen lassen.“ (JÄCKLE u. a. 2012: 100)

„Demokratie“ wird dabei ganz explizit als normatives Konzept verstanden und ‚Demokratisierung‘ als ein grundsätzlich wünschenswerter Prozess. Um aber beispielsweise besser bestimmen zu können, ob in den letzten zehn Jahren bestimmte Staaten, Weltregionen oder auch die Welt im Ganzen demokratischer oder autokratischer geworden sind, müssen Staaten anhand von beobachtbaren Kriterien als mehr oder weniger demokratisch bestimmt werden können. Als Antwort auf diesen Forschungsbedarf sind über die letzten Jahre eine ganze Reihe von Demokratie-Indizes entstanden, die auf das Konzept der normativen Messung zurückgreifen. Zu den bekanntesten zählt hierbei etwa der jährlich erscheinende Freedom House Index. Die jeweiligen Unterschiede in den Ergebnissen der einzelnen Ansätze zur Demokratiemessung befeuerten dann wiederum die Methodendebatten, sodass epistemologische und methodische Fragen mittlerweile zunehmend ins Zentrum dieses Forschungsfeldes rücken (JÄCKLE u. a. 2012: 100).¹⁴⁶

¹⁴⁶ Ich greife in meiner Darstellung vor allem auf Sebastian Jäckle, Uwe Wagschal und Rafael Bauschke (2012), auf Thomas Müller und Susanne Pickel (2007), auf Hans-Joachim Lauth (2007) sowie in besonderer Weise auf Gerardo Munck und Jay Verkuilen (2002) zurück.

Ausarbeitung normativer Konzeptbäume

Obwohl man die Qualität einer Demokratie nicht direkt anhand offensichtlicher Gegebenheiten einschätzen kann, haben die meisten Menschen dennoch eine gewisse Vorstellung davon, was es bedeutet, wenn wir sagen, „In Norwegen sind Politik und Gesellschaft demokratischer als in Südafrika“. Auf ein solches Verständnis bauen Ansätze zur Demokratiemessung auf, wobei ein erster methodischer Schritt in einer begrifflichen Ausarbeitung der Bedeutung von ‚Demokratie‘ in sogenannten Konzeptbäumen (MUNCK & VERKUILEN 2002: 7, 12–14) besteht: Zunächst werden dabei bestimmende Eigenschaften von Demokratien identifiziert und diese anschließend über hierarchische Ebenen semantisch ausdifferenziert. Abb. 7 zeigt einen Ausschnitt davon, wie dies im Falle des Demokratiebarometers gelöst wurde: Die drei bestimmenden Eigenschaften (im Original: *principles*) Freiheit, Kontrolle und Gleichheit werden in jeweils drei relevante Bedeutungsaspekte (*functions*) ausdifferenziert. Für jeden dieser Bedeutungsaspekte werden im Demokratiebarometer dann nochmals je zwei semantische Unteraspekte (*components*) identifiziert, die dann als Bewertungskriterien dienen sollen (BÜHLMANN u. a. 2012).

Ziel solcher semantischen Ausdifferenzierungen in der Demokratiemessung ist es dabei, das abstrakte Konzept ‚Demokratie‘ über mehrere Schritte zu konkretisieren. Für eine Bewertung kann dann zunächst von deutlich weniger abstrakten Bedeutungsgehalten ausgegangen werden. Als ein zentrales Qualitätsmerkmal hat sich hierbei etabliert, dass die semantischen Ausdifferenzierungen im Konzeptbaum möglichst vollständig den relevanten Bedeutungsgehalt von ‚Demokratie‘ beschreiben, zugleich aber auch nicht darüber hinausgehen sollen (MUNCK & VERKUILEN 2002: 7–11). Ansonsten bleiben wichtige Aspekte unberücksichtigt oder es werden Aspekte einbezogen, die gar nichts mit der Demokratiequalität zu tun haben. Zudem muss die Ausdifferenzierung begrifflich möglichst trennscharf funktionieren, denn überlappen sich Unteraspekte, so müssten einzelne Bewertungskriterien im Konzeptbaum mehrfach auftreten – was in den folgenden Bewertungsschritten dann konstant mitberücksichtigt werden müsste, damit Einzelaspekte nicht unversehens übermäßiges Gewicht erhalten (MUNCK & VERKUILEN 2002: 13).

Die methodische Nähe einer solchen Ausdifferenzierung von normativen Begriffen in Konzeptbäumen zu der Erarbeitung von Wertbäumen sowie zu den von mir in Kapitel 3 entwickelten Problemtypiken ist freilich kaum übersehbar. Betrachtet man aber etwas näher, mit welchen Begrifflichkeiten etwa Munck und Verkuilen (2002: 13) die verschiedenen Ebenen in den Konzeptbäumen bezeichnen, so wird deutlich, wie in der normativen Messung der Bedeutungsgehalt normativer Konzepte aufgefasst wird. Verbreitet sind Ebenen-Bezeichnungen wie Attribute, Attribut-Komponenten und Unterkomponenten. Diese Begrifflichkeiten verweisen dabei auf die implizite Vorannahme, dass die Seman-

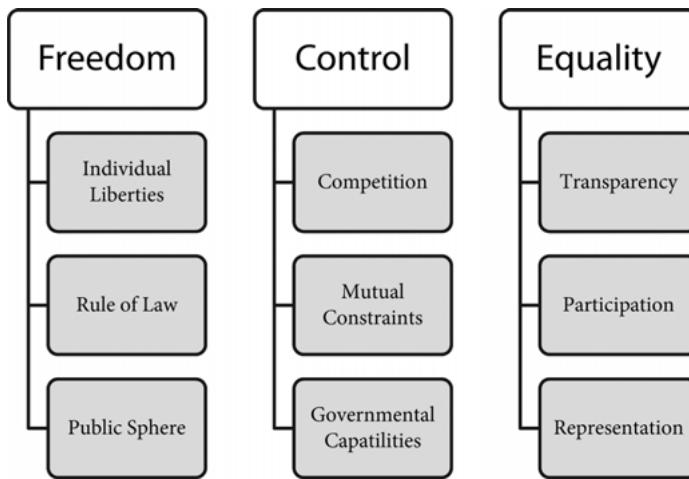


Abb. 7: Eigene Darstellung des Konzeptbaums im Demokratiebarometer (BÜHLMANN u. a. 2012: 523).

tik von abstrakten normativen Begriffen wie ‚Demokratie‘ in immer kleinere Teilbedeutungen auseinandergelegt werden kann. Entsprechend wird die schrittweise Konkretisierung des Begriffs auch als ‚Disaggregation‘ bezeichnet (MUNCK & VERKUILEN 2002: 13).

Die Nutzung dieser offenbar problematischen, recht positivistisch anmutenden Fachbegriffe darf allerdings auch nicht überbewertet werden. Zum einen werden parallel durchaus auch andere Metaphern gebraucht: „Das abstrakte Konzept der Demokratie – versinnbildlicht als Stamm eines Baumes – lässt sich demnach in einzelne Äste aufteilen, die wiederum in Zweige ausdifferenziert werden können“ (JÄCKLE u. a. 2012: 103). Dass solche normativen Konzepte dabei im Anschluss an Deweys Moralpragmatismus auch als Problemtypiken gefasst werden können, habe ich in Kapitel 3 bereits gezeigt. Aus moralpragmatischer Sicht ist die Tauglichkeit solcher semantischen Ausdifferenzierungen freilich nicht von ‚vorfindlichen‘ normativen Entitäten abhängig, sondern sie erweist sich vielmehr mit Blick auf ein an die jeweilige Problemsituation rückgebundenes Erkenntnisinteresse. Zentral für einen Anschluss der Problemtypiken an operationalisierte Bewertungen bleiben aber dennoch die oben erwähnten Qualitätskriterien, nämlich dass die Typiken tatsächlich alle relevanten Aspekte abdecken, dass sie keine nicht-relevanten Problemkonstellationen berücksichtigen sowie dass die Ausdifferenzierung eine trennscharfe Zuordnung von Wertungskonflikten zu Problemtypen ermöglicht. Denn nur dann ist tatsächlich sichergestellt, dass die identifizierten Problemkonstellationen als möglichst konkrete, situativ verortete Orientierungspunkte für die jeweilige Problemkategorie dienen können.

Wie ich in Kapitel 3 bereits ausgeführt habe, müssen alle ausdifferenzierten Teilespekte bei einer normativen Messung zudem grundsätzlich einer Einschät-

zung im Sinne von ‚mehr oder weniger‘ zugänglich sein. Ansonsten könnte eine qualitative Bewertung nicht funktionieren. Was in der positivistisch anmutenden Begrifflichkeit der normativen Messung implizit angelegt ist, durchaus aber expliziert werden sollte, ist dabei, dass wenn ein ‚Mehr‘ bei einem der konkretisierten Teilespekte angenommen wird, dies – *ceteris paribus* – immer auch ein ‚Mehr‘ für das normative Gesamtkonzept bedeutet. Anhand von Abb. 7 lässt sich entsprechend ablesen, dass im Demokratiebarometer angenommen wird, dass ein mehr an politischer Partizipation auch ein mehr an Demokratie bedeutet. Für pragmatische Problemtypiken impliziert dies folglich die Anforderung, dass diese so konstruiert werden, dass eine Verschärfung eines konkretisierten Wertungskonflikts auch als Verschärfung des jeweiligen Problemtypus und dann auch als Verschärfung der übergreifenden Problemkategorie verstanden werden kann.¹⁴⁷

Bewertung der normativen Teilespekte im Konzeptbaum

Der zweite Schritt einer normativen Messung besteht in einer Bewertung der ausdifferenzierten Aspekte des Konzeptbaums. Hierbei werden die Teilbedeutungen auf der konkretesten Ebene als Bewertungskriterien herangezogen. Im Kern geht es bei diesem Schritt um die Zuordnung bestimmter Sachverhalte, die als wertungsrelevant aufgefasst werden, zu Bewertungsskalen – gewissermaßen das Herzstück der eigentlichen operationalisierten Messung. Aus methodischer Sicht sind hierbei drei Teilschritte zentral: (1) Die Identifizierung von beobachtbaren Sachverhalten, die als wertungsrelevant begriffen werden und gut zu den einzelnen Bewertungskriterien passen; sodann (2) die Wahl einer passenden Skala, die qualitative oder quantitative Unterschiede adäquat ausdrücken kann; und schließlich (3) die Wahl passender Zuordnungsregeln, welche für konkrete Fälle die beobachtbaren, wertungsrelevanten Sachverhalte einem Wert auf der Skala zuordnen (MUNCK & VERKUILEN 2002: 15–22). Abb. 8 zeigt diese drei Teilschritte in ihrem logischen Zusammenhang: Für jedes der ausdifferenzierten Bewertungskriterien werden beobachtbare, wertungsrelevante Sachverhalte definiert, die dann zur konkreten Bewertung von Fällen über Zuordnungsregeln auf einer Skala abgetragen werden.

In der Demokratiemessung gehen die meisten Autoren bei der Entwicklung eines Frameworks zur normativen Messung auch gemäß diesem Schema vor. Zunächst sollen zur Bemessung der ausdifferenzierten Bewertungskriterien (*leaves of the concept tree*) beobachtbare Sachverhalte (*indicators*) definiert werden: ‐The first decision in the formation of measures is the selection of indicators that operationalize the leaves of the concept tree. Because there are no hard and fast rules for choosing valid indicators, this is one of the most elusive goals in the so-

¹⁴⁷ Frühe Entwürfe meiner in Kapitel 3 entwickelten Typiken habe ich natürlich bereits mit Blick auf diese Anforderung weiterentwickelt. Meine These ist also, dass diese Anforderung für die drei pragmatischen Problemtypiken bereits erfüllt wird.

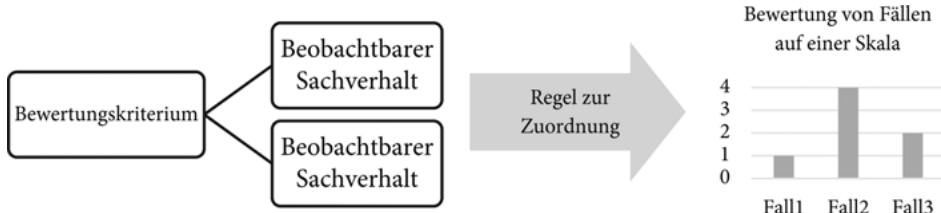


Abb. 8: Methodisches Schema der normativen Messung zur Bewertung von Fällen auf Skalen anhand von Sachverhalten

cial sciences" (MUNCK & VERKUILEN 2002: 15). Hierin liegt entsprechend eine der zentralen Herausforderungen bei der Demokratiemessung.

Für jedes Bewertungskriterium müssen beobachtbare Sachverhalte definiert werden, die über verschiedene Kulturen und Regierungssysteme sowie über möglichst lange Zeiträume hinweg immer in vergleichbarer Art und Weise dieses Kriterium adäquat ausdrücken. Wie Lauth (2007: 234) ausführt, kommt den gewählten beobachtbaren Sachverhalten in der normativen Messung dabei eine Doppelfunktion zu: Einerseits müssen sie den Bedeutungsgehalt des Bewertungskriteriums „valide“ erschließen und dieses andererseits zugleich einer empirischen Untersuchung in unterschiedlichsten kulturellen Kontexten zugänglich machen. Für die Demokratiemessung bedeutet dies freilich eine große methodische Herausforderung, die allerdings nicht immer ausreichend berücksichtigt wird: "Overall, democracy indices reflect insufficient sensitivity to the key issues involved in the choice of indicators" (MUNCK & VERKUILEN 2002: 17).

Diese in der Methodendebatte zur Demokratiemessung umfassend diskutierten Probleme lassen sich aus der epistemologischen Position des Pragmatismus gut nachvollziehen. Denn die theoretische Abstraktion von den konkreten institutionellen, kulturellen und zeitlichen Unterschieden macht es schwierig, konkrete Sachverhalte zu identifizieren, die durchgehend für die jeweils recht unterschiedlichen Demokratien in gleicher Weise relevant sind. Zudem sehen sich Demokratieindizes aus pragmatischer Sicht mit einem weiteren epistemologischen Fallstrick konfrontiert, denn zu welchem Zweck die Bewertung letzten Endes erfolgt, ist keineswegs immer klar. So kann eine Messung der qualitativen Unterschiede in Demokratien dafür eingesetzt werden, anschließend ganz andere Forschungsfragen zu bearbeiten – etwa die Frage, ob demokratischere Strukturen auch eine höhere Zufriedenheit in der Bevölkerung implizieren. Eine solche durchaus gewünschte Weiterverwendung von Demokratieindizes macht es schwer, das jeweilige Erkenntnisinteresse genau zu spezifizieren.

Mit Blick auf eine dezidiert moralpragmatisch ausgerichtete Operationalisierung stellt sich aber gerade *nicht* die (positivistisch anmutende) Frage, ob die beobachtbaren Sachverhalte einen normativen Bedeutungsgehalt „valide“ erschließen, sondern ob die gewählten Sachverhalte *handlungsrelevante Konsequenzen* betreffen, durch die *das konkrete Erkenntnisinteresse* adäquat ausgedrückt wird. Für eine moral-

pragmatisch entwickelte Operationalisierung ist Lauths Beschreibung der Doppel-funktion von beobachtbaren Sachverhalten folglich aus epistemologischen Gründen unpassend. Deren Funktion muss es stattdessen sein, relevante beobachtbare Sachverhalte mit Blick auf genau jene qualitativen Unterschiede zu identifizieren, die durch die Bewertung am Ende auch zum Ausdruck gebracht werden sollen.¹⁴⁸ Ausgangspunkt für eine moralpragmatische Operationalisierung ist also nicht die Identifizierung ‚valider‘ Sachverhalte, sondern die Spezifizierung des Erkenntnisinteresses über eine geeignete Bewertungsskala. Erst über sie lassen sich die handlungsrelevanten Konsequenzen ausmachen, die für das je situative Erleben der auszudrückenden qualitativen Unterschiede auch wirklich relevant sind.

Ich hatte in Abschnitt 4.2.1 bei der Diskussion der Wertbaumanalyse bereits angedeutet, dass es unterschiedliche Typen von Bewertungsskalen gibt. Dabei darf man nicht den Fehler machen, von einem allzu engen Verständnis von ‚Messung‘ auszugehen. Es ist das große Verdienst einer Forschergruppe um den amerikanischen Psychologen Stanley Smith Stevens, maßgeblich zur Klärung dessen beigetragen zu haben, was durch verschiedene Messskalen überhaupt ausgedrückt werden kann (STEVENS 1946).

Gängigerweise verstehen wir unter ‚messen‘ einen Vorgang, bei dem wir von einem absoluten Nullpunkt ausgehen und das zu Messende dann einer Skala mit festen Abständen zuordnen. Ein Beispiel für eine solche sogenannte Verhältnisskala ist das Metermaß, über das wir die räumliche Ausdehnung eines Gegenstands von einem Nullpunkt aus in zueinander immer gleichen Schritten (ein Hundertstel der Standardausdehnung Meter) ‚abmessen‘ können. Gibt es keinen solchen absoluten Nullpunkt, aber klar definierte Abschnitte auf der Skala, welche zueinander in Verhältnis gesetzt werden können – eine Stunde heute und eine Stunde morgen – so spricht man von einer Intervallskala.

Es ist nun wichtig zu bemerken, dass beim Schritt zur Intervallskala nicht mehr die gleichen mathematisch-statistischen Untersuchungen durchgeführt werden können wie noch bei der Verhältnisskala (STEVENS 1946: 678). Dies wird besonders deutlich beim Schritt zur Rangskala¹⁴⁹, die ich für meinen Ansatz nutzen werde. Hier kann lediglich ein Mehr oder Weniger zwischen Abschnitten auf der Skala bestimmt werden; die Abschnitte der Skala können jedoch nicht mehr mathematisch zueinander ins Verhältnis gesetzt werden. Rangskalen werden häufig für psychologische Umfragen genutzt, wenn gefragt wird, ob man eine Aussage als ‚völlig zutreffend‘, ‚zutreffend‘, ‚unzutreffend‘ oder ‚völlig unzutreffend‘ bewerten würde. Hier wäre es nun unzulässig anzunehmen, dass, wenn 25 Personen eine Aussage als ‚völlig zutreffend‘ einschätzen und 25 Personen als ‚unzutreffend‘, die 50 befragten Personen die Aussage im Durchschnitt als ‚zutreffend‘ bewerten.

¹⁴⁸ Dies gilt natürlich nicht nur für die drei in dieser Arbeit relevanten Typiken, sondern kann als allgemeine moralpragmatische Kritik an vielen Ansätzen zur normativen Messung betrachtet werden.

¹⁴⁹ Rangskalen werden oft auch als Ordinalskalen bezeichnet.

Ist auch eine Hierarchisierung nicht mehr möglich, so haben wir es mit einer sogenannten Nominalskala zu tun, über die lediglich die Häufigkeit von Gruppenzuordnungen bestimmt werden kann – etwa ‚Demokratie‘ und ‚Autokratie‘. Allerdings wird eine solche bloß zählende Zuordnung gemeinhin meistens nicht mehr als ‚messen‘ sondern als ‚klassifizieren‘ bezeichnet. Der oben beschriebene Trend in der empirischen Demokratieforschung weg von Klassifikationen und hin zu Demokratieindizes ist – so lässt sich hierdurch besser verstehen – vor allem mit einem Wechsel zu einer aussagekräftigeren Messskala verbunden, zu meist hin zu einer Rang- oder Intervallskala (MUNCK & VERKUILEN 2002: 10). Dabei gibt es auch Ansätze zur Demokratiemessung, die sich auf allen Analyseebenen vollständig auf Rangskalen beschränken. Im Gegensatz zu Ansätzen, die den Grad der Demokratie in Zahlenwerten auszudrücken versuchen, verzichten diese „auf eine Quantifizierung ihrer Befunde ... Wir betreten somit den engeren Bereich der qualitativen Demokratiemessung“ (LAUTH 2007: 280). Eine solche rein qualitativ vorgehende Bewertung werde auch ich anstreben.

Ein Vorteil rein qualitativer Ansätze besteht darin, dass sich hier sehr transparent *argumentieren* lässt, welcher konkrete Rang auf der Bewertungsskala welche qualitativen Unterschiede ausdrücken und daher mit welchen beobachtbaren Sachverhalten korrespondieren soll. Quantitative Ansätze in der Demokratiemessung greifen hierbei hingegen oft auf mathematisch-statistische, das heißt: auf *berechnende* Methoden zurück. Hier ist es zum Teil sehr schwierig zu rechtfertigen, warum konkrete Berechnungsformeln oder Durchschnittswerte auch tatsächlich genau diejenigen qualitativen Unterschiede ausdrücken, die sie laut der gewählten Skala darstellen sollen.

Munck und Verkuilen (2002: 19–21) sehen hier große Defizite bei den aktuellen Demokratieindizes, weil oftmals keine argumentative Begründung für die gewählten mathematischen Berechnungsformeln (*coding rules*) geliefert wird. Zudem sei oft nicht klar, wie der Prozess durch die beteiligten Wissenschaftler konkret durchgeführt wurde, was impliziert, dass die Ergebnisse zum Teil sehr schwer zu reproduzieren sind. Meine in Fußnote 148 angedeutete Kritik an den eher positivistisch anmutenden Ansätzen zur normativen Messung lässt sich an dieser Stelle weiter spezifizieren: Aus pragmatischer Sicht kann es bei einer operationalisierten Bewertung moralischer Konfliktsituationen nicht darum gehen, vorfindliche normative Konzepte semantisch auseinanderzulegen und die begrifflichen Einzelteile anhand von situationsunabhängigen, universell übertragbaren Sachverhalten auf möglichst detaillierten Skalen abzutragen. Vielmehr muss eine Bewertungsskala gefunden werden, welche genau jene qualitativen Unterschiede zum Ausdruck bringen kann, die für das Erkenntnisinteresse relevant sind. Erst anhand dieses konkretisierten Erkenntnisinteresses lassen sich dann situativ relevante Sachverhalte finden, welche diese qualitativen Unterschiede typischerweise im geteilten Erleben ausmachen.

Zusammenfassung der Bewertungen der Teilespekte zu einer Gesamtwertung

Dass eine große Schwierigkeit bei quantitativen Ansätzen darin besteht, mathematische Formeln zu finden, die normative Verhältnisse adäquat ausdrücken, gilt umso mehr auch für den dritten und letzten methodischen Schritt in der Demokratiemessung. Hier sollen die Einzelbewertungen der ausdifferenzierten Teilespekte zu übergreifenden Wertungen zusammengefasst werden. Dieser Schritt ist somit gegenläufig zum ersten Schritt der Ausdifferenzierung eines normativen Konzeptes in einem Konzeptbaum. Entsprechend wird der Schritt in Ansätzen der normativen Messung oft auch als (Re-)Aggregation bezeichnet, welche die Messung über die verschiedenen Verästelungen nach und nach „wieder zurück zum Stamm des Konzeptbaums“ führen soll (JÄCKLE u. a. 2012: 104). Den sich somit ergebenden methodischen Dreischritt für Ansätze der Demokratiemessung stellt Abb. 9 schematisch dar.

Demokratiemessungen, die auf quantitative Methoden zurückgreifen, definieren dabei auch für die Zusammenfassung der Einzelbewertungen mathematische Formeln. Auch hier gilt wieder, dass die Formeln so gewählt werden müssen, dass sie das Verhältnis der normativen Teilespekte zueinander adäquat ausdrücken. Wählt eine Forscherin zur Zusammenfassung der Wertungen auf einer Ebene etwa einen einfachen Durchschnittswert, so würde dies implizieren, dass alle Teilwertungen auf dieser Ebene für die Gesamtwertung gleichermaßen wichtig sind und dass schlechte Teilwertungen problemlos durch gute Teilwertungen ausgeglichen werden können: „Defizite etwa bei der Freiheit können durch höhere Werte bei (sozialer) Gerechtigkeit kompensiert werden“ (JÄCKLE u. a. 2012: 104). Bei multiplikativen Formeln bedingen sich die Teilwertungen hingegen stärker gegenseitig, „sie sind jeweils notwendige Bedingungen für eine Demokratie. Ist der Wert einer Teilkomponente (z.B. Freiheit) ‚null‘ dann ist auch der Gesamtindex null“ (JÄCKLE u. a. 2012: 104). Eine als ‚wichtiger‘ erachtete Teilbewertung kann über die Formeln für die Gesamtwertung zudem mathematisch auch stärker gewichtet werden.

Die implizite Annahme ist hierbei häufig, dass sich die methodischen Regeln für die Zusammenfassung der Teilwerte so definieren lassen, dass sie die Verhältnisse der Teilwertungen untereinander und deren Rolle für das Gesamtkonzept „Demokratie“ adäquat widerspiegeln. Zumeist wird dies spiegelbildlich als Formalisierung dessen verstanden, was im ersten Schritt der Ausdifferenzierung im Konzeptbaum bereits weitestgehend angelegt ist (MUNCK & VERKUILEN 2002: 24; JÄCKLE u. a. 2012: 104).

Wie schon für den letzten Schritt muss diese Sichtweise allerdings aus pragmatischer Perspektive kritisiert werden, denn die Adäquatheit der Aggregationsregeln erweist sich demnach nicht spiegelbildlich zu einer ‚Auseinanderlegung‘ vorfindlicher normativer Konzepte, sondern anhand eines vertieften Verständnisses der Problemsituation und des konkreten Erkenntnisinteresses. Wie schon für die Definition der Bewertungsregeln zuvor, muss also auch für diesen Schritt

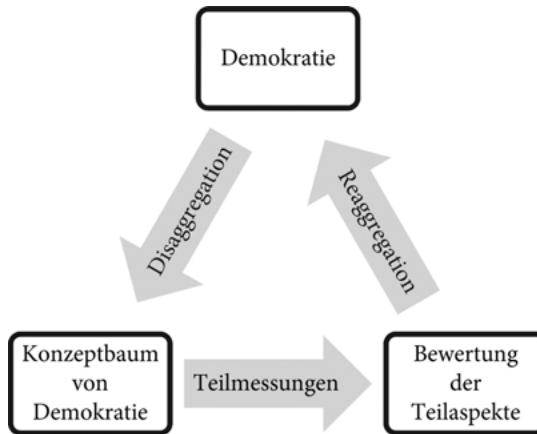


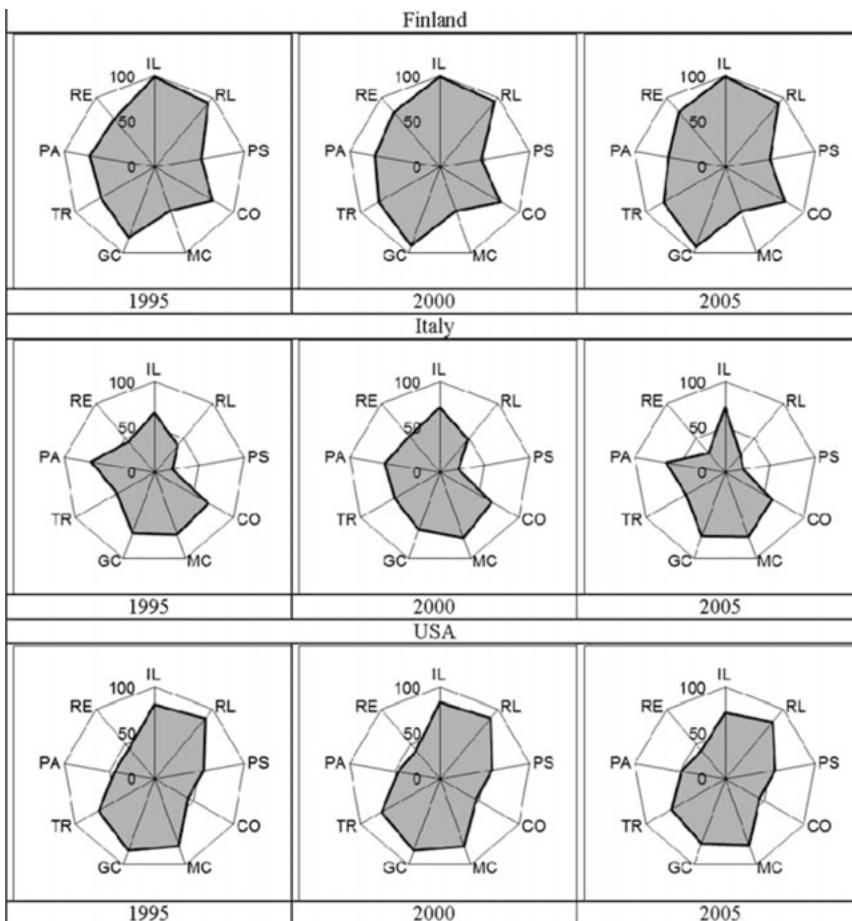
Abb. 9: Schematische Darstellung des methodischen Dreischritts in der Demokratiemessung

der methodische Ausgangspunkt die Frage sein, welche qualitativen Unterschiede die Bewertung letzten Endes ausdrücken soll und inwiefern diese sich in handlungsrelevanten Konsequenzen niederschlagen – und hierauf müssen die drei methodischen Schritte einer Bewertung auch immer bezogen bleiben.

Dass sich eine ‚Re-Aggregation‘ der Teilwertungen grundsätzlich am Erkenntnisinteresse ausrichten muss, wird in der Methodendebatte zur Demokratiemessung freilich durchaus berücksichtigt. So plädieren Munck und Verkuilen (2002: 22–23) dafür, stärker zu reflektieren, bis zu welcher Ebene die Zusammenfassung von Teilwerten durchgeführt werden *soll* – und fordern, diese gegebenenfalls auch nicht bis zum ‚Ende‘ durchzuführen. Denn jede Zusammenfassung von Teilaspekten auf einer Ebene des Konzeptbaums bringt auch unweigerlich einen Informationsverlust mit sich:

“The standard practice with regard to the selection of the level of aggregation has been to proceed as though parsimony were the only consideration, fully warranting a decision to push the process of aggregation to the highest level possible so as to reduce the disaggregate data into one single score. Index creators have thus done little to prevent a loss of information.” (MUNCK & VERKUILEN 2002: 22–23)

Als lobenswerte Ausnahme kann hier das oben bereits erwähnte Demokratiebarmeter genannt werden. Bei diesem werden die Messungen zur Demokratiequalität zwar ebenfalls in einem einzigen Wert zwischen 0 und 100 zusammengefasst, gleichzeitig wird es aber auch ermöglicht, verschiedene Demokratien auf Ebene der neun Bedeutungsaspekte (siehe Abb. 7 auf Seite 239) über verschiedene Zeitpunkte hinweg qualitativ miteinander zu vergleichen. Abb. 10 zeigt dies für die Länder Finnland, Italien und die USA zu den Jahren 1995, 2000 und 2005. Durch die Darstellung lässt sich ein deutlich besseres Bild davon gewinnen, wie das Konzept Demokratie in den drei Staaten qualitativ ausgestaltet ist,



IL: Individual Liberties; RL: Rule of Law; PS: Public Sphere; CO: Competition; MC: Mutual Constraints; GC: Governmental Capability; TR: Transparency; PA: Participation; and RE: Representation

Abb. 10: Vergleich der Demokratiequalität dreier Länder (BÜHLMANN u. a. 2012: 529)

in welchen Hinsichten Demokratie jeweils stärker oder schwächer ausgeprägt ist und ob diese Ausgestaltungen über die Zeit stabil geblieben sind.

Über eine solche multiperspektivische Darstellung können normative Messungen der Komplexität eines Begriffs wie Demokratie somit deutlich besser gerecht werden, als wenn nur ein einziger Wert die Qualität bemessen soll. Dabei zeigt sich, dass es keinesfalls problematisch, sondern für ein Verständnis komplexer Wertungssituationen sogar sehr hilfreich ist, verschiedene Wertungen gleichberechtigt nebeneinander bestehen zu lassen. Entsprechend lässt sich pragmatisch argumentieren, dass es mit Blick auf die drei in Kapitel 3 entwickelten Problemtypiken auch in dieser Hinsicht sinnvoll ist, diese gleichberechtigt

nebeneinander bestehen zu lassen, statt Versuche zu unternehmen, diese trotz der oben geschilderten Bedenken hinsichtlich ihrer Inkommensurabilität unter einem einzigen normativen Begriff zu versammeln.

In der Folge werde ich darlegen, wie man den dreischrittigen Ansatz der Demokriemessung adaptieren kann, um die in den Kapiteln 2 und 3 entwickelte moralpragmatische Landkarte für die Organisation und Implementierung von Fluggastkontrollen um eine vergleichende Bewertungsmethode zu ergänzen. Mit Hilfe dieser Adaption lässt sich, so meine These, ein Framework zur Abschätzung entwickeln, wie scharf sich die identifizierten typischen Problemkonstellationen für verschiedene Formen der Fluggastkontrollen voraussichtlich manifestieren werden. Kompromisse und Trade-Offs würden sich für die drei Problemkategorien gut darstellen lassen, etwa wenn ein voraussichtlich geringerer Eingriff in private Lebensbereiche dadurch erkauft wird, dass sich Diskriminierungsprobleme verschärfen oder der Beitrag zur Restriktion offener Gesellschaften verstärkt.

4.2.3 Moralpragmatische Adaption des Bewertungsansatzes für Fluggastkontrollen

Bewertung von Wegen durch den Checkpoint

Für meine moralpragmatische Adaption der Grundidee der normativen Messung kann, so habe ich oben bereits erläutert, die methodische Rolle des Wert- bzw. Konzeptbaums durch die drei Problemtypiken übernommen werden, die ich in Kapitel 3 entwickelt habe. Über sie lassen sich insgesamt zwölf typische sozial-ethische Problemkonstellationen identifizieren, mit denen beim Entwurf und bei der Implementierung von Prozessen der Fluggastkontrolle gerechnet werden muss. Neben der in Kapitel 3 geelisteten hermeneutischen Klärung wurden für diese Problemkonstellationen dabei auch jeweils zwei bzw. vier konkrete Modi erarbeitet, in denen sich die Problemtypen bei Fluggastkontrollen manifestieren. Zudem habe ich deutlich gemacht, dass alle Problemtypen einer graduellen Bewertung im Sinne eines ‚Mehr oder Weniger‘ zugänglich sind.¹⁵⁰

Wie ich im letzten Abschnitt ausgeführt habe, kann eine pragmatisch ausgerichtete Operationalisierung dieser drei Problemtypiken nicht einfach von den dort entwickelten Bewertungskriterien ausgehen – so als handle es sich bei diesen gewissermaßen um begriffliche Entitäten, die nun quasi-physisch bemessen wer-

¹⁵⁰ Wie in Fußnote 147 bereits angedeutet, wurden die drei Problemtypiken bei der Genese dieser Arbeit bereits mit Blick auf jene Qualitätskriterien entwickelt, die auch für eine pragmatische Adaption des methodischen Ansatzes ‚normative Messung‘ gelten. Meine These ist dabei, dass die Typiken alle relevanten Aspekte abdecken, dass sie keine nicht-relevanten Konstellationen berücksichtigen sowie dass die Ausdifferenzierung eine trennscharfe Zuordnung zu Problemkonstellationen ermöglichen. Zudem ist meine These, dass für jede der drei Problemtypiken gilt, dass eine Verschärfung eines Problemtypus auch (ce-teris paribus) als Verschärfung der Problemkategorie insgesamt verstanden werden kann.

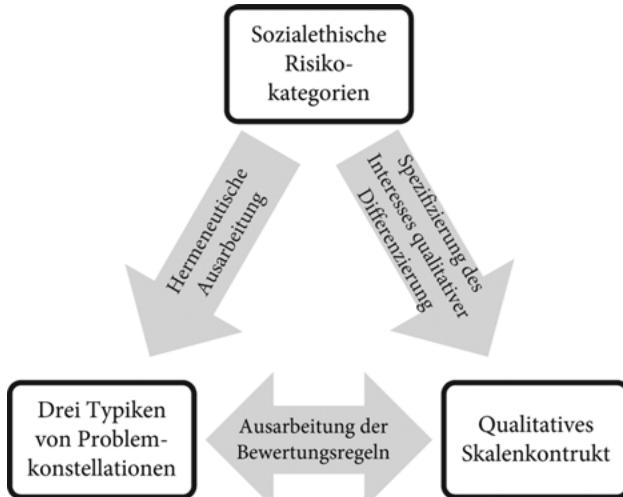


Abb. 11: Schema der moralpragmatischen Adaption des methodischen Dreischritts in der Demokratiemessung

den könnten. Vielmehr muss sich eine moralpragmatische Operationalisierung am situativen Erkenntnisinteresse orientieren, das heißt an der Frage, welche qualitativen Unterschiede mit Blick auf die identifizierten sozialethischen Risiken von Fluggastkontrollen *als Planungsgrößen* konkret dargestellt werden sollen.¹⁵¹ Dies bedeutet, dass sich – entgegen dem methodischen Vorgehen in Ansätzen der normativen Messung – an die hermeneutische Ausdifferenzierung der sozialethischen Risikokategorien nun zunächst die Frage anschließt, *welche qualitativen Differenzierungen konkret zum Ausdruck gebracht werden sollen*. Erst im letzten Schritt sollen dann für die Bewertungskriterien beobachtbare Sachverhalte gesucht werden, die diese qualitativen Unterschiede zum Ausdruck bringen. Abb. 11 zeigt diesen leicht veränderten Dreischritt für die moralpragmatische Adaption der normativen Messung.

Es ist wichtig zu bemerken, dass sich die hier dargestellten Änderungen lediglich auf die *Entwicklung* des Frameworks, genauer gesagt *auf die Entwicklung des Skalenkonstrukt und der Bewertungsregeln* beziehen, nicht aber auf den späteren Vorgang der Bewertung tatsächlicher Checkpoints mit Hilfe dieses Frameworks. Die Bewertung folgt auch weiterhin dem Dreischritt der Demokratiemessung: Die ausdifferenzierten Problemtypen werden über Bewertungsregeln einzeln bewertet und diese Teilwertungen werden sodann anhand von weiteren Regeln über das Skalenkonstrukt zu übergreifenden Wertungen zusammengefasst.

¹⁵¹ Mit diesem Wechsel des Fokus hin zum situativen Erkenntnisinteresse ist damit auch wieder die in Kapitel 1 entwickelte Aufgabe einer anwendungsorientierten Technikethik angegangen, nämlich Entscheidungen zu technischem Handeln hinsichtlich ihrer sozialethischen Implikationen besser zu informieren und anzuleiten.

Im letzten Abschnitt hatte ich dafür argumentiert, dass es zur besseren Darstellung der normativen Komplexität von Fluggastkontrollen lohnenswert ist, die drei Problemtypiken gleichberechtigt nebeneinander zu stellen, statt sie auf ein noch abstrakteres Konzept zu beziehen. Ziel des hier zu entwickelnden Frameworks ist es demnach, für jede der drei Problemkategorien eine eigenständige, graduelle Bewertung zu ermöglichen, wie ‚scharf‘ sie sich für einen konkreten Flughafen-Checkpoint voraussichtlich manifestieren werden. Dabei bezieht sich diese Bewertung – wie oben ausgeführt – einerseits auf die Frage *wie häufig* sich bestimmte Problemkonstellationen ergeben, sowie andererseits darauf, *wie ‚stark‘ oder ‚brisant‘* diese voraussichtlich ausfallen werden.

Der Aspekt der Brisanz eines Wertungskonflikts verweist zunächst auf jene Modi, in denen sich die Typen von sozialethischen Risiken bei Fluggastkontrollen manifestieren können. Dabei hängt eine Einschätzung dessen, wie stark beispielsweise die körperliche Privatheit von Passagieren durch Berührungen negiert wird, davon ab, als wie problematisch wir bestimmte Formen von Berührungen einschätzen: So werden Berührungen im Intimbereich, etwa im Schritt, *gemeinhin* deutlich negativer bewertet als Berührungen an den Oberarmen. Ebenso wird ein Abtasten des gesamten Körpers negativer bewertet als das gezielte Abtasten einer begrenzten Körperregion.

Dass eine solche graduelle Abstufung der Brisanz als *gemeinhin gültig* bezeichnet werden kann, ist dabei nicht das Resultat bloß zufälliger Überzeugungen, sondern gründet im oben beschriebenen Konzept der Privatheit, sowie in der Tatsache, dass bei Fluggastkontrollen der Eingriff in private Lebensbereiche als relevanter Wertungskonflikt erlebt wird – und zwar im Deweyschen Sinne eines *sozial geteilten* Erlebens. Entsprechend verweist eine solche graduelle Abstufung nicht auf vereinzelt gemachte Erfahrungen von Passagieren, sondern auf kulturell verankerte Erwartungshaltungen, auf die wir im Sinne von Planungsgrößen zurückgreifen dürfen. Eine Aussage wie „Das Abtasten im Schritt respektiert meine Privatsphäre *besser* als das Abtasten am Oberarm“ dürfte entsprechend, auch wenn sie ernst gemeint wäre, *zu Recht* als sehr ungewöhnliche und sozialethisch nicht weiter relevante Einzelwertung gelten.

Wenn dies der Fall ist, wenn also die sozial geteilte Brisanz eines Wertungskonflikts über den Modus zugänglich wird, in dem sich dieser manifestiert, dann ergibt sich im Falle der Fluggastkontrollen eine interessante Möglichkeit zur Bewertung der Folgewirkungen. Denn die Prozesse der Fluggastkontrollen sind, wie ich in Abschnitt 2.2.2 ausgeführt habe, sehr stark formalisiert. Sie gliedern sich in technische Einzelschritte, die über kaskadenartigen Entscheidungslogiken (zumeist anhand von ‚Alarm/kein Alarm‘) zu einem Gesamtprozess zusammengefasst sind. Selbst solche Techniken, die vom Kontrollpersonal manuell durchgeführt werden – etwa das Abtasten von Passagieren –, sind hoch formalisiert und dem Personal vor Ort bleibt wenig selbst überlassen. Die Techniken müssen entsprechend auch zunächst in Schulungen vermittelt werden. Ob in einem Kon-

trollschritt also beispielsweise der gesamte Körper eines Passagiers abgetastet wird, oder nur eine spezifische Körperregion, und ob es sich dabei um ein intensives Abtasten auch der Intimregionen handelt, ist in der Definition der Kontrollprozesse bereits detailliert festgelegt.

In Abschnitt 2.2.2 habe ich zudem dargelegt, dass sich bei Fluggastkontrollen die konkreten Kontrollschrifte, die ein Passagier durchlaufen muss, als ‚Wege durch den Checkpoint‘ darstellen lassen, nämlich je nachdem, bei welchen Techniken konkret ein Alarm entsteht und über welche Folgeschritte dieser Alarm wieder aufgelöst wird (siehe hierzu Abb. 1 auf Seite 130). Weil die Kontrollschrifte nun derart stark formalisiert sind, lässt sich gut vorhersagen, auf welchen Wegen durch den Checkpoint welche Techniken in welcher Form zum Einsatz kommen: So lässt sich anhand der formal definierten Entscheidungslogiken für die Kontrollen beispielsweise eindeutig bestimmen, ob das Kontrollpersonal auf einem bestimmten Weg durch Abtasten in die körperliche Privatheit der Passagiere eingreift, und wenn ja, ob dies ein Abtasten des gesamten Körpers impliziert oder nicht. Entsprechend lässt sich für eine vergleichende Bewertung über jene formal definierten Kontrollprozesse, die wir als ‚Wege durch den Checkpoint‘ bezeichnet haben, auf die *Brisanz* eines Wertungskonflikts zugreifen.

Wie häufig sich ein solcher Wertungskonflikt voraussichtlich manifestieren wird, hängt dann aber ebenfalls von der Definition der Kontrollprozesse ab: Wie ich im Exkurs über die Signalentdeckungstheorie in Abschnitt 2.2.2 ausgeführt habe, lassen sich für die spezifischen Kontrolltechniken nicht nur die Trefferquoten, sondern auch die Alarmraten bestimmen. Entsprechend lässt sich auch gut abschätzen, welcher Anteil an Passagieren welche Wege durch den Checkpoint nehmen wird. Die formalisierten ‚Wege durch den Checkpoint‘ erlauben demnach einen Zugang sowohl zur Brisanz von Wertungskonflikten, wie auch dazu, wie häufig sich diese manifestieren. Für eine Operationalisierung der moralpragmatischen Bewertung von Fluggastkontrollen bilden diese entscheidungslogischen Wege entsprechend eine exzellente Basis.¹⁵²

Wie häufig welche Wege durch den Checkpoint zum Tragen kommen, hat freilich nicht nur einen großen Einfluss auf die sozialethischen Folgewirkungen, sondern ebenso auf sicherheitstechnische und ökonomische Fragen. Nur über diese Wege lässt sich die Erkennungsrate verbotener Gegenstände, wie auch der Personal- und Platzaufwand für einen Checkpoint verlässlich abschätzen. Entsprechend kennen verantwortliche Personen bei den Flughäfen und in Polizeibehörden in aller Regel diese Quoten für ihre Checkpoints recht genau und validieren ihre Berechnungen auch immer wieder empirisch im Betrieb. Auch der Gesetzgeber kennt solche Quoten und gibt sie – wie oben bereits erwähnt – in

¹⁵² Für die Übertragbarkeit meines Frameworks auf andere Kontexte ist dies eine wichtige Einschränkung, denn ein derart hoher Formalisierungsgrad von Sicherheitsmaßnahmen, wie er bei Fluggastkontrollen vorherrscht, kann nicht für andere Bereiche der öffentlichen Sicherheit vorausgesetzt werden.

Form von Stichprobenquoten oder technischen Zulassungskriterien zum Teil sogar vor. Die Kenngrößen, die wir für einen Zugriff auf die Brisanz und die Häufigkeit von Wertungskonflikten durch das Framework benötigen, werden also auch aus anderen Gründen erhoben. Wie ebenfalls bereits erwähnt, sind diese Quoten allerdings nicht öffentlich zugänglich. Für den Zweck des Frameworks, nicht nur als ein intellektuelles Instrument zur Politik- sondern auch zur Gesellschaftsberatung zu fungieren, bedeutet dies eine wichtige Einschränkung. Allerdings wird sich im Folgenden dennoch zeigen, dass bereits grobe Einschätzungen genügen, um einen guten Eindruck davon zu gewinnen, wie sich typische Problemkonstellationen für spezifische Kontrollprozesse manifestieren.

Im folgenden Abschnitt werde ich das qualitative Skalenkonstrukt für ein operationalisiertes Framework entwickeln. Dies wird in zwei Schritten geschehen, die sich zum einen auf die Häufung von sozialethischen Problemkonstellationen beziehen, sowie zum anderen auf deren Brisanz. Damit das Skalenkonstrukt anschließend genutzt werden kann, um Teilwertungen zu den sozialethischen Risikotypen zusammenzufassen, werde ich zudem auch für beide Schritte entsprechende Regeln vorschlagen. Diese Regeln zur Zusammenfassung von Teilwertungen übernehmen in meinem Framework dabei jene Rolle, die bei der Demokratiemessung den Regeln zur Re-Aggregation zukommt (siehe Abb. 9 auf Seite 245).

Das qualitative Differenzierungsinteresse als Skalenkonstrukt

In der Technikfolgenforschung werden zur operationalisierten Bewertung oft Ampelmodelle eingesetzt, die über einige wenige Ränge qualitative Unterschiede ausdrücken sollen. Allerdings sind klassische Ampelmodelle für ein Framework zur Bewertung von Fluggastkontrollen eher problematisch. Wie ich in Abschnitt 4.1 dargelegt habe, schwingt bei der Abstufung über grün, gelb und rot zumeist auch der Bedeutungsgehalt ‚problemlos‘, ‚problembehaftet‘, ‚unzulässig‘ mit. Auf eine Bewertung dieser Art kann das hier zu entwickelnde Framework allerdings nicht abzielen, denn mit Blick auf die Grundfunktion der Aufdeckung können wir sagen, dass jede Form von Fluggastkontrollen eine Verletzung privater Lebensbereiche impliziert, insofern also auch *immer* problembehaftet ist. Vielmehr soll die Operationalisierung der Problemtypiken eine *vergleichende* Einschätzung zwischen verschiedenen Kontrollprozessen erlauben. Der heuristische Wert eines solchen Frameworks liegt dabei in einer abgestuften qualitativen Differenzierung dessen, ‚wie scharf‘ sich eine Problemkonstellation voraussichtlich manifestieren wird. Wie in Abschnitt 4.1 ebenfalls bereits angedeutet, wird sich eine solche vergleichende Einschätzung auf einer qualitativen Abstufung zwischen den Extremen eines realistischen ‚*Best-case*‘- und eines realistischen ‚*Worst-case*‘-Szenarios bewegen.

Die Zahl an qualitativen Abstufungen sollte hierbei nicht zu gering sein – eine einfache Zweiteilung nach dem Schema ‚näher am *Best-case*‘ und ‚näher am

Worst-case wäre für eine informierte Entscheidung über die Ausgestaltung der Kontrollprozesse sicher zu grob. Gleichzeitig dürfen durch das Framework aber auch nicht mehr Detailgrade angeboten werden, als tatsächlich auch adäquat unterschieden werden können. Insofern ist es durchaus sinnvoll, sich auf einige wenige qualitative Abstufungen zu beschränken. Ich möchte deshalb zur qualitativen Bewertung der drei sozialethischen Problemkategorien jeweils eine vierstufige Rangskala von ‚A‘ bis ‚D‘ vorschlagen. Dabei soll die schlechteste Wertung ‚D‘ jeweils ausdrücken, dass es sich hierbei um einen Fall handelt, der mit Blick auf die jeweilige Problemkategorie recht nah am *Worst-case* liegt. Umgekehrt soll der Rang ‚A‘ ausdrücken, dass die Kontrollprozesse so gestaltet sind, dass sie hinsichtlich der jeweiligen sozialethischen Problemkategorie recht nahe am Bestfall liegen. Die beiden Wertungsstufen ‚B‘ und ‚C‘ sollen dann folglich Zwischenstufen ausdrücken, bei denen sich Kontrollprozesse hinsichtlich der Qualität einer Problemkategorie deutlich von den beiden Extrempfälzen abheben, aber entweder näher am *Best-case* oder, respektive, näher am *Worst-case* liegen.

Da hier mit einer reinen Rangskala gearbeitet werden soll, kann allerdings nicht angenommen werden, dass die Abstände zwischen den einzelnen Wertungsstufen miteinander vergleichbar sind – es darf also nicht angenommen werden, dass der Schritt von der Wertungsstufe ‚D‘ zu ‚C‘ eine vergleichbare Verbesserung impliziert wie der Schritt von ‚C‘ zu ‚B‘. Weil sich die Wertungsstufe ‚C‘ dabei aber qualitativ deutlich von der schlechtesten Rangstufe unterscheiden soll, kommt ihr noch eine besondere Rolle zu, die eine bessere Orientierung auf der Rangskala erlaubt: Um sich deutlich vom *Worst-case* etwa hinsichtlich der Verletzung privater Lebensbereiche zu unterscheiden, muss bei Kontrollprozessen ein deutlich erkennbares Mindestmaß an Schutz vorliegen, um noch die Wertung ‚C‘ zu erhalten. Abb. 12 stellt das von mir vorgeschlagene Skalenkonzept zur Bewertung der drei sozialethischen Problemkategorien von Fluggastkontrollen in vier qualitativen Rangabstufungen dar. Dabei werden die qualitativen Abstufungen innerhalb einer Risikokategorie über die graduellen Einschränkungen ‚moderat‘, ‚geringfügig‘ und ‚kaum‘ bezeichnet.

Die Bezeichnungen der qualitativen Unterschiede sind dabei noch bewusst vage gehalten. Wie ich oben argumentiert habe, beziehen diese sich einerseits auf die Brisanz des jeweiligen Wertungskonflikts auf einem bestimmten ‚Weg durch den Checkpoint‘, andererseits aber auch darauf, wie groß der Anteil der jeweils betroffenen Passagiere ist. In beiderlei Hinsicht werden die durch die Skala bezeichneten qualitativen Unterschiede noch zu spezifizieren sein. Doch lässt sich bereits jetzt erkennen, wie sich die dargestellte Rangskala anwenden lässt: Für jede Wertungsstufe können bestimmte Qualitäten als Anforderungen definiert werden. Bei einer konkreten Wertung auf dieser Skala wird zunächst geprüft, ob alle Anforderungen für die Bestwertung ‚A‘ erfüllt werden. Ist dies nicht der Fall, wird geprüft, ob die Anforderungen für die Wertung ‚B‘ erfüllt werden. Werden auch die Anforderungen für die Wertung ‚C‘ nicht erfüllt, so ergibt sich die Wertung ‚D‘.

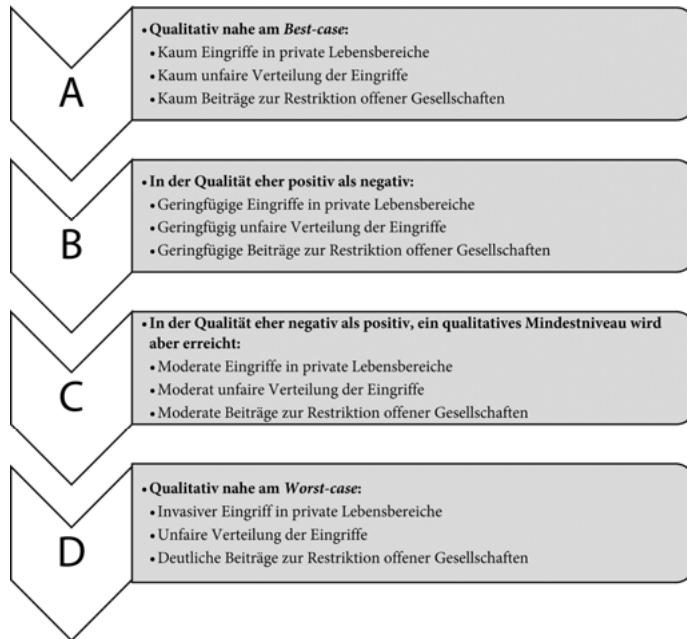


Abb. 12: Konzept des Skalenkonstrukts für Bewertung der sozialethischen Konfliktlinien

Bei der Frage, wie die qualitativen Unterschiede mit Blick auf Häufung und Brisanz von Wertungskonflikten nun weiter spezifiziert werden können, lassen sich weitere Orientierungshilfen anhand von Tauglichkeitskriterien aus dem Erkenntnisinteresse der Arbeit entwickeln. Denn das Skalenkonstrukt kann nur dann seinen Zweck erfüllen, wenn es tatsächlich erlaubt, implementierte Formen von Fluggastkontrollen über alle vier Stufen zu unterscheiden. Würden sich alle bekannten Checkpoints in lediglich zwei oder drei Wertungsstufen gruppieren, so erwiese sich die Skala letztlich als untauglich: Zum einen gäbe sie ein größeres qualitatives Differenzierungspotential vor, als tatsächlich eingelöst werden kann. Zugleich liefe die Operationalisierung im Ganzen aber auch in Gefahr, über die verbleibenden, tatsächlich genutzten Abstufungen zu grob zu differenzieren und so einer informierten Entscheidungsfindung nicht mehr dienlich zu sein.

Dabei lässt sich anhand der Funktionsweise von Fluggastkontrollen sagen, dass es für *alle* Flughafen-Checkpoints eine Reihe von Passagieren geben wird, für welche die Kontrollen verhältnismäßig scharf ausfallen. Der Grund hierfür liegt darin, dass, wie in Anschnitt 2.2 ausgeführt, jeder durch eine Kontrollmaßnahme ausgelöste Alarm weitere Kontrolltechniken nach sich ziehen muss, bis alle Alarne für das Kontrollpersonal zufriedenstellend aufgelöst sind. An jedem Checkpoint wird es dabei Grenzfälle geben, in denen das Kontrollpersonal sehr genau kontrolliert, beispielsweise aufgrund einer zufälligen Häufung von Fehlalarmen bei einzelnen Passagieren. Die Anforderungen an eine ‚A‘-Wertung bei

der Zusammenfassung von Teilwertungen beispielsweise mit Blick auf Privatheitsprobleme dürfen also nicht so strikt gestaltet werden, dass invasive Eingriffe in private Lebensbereiche überhaupt nicht vorkommen dürfen. Sie müssen aber auf einen *im Vergleich zu anderen Checkpoints* außergewöhnlich geringen Anteil begrenzt bleiben; und auch sonst müssen ungewöhnlich viele Passagiere so kontrolliert werden, dass sie kaum oder nur geringe Eingriffe erdulden müssen. Umgekehrt darf eine ‚D‘-Wertung nicht erst dann erfolgen, wenn alle oder fast alle Passagiere über invasive Techniken kontrolliert werden, sondern bereits dann, wenn dies für einen ungewöhnlich hohen Anteil gilt. Die qualitativen Unterscheidungen sollen demnach also keine absoluten Abstufungen ausdrücken, sondern sie sollen Abstufungen relativ zum Spektrum dessen bezeichnen, was im Kontext der Fluggastkontrollen als erwartbar gelten kann.

Dass Wertungsskalen auf das Spektrum der tatsächlich erwartbaren Ergebnisse bezogen sind und dass sie die empirischen Fälle dabei auch möglichst über alle Wertungsstufen auffächern können müssen, wird auch in der Methodendebatte der Demokratiemessung so formuliert (MUNCK & VERKUILEN 2002: 8). Für Demokratiemessungen, die bereits eine Vielzahl politischer Systeme über verschiedene Zeitpunkte miteinander verglichen haben, lässt sich diese Differenzierungsfähigkeit naturgemäß sehr gut darstellen. Solche umfangreichen Sammlungen von Messergebnissen in sogenannten ‚Demokratie-Indizes‘ sind dem aktuellen Stand des hier zu entwickelnden Frameworks klar überlegen. Denn aktuell liegen noch zu wenige Bewertungsergebnisse vor, um die Tauglichkeit der Differenzierungsleistung auch wirklich aufzuzeigen zu können. Aktuell werde ich mich bei der Definition der konkreten Anforderungen für die Zusammenfassung von Teilwertungen also noch auf einige wenige, im Rahmen des EU-Projekts XP-DITE durchgeführte Bewertungen sowie auf meine hierbei gesammelten Erfahrungen stützen.

Diese Einschränkung mit Blick auf den Nachweis der Tauglichkeit des Frameworks ist der noch recht frühen Entwicklungsphase der Operationalisierung geschuldet. Grundsätzlich muss *jeder* Ansatz zur normativen Messung bei der Entwicklung zunächst auf die gesammelten Erfahrungen der Forschenden zurückgreifen. In den Forschungsdebatten zur Demokratiemessung hat sich aber bereits herausgestellt, wie konstruktiv hierbei auf die Identifizierung konkreter Schwächen eingegangen werden kann. Denn Kritik, beispielsweise an den spezifischen Regeln zur Zusammenfassung von Teilwertungen, lässt sich zumeist produktiv wenden, um Schwächen auszubessern. Entsprechend ist die Vielzahl an Demokratieindizes, die über die letzten Jahre entwickelt wurde, auch als Zeichen einer zunehmenden Verfeinerung der Bewertungsmethode zu interpretieren.

Für die Zusammenfassung der Teilwertungen zu den einzelnen Wegen durch den Checkpoint *anhand der Häufigkeit* schlage ich die in Tabelle 1 dargestellte qualitative Abstufung vor, für die ich entsprechende Regeln in Form von Anforderungen formuliert habe. Später werde ich dann noch ähnliche Regeln zur Zusammenfassung von Teilwertungen *anhand der Brisanz* formulieren.

Tabelle 1: Qualitative Differenzierung von Flughafen-Checkpoints anhand der Häufigkeit von Wertungskonflikten

	Qualität des Eingriffs in die Privatheit der Passagiere	Qualität der Schlechterstellung verletzlicher Gruppen	Qualität des Beitrags zur Restriktion offener Gesellschaften	Anforderungen bei der Zusammenfassung von Wertungen
A	Nur ein <i>minimaler Anteil</i> der Passagiere wird über eine invasive Kombination von Techniken nahe dem <i>Worst-case</i> kontrolliert. Für die <i>meisten</i> Passagiere bedeuten die Kontrollen geringfügige Eingriffe. Ein <i>großer Anteil</i> wird zudem über eine kaum eingreifende Kombination von Techniken nahe am <i>Best-case</i> Szenario kontrolliert.	Nur ein <i>minimaler Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, die sie als Angehörige einer als verletzlich geltenden Gruppe nahe dem <i>Worst-case</i> schlechterstellen. Für die <i>meisten</i> Passagiere bedeuten die Kontrollen eine nur geringfügig unfaire Verteilung der Eingriffe. Ein <i>großer Anteil</i> wird zudem über eine kaum unfaire Kombination von Techniken nahe am <i>Best-case</i> Szenario kontrolliert.	Nur ein <i>minimaler Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, durch welche deutliche Beiträge zur Restriktion offener Gesellschaften nahe dem <i>Worst-case</i> geleistet werden. Die <i>meisten</i> Passagiere werden über eine Kombination von Techniken kontrolliert, die nur geringfügige Beiträge leisten. Ein <i>großer Anteil</i> wird zudem über Techniken kontrolliert, deren Beiträge nahe am <i>Best-case</i> Szenario liegen.	Für max. 0,3% der Passagiere ist die Wertung „D“ UND Für max. 15% der Passagiere ist die Wertung „C“ UND Für mind. 30% der Passagiere ist die Wertung „A“
B	Nur ein <i>sehr kleiner Anteil</i> der Passagiere wird über eine invasive Kombination von Techniken nahe dem <i>Worst-case</i> kontrolliert. Für die <i>meisten</i> Passagiere bedeuten die Kontrollen nur geringfügige Eingriffe.	Nur ein <i>sehr kleiner Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, die sie als Angehörige einer als verletzlich geltenden Gruppe nahe dem <i>Worst-case</i> schlechterstellen. Für die <i>meisten</i> Passagiere bedeuten die Kontrollen eine nur geringfügig unfaire Verteilung der Eingriffe.	Nur ein <i>sehr kleiner Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, durch welche deutliche Beiträge zur Restriktion offener Gesellschaften nahe dem <i>Worst-case</i> geleistet werden. Die <i>meisten</i> Passagiere werden über eine Kombination von Techniken kontrolliert, die nur geringfügige Beiträge leisten.	Für unter 0,7% der Passagiere ist die Wertung „D“ UND Für max. 25% der Passagiere ist die Wertung „C“
C	Nur ein <i>kleiner Anteil</i> der Passagiere wird über eine invasive Kombination von Techniken nahe dem <i>Worst-case</i> kontrolliert.	Nur ein <i>kleiner Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, die sie als Angehörige einer als verletzlich geltenden Gruppe nahe dem <i>Worst-case</i> schlechterstellen.	Nur ein <i>kleiner Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, durch welche deutliche Beiträge zur Restriktion offener Gesellschaften nahe dem <i>Worst-case</i> geleistet werden.	Für unter 2% der Passagiere ist die Wertung „D“
D	Ein <i>deutlicher Anteil</i> der Passagiere wird über eine invasive Kombination von Techniken nahe dem <i>Worst-case</i> kontrolliert.	Ein <i>deutlicher Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, die sie als Angehörige einer als verletzlich geltenden Gruppe nahe dem <i>Worst-case</i> schlechterstellen.	Ein <i>deutlicher Anteil</i> der Passagiere wird über eine Kombination von Techniken kontrolliert, durch welche deutliche Beiträge zur Restriktion offener Gesellschaften nahe dem <i>Worst-case</i> geleistet werden.	Für mind. 2% der Passagiere ist die Wertung „D“

In der ersten Spalte ist dabei die Wertungsstufe von der Bestwertung ‚A‘ bis zur schlechtesten Wertung ‚D‘ angegeben, in der letzten Spalte die korrespondierenden Anforderungen an die Häufigkeit des Auftretens von Teilwertungen zur Brisanz der Wertkonflikte (beides grau hinterlegt). Zwischen der Wertungsstufe und den Anforderungen ist zudem für jede der drei sozialethischen Konfliktkategorien in einer eigenen Spalte noch eine Beschreibung dessen beigegeben, was über die Differenzierung konkret qualitativ ausgedrückt werden soll.

Die dargestellten qualitativen Abstufungen von ‚A‘ bis ‚D‘ spezifizieren die Anforderungen für eine Wertungsstufe als konkrete Regeln zur Zusammenfassung von Teilwertungen, die für die einzelnen ‚Wege durch den Checkpoint‘ gelten. Sie verweisen dabei auf weitere qualitative Abstufungen dazu, wie brisant sich die jeweilige Problemkategorie voraussichtlich auf diesem Weg manifestieren wird. Diese Brisanz ist entsprechend für jeden der Wege ebenfalls noch auf einer vierstufigen Skala von ‚A‘ bis ‚D‘ zu bewerten. Da aber für jeden Weg bestimmt werden kann, wie häufig er für Passagiere zum Tragen kommt, kann somit auch bestimmt werden, wie häufig ein Passagier auf einem beispielsweise mit ‚A‘ bewerteten Weg kontrolliert wird. Zur Bestimmung dieser Häufigkeiten werden dann alle gleich bewerteten Wege, also etwa alle, die mit ‚A‘ bewertet wurden, zusammengenommen. Wie oben beschrieben, kann zur Zusammenfassung der Teilwertungen nun geprüft werden, ob dabei beispielsweise hinsichtlich des Eingriffs in die Privatheit der Passagiere die Anforderungen für eine ‚A‘-Wertung erfüllt werden – was einen Checkpoint nahe dem Bestfall bezeichnen würde. Ist dies nicht der Fall, werden die Anforderungen für die ‚B‘- und ‚C‘-Wertungen entsprechend weiter geprüft. Können auch diese Anforderungen nicht erfüllt werden, so wird der Checkpoint hinsichtlich dieser Problemkategorie mit ‚D‘ bewertet, also als invasiv nahe dem *Worst-case*.¹⁵³

Ein Blick auf die Formulierung dieser Anforderungen lässt erkennen, dass diese Regeln formallogischen Ausdrücken ähneln und sich leicht in eine entsprechende Form¹⁵⁴ bringen lassen. Dies ist durchaus gewollt, denn es unterstützt die intersubjektive Nachvollziehbarkeit wie auch die präzise Replizierbarkeit dieses Bewertungsschritts durch Andere – was, wie oben bereits erwähnt, letztlich eines der Kernziele bei einer Operationalisierung ist.¹⁵⁵ Zusammengenommen

¹⁵³ Logisch betrachtet ist die Anforderungen an eine ‚D‘-Wertung also eine Inversion der Anforderungen an eine ‚C‘-Wertung. Diese Anforderung wird freilich nie geprüft, denn wann immer keine ‚C‘-Wertung gegeben werden kann, impliziert dies in der Operationalisierung eine ‚D‘-Wertung. Dennoch ist es wichtig für ein besseres Verständnis der Differenzierungsleistung, auch für diese Wertungsstufe die qualitative Beschreibung zusammen mit einer formalen Anforderung auszuführen.

¹⁵⁴ Je nach formallogischem System und je nach der gewählten Zeichensprache wird diese Form im Detail dann freilich unterschiedlich ausfallen.

¹⁵⁵ Eine solche formallogische Form eröffnet zudem auch die Möglichkeit einer computerunterstützten Auswertung, wie dies in sozialwissenschaftlichen *Mixed-methods*-Ansätzen, so etwa bei der Qualitativen Inhaltsanalyse, nicht unüblich ist. Als Teil der im EU-Projekt XP-DITE

fungieren diese Anforderungen im Framework dabei als *Regeln für die Zusammenfassung einzelner Bewertungen der Wege durch den Checkpoint hinsichtlich der Brisanz sozialethischer Problemkategorien*.

Für die weitere Spezifizierung des Skalenkonstrukts stellt sich nun also die Frage, nach welchen Regeln die hierbei vorausgesetzten Bewertungen vollzogen werden sollen. Analog zur Methode der Demokratiemessung müssen diese Bewertungen der Brisanz sozialethischer Problemkategorien auf die ausdifferenzierten Problemtypen verweisen. Entsprechend wird auch dieser Schritt in der Operationalisierung Teilwertungen nach möglichst präzisen Regeln zu übergreifenden Wertungen zusammenfassen. Und auch bei diesen Regeln ist das zentrale Tauglichkeitskriterium, dass sie anhand des dargestellten Skalenkonzepts eine qualitative Differenzierung über alle vier Ränge leisten können.¹⁵⁶ Tabelle 2 stellt meinen Vorschlag für eine solche regelbasierte qualitative Differenzierung vor. Die Wertungsstufen von der Bestwertung ‚A‘ bis hin zur schlechtesten Wertung ‚D‘ sind wieder in der ersten, die korrespondierenden Anforderungen in der letzten Spalte grau hinterlegt dargestellt. Dazwischen ist wieder für jede der drei sozialethischen Risikokategorien eine Beschreibung der intendierten Differenzierung beigegeben.

Auch hier erfolgt die konkrete Zusammenfassung von Einzelwertungen nach dem oben dargelegten Schema: Von der Bestwertung ausgehend wird überprüft, ob die jeweiligen Anforderungen für eine ‚A‘-, ‚B‘- oder ‚C‘-Wertung erfüllt wurden. Auch diese Anforderungen ähneln formallogischen Ausdrücken, wobei neben dem bereits verwendeten Junktor ‚UND‘ nun auch ‚ODER‘ sowie Klammern zur Bezeichnung des logischen Vorrangs zum Einsatz kommen. Dabei basieren die Anforderungen für die jeweilige Wertungsstufe hinsichtlich der Brisanz, mit der sich eine sozialethische Problemkategorie manifestieren wird, auf den Wertungen zu den einzelnen Problemtypen.

Es ist wichtig, zu bemerken, dass bei der Zusammenfassung der Teilwertungen zu den Problemtypen zwischen diesen nicht genauer unterschieden wird. Die Regeln berücksichtigen, *wie viele* Problemtypen *wie* auf der vierstufigen Skala bewertet wurden, nicht aber welche. Dies liegt darin begründet, dass die Problemtypen gleichrangig behandelt werden sollen. Beispielsweise soll also die Problemkategorie des Eingriffs in die Privatheit von Passagieren nicht derart begriffen werden, dass bestimmte Problemtypen eine wichtigere oder zentralere Rolle spielen als andere, sondern so, dass jeder Problemtypus einen gleichwichtigen Teilaspekt der Problemkategorie abdeckt. Die Anforderungen für die qualitativen

entwickelten *Shared Evaluation Platform* wird angestrebt, eine solche computerunterstützte Auswertung auf Grundlage des hier dargestellten Frameworks auch tatsächlich umzusetzen.

¹⁵⁶ Es gilt wie oben die Einschränkung, dass aktuell noch zu wenige Bewertungsergebnisse vorliegen, um die Tauglichkeit der Differenzierungsleistung anhand einer Vielzahl von Bewertungen nachweisen zu können. Entsprechend beruhen die Bewertungsregeln derzeit auf den im EU-Projekt XP-DITE durchgeführten Bewertungen, insbesondere aber auf meinen eigenen Erfahrungen.

Tabelle 2: Qualitative Differenzierung von ‚Wegen durch den Checkpoint‘ anhand der Brisanz von Wertungskonflikten

Qualität des Eingriffs in die Privatheit der Passagiere	Qualität der Schlechterstellung verletzlicher Gruppen	Qualität des Beitrags zur Restriktion offener Gesellschaften	Anforderungen zur Zusammenfassung
<p>A</p> <p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die kaum in private Lebensbereiche eingreift.</p> <p>Mit Blick auf mindestens einen Problemtypus sind die Kontrollen <i>nahe am Best-case-Szenario</i>.</p> <p>Durchweg bedeuten die Kontrollen aber <i>nur geringfügige</i> Eingriffe in private Lebensbereiche.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die verletzliche Gruppen kaum schlechterstellt.</p> <p>Mit Blick auf mindestens einen Problemtypus sind die Kontrollen <i>nahe am Best-case-Szenario</i>.</p> <p>Durchweg bedeuten die Kontrollen aber eine <i>nur geringfügig</i> unfaire Verteilung der Eingriffe in private Lebensbereiche.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die kaum Beiträge zur Restriktion offener Gesellschaften leisten.</p> <p>Mit Blick auf mindestens einen Problemtypus sind die Kontrollen <i>nahe am Best-case-Szenario</i>.</p> <p>Durchweg bedeuten die Kontrollen aber <i>nur geringfügige</i> Beiträge.</p>	<p>Mindestens ein Problemtypus wurde mit ‚A‘ bewertet</p> <p>UND</p> <p>Kein Problemtypus wurde schlechter als ‚B‘ bewertet</p>
<p>B</p> <p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die nur geringfügig in private Lebensbereiche eingreift.</p> <p>Entweder bedeuten die Kontrollen mit Blick auf meisten Problemtypen <i>nur geringfügige</i> Eingriffe in private Lebensbereiche, oder die Kontrollen sind hinsichtlich mancher Problemtypen <i>nahe am Best-case-Szenario</i>;</p> <p>ansonsten sind sie aber höchstens <i>moderat</i>.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die verletzliche Gruppen nur geringfügig schlechterstellt.</p> <p>Entweder bedeuten die Kontrollen mit Blick auf meisten Problemtypen eine <i>nur geringfügig</i> unfaire Verteilung der Eingriffe, oder die Kontrollen sind hinsichtlich mancher Problemtypen <i>nahe am Best-case-Szenario</i>;</p> <p>ansonsten sind die aber höchstens <i>moderat</i> unfair.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die nur geringfügige Beiträge zur Restriktion offener Gesellschaften leisten.</p> <p>Entweder bedeuten die Kontrollen mit Blick auf meisten Problemtypen einen <i>nur geringfügigen</i> Beitrag, oder die Kontrollen sind hinsichtlich mancher Problemtypen <i>nahe am Best-case-Szenario</i>;</p> <p>ansonsten leisten sie aber höchstens einen <i>moderaten</i> Beitrag</p>	<p>Kein Problemtypus wurde schlechter als ‚B‘ bewertet</p> <p>ODER</p> <p>(mindestens ein Problemtypus wurde mit ‚A‘ bewertet</p> <p>UND höchstens ein Problemtypus wurde mit ‚C‘ bewertet).</p> <p>UND</p> <p>Kein Problemtypus wurde schlechter als ‚C‘ bewertet</p>

Qualität des Eingriffs in die Privatheit der Passagiere	Qualität der Schlechterstellung verletzlicher Gruppen	Qualität des Beitrags zur Restriktion offener Gesellschaften	Anforderungen zur Zusammenfassung
<p>C Die Passagiere werden über eine Kombination von Techniken kontrolliert, die ein Mindestmaß an Schutz für private Lebensbereiche garantiert.</p> <p>Mit Blick auf die verschiedenen Problemtypen bedeuten die Kontrollen höchstens <i>moderate</i> Eingriffe.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die verletzlichen Gruppen ein Mindestmaß an Schutz vor Schlechterstellung gewährt.</p> <p>Mit Blick auf die verschiedenen Problemtypen bedeuten die Kontrollen eine höchstens <i>moderat</i> <i>unfaire</i> Verteilung der Eingriffe.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die ein Mindestmaß an Schutz vor der Restriktion offener Gesellschaften leisten.</p> <p>Mit Blick auf die verschiedenen Problemtypen bedeuten die Kontrollen höchstens <i>moderate</i> Beiträge.</p>	<p>Kein Problemtypus wurde schlechter als ,C‘ bewertet</p>
<p>D Die Passagiere werden über eine invasive Kombination von Techniken kontrolliert.</p> <p>Mit Blick auf die verschiedenen Problemtypen bedeuten die Kontrollen in mindestens einer Hinsicht einen Eingriff <i>nähe dem Worst-case</i>.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die verletzliche Gruppen schlechterstellt.</p> <p>Mit Blick auf die verschiedenen Problemtypen bedeuten die Kontrollen in mindestens einer Hinsicht eine <i>unfaire</i> Verteilung der Eingriffe <i>nähe dem Worst-case</i>.</p>	<p>Die Passagiere werden über eine Kombination von Techniken kontrolliert, die deutliche Beiträge zur Restriktion offener Gesellschaften leistet.</p> <p>Mit Blick auf die verschiedenen Problemtypen bedeuten die Kontrollen in mindestens einer Hinsicht einen Beitrag <i>nähe dem Worst-case</i>.</p>	<p>Mindestens ein Problemtypus wurde schlechter als ,C‘ bewertet</p>

Abstufungen sind dabei so formuliert, dass ein gewisser Ausgleich von guten und schlechten Wertungen ermöglicht wird. So erlaubt die Wertungsstufe ‚B‘ etwa, ‚C‘-Wertungen mit ‚A‘-Wertungen auszugleichen – ein sehr hoher Schutz für körperliche Privatheit, so das implizite Argument, kann also etwa einen nur moderaten Schutz für die räumliche Privatheit ein Stück weit aufwiegen. Diese Möglichkeit zum Ausgleich findet seine Grenze aber darin, dass schon für eine ‚C‘-Wertung die Bedingung gilt, dass ein Mindestmaß an Schutz für alle Aspekte gewährleistet sein muss. Dieser durchgängige Mindestschutz wird somit als notwendige Bedingung dafür konzipiert, eine konkrete Kombination von Kontrollmethoden vom *Worst-case* zu unterscheiden. Auch ein sehr hoher Schutz vor öffentlicher Preisgabe, so das hierbei implizite Argument, kann einen invasiven Eingriff in die Privatheit individueller Lebensführung nicht ausgleichen.

Mit dem zweiten Schritt der Ausdifferenzierung des Skalenkonstrukts sind wir auf einer Ebene angelangt, die sich auf die ausgearbeiteten drei Problemtypiken beziehen lässt. Im letzten Schritt der Operationalisierung müssen nun also noch jene Regeln ausgearbeitet werden, welche es erlauben, für einen Problemtypus beobachtbare Sachverhalte zu finden, die einem Rang auf der vierstufigen Skala gemäß dem obigen Skalenkonzept zugeordnet werden können.

Vermittelnde Ausarbeitung der Bewertungsregeln

Der letzte Schritt der Operationalisierung besteht in der Ausarbeitung der Bewertungsregeln. Die Ausarbeitung wird dabei zwischen den ausdifferenzierten Problemtypen und den Wertungsstufen auf dem Skalenkonzept vermitteln müssen. In diesem Abschnitt möchte ich vor allem die methodischen Überlegungen herausstellen, die in die Operationalisierung der zwölf Problemtypen eingeflossen sind. Um die Lesbarkeit meiner Arbeit zu erhöhen und um eine sich für alle zwölf Problemtypen wiederholende Struktur zu vermeiden, werde ich die vollständige Dokumentation des Frameworks der Arbeit im Anhang beigeben,¹⁵⁷ mich hier statt dessen aber auf eine möglichst konzise Darstellung der methodischen Grundideen anhand ausgewählter Teile der Problemtypologie beschränken. Aufgrund der recht hohen Anschaulichkeit sowie aufgrund der medialen Aufmerksamkeit für diesen Problemtypus in der sogenannten Nacktscanner-Debatte, werde ich mich hierfür zunächst dem Problemtypus ‚Verletzung körperlicher Aspekte von Privatheit‘ zuwenden.

In Abschnitt 3.2.2 hatte ich diesen Problemtypus genauer beschrieben und dabei vier Modi identifiziert, in denen sich ein Eingriff in körperliche Privatheit bei Fluggastkontrollen typischerweise manifestieren kann. Zum einen kann dies dadurch geschehen, dass die Kontrollen eine Entfernung ‚bedeckender Kleidung‘ nötig machen – sei es physisch oder durch bildgebende Techniken. Zweitens kann sich ein solcher Eingriff dadurch manifestieren, dass Passagiere bei den Kontrollen durch das Personal berührt werden. Ein dritter Modus des Eingriffs in körperliche Privatheit besteht in der Erhebung von Informationen über körperliche Aspekte, ein vierter schließlich in der Erhebung und Speicherung biometrischer Merkmale zu den Passagieren.

¹⁵⁷ Die Dokumentation der beobachtbaren Sachverhalte zu den Bewertungskriterien sowie die Bewertungsregeln zur Zuordnung auf einer Skala von ‚A‘ bis ‚D‘ finden sich in Anhang A. Zusammen mit den Regeln zur Zusammenfassung von Teilwertungen in Tabelle 1 auf Seite 255 und Tabelle 2 auf Seite 258 ergibt sich somit eine vollständige Dokumentation der Operationalisierung. Als Bewertungsframework, das letztlich auf ein vertieftes Verständnis von gesellschaftspolitischen Wertungskonflikten abzielt, muss diese Operationalisierung freilich immer im Kontext der moralischen Landkarte gelesen werden, die ich in den Kapiteln 2 und 3 erarbeitet habe.

Durch diese Spezifizierung der Eingriffs-Modi fällt es nun verhältnismäßig einfach, einige beobachtbare Sachverhalte zu identifizieren, die mit diesen Modi korrespondieren. So ließe sich für den zweiten Modus, der Berührung von Passagieren durch das Kontrollpersonal, schlicht die Frage stellen, ob dies auf einem bestimmten Weg durch den Checkpoint tatsächlich der Fall ist. Wie ich in Abschnitt 2.2.2 mit Blick auf die kaskadenartigen Kontrollprozesse ausgeführt habe, ist dies keineswegs immer der Fall. Denn eine gezielte Berührung des Körpers von Passagieren durch das Kontrollpersonal kommt letztlich nur dann zustande, wenn diese tatsächlich über die Technik des manuellen Abtastens kontrolliert werden. Dies geschieht, wie etwa in Abb. 1 auf Seite 130 deutlich wird, bei gängigen Kontrollen aber nur dann, wenn bei vorangegangen Schritten ein Alarm ausgelöst wurde. Passagiere, bei denen entweder in den Primärkontrollen mittels der Metalldetektorschleuse oder bei den Nachkontrollen durch den Körperscanner kein Alarm ausgelöst wurde, werden hingegen nicht abgetastet. Ein Eingriff in körperliche Privatheit über den Modus der Berührung durch das Kontrollpersonal manifestiert sich in diesen Fällen demnach *beobachtbar nicht*.

Ganz ähnlich verhält es sich mit dem Modus der Entfernung bedeckender Kleidung. Wenn die Regeln zum Ablegen von Jacken und Handgepäck nicht vorschreiben, dass entgegen der gängigen Praxis in der EU auch bedeckende Kleidung wie Kopftücher, Hemden, Blusen und Ähnliches abgelegt werden müssen, dann manifestiert sich im obigen Beispiel für alle Passagiere, bei denen die Kontrollen in der Metalldetektorschleuse keinen Alarm auslösen, ein entsprechender Eingriff ebenfalls *beobachtbar nicht*. Sofern bei den Kontrollen moderne Körperscanner ohne Erzeugung eines Bildes genutzt werden und sofern die spezifischen Regeln beim Abtasten ein Ablegen von religiösen Kopfbedeckungen, Burkas etc. nicht voraussetzen, kann dies sogar für alle Passagiere gelten, denen der Übertritt auf die Luftseite gestattet wird.

Durch derartige Körperscanner werden allerdings, so habe ich in Abschnitt 2.2.2 ausgeführt, durchaus körperbezogene Informationen erhoben und über eine sogenannte Anomalie-Detektion ausgewertet, weshalb sich in diesen Fällen der entsprechende Modus des Eingriffs *beobachtbar* manifestieren würde. Dies würde auch dann der Fall sein, wenn statt Techniken der Anomalie-Detektion beispielsweise Sprengstoffspuren-Detektoren eingesetzt würden. Da im obigen Beispiel keinerlei biometrische Techniken eingesetzt werden, kann zudem *beobachtbar* ausgeschlossen werden, dass biometrische Daten erhoben und gespeichert werden.

Eine solche erste Identifikation von wertungsrelevanten, beobachtbaren Sachverhalten ließe sich demnach darüber leisten, dass für die einzelnen Wege¹⁵⁸

¹⁵⁸ Die Zusammenfassung der Einzelwertungen zu den Problemtypiken bezieht sich, wie oben deutlich wurde, zunächst nur auf die einzelnen Wege durch den Checkpoint. Im letzten Schritt der Zusammenfassung werden dann anhand der Häufigkeiten der über vier Wertungsrange abgestuften Brisanz die Einzelwertungen der Wege durch den Checkpoint

durch den Checkpoint anhand der kombinierten Kontrolltechniken festgestellt wird, ob sich ein Eingriff über einen bestimmten Modus überhaupt manifestiert. Für die Operationalisierung lässt sich ein solcher beobachtbarer Sachverhalt recht gut über Entscheidungsfragen definieren, also über Fragen, die mit „ja“ oder „nein“ beantwortet werden können. Ein Beispiel wäre etwa die Frage „Berührt das Kontrollpersonal den Körper der Passagiere?“, welche für die einzelnen Wege dann anhand der stark formalisierten Kontrollprozesse beantwortet werden kann. Zur besseren Übersicht und Handhabbarkeit im Framework ist es dabei hilfreich, diese Entscheidungsfragen so aufeinander abzustimmen, dass die Antworten immer in gleicher Weise auf die Manifestation eines Wertungskonflikts hinweisen. Wie für die Frage nach einer Berührung durch das Kontrollpersonal bietet es sich also beispielsweise an, auch die Frage nach der Erhebung und Speicherung biometrischer Daten so zu formulieren, dass die Antwort „Ja“ eine Manifestation des Wertungskonflikts impliziert. Eine solche Formulierung wäre etwa „Werden bei den Kontrollen biometrische Daten erhoben und gespeichert?“.

Bereits jetzt lässt sich dabei festhalten, dass eine Kombination von Kontrolltechniken, bei der all diese Fragen mit „Nein“ beantwortet werden, mit der Bestnote ‚A‘ bewertet werden muss – denn keiner der identifizierten Wertungskonflikte würde sich demnach manifestieren. Freilich ist die bloße Feststellung, *dass* sich bei einer konkreten Kombination von Kontrolltechniken ein Wertungskonflikt in bestimmten Modi manifestieren wird, zwar notwendig, aber keinesfalls hinreichend, um eine differenzierte Bewertung der Brisanz dieses Konflikts auf einer vierstufigen Skala durchzuführen. Vielmehr müssen hierfür weitere beobachtbare Sachverhalte identifiziert werden, die es erlauben, die Brisanz dieses Konflikts abzustufen.

Wird also beispielsweise die Entscheidungsfrage „Berührt das Kontrollpersonal den Körper der Passagiere?“ mit „Ja“ beantwortet, so bedarf es weiterer Folgefragen, die eine gewisse Abstufung zwischen ‚kaum eingreifend‘ bis ‚invasiv‘ erlauben. Ein Beispiel hierfür wäre etwa eine Folgefrage danach, ob dieses Abtasten den gesamten Körper eingeschlossen Intimbereiche betrifft. Für Berührungen durch das Kontrollpersonal können hierdurch gezielte Nachkontrollen, wie sie etwa häufig auf Grundlage eines lokalisierten Alarms in einem Körperscanner durchgeführt werden, von Nachkontrollen des gesamten Körpers unterschieden werden. Auch die in Abschnitt 3.2.2 angesprochenen Aspekte, ob gezielte Berührungen unterhalb von bedeckender Kleidung möglich sind und ob die Kontrollen gleichgeschlechtlich durchgeführt werden, lassen sich über Entscheidungsfragen zur qualitativen Differenzierung einsetzen. Dies betrifft auch die Frage, ob Prothesen oder medizinische Geräte bei einer konkreten Kombination von Kontrolltechniken einer ‚Handhabung‘ durch das Personal unterworfen werden,

zu einer übergreifenden Bewertung der Problemkategorie zusammengefasst. In Abschnitt 4.3 werde ich dies anhand eines Beispiels veranschaulichen.

denn auch dieser Umgang ist mittlerweile stark formalisiert.¹⁵⁹ Auch zu dem gegebenenfalls festgestellten Sachverhalt, dass bei einem Kontrollprozess biometrische Daten zu den Passagieren erhoben und gespeichert werden, lassen sich entsprechende Folgefragen formulieren. Qualitative Abstufungen hierzu werden etwa über Entscheidungsfragen zu den beobachtbaren Sachverhalten ermöglicht, ob diese Daten auch nach Beendigung des Kontrollvorgangs gespeichert bleiben und ob dies durch den Passagier kontrolliert werden kann.

Dabei ist es hilfreich, bei der Formulierung der Fragen die verschiedenen Kontrolltechniken und ihre Eigenheiten immer im Blick zu behalten, sowohl die aktuell eingesetzten wie auch mögliche zukünftige Techniken. Dies erlaubt es, die Fragen so zu stellen, dass wichtige qualitative Unterschiede und Gemeinsamkeiten treffend berücksichtigt werden. Zugleich sollten die Fragen aber auch nicht zu technikspezifisch formuliert werden, sondern sich ihrem Fragegehalt hauptsächlich auf den jeweiligen Modus des Problemtypus beziehen, sodass sichergestellt bleibt, dass die beobachtbaren Sachverhalte *problemorientiert* sind und sich die Operationalisierung nicht unversehens zu sehr an rein technischen Eigenheiten ausrichtet.

In jedem Fall wäre es zur Handhabung des operationalisierten Frameworks auch bei der Formulierung dieser Folgefragen sinnvoll darauf zu achten, dass ein „Ja“ durchgängig eine Verschärfung des Wertungskonflikts impliziert. Durch die gewählte Form der Entscheidungsfrage lassen sich die beobachtbaren Sachverhalte dann leichter in eine formallogische Form bringen, denn die binäre Entscheidungsmöglichkeit zwischen „Ja“ und „Nein“ lässt sich dann recht einfach als strukturierte Wertetabellen darstellen. Dabei können die beobachtbaren Sachverhalte mit den einzelnen Kontrollschriften eines Weges durch den Checkpoint abgeglichen werden. Wann immer für eine konkrete Kombination von Kontrolltechniken die Antwort auf eine der Entscheidungsfragen mit „Ja“ beantwortet werden muss, wechselt auch der Wert des entsprechenden beobachtbaren Sachverhalts in der Tabelle für diese Kombination auf „Ja“. Entsprechend können darüber auch verschiedene Wege durch den Checkpoint verglichen werden.

Zur Bewertung der Brisanz eines Problemtypus muss dabei nicht *für jeden identifizierten Modus* eine Abstufung über alle vier Ränge ermöglicht werden. Es ist ausreichend, wenn die Differenzierung der Brisanz eines Wertungskonflikts unter Berücksichtigung *aller identifizierten Modi gemeinsam* möglich ist. Auch hier kann eine Bestwertung aber nicht ausschließlich dann vergeben werden, wenn keiner der beobachtbaren Sachverhalte auf eine Manifestation des Problemtypus hinweist. Wie bereits erwähnt wird jede Form von Fluggastkontrollen private Lebensbereiche verletzen, sofern diese die Grundfunktion der Aufdeckung für jeden Passagier erfüllen sollen, also dem bislang geltenden Grundsatz folgen, dass wirklich *alle* Passagiere

¹⁵⁹ Beispielsweise schließt die US-TSA nach einigen hochproblematischen Vorfällen mittlerweile aus, dass orthopädische Schuhe, Prothesen oder medizinische Geräte entfernt oder berührt bzw. entblößt werden müssen (ALLEN 2010).

kontrolliert werden müssen. Entsprechend muss die qualitative Differenzierung für die Bestwertung ‚A‘ wie auch für die schlechteste Wertung ‚D‘ daran ausgerichtet werden, was realistische *Best-case*- und *Worst-case*-Szenarien bei der Kombination von Kontrolltechniken sind, sowie für die Wertungsstufe ‚C‘ daran, was für alle identifizierten Modi sinnvolle Mindestkriterien für einen Schutz körperlicher Privatheit sein könnten, der sich deutlich vom *Worst-case* unterscheidet.

Tabelle 3 zeigt hierzu ein vereinfachtes Beispiel für zwei dieser Wege, das sich anhand der bereits wieder aufgegriffenen Darstellung in Abb. 1 auf Seite 130 illustrieren lässt. Nach dem Ablegen von Jacken und Handgepäck (nicht aber von bedeckender Kleidung) werden die Passagiere der Primärkontrolle durch eine Metalldetektorschleuse unterzogen. Metalldetektorschleusen nutzen eine Anomalie-Detektion auf Basis des gemessenen Magnetfeldes. Metallische Implantate und Piercings verursachen dabei eine Häufung von Alarmen, deren Zuordnung in diesem Fall aber dadurch erschwert wird, dass der Metalldetektor für einen bestimmten Anteil der Passagiere einen zufälligen Alarm ausgibt, um stichprobenartige Nachkontrollen zu ermöglichen. Weg i stellt jene Kombination von Kontrolltechniken dar, bei der die Metalldetektorschleuse *keinen* Alarm ausgibt und die Passagiere ohne weitere Behelligung auf die Luftseite überreten können.¹⁶⁰

Weg ii hingegen soll einen Weg durch den Checkpoint beschreiben, bei dem der Metalldetektor einen Alarm ausgibt und die Passagiere einer Sekundärkontrolle durch einen modernen Körperscanner unterzogen werden. Dieser arbeitet ebenfalls auf Basis einer automatisierten Anomalie-Detektion und gibt entsprechend kein Bild aus. Diese Kontrolltechnik löst jedoch bei verbreiteten medizinischen Produkten wie Schutzhosen bei Inkontinenz vermehrt Fehlalarme aus. Da die Alarmrate bei diesen Geräten aber ohnehin sehr hoch ist und *aus verschiedensten Gründen* Fehlalarme ausgegeben werden, kommt es hierbei zum gleichen Effekt wie beim Metalldetektor: Von einem Alarm kann nicht ohne weiteres auf die Ursache rückgeschlossen werden. Auf Weg ii soll auch diese Kontrolltechnik einen Alarm ausgeben. Weitere Nachkontrollen erfolgen dann in Form eines manuellen, klassischen Abtastens des gesamten Körpers über der Kleidung durch gleichgeschlechtliches Kontrollpersonal. Dieser Kontrollschnitt kann den Alarm auflösen und die Passagiere, die auf diesem Weg kontrolliert werden, können auf die Luftseite überreten.

In Tabelle 3 lassen sich die beiden Wege durch den Checkpoint anhand beobachtbarer Sachverhalte hinsichtlich des Eingriffs in körperliche Aspekte von

¹⁶⁰ Wie in Abschnitt 2.2.2 dargestellt, trifft dies für alle Passagiere zu, bei denen der Detektor *nicht anschlägt* – also Passagiere *mit und ohne* metallische Gegenstände, bei denen der durch das Alarmkriterium definierte Schwellenwert bei der Messung durch den Metalldetektor nicht überschritten wurde. Entsprechend gilt für Weg ii, dass hier dieser Schwellenwert überschritten wurde und der Metalldetektor somit einen Alarm ausgibt – sei es, weil der Passagier tatsächlich einen metallischen Gegenstand bei sich führt, oder sei es aufgrund eines falschen Alarms.

Privatheit nun miteinander vergleichen. Dabei ist gut ersichtlich, dass für Weg i keine der Entscheidungsfragen mit „Ja“ beantwortet wurde, die nicht auch für Weg ii mit „Ja“ beantwortet wurde. Gleichzeitig weisen für Weg ii weitere beobachtbare Sachverhalte auf eine Manifestation des Problemtypus durch den Modus der Berührung hin. Dies mag als eine recht triviale Beobachtung anmuten, denn die Kontrolltechnik, die Weg i ausmacht, ist auch Teil des Kontrollprozesses auf Weg ii, welche wiederum ein Abtasten mit einschließt. Dennoch lässt

Tabelle 3: Beobachtbare Sachverhalte zur Bewertung der Verletzung körperlicher Privatheit für ein einfaches Beispiel

Modus	Beobachtbarer Sachverhalt als Entscheidungsfrage ¹⁶¹	Weg i	Weg ii
Berührung von Passagieren	Does the screener touch the passenger's body?	Nein	Ja
	If the screener touches the passenger's body, do they touch beneath covering clothes or is it an enhanced frisk search?	Nein	Nein
	If the screener touches the passenger's body, is it a full search or are intimate zones included in a directed search?	Nein	Ja
	If the screener touches the passenger's body, does the screener have a different sex than the passenger?	Nein	Nein
	Does the screener handle prostheses or other body extensions? ¹⁶²	Nein	Nein
Entfernung von bedeckender Kleidung	Is the passenger required to divest covering clothes?	Nein	Nein
	Is an image of the body made visible to the screener?	Nein	Nein
	If an image of the body is made visible to the screener, are intimate zones included in the image (e.g. full-body image)?	Nein	Nein
	If an image of the body is made visible to the screener, can the screener see details intimate areas?	Nein	Nein

¹⁶¹ Wie an einigen Stellen bereits erwähnt war die Entwicklung dieses Frameworks Teil des EU-Projekts XP-DITE, dessen Arbeitssprache Englisch ist. Entsprechend wurden die insgesamt über 150 Entscheidungsfragen sowie die Bewertungsregeln auf Englisch formuliert. Auch an dieser Stelle lässt sich gut einsehen, dass eine durchgängig formallogische Darstellung der Regeln zur Bewertung sowie zur Zusammenfassung von Teilstudien es ermöglicht, die Auswertung auch computerunterstützt durchzuführen. Gerade angesichts der großen Zahl an zu beantwortenden Entscheidungsfragen würde sich eine solche Unterstützung, wie sie im Rahmen von XP-DITE angedacht ist und aktuell entwickelt wird, als sehr hilfreich erweisen. Die Beantwortung zumindest eines großen Teils der Entscheidungsfragen könnte dann automatisiert über eine Abfrage der eingesetzten Techniken beantwortet werden, zu denen hinterlegt ist, für welche beobachtbaren Sachverhalte sie den Wert von „Nein“ zu „Ja“ verändern.

¹⁶² Ich nehme hier den verbreiteten Fall an, dass für Personen mit Prothesen und medizinischen Geräten angepasste Kontrollprozesse zur Verfügung stehen, weshalb diese Frage hier mit „Nein“ beantwortet werden kann. Ob das Personal bei diesen angepassten Kontrollprozessen beispielsweise Prothesen berührt oder gar wie Handgepäck „handhabt“, würde dann für diese Kontrollprozesse gesondert bewertet werden. Ob dadurch diskriminierende Effekte entstehen, wird hingegen über die Problemkategorie ‚Schlechterstellung gesellschaftlicher Gruppen‘ berücksichtigt.

Modus	Beobachtbarer Sachverhalt als Entscheidungsfrage ¹⁶¹	Weg i	Weg ii
Erheben von körperbezogenen Informationen	Is trace detection ¹⁶³ used on the body?	Nein	Nein
	If trace detection is used on the body, do certain known medicinal products or drugs cause an alarm?	Nein	Nein
	If trace detection is used on the body, is the alarm rate below 15% (taking a combination with a random alarm into account)?	Nein	Nein
	Is anomaly detection ¹⁶⁴ used on the body?	Ja	Ja
	If anomaly detection is used on the body, will medical implants, piercings, amputations, etc. cause an alarm?	Ja	Ja
	If anomaly detection is used on the body, is the alarm rate below 15% (taking a combination with a random alarm into account)?	Nein	Nein
Erheben biometrischer Informationen	Is biometric screening data stored?	Nein	Nein
	If biometric screening data is stored, is it not stored in a passenger-controlled (e.g. token based) way?	Nein	Nein
	If biometric screening data is stored, is it stored after the passenger is cleared in a non-passenger controlled way?	Nein	Nein

sich hier gut aufzeigen, dass ein vergleichender Zugriff auf wertungsrelevante beobachtbare Sachverhalte über solche Wertetabellen für eine Operationalisierung logisch formalisiert werden kann. Zudem kann auf Grundlage dieser formalisierten Informationen festgestellt werden, dass Weg ii hinsichtlich der Verletzung körperlicher Aspekte von Privatheit nicht besser bewertet werden kann als Weg i.

Weg i, so lässt sich mit Blick auf realistische Kontrollszenarien feststellen, liegt hinsichtlich des Eingriffs in körperliche Aspekte von Privatheit recht nahe am Bestfall: Die Grundfunktion der Aufdeckung beschränkt sich hierbei ausschließlich auf die Anomaliedetektion, wobei aufgrund der hohen Alarmraten kaum verlässliche Rückschlüsse auf Gründe für einen Alarm gezogen werden können. Entsprechend könnte für diese Kombination von Kontrolltechniken die Bestwertung ‚A‘ vergeben werden.

Für Weg ii kann dies aufgrund des manuellen Abtastens am ganzen Körper, inklusive der Intimregionen, sicherlich nicht gelten. Doch liegt dieser Weg auch nicht nahe am *Worst-case*, denn Passagiere werden weder unterhalb bedeckender Kleidung berührt, noch müssen sie das intensivere Abstreichen aller Körper- und Intimregionen mit der Hand (*frisk search*) erdulden. Entsprechend kommt auch die Wertungsstufe ‚D‘ nicht in Frage. Ein Mindestmaß an Schutz gegen Eingriffe in körperliche Privatheit könnte entsprechend im Ausbleiben solcher invasiven Kontrolltechniken gesehen werden; hier ließe sich also recht gut für die Wertungsstufe ‚C‘ argumentieren, was eine moderat eingreifende Kombination von Techniken bezeichnet. Die Wertungsstufe ‚B‘ ließe sich dann für Fälle reservieren,

¹⁶³ Sprengstoffspuren-Detektion, vergleiche hierzu meine Ausführungen in Abschnitt 2.2.2.

¹⁶⁴ Anomalie-Detektion, vergleiche hierzu meine Ausführungen in Abschnitt 2.2.2.

in denen der Körper der Passagiere zwar berührt wird, in denen diese aber gleichzeitig weder Berührungen unterhalb bedeckender Kleidung noch Berührungen am ganzen Körper (auch der Intimregionen) erdulden müssen. Ähnliche Überlegungen lassen sich auch zu den anderen beobachtbaren Sachverhalten darlegen.

Auf Grundlage von in Tabellen zusammengestellten ‚Ja‘- und ‚Nein‘-Werten zu den beobachtbaren Sachverhalten lassen sich demnach Bewertungsregeln sehr genau definieren. Wie schon bei den Regeln zur Zusammenfassung von Teilwertungen, lassen sich zu diesem Zweck für jede der vier Bewertungsstufen Anforderungen formulieren, und zwar unter Rückgriff auf die formallogischen Junktoren ‚UND‘ sowie ‚ODER‘, wobei ein logischer Vorrang wieder über eine Klammer ausgedrückt werden kann. So formulierte Regeln wirken zwar ein wenig sperrig und sind nicht ganz einfach zu lesen, sind aber eindeutig und erlauben sogar eine computerunterstützte Auswertung, wie ich sie in den Fußnoten 155 und 161 bereits angedeutet habe.

Dass bei der Formulierung solcher Bewertungsregeln (wie schon zuvor auch von Regeln zur Zusammenfassung von Teilwertungen) ein gewisses subjektives Element mitschwingt, ist unübersehbar. Dies gilt im Speziellen für die Frage danach, worin wir – abgesehen von einem deutlichen qualitativen Unterschied in der Brisanz zum *Worst-case* – Anforderungen für einen ‚Mindestschutz‘ sehen können.¹⁶⁵ Entsprechend stellt sich auch an dieser Stelle die Frage, wie die Tauglichkeit der hier entwickelten Bewertungsregeln aufgezeigt werden kann. Dabei gilt entsprechend wieder die oben gemachte Einschränkung, dass angesichts der noch recht frühen Entwicklungsphase der Operationalisierung sich die Tauglichkeit erst in der Anwendung in möglichst vielen Fällen erweisen wird. Allerdings lässt sich hier nochmals festhalten, dass das Framework grundsätzlich offen ist für die Identifizierung neuer beobachtbarer Sachverhalte in der Form von zusätzlichen Entscheidungsfragen, wie auch für Kritik hinsichtlich der Bewertungsregeln. Tabelle 4 zeigt meinen Vorschlag für die entsprechenden Bewertungsregeln zum Eingriff in körperliche Aspekte von Privatheit auf Basis der im Rahmen des EU-Projekts XP-DITE durchgeführten vergleichenden Bewertungen sowie auf Basis meiner hierbei gesammelten Erfahrungen.

Derartige auf Entscheidungsfragen beruhende Bewertungsregeln lassen sich auch für die Problemtypen der anderen Problemkategorien formulieren, wobei es noch einige Besonderheiten zu beachten gilt. Zum einen wird es einige Fragen geben, die für alle Passagiere – unabhängig vom Weg durch den Checkpoint –

¹⁶⁵ So könnte man etwa kritisch fragen, ob nicht schon das Abtasten am ganzen Körper dieses Mindestmaß an Schutz unterläuft. Schließlich werden bei dieser Kontrolltechnik auch Intimregionen durch das Kontrollpersonal berührt. Mir schien es an dieser Stelle aber wichtiger, den Unterschied zur Kontrolltechnik eines sehr zudringlichen Abstreichen des Körpers in den Wertungsstufen widerzuspiegeln (diese Kontrolltechnik löste, wie ich in Fußnote 81 erwähnt habe, bei der Einführung in den USA heftige Proteste aus). Zudem kommt diese Kontrolltechnik so verbreitet zum Einsatz, dass eine Differenzierung von Systemen zur Fluggastkontrolle nicht mehr wie intendiert möglich würde.

Tabelle 4: Regeln zur operationalisierten Bewertung der Brisanz des Eingriffs in körperliche Privatheit

Qualität des Eingriffs in die körperliche Privatheit der Passagiere	Anforderungen für die Wertungsstufen
A Die Passagiere werden über eine Kombination von Techniken nahe am <i>Best-case</i> kontrolliert, die für Passagiere <i>kaum</i> Eingriffe in körperliche Aspekte von Privatheit bedeuten.	<ul style="list-style-type: none"> The passengers' body is NOT touched AND body extensions are NOT handled by the screener AND passengers are NOT required to divest covering clothes AND an image of the body is NOT visible to the screener AND (trace detection is NOT used OR (used AND (certain medicinal products or drugs are NOT known to cause alarms OR a combination with a random alarm is NOT absent))) AND (anomaly detection is NOT used OR (used AND (common body modifications do NOT cause an alarm OR a combination with a random alarm is NOT absent))) AND biometric data is NOT stored.
B Die Passagiere werden über eine Kombination von Techniken kontrolliert, die für Passagiere nur <i>geringfügige</i> Eingriffe in körperliche Aspekte von Privatheit bedeuten.	<ul style="list-style-type: none"> The passengers' intimate zones are NOT touched AND they're NOT frisked or touched beneath covering clothes AND the screener does NOT have a different sex AND body extensions are NOT handled by the screener AND passengers are NOT required to divest covering clothes AND an image of the body is NOT visible to the screener AND (trace detection is NOT used OR (used AND (certain medicinal products or drugs are NOT known to cause alarms OR a combination with a random alarm is NOT absent))) AND (anomaly detection is NOT used OR (used AND (common body modifications do NOT cause an alarm OR a combination with a random alarm is NOT absent))) AND biometric data is NOT stored after the screening in a non-passenger-controlled way.
C Die Passagiere werden über eine Kombination von Techniken kontrolliert, die ein Mindestmaß an Schutz für körperliche Aspekte von Privatheit bieten und höchstens <i>moderate</i> Eingriffe bedeuten.	<ul style="list-style-type: none"> Passengers are NOT frisked or touched beneath covering clothes AND the screener does NOT have a different sex AND body extensions are NOT handled by the screener AND the passenger is NOT required to divest covering clothes AND the screener can NOT see details of intimate areas AND biometric data is NOT stored after the screening in a non-passenger-controlled way.

Qualität des Eingriffs in die körperliche Privatheit der Passagiere	Anforderungen für die Wertungsstufen
<p>D Die Passagiere werden über eine Kombination von Techniken nahe dem <i>Worst-case</i> kontrolliert, die einen <i>invasiven</i> Eingriff in körperliche Aspekte von Privatheit bedeuten.</p>	<ul style="list-style-type: none"> • Passengers are frisked or touched beneath covering clothes • OR (they are touched AND the screener has a different sex) • OR body extensions are handled by the screener • OR the passenger is required to divest covering clothes • OR the screener can see details of intimate areas • OR biometric data is stored after the screening in a non-passenger-controlled way.

gleich beantwortet werden. Dies betrifft etwa Fragen nach der Wartezeit der Passagiere, die für Wertungen hinsichtlich der Behelligung von Reisenden (Abschnitt 3.4.1) wichtig ist. Für die Bewertung stellt dies aber keine Hürde dar, denn die entsprechenden Sachverhalte können jeweils der Methode entsprechend bei allen Wegen berücksichtigt werden.

Weiterhin habe ich in Abschnitt 3.2 ausgeführt, dass Fragen nach diskriminierenden Effekten einen gruppenspezifischen Vergleich implizieren, etwa hinsichtlich der Frage, ob gehbehinderte Passagiere, die über gesonderte Prozesse kontrolliert werden, beim Eingriff in private Lebensbereiche schlechter gestellt werden als die Menge der anderen Passagiere. Derartige Vergleiche lassen sich auf Basis der Einzelwertungen zu den unterschiedlichen Wegen durchführen. So lässt sich beispielsweise die Frage stellen, ob alternative Kontrollprozesse stärker in die Privatheit eingreifen als die Standardprozesse für all jene, die nicht gehbehindert sind. Für die Operationalisierung heißt dies, dass die Wege für die Standardkontrollen und die alternativen Kontrollen hinsichtlich des Eingriffs in private Lebensbereiche zunächst getrennt bewertet und die Einschätzungen sodann verglichen werden. Wurden die alternativen Kontrollprozesse schlechter bewertet als die Standardkontrollprozeduren, so wird die entsprechende Frage nach einem stärkeren Eingriff mit „Ja“ beantwortet. Folgefragen können dann klären, ob die Prozesse sich um zwei oder gar drei Wertungsstufen unterscheiden (siehe hierzu etwa Tabelle 11 im Anhang auf Seite 304).

Für die Operationalisierung der sozialethischen Risikokategorie der Schlechterstellung verletzlicher Gruppen stellt sich dabei eine überaus wichtige Frage: Auf welchen Wegen durch den Checkpoint manifestiert sich die Schlechterstellung und wo soll sie entsprechend für die Wertung berücksichtigt werden? Für wen manifestiert sich, so könnte man die Frage noch klarer formulieren, das sozialethische Risiko der Schlechterstellung? Zunächst könnte man versucht sein, eine solche Manifestation nur für diejenigen Passagiere anzuerkennen, die auch wirklich den alternativen Kontrollprozessen unterzogen werden. Aus einer rein utilitaristischen Perspektive schiene dies sogar fast unumgänglich, denn nur diese Personen würden die für diesen Ansatz relevante ‚Unlust‘ auch tatsächlich erfahren.

Über den moralpragmatischen Ansatz hatten wir eine solche Problemkonstellation jedoch als sozialethisches Risiko bezeichnet, entsprechend geht es hierbei um

einen Wertungskonflikt, der *sozial geteilt* erlebt wird. Nicht die konkrete Lust oder Unlust bei einzelnen Personen ist für diesen Wertungskonflikt relevant, sondern die *gemeinbin* geteilte wertungsrelevante Überzeugung, dass staatliche Maßnahmen zum Schutz von Passagieren einzelne verletzliche Gruppen nicht schlechterstellen dürfen – zumal wenn diese ohnehin als verletzlich gelten.¹⁶⁶ Obwohl die Schlechterstellung selbst also zweifelsohne nur jene als verletzlich geltende gesellschaftliche Gruppe betreffen würde, betrifft der hier relevante Wertungskonflikt durchaus die Gruppe der Passagiere in ihrer Gesamtheit. Sofern Standardkontrollverfahren im konkreten Fall also besser bewertet werden als mögliche alternative Kontrollprozesse, sollte die Operationalisierung der Schlechterstellung also im Ergebnis eine Herabstufung der *Standardkontrollen* vorsehen.

Aus den oben genannten Gründen werde ich darauf verzichten, die Bewertungsregeln zu jedem einzelnen Problemtypus an dieser Stelle genauer darzulegen und verweise hierfür auf die Dokumentation in Anhang A sowie auf den nächsten Abschnitt, der die Bewertungsprozedur anhand eines Anwendungsbeispiels anschaulich macht.

4.3 Das Framework in der gesellschaftspolitischen Anwendung

4.3.1 Bewertung eines einfachen Beispielcheckpoints

Kontrollprozesse eines einfachen Beispielcheckpoints

In diesem Abschnitt möchte ich anhand eines recht einfach aufgebauten Checkpoints darstellen, welche Informationen zu den in Kapitel 3 identifizierten sozial-ethischen Wertungskonflikten sich über das Framework konkret ausweisen lassen. Dabei werde ich zeigen, dass man von den Bewertungsergebnissen ausgehend über das Framework zentrale Wertungskonflikte identifizieren und analysieren kann. Im nächsten Abschnitt werde ich dann darstellen, wie sich diese Informationen für gesellschaftspolitische Entscheidungssituationen nutzen lassen, um konkrete Handlungsempfehlungen zur Entschärfung der sozialetischen Risiken zu gewinnen. Hierbei wird sich auch nochmals zeigen, dass sich Wertungskonflikte nicht einfach auflösen lassen. Was aber über das Framework durchaus gelingen

¹⁶⁶ Ganz ähnlich lässt sich im Moralpragmatismus recht einleuchtend argumentieren, dass von der Diskriminierung religiöser Minderheiten nicht nur die Angehörigen dieser Minderheiten selbst betroffen sind, sondern wir alle als Gesellschaft. Entsprechend steht hierbei auch unser aller Religionsfreiheit auf dem Spiel, nicht nur die Freiheit dieser Minderheiten. Dies lässt sich auch daran erkennen, dass einzelne Personen bei solchen Diskriminierungen auch immer ein Stück weit in ihrer Freiheit eingeschränkt werden, sich durch eine Religion bekehren zu lassen, wenn diese hierfür auch eine Reihe von Schlechterstellungen in Kauf nehmen müssen – unabhängig davon wie plausibel oder unplausibel eine Bekehrung für den Einzelnen konkret sein mag.

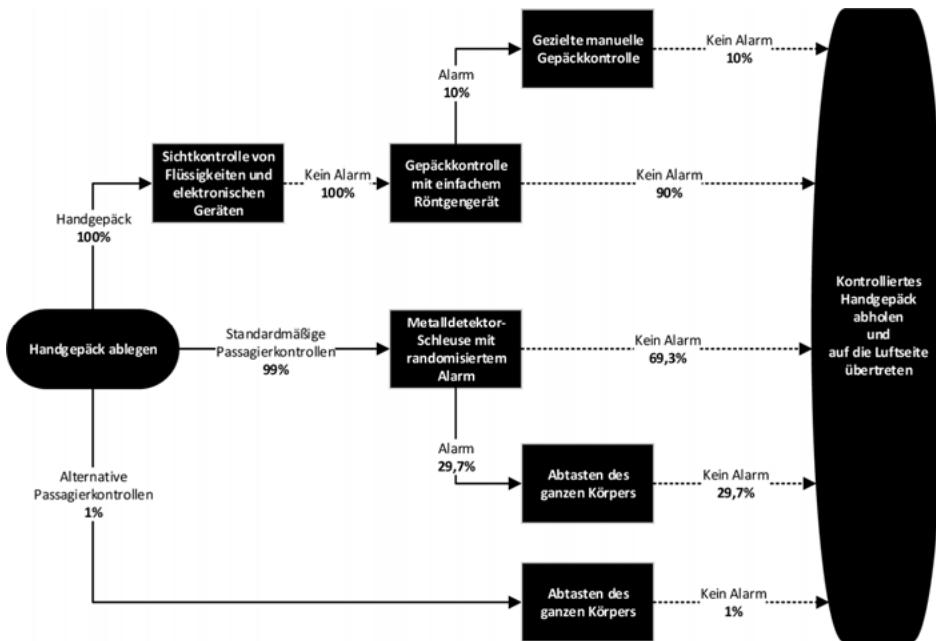


Abb. 13: Logik der Kontrollprozesse eines einfachen Checkpoints mit Alarmraten

kann, ist eine Entscheidungsfindung aus sozialethischer Perspektive deutlich besser zu informieren und implizite Trade-offs hierbei explizit zu machen.

Abb. 13 stellt die Entscheidungslogik der Kontrollprozesse eines einfachen Checkpoints dar. Mit angegeben sind auch die Anteile der jeweils betroffenen Passagiere.¹⁶⁷ Der Kontrollprozess beginnt damit, dass alle Passagiere zunächst Jacken und Handgepäck ablegen. Dabei sollen Laptops und Tablets sowie Flüssigkeiten aus den Gepäckstücken herausgenommen werden. Medizinische Geräte können im Gepäck verbleiben, flüssige Medikamente müssen aber zusammen mit anderen Flüssigkeiten separat abgelegt werden. Bedeckende Kleidung muss nicht abgelegt werden, wohl aber Gürtel mit metallischer Schnalle. Die meisten Passagiere (99%) werden nun im Standard-Kontrollverfahren zunächst über eine Metalldetektorschleuse kontrolliert. Für den Großteil der so kontrollierten Passagiere (70%) gibt der Detektor dabei keinen Alarm aus. Für sie ist die Personenkontrolle somit bereits beendet. Hieraus ergibt sich für die Personenkontrollen bereits ein erster Weg durch den Checkpoint, der von einem großen Teil der Fluggäste ($99\% \times 70\% = 69,3\%$) durchlaufen wird. Für einen kleineren Teil der Fluggäste (30%) gibt die Metalldetektorschleuse aber einen Alarm aus. Hierin sind nicht nur Fehlalarme eingeschlossen, sondern auch auf zufälliger Basis aus-

¹⁶⁷ Die hierbei zugrunde gelegten Alarmraten wurden von mir beliebig festgelegt, da die tatsächlichen Alarmraten der verschiedenen Technologien als vertraulich eingestuft und entsprechend nicht öffentlich zugänglich sind.

gegebene Alarme, über die stichprobenartige Nachkontrollen veranlasst werden. Die Nachkontrollen bestehen aus einem manuellen Abtasten des ganzen Körpers. Dies betrifft auch beispielsweise religiöse Kopfbedeckungen. Für die Personenkontrollen lässt sich in dieser Kombination aus Metalldetektor und Abtasten demnach ein zweiter Weg durch den Checkpoint identifizieren, auf dem ein weiterer Teil der Fluggäste ($99\% \times 30\% = 29,7\%$) kontrolliert wird.

Die so definierten Standard-Kontrollverfahren sind allerdings nicht für alle Passagiere geeignet. Fluggäste, die nicht über eine Metalldetektorschleuse kontrolliert werden können, müssen entsprechend über alternative Prozesse kontrolliert werden (1%). Dies betrifft beispielsweise Personen mit Krücken oder Rollstuhl, Personen mit Herzschrittmacher, wie auch Personen mit größeren Prothesen oder medizinischen Geräten. Diese Passagiere werden durchweg über ein manuelles Abtasten des ganzen Körpers kontrolliert, worin mit Blick auf die Personenkontrollen somit ein dritter Weg durch den Checkpoint besteht.¹⁶⁸

Parallel zur Personenkontrolle auf diesen drei Wegen wird auch noch das abgelegte Handgepäck kontrolliert. Zunächst unterzieht das Kontrollpersonal die aus den Taschen herausgenommenen Flüssigkeiten und elektronischen Geräte einer Sichtkontrolle, wobei es diese auch zum Teil auch berührt und näher untersucht. Anschließend wird das Handgepäck über ein einfaches Röntgengerät kontrolliert. Zumeist wird dabei kein Alarm ausgelöst, weshalb ein Großteil der Fluggäste (90%) das kontrollierte Gepäck abholen und auf die Luftseite übertrreten kann. Für einen Teil der Passagiere (10%) wird vom Kontrollpersonal bei den Handgepäckkontrollen aber ein Gegenstand als verdächtig deklariert und es wird entsprechend ein Alarm ausgelöst. Zur Auflösung solcher Alarme werden die entsprechenden Gepäckstücke dann manuell gezielt nach den verdächtigen Gegenständen durchsucht. Insgesamt ergeben sich somit sechs mögliche Wege durch den Checkpoint, denn an jeden der drei Wege für die Personenkontrollen schließen sich jeweils noch zwei mögliche Wege für die Gepäckkontrollen an: Etwa ist ein Teil der Passagiere, die über die alternativen Kontrollverfahren kontrolliert werden, von den gezielten manuellen Nachkontrollen des Handgepäcks betroffen ($1\% \times 10\% = 0,1\%$), ein anderer Teil hingegen nicht ($1\% \times 90\% = 0,9\%$).

Der hier beispielhaft definierte Checkpoint ist freilich recht einfach aufgebaut, spiegelt aber im Großen und Ganzen die weithin bekannten Kontrollprozesse wider, die bis zu den Anschlägen vom 11. September weltweit Anwendung fanden und zum Teil noch heute finden. Vergleicht man die schematische Darstellung der Kontrollprozesse in Abb. 13 mit jener in Abb. 1 auf Seite 130, so wird deutlich, dass die Prozesse des Beispielcheckpoints in einer Hinsicht vereinfacht wurden: Weder für das manuelle Abtasten der Passagiere, noch für die manuelle

¹⁶⁸ Ich nehme hier den gängigen Fall an, dass auch bei den alternativen Kontrollprozessen Prothesen und medizinische Geräte so kontrolliert werden, dass diese nicht berührt werden müssen.

Gepäckkontrolle wurde berücksichtigt, dass auch hier Alarme ausgelöst werden können. Diese Alarme sind freilich eher selten und resultieren, wie in Abschnitt 2.2.1 ausgeführt, zumeist darin, dass ein verbotener Gegenstand konfisziert und die Person bzw. das Gepäckstück erneut manuell kontrolliert wird. Gegebenenfalls werden solche Kontrollen dann auch sehr genau durchgeführt. In sehr seltenen Fällen kann es zudem auch passieren, dass ein Angriff auf die zivile Luftfahrt oder ein anderes kriminelles Vergehen nicht zufriedenstellend ausgeschlossen werden kann. Die entsprechende Person wird vom Kontrollpersonal dann an Polizeikräfte übergeben, die danach gegebenenfalls Ermittlungen aufnehmen. Der Kontrollprozess endet in diesem Fall also, ohne dass ein Übertritt auf die Luftseite gestattet wird. Für unseren Beispielcheckpoint nehmen wir der Übersichtlichkeit halber den durchaus realistischen Fall an, dass intensive Nachkontrollen und Übergaben an Polizeikräfte so selten auftreten, dass sie für die Wertungen über das Framework nicht weiter ins Gewicht fallen.¹⁶⁹ Grundsätzlich ließen sich solche Fälle aber dennoch über weitere Wege durch den Checkpoint darstellen und für die Bewertung entsprechend einbeziehen.

Für die operationalisierte Bewertung dieses Beispielcheckpoints – insbesondere hinsichtlich der Kategorie des Beitrags zur Restriktion offener Gesellschaften – müssen allerdings noch einige weitere Sachverhalte definiert werden, die nicht nur von den konkreten Kontrolltechniken, sondern ebenso vom Passagieraufkommen und von Regelungen am Flughafen abhängig sind: So nehmen wir für unseren einfachen Beispielcheckpoint an, dass Messungen im Betrieb für die Passagiere durchschnittlich eine Wartezeit von etwa acht Minuten ergeben haben. Passagiere, die nicht über die Standardverfahren kontrolliert werden, können hingegen an den Schlangen vorbeigehen und werden ohne nennenswerte Verzögerung kontrolliert. Für die meisten Passagiere sind die Kontrollen dabei in unter 2 Minuten beendet. Für Passagiere, bei denen eine manuelle Gepäckkontrolle durchgeführt wird, dauert der Kontrollprozess durchschnittlich etwa 5 Minuten. Aufgrund dessen, dass das Kontrollpersonal entsprechend Rücksicht auf die jeweiligen Bedürfnisse des Passagiers nehmen muss, gilt dies ebenso für Passagiere, die über die alternativen Wege kontrolliert werden. Finden für diese zudem noch Nachkontrollen des Gepäcks statt, summiert sich die Dauer der Kontrollen auf durchschnittlich 11 Minuten, denn für das Kontrollpersonal gilt die Vorschrift, dass Nachkontrollen des Gepäcks nur im Beisein der Passagiere durchgeführt werden dürfen.¹⁷⁰

¹⁶⁹ Aus den Angaben in Tabelle 1 auf Seite 255 wird ersichtlich, dass dies immer dann der Fall ist, wenn weniger als 0,3% der kontrollierten Passagiere betroffen sind. Begründet habe ich dies an obiger Stelle damit, dass selbst bei Kontrollen nahe dem *Best-case* aufgrund der Grundfunktion des Aufdeckens bei den Kontrollen immer mit einigen wenigen Fällen intensiver Nachkontrollen gerechnet werden muss.

¹⁷⁰ Weitere wertungsrelevante Sachverhalte zu dem einfachen Beispielcheckpoint können über die Tabellen in Anhang B erschlossen werden. Da ich auf diese aber in der Folge nicht näher eingehie, sehe ich von einer weiteren Ausführung an dieser Stelle ab.

Tabelle 5: Gesamt- und Teilwertungen zu den sozialethischen Risiken für den einfachen Beispielpoint

Weg	Kombination der Kontrolltechniken	Passagier-Anteil	Teilwertung zur Brisanz der Problemkategorien und Problemtypen
i	Metalldetektorschleuse; Röntgengerät	62,37 %	Eingriff in Privatheit: A (BABB) Schlechterstellung: B (BCAA) Restriktivität: B (CBAB)
ii	Metalldetektorschleuse; Röntgengerät, gezielte manuelle Gepäckkontrolle	6,93 %	Eingriff in Privatheit: B (CABB) Schlechterstellung: C (CCAA) Restriktivität: B (CBAB)
iii	Metalldetektorschleuse, Abtasten; Röntgengerät	26,73 %	Eingriff in Privatheit: C (BCBB) Schlechterstellung: D (DAAA) Restriktivität: B (CBAB)
iv	Metalldetektorschleuse, Abtasten; Röntgengerät, gezielte manuelle Gepäckkontrolle	2,97 %	Eingriff in Privatheit: C (CCBB) Schlechterstellung: D (DAAA) Restriktivität: B (CBAB)
v	Abtasten (alternative Kontrollen); Röntgengerät	0,09 %	Eingriff in Privatheit: C (BCBB) Schlechterstellung: D (DAAA) Restriktivität: B (CBAB)
vi	Abtasten (alternative Kontrollen); Röntgengerät, gezielte manuelle Gepäckkontrolle	0,01 %	Eingriff in Privatheit: C (CCBB) Schlechterstellung: D (DAAA) Restriktivität: B (CBAB)
Gesamter Checkpoint:		100 %	Eingriff in Privatheit: C Schlechterstellung: D Restriktivität: B

Tabelle 5 zeigt eine Übersicht über die sechs oben ausgearbeiteten Wege dieses einfachen Checkpoints. Die letzte Spalte gibt zudem für jeden Weg die Wertungen hinsichtlich der Brisanz der drei sozialethischen Problemkategorien an. In Klammern sind auch die Wertungen zu den ausdifferenzierten Problemtypen beigegeben, auf denen die Bewertungen der Kategorien jeweils beruhen. Wo ich mich in der Folge auf eine dieser Teilwertungen beziehe, werde ich im Text explizit machen, um welche Teilwertung es sich dabei konkret handelt.¹⁷¹ Die letzte Zeile zeigt

¹⁷¹ Die Reihenfolge der Wertungen entspricht dem Aufbau von Kapitel 3 sowie der Darstellung in Abb. 7 auf Seite 239. Für die Wertungskategorie ‚Eingriff in private Lebensbereiche‘ sind die Wertungen also in folgender Reihung angegeben: Verletzung räumlicher Aspekte von Privatheit, Verletzung körperlicher Aspekte von Privatheit, Verletzung der Privatheit individueller Lebensführung, Öffentliche Preisgabe von Privatheit. Für die Wertungskategorie ‚Schlechterstellung gesellschaftlicher Gruppen‘ sind die Werte in folgender Reihung angegeben: Benachteiligung aufgrund unpassender Standardkontrollverfahren, Benachteiligung aufgrund wiederkehrender Falschalarme, Mangelnder Schutz vor Benachteiligung aufgrund falscher Daten, Verletzung der Privatheit von Nicht-Passagieren. Für die Wertungskategorie ‚Beitrag zur Restriktion offener Gesellschaften‘ sind die Wertungen in folgender Reihung angegeben: Behelligung von Reisenden, Mangelnde Rechenschaft, Mangelnder Schutz vor Datenmissbrauch, Mangelnde Transparenz.

schließlich das Ergebnis der Gesamtwertung für den Beispielcheckpoint für jede der drei Problemkategorien. Die Details zu den einzelnen Wertungsschritten können dem Anhang B entnommen werden, doch möchte ich in Folge einige Punkte explizieren, die sich über eine Analyse der Bewertung herausarbeiten lassen.

Explikation der Bewertung zum Eingriff in die Privatheit der Passagiere

Bei diesem Checkpoint wird der Großteil der Passagiere auf den Wegen i und ii über eine Kombination von Techniken kontrolliert, deren Eingriff in private Lebensbereiche über das operationalisierte Framework als höchstens geringfügig eingestuft wurde. Dies gilt insbesondere für den großen Anteil an Passagieren, die über Weg i kontrolliert werden – mit Blick auf die *körperliche* Privatheit sind die Kontrollen hier sogar nahe am Bestfall (für diesen Weg kann demnach auch zusammengekommen die Bestwertung vergeben werden). Was dies konkret bedeutet, lässt sich unter Rückgriff auf die hermeneutische Ausarbeitung in Abschnitt 3.2.2 situativ für Fluggastkontrollen ausbuchstabieren: Im Verhältnis zu anderen Formen der Fluggastkontrolle werden bei dieser Kombination durch das Kontrollpersonal kaum körperliche Aspekte enthüllt, die der Passagier nicht von sich aus bereits offenbart. Die Entscheidung der Passagiere darüber, welche körperlichen Einblicke sie der Öffentlichkeit gewähren wollen, werden bei den Kontrollen also weitgehend respektiert. Gleiches gilt auch für körperbezogene Informationen etwa zu gesundheitlichen oder ästhetischen Aspekten, die weitgehend unbestimmt bleiben. Der Passagier wird auch nicht aus der Anonymität der Masse herausgehoben, kann also – sofern er dies wünscht – in der Menge der Passagiere untertauchen. Zuletzt hält diese Kombination von Kontrolltechniken auch respektvollen Abstand zum Körper, weil bewusste Berührungen des Passagiers vermieden werden. Dieser hohe Schutz der körperlichen Privatheit ist auf diesem Weg vor allem deshalb gewährleistet, weil der Fluggast keinen Nachkontrollen unterzogen wird.

Auch die räumliche Privatheit des Handgepäcks, die sich über die hermeneutischen Bestimmungen in Abschnitt 3.2.1 weiter ausführen lässt, wird auf diesem Weg weitgehend respektiert, denn Nachkontrollen finden auf Weg i ebenfalls nicht statt. Passagiere müssen hier vergleichsweise geringfügige Eingriffe erdulden, was bedeutet, dass die Abschirmung des Gepäcks vor Blicken und Zugriffen, sowie vor der Erhebung und Speicherung von Informationen zumindest weitgehend in Takt bleibt. Auf Weg ii finden hingegen Nachkontrollen des Handgepäcks statt und die ‚C‘-Wertung zur räumlichen Privatheit weist hierbei auf moderate Eingriffe hin. Zwar wird hierbei die Abschirmung des Gepäcks vor der Öffentlichkeit vergleichsweise stark unterminiert, doch bleiben die Nachkontrollen darauf beschränkt, gezielt verdächtige Gegenstände zu kontrollieren. Derartige Nachkontrollen unterscheiden sich also dennoch deutlich vom *Worst-case* einer nahezu vollständigen Negation dieser Abschirmung – etwa bei einer manuellen Nachkontrolle des gesamten Gepäcks, bei dem die mitgeführte Kleidung sowie

andere Gegenstände, Medikamente etc. vollständig einer Sichtkontrolle durch das Kontrollpersonal unterzogen und dabei in aller Regel auch berührt werden.

Ein solch starker Eingriff in die räumliche Privatheit der Passagiere könnte dann auch nicht mehr, wie es für Weg ii noch der Fall ist, durch die Bestwertung hinsichtlich des Eingriffs in körperliche Privatheit ausgeglichen werden. Für Weg ii werden die Kontrolltechniken mit Blick auf die übergreifende Kategorie des Eingriffs in die Privatheit aber noch auf der Wertungsstufe ‚B‘ eingruppiert, was auf vergleichsweise geringfügige Eingriffe verweist. Damit soll ausgedrückt werden, dass ein Mindestniveau an Respekt für private Lebensbereiche durchgehend erreicht wurde und sich die moderaten Eingriffe auf einige private Lebensbereiche beschränken, dass die Privatheit anderer Lebensbereiche aber gleichzeitig außergewöhnlich deutlich respektiert wird.

Wie sich anhand der untersten, in grau hervorgehobenen Zeile erkennen lässt, wurden die Kontrollprozesse dieses einfachen Beispielcheckpoints *insgesamt* hinsichtlich ihres Eingriffs in private Lebensbereiche der Passagiere aber dennoch auf der Wertungsstufe ‚C‘ eingruppiert. Über das Framework soll damit ausgedrückt werden, dass die Kontrollprozesse dieses Checkpoints für die Passagiere insgesamt vergleichsweise moderate Eingriffe in private Lebensbereiche bedeuten – wobei ein Mindestschutz dennoch gewährleistet ist.¹⁷² Diese Bewertung liegt darin begründet, dass ein recht großer Anteil der Passagiere über eine Kombination von Techniken kontrolliert wird, welche deutlich weniger Respekt vor privaten Lebensbereichen wahren. Für vergleichsweise viele Passagiere bedeuten die Kontrollen also eine ungewollte Offenlegung von Lebensbereichen, deren Abschirmung gegenüber der Öffentlichkeit (sowohl gegenüber anderen Passagieren, wie auch gegenüber dem staatlichen Kontrollpersonal) gemeinhin als schützenswert begriffen wird.

An diese Ausführungen zur Bewertung des einfachen Beispielcheckpoints als System zur Fluggastkontrolle, das für Passagiere moderate Eingriffe in private Lebensbereiche impliziert, schließt sich nun freilich die Frage an, wie denn eine bessere Gesamtwertung möglich wäre. Grundsätzlich, so wurde bereits deutlich, muss hierfür jener Anteil der Passagiere deutlich reduziert werden, der über eine Kombination von Techniken kontrolliert wird, die mit ‚C‘ bewertet wurde, also moderate Eingriffe in private Lebensbereiche bedeutet. Dies betrifft die Wege iii bis vi, wobei Weg iii aufgrund dessen, dass mehr als ein Viertel der Passagiere über ihn kontrolliert wird, hierbei besonders im Fokus der Aufmerksamkeit stehen sollte.

Eine Verbesserung könnte dabei auf zwei Arten geschehen: Entweder müsste der Anteil der Passagiere reduziert werden, die auf diesem Weg kontrolliert werden – etwa indem eine Metalldetektorschleuse eingesetzt wird, die in Kombina-

¹⁷² Wie oben erläutert habe ich zur Vereinfachung festgelegt, dass nur ein minimaler Anteil der Fluggäste dieses Checkpoints über eine invasive Kombination von Techniken nahe dem *Worst-case* kontrolliert wird – was diesen Mindestschutz einlöst.

tion mit den zufällig generierten Alarmen deutlich seltener Nachkontrollen erforderlich macht. Alternativ können aber auch die Techniken für die Nachkontrolle so modifiziert werden, dass sie zumeist geringere Eingriffe in private Lebensbereiche implizieren. Für eine solche Modifikation stellt sich dann aber die Frage, worin sich für Weg iii die Herabstufung auf die Wertungsstufe ‚C‘ hinsichtlich des Eingriffs in private Lebensbereiche begründet. Dies kann anhand der Teilwertungen zu diesem Weg deutlich gemacht werden, denn die Brisanz von drei Problemtypen wird mit der Wertung ‚B‘ als vergleichsweise geringfügig eingeschätzt. Lediglich der Eingriff in körperliche Privatheit wurde über die Wertungsstufe ‚C‘ als brisanterer, moderater Eingriff eingestuft. Dies liegt freilich in den Nachkontrollen begründet, bei denen Passagiere am gesamten Körper abgetastet werden. Wir kennen diese Kombination aus der Metalldetektion und Nachkontrolle durch Abtasten des gesamten Körpers bereits aus Abschnitt 4.2.3, wo es als Beispiel zur Ausarbeitung der Bewertungsregeln gedient hat (Tabelle 3 auf Seite 265). Für die Wege i und ii, bei denen dieses Abtasten nicht zur Anwendung kommt und lediglich über die Metalldetektion einige wenige körperbezogene Informationen erhoben werden, wurde der Eingriff in körperliche Aspekte von Privatheit hingegen mit der Bestwertung ‚A‘ als Kombination nahe dem *Best-case* eingestuft. Damit lässt sich aus dem Framework beispielsweise folgender Hinweis generieren: Zur Entschärfung des sozialethischen Wertungskonflikts zwischen der Gewährleistung von Schutz vor Angriffen auf die zivile Luftfahrt und dem Eingriff in private Lebensbereiche würde es sich bei dem hier vorgestellten Beispielcheckpoint lohnen, bei den Nachkontrollen der Fluggäste das Abtasten des gesamten Körpers zu minimieren bzw. es für möglichst viele Passagiere durch weniger invasive Kontrolltechniken zu ersetzen.

Weiterhin ist bei diesem Checkpoint auffällig, dass selbst dann, wenn das Handgepäck *keinen* Nachkontrollen unterzogen wird, der Eingriff in die räumliche Privatheit nicht besser als ‚B‘ bewertet wird. Dies liegt insbesondere daran, dass *alle* Passagiere beim Ablegen Flüssigkeiten, auch flüssige Medikamente und Ähnliches, sowie elektronische Geräte aus ihren Gepäckstücken herausnehmen, separat ablegen und per Sichtkontrolle inspizieren lassen müssen.¹⁷³ Auch hier wäre Raum zur Verbesserung, etwa indem Techniken genutzt werden, welche solche Eingriffe zumindest für jene Passagiere minimieren, die keinen Nachkontrollen unterzogen werden.

Aus diesen Überlegungen heraus wird ein großer Vorteil der für diese Arbeit gewählten Methode zur Operationalisierung deutlich: Das in Abschnitt 4.2 entwickelte Framework lässt sich sowohl zur *Bewertung* von geplanten oder tatsächlich eingesetzten Systemen zur Fluggastkontrolle einsetzen, als auch dazu, entscheidende beobachtbare Sachverhalte zu benennen, die zu dieser Bewertung ge-

¹⁷³ Für die beobachtbaren Sachverhalte zur Verletzung räumlicher Privatheit siehe Tabelle 19 auf Seite 314 im Anhang B.

führt haben, und deren situative Bedeutung für relevante, sozialethische Risiken hermeneutisch zu erschließen. Hierüber lassen sich dann auch *Handlungsanweisungen für eine Entschärfung* dieser Risiken generieren.

Darauf, dass dies zumeist keine Auflösung des zugrundeliegenden Wertungskonflikts zwischen der Gewährleistung von Schutz gegen Angriffe auf die zivile Luftfahrt, ökonomischen Faktoren und den hier identifizierten sozialethischen Risiken bedeutet, habe ich bereits mehrfach hingewiesen. So würde etwa mit Blick auf den hohen Anteil der Passagiere, der über Weg iii kontrolliert wird und dabei moderate Eingriffe in private Lebensbereiche impliziert, eine simple Reduktion der stichprobenartigen Nachkontrollen, wie ich sie oben bereits angedacht habe, gegebenenfalls zu einer besseren Wertung führen. Gleichzeitig würde dies aber auch implizieren, das Sicherheitsniveau des Checkpoints zumindest für solche Szenarien zu reduzieren, in denen ein Angreifer einen nicht-metallischen verbotenen Gegenstand bei sich führt.

Obwohl diese Wertungskonflikte nicht vollständig auflösbar sind, ermöglicht das Framework dennoch deren Bearbeitung, indem Handlungsanweisungen für die Entschärfung der sozialethischen Risiken generiert werden können, die dann wiederum in Relation gesetzt werden können zum Gewinn oder Verlust an Schutz vor Angriffen, sowie in Relation zu den Kosten für die Umsetzung. Ich möchte dies zunächst noch anhand einiger beispielhaft ausgewählter Aspekte zur Bewertung der anderen beiden Risikokategorien darlegen. Im nächsten Abschnitt werde ich dann zeigen, wie diese Überlegungen für Verbesserung des einfachen Beispielcheckpoints dahingehend eingesetzt werden können, die implizierten sozialethischen Risiken zu entschärfen.

Explikation der Bewertung zur Schlechterstellung verletzlicher Gruppen

Mit Blick auf die Risikokategorie der Schlechterstellung verletzlicher Gruppen wurde das System zur Fluggastkontrolle auf der Wertungsstufe ‚D‘ eingruppiert, was einen Checkpoint nahe dem *Worst-case* beschreiben soll. Zwar wird der Großteil der Passagiere auf den Wegen i und ii über eine Kombination von Techniken kontrolliert, bei denen ein Mindestmaß an Schutz vor Schlechterstellung gewährleistet wurde, doch ist der Anteil der Passagiere, für die dies nicht der Fall ist (Wege iii bis vi), deutlich zu groß für eine bessere Wertung. Bei genauerer Analyse von Tabelle 5 auf Seite 274 zeigt sich, dass sich hier keineswegs alle sozialethischen Problemtypen der Schlechterstellung manifestieren. Vielmehr lässt sich als zentrales Problem eine Schlechterstellung aufgrund unpassender Standardkontrollverfahren ausmachen – alle anderen Problemtypen wurden auf diesen Wegen mit der Bestwertung eingestuft.¹⁷⁴

¹⁷⁴ Die Bestwertung hinsichtlich einer Benachteiligung aufgrund falscher Daten liegt darin begründet, dass bei diesem Beispielcheckpoint keinerlei Passagierdifferenzierung vorge-

In der hermeneutischen Ausarbeitung in Abschnitt 3.3.1 wurden dabei zwei Modi identifiziert, in denen sich dieser Problemtypus bei Fluggastkontrollen situativ manifestieren kann: Zum einen dadurch, dass die Standardkontrollverfahren unpassend sind und alternative Kontrollverfahren die Privatheit der betroffenen Passagiere deutlich stärker verletzten. Zum anderen können die Standardkontrollverfahren aber auch für Angehörige einer als verletzlich geltenden gesellschaftlichen Gruppe einen deutlich größeren Eingriff in deren private Lebensbereiche bedeuten. Hier lohnt es sich, die Details zur Bewertung des einfachen Beispielcheckpoints zu Rate zu ziehen, die ich in Anhang B dieser Arbeit beigegeben habe. Tabelle 23 auf Seite 318 zeigt dabei, dass der zuerst genannte Modus, also die Schlechterstellung durch stärker eingreifende, alternative Kontrollprozesse, hier nicht entscheidend sein kann – keiner der beobachtbaren Sachverhalte weist auf eine Manifestation des Problemtypus in diesem Modus hin. Dies erklärt sich dadurch, dass für die alternativen Kontrollen der Eingriff in private Lebensbereiche gleich bewertet wird wie für die Standardkontrollverfahren: Eine separate Bewertung des Eingriffs in private Lebensbereiche auf den Wegen i bis iv einerseits und auf den Wegen v bis vi andererseits ergibt in beiden Fällen also eine ähnliche Einschätzung, nämlich eine Eingruppierung auf der Wertungsstufe ‚C‘.¹⁷⁵ Mit Blick auf die Ausführungen in Abschnitt 3.3.1 ließe sich also schließen, dass eine Schlechterstellung von Personen mit reduzierter Mobilität vermeint werden kann, weil die Standardkontrollverfahren letztlich ähnlich schlecht eingeschätzt werden.

Für den zweitgenannten Modus weisen die Details zur Bewertung des einfachen Beispielcheckpoints in Tabelle 23 auf Seite 318 für die Wege iii bis vi hingegen

nommen wird und die Kontrollen entsprechend auch nicht von externen Daten abhängen. Zudem kann weitgehend ausgeschlossen werden, dass Nicht-Passagiere Kontrollen unterzogen werden, zumal keine vorgelagerten Kontrollprozesse angewandt werden. Beide Problemtypen werden für alle Wege entsprechend durchgängig mit der Bestwertung ‚A‘ eingestuft. Für die Wege iii bis vi, die hier zu Debatte stehen, ergibt sich zudem eine weitere Bestnote, weil eine Schlechterstellung durch eine Häufung von Nachkontrollen aufgrund wiederkehrender Falschalarme ausgeschlossen werden kann: Entweder dies ist bereits die letzte Stufe der Nachkontrollen oder Nachkontrollen mit einem höheren Eingriff in private Lebensbereiche häufen sich nicht für als verletzlich geltende gesellschaftliche Gruppen. Auf den Wegen i und ii ist dies hingegen nicht der Fall, diese wurden entsprechend auch auf der Wertungsstufe ‚C‘ eingruppiert. Denn hier wird mit dem Metalldetektor eine Kontrollmethode angewandt, die etwa für Menschen mit metallischen Implantaten deutlich häufiger Alarne auslösen – was dann zu Nachkontrollen mit einem stärkeren Eingriff in private Lebensbereiche führt.

¹⁷⁵ Dies lässt sich anhand der Angaben in Tabelle 5 auf Seite 274 zeigen: Die alternativen Kontrollprozeduren auf den Wegen v und vi würden für sich genommen zusammen auf der Wertungsstufe ‚C‘ eingruppiert. Gleches gilt für die Standardkontrollprozesse auf den Wegen i bis iv aufgrund des großen Anteils der Passagiere, die über Weg iii moderate Eingriffe in ihre Privatheit erdulden müssen. Konkret bedeutet dies, dass der Eingriff in private Lebensaspekte bei den alternativen Kontrollen ähnlich stark bewertet wird wie bei den Standardkontrollprozessen. Eine deutliche Schlechterstellung etwa von Passagieren mit eingeschränkter Mobilität durch alternative Kontrolltechniken kann durch das Framework also nicht ausgewiesen werden.

gen einige Probleme aus, wobei der zentrale beobachtbare Sachverhalt die Frage betrifft, ob das Kontrollpersonal religiös bedeutsame Kopfbedeckungen berührt. Wie ich oben expliziert habe, soll bei diesem einfachen Beispielcheckpoint das Abtasten des gesamten Körpers auch religiöse Kopfbedeckungen mit einschließen. Auf allen Wegen durch den Checkpoint, auf denen ein solches Abtasten vorgesehen ist, wird entsprechend eine solche Schlechterstellung einer als verletzlich geltenden, gesellschaftlichen Gruppe (religiöse Minderheit) identifiziert. Die Brisanz eines solchen Eingriffs ist – wie ich in Abschnitt 3.3.1 dargelegt habe – für manche religiöse Gruppen dabei so groß, dass nicht mehr davon ausgegangen werden kann, die entsprechenden Kombinationen von Kontrolltechniken böten ein Mindestmaß an Schutz vor Schlechterstellung.

Wie schon für die privatheitsbezogenen Risikotypen wird also auch hier deutlich, dass sich über das Framework die für die sozialethischen Problemkonstellationen zentralen Sachverhalte recht gut identifizieren lassen. Personen mit Entscheidungskompetenzen darüber, wie die Prozesse zur Fluggastkontrolle gestaltet werden sollen, können also nicht nur eine Bewertung verschiedener Entwürfe solcher Prozesse durchführen, sondern sie können über das Framework jene Sachverhalte und deren moralische Relevanz herausarbeiten, die für eine Bearbeitung und Entschärfung dieses Wertungskonflikts zentral sind. Im Fall dieses einfachen Beispielcheckpoints würde sich entsprechend die Frage stellen, wie die Kontrollen auch weiterhin zuverlässig versteckte verbotene Gegenstände aufdecken können, ohne dass religiöse Kopfbedeckungen berührt werden müssen.¹⁷⁶

Explikation der Bewertung zum Beitrag zur Restriktion offener Gesellschaften

Die schlechte Gesamtwertung des einfachen Beispielcheckpoints hinsichtlich möglicher Schlechterstellungen wiederholt sich mit Blick auf die letzte Risikokategorie nicht. Über das Framework kann das System zur Fluggastkontrolle auf der Wertungsstufe ‚B‘ eingruppiert werden, was einen Checkpoint beschreibt, der sich zwar deutlich vom *Best-case* unterscheidet, aber dennoch nur geringfügig zu einer Restriktion offener Gesellschaften beiträgt. Zur Operationalisierung greift das Framework dabei auf einige beobachtbare Sachverhalte zurück, die den Checkpoint zu meist als Ganzen betreffen, also in aller Regel für alle Wege gleich beantwortet werden. Zwar ist es durchaus denkbar, dass dies nicht immer der Fall sein muss – gerade mit Blick auf risikobasierte Fluggastkontrollen – doch sind solche Kontroll-

¹⁷⁶ Darauf, dass die Berührung von religiösen Kopfbedeckungen durchaus große gesellschaftliche Relevanz hat, habe ich in Abschnitt 3.3.1 bereits hingewiesen. Der heftige Protest der Sikh-Gemeinde in Großbritannien (BRANAGH 2011; NEIYYAR 2011) hat dies deutlich vor Augen geführt. In Reaktion haben die Sicherheitsbehörden alternative Kontrolltechniken definiert, die hierauf Rücksicht nehmen und auf die im nächsten Abschnitt zur Entschärfung der Schlechterstellung ebenfalls zurückgegriffen wird.

konzepte in dem einfachen Beispielcheckpoint nicht vorgesehen. Entsprechend werden auch alle Wege durchgängig auf der Wertungsstufe ‚B‘ eingruppiert.

Bei genauerer Analyse von Tabelle 5 auf Seite 274 zeigt sich, dass die Herabstufung von der Bestwertung im Framework darauf zurückzuführen ist, dass die Kontrollen die Passagiere recht deutlich behelligen und so moderat zur Restriktion der Bewegungsfreiheit beitragen. Dieser Beitrag zu einer Restriktion offener Gesellschaften wird dadurch ein Stück weit ausgeglichen, dass bei diesem Checkpoint vollständig auf den Rückgriff auf und die Speicherung von personenbezogenen Daten verzichtet wird. Entsprechend kann durchweg für alle Kombinationen von Kontrolltechniken der Bestfall eines sehr hohen Schutzes vor Datenmissbrauch garantiert werden.

Unter Rückgriff auf die hermeneutische Ausarbeitung in Abschnitt 3.4.1 lassen sich für die Behelligung von Passagieren durch Fluggastkontrollen dabei zwei relevante Modi identifizieren. Die Kontrollen können sich zum einen dadurch restriktiv auswirken, dass sie einen großen Zeit- oder Bequemlichkeitsverlust implizieren, sodass Passagiere für die Kontrollen mehr Zeit einplanen oder früher am Flughafen ankommen müssen. Doch selbst wenn Systeme zur Fluggastkontrolle die Möglichkeiten vorsehen, Passagieren etwa geringe Wartezeiten und kurze Kontrollen zu garantieren, können Checkpoints dadurch zu einer Restriktion offener Gesellschaften beitragen, dass sie einen hohen Eingriff in private Lebensbereiche zu einer Bedingung des Reisens per Flugzeug machen oder eine Abmilderung der Behelligung von Reisenden an Geldleistungen knüpfen, denn auch solche Praktiken stehen der Idee der Bewegungsfreiheit deutlich entgegen.

Ein Blick in Tabelle 27 auf Seite 324, wo die Details zur Bewertung des einfachen Beispielcheckpoints zusammengestellt sind, macht dabei deutlich, dass sich dieser Problemtypus über beide Modi manifestiert. Neben dem Zeitverlust tragen hier insbesondere die aufwändigen Richtlinien zum Ablegen des Handgepäcks, bei denen einige Gegenstände auch aus den Koffern und Taschen herausgenommen werden müssen, zu einem deutlichen Bequemlichkeitsverlust bei. Zwar ist es bei diesem Checkpoint nicht vorgesehen, dass sich Passagiere von Restriktionen durch Geldleistungen freikaufen können, doch macht der Checkpoint, wie oben bereits deutlich wurde, zumindest für manche Kombinationen von Kontrolltechniken deutliche Eingriffe in private Lebensbereiche zur Bedingung einer Flugreise.

Solche Beiträge für eine Restriktion offener Gesellschaften durch die Behelligung von Reisenden können demzufolge erst dann entschärf werden, wenn es gelingt, sowohl den Bequemlichkeitsverlust beim Ablegen des Handgepäcks als auch den Eingriff in private Lebensbereiche substantiell abzumildern. In der Folge möchte ich darlegen, wie diese durch die Operationalisierung zugänglichen Informationen über den einfachen Beispielcheckpoint genutzt werden können, um diesen hinsichtlich der sozialethischen Risiken zu verbessern.

4.3.2 Handlungsempfehlungen zur Entschärfung der sozialethischen Risiken

Identifikation von Chancen zur Entschärfung sozialethischer Konflikte

Für den im letzten Abschnitt vorgestellten einfachen Beispielcheckpoint konnten über das Framework typische sozialethische Problemkonstellationen benannt werden, die sich für dieses konkrete System von Kontrollprozessen aller Voraussicht nach manifestieren werden. Der hierzu entwickelte Ansatz zur Operationalisierung hat es dabei nicht nur ermöglicht, die relative ‚Schwere‘ dieser Problemkonstellationen einzuschätzen, sondern es zudem auch erlaubt, eine Reihe von relevanten Sachverhalten zu identifizieren, die für eine Entschärfung der sozialethischen Probleme zentral sind.

Mit Blick auf Eingriffe in private Lebensbereiche betrifft dies vor allem eine deutliche Reduzierung desjenigen Anteils an Passagieren, der moderate Eingriffe in ihre körperliche Privatheit durch das Abtasten am ganzen Körper über sich ergehen lassen muss. Weiterhin betrifft dies auch Eingriffe in die räumliche Privatheit durch das separate Ablegen von Flüssigkeiten und einigen elektronischen Geräten. Hinsichtlich der Schlechterstellung von gesellschaftlichen Gruppen, die ohnehin bereits als verletzlich gelten, ließ sich insbesondere das Problem ausmachen, dass ein recht großer Anteil der Passagiere über eine Kombination von Techniken kontrolliert wird, die das Abtasten religiöser Kopfbedeckungen vor sieht – was aus Sicht bestimmter religiöser Überzeugungen inakzeptabel ist und einen deutlich stärkeren Eingriff in private Lebensbereiche bedeutet als das Berühren von Hüten, Mützen und dergleichen. Die alternativen Kontrollprozesse etwa für Passagiere im Rollstuhl waren in der Schwere des Eingriffs hingegen im Ergebnis vergleichbar mit den Standardkontrollverfahren, was nach einer Veränderung der Prozesse auch weiterhin der Fall bleiben sollte. Schließlich zeigte sich in Bezug auf mögliche Beiträge zur Restriktion offener Gesellschaften, dass der Checkpoint hier recht gut abschneidet. Eine Verringerung solcher Beiträge wäre aber durchaus noch dadurch möglich, dass der Bequemlichkeitsverlust durch die recht weitgehenden Regeln zum Ablegen sowie der Eingriff in private Lebensbereiche für möglichst viele Passagiere noch reduziert wird.

Die derzeit gültigen, recht weitgehenden Regeln zum separaten Ablegen von Flüssigkeiten haben dabei den Hintergrund, dass Flüssigsprengstoffe über die verbreiteten Röntgengeräte kaum von anderen Flüssigkeiten unterschieden werden können. Müssten Flüssigkeiten bei den Kontrollprozessen dieses Checkpoints nicht separat abgelegt werden, so könnten Flüssigsprengstoffe ausschließlich dann gefunden werden, wenn diese etwa bei ohnehin veranlassten, gezielten Nachkontrollen des Gepäcks Verdacht erregen. Die sich daraus ergebende recht geringe Erkennungsrate soll über eine Limitierung der mitgeführten Flüssigkeitsmenge

sowie über die Pflicht zum separaten Ablegen verbessert werden.¹⁷⁷ Ob die Detektion solcher Sprengstoffe hierdurch tatsächlich verlässlicher gelingt, ist umstritten. Entsprechend wird, wie ich in Abschnitt 2.2.2 erwähnt habe, seit längerem mit Hochdruck an der Einführung von Flüssigsprengstoff-Detektoren gearbeitet.

Da die durch die aktuellen Regelungen entstehenden Verzögerungen bei der Abfertigung von Passagieren für die Flughäfen zum Teil deutliche höhere Kosten implizieren, sind aus ökonomischer wie auch ethischer Perspektive insbesondere solche Geräte interessant, die es nicht mehr erforderlich machen, Flüssigkeiten zur Detektion von Sprengstoffen aus den Gepäckstücken herauszunehmen (*Liquid Explosive Detection System*, LEDS, Typ D). Noch vielversprechender sind hierbei Geräte, bei denen zudem auch komplexere elektronische Geräte wie Laptops im Gepäck verbleiben können (LEDS Typ D+), sodass es nicht mehr nötig ist, Gegenstände aus dem Gepäck herauszunehmen und separat abzulegen (ECAC 2016: 25). Einige dieser noch recht neuen Geräte sind durch die entsprechenden europäischen Behörden mittlerweile für Kontrollen an europäischen Flughäfen zugelassen (ECAC 2016: 1, 12–13), sodass deren Einsatz eine gute Möglichkeit zur Entschärfung einiger sozialethischer Problemkonstellationen ist: Zum einen mildert dies den Eingriff in räumliche Aspekte von Privatheit für eine große Anzahl an Passagiere ab, zum anderen verringert es auch den Aufwand und den Bequemlichkeitsverlust, der die Bewegungsfreiheit der Passagiere einschränkt.

Wie bereits erwähnt werden seit einiger Zeit auch Körperscanner mit automatisierter Bedrohungserkennung (*automatic threat detection*, ATD) eingesetzt, die ohne die Ausgabe eines Körperbilds arbeiten. Ähnlich wie die Metalldetektoren führen auch die Körperscanner bei bestimmten medizinischen Gegebenheiten zu einer Häufung von Fehlalarmen. Im Unterschied zu den Metalldetektorschleusen erlauben es die Körperscanner aber, wie in Abschnitt 2.2.2 dargelegt, lokalisierte Alarme auszugeben, also jene Körperzonen anhand einer Silhouette anzuzeigen, bei denen vermeintlich ein Gegenstand unter der Kleidung erkannt wurde. Nachkontrollen durch das manuelle Abtasten können entsprechend auf bestimmte Körperregionen beschränkt werden. Derartige gezielte Nachkontrollen sparen dabei nicht nur Zeit und somit Kosten, sondern verringern auch den Eingriff in die körperliche Privatheit der Passagiere deutlich – zumal das Abtasten von Intimbereichen so deutlich reduziert werden kann. Aus einer Sicherheitsperspektive mag es freilich problematisch sein, auf die manuellen Nachkontrollen am ganzen Körper vollständig zu verzichten. Je nachdem, wie zuverlässig die Körperscanner arbeiten, das heißt je nachdem wie hoch die Verpasser-Rate¹⁷⁸ für verschiedene Angriffsszenarien ist, kann das Abtasten auch am ganzen Körper

¹⁷⁷ Wie ich in Abschnitt 2.2.2 ausgeführt habe, wurde diese Maßnahme in Reaktion auf einen im Juni 2006 vereiteten Angriff mit Flüssigsprengstoff auf sieben Flugzeuge eingeführt.

¹⁷⁸ Wie in Abschnitt 2.2.2 erläutert wird die Verpasser-Rate – auch Falsch-negativ-Rate genannt – ermittelt, indem die relative Häufigkeit bestimmt wird, bei der ein verbotener Gegenstand anwesend ist, aber kein Alarm ausgegeben wird.

einschließlich der Intimbereiche als zusätzliche Maßnahme stichprobenartig eingesetzt werden. Solange dies deutlich weniger Passagiere betrifft als bei dem einfachen Beispielcheckpoint, würde sich der Eingriff in die körperliche Privatheit durch das Gesamtsystem dennoch deutlich verringern lassen.

Eine solche Verbesserung sollte nach Möglichkeit dann aber auch für die alternativen Kontrollprozeduren vorgesehen werden, sodass beispielsweise Passagiere im Rollstuhl nicht gegenüber der Mehrheit der Passagiere schlechtergestellt werden. Körperscanner auch für die alternativen Kontrollen vorzusehen, wird in diesem Fall aber nicht funktionieren; zwar können Personen mit Herzschrittmacher die Körperscanner problemlos nutzen, für viele andere Personengruppen – etwa Rollstuhlfahrer – sind die Geräte aber nicht geeignet. Gelingt es im Einzelfall nicht, die alternativen Kontrollen in ähnlicher Weise zu verbessern, so kann man von einem Trade-off zwischen verschiedenen sozialethischen Problemtypen sprechen: In einem solchen Fall kann der Eingriff in private Lebensbereiche der Passagiere reduziert werden, doch kommen all jene, für die die Standardkontrollen unpassend sind, nicht im gleichen Maße in den Genuss dieser Verbesserung.

Dies bedeutet aus moralpragmatischer Perspektive natürlich nicht, dass ein solcher Trade-off grundsätzlich unzulässig wäre. Über den in dieser Arbeit entwickelten Ansatz lassen sich solche Wertungskonflikte aber benennen und hermeneutisch ausbuchstabieren. Die wertungsrelevanten Konsequenzen der Entscheidung, diesen Trade-off einzugehen, können so analysiert und strukturiert in die Abwägungen einbezogen werden. Die letztendliche Entscheidung könnte sich – wenn man Dewey diesbezüglich folgt – an der Überlegung orientieren, ob wir unserer Gesellschaft einen Charakter geben wollen, der dem Respekt vor privaten Lebensbereichen einen großen Stellenwert einräumt, aber dabei zugleich bewusst in Kauf nimmt, dass von einer Verbesserung nicht alle gesellschaftlichen Gruppen gleichermaßen profitieren. Dass ein solcher Wertungskonflikt nicht einfach rational aufgelöst werden kann, ist einer der Kerngedanken des Moralpragmatismus nach Dewey und damit auch des in dieser Arbeit entwickelten ethischen Ansatzes.

Allerdings verweisen derartige moralische Dilemmata zumeist auch auf die Frage, ob es nicht noch andere technische Lösungsmöglichkeiten gibt, um eine ähnliche Verbesserung hinsichtlich des Respekts vor privaten Lebensbereichen auch für die alternativen Kontrollprozesse zu erreichen. Eine mögliche Option besteht dann etwa darin, dass Rollstuhlfahrer sich vor den Kontrollen in flughafeneigene speziell angepasste Rollstühle umsetzen – was zumeist ohnehin der Fall ist, wenn der eigene Rollstuhl mitgenommen werden soll.¹⁷⁹ Für die eigentlichen Personenkontrollen kann dann, wie es zum Teil bereits geschieht, auf ein Abtasten des gesamten Körpers verzichtet und auf Techniken wie die Sprengstoffspuren-Detektion zurückgegriffen werden. In den USA wird diese Form der

¹⁷⁹ Für das Boarding werden in aller Regel spezielle Rollstühle genutzt, der eigene Rollstuhl kann dann eingekennzeichnet und als Fracht im Flugzeug mitgeführt werden.

Detektion zum Teil auch für medizinische Geräte am Körper der Passagiere, Urinbeutel etc. angewandt, bei der sich Personen verdeckt „selbst abtasten“ und anschließend deren Hände auf Sprengstoffspuren kontrolliert werden (US TSA 2016a). Gegebenenfalls kann diese Technik auch mit dem Einsatz von Metalldetektor-Handgeräten kombiniert werden, wie es in der EU für alternative Kontrollverfahren explizit vorgesehen ist (EU 2015: 15).

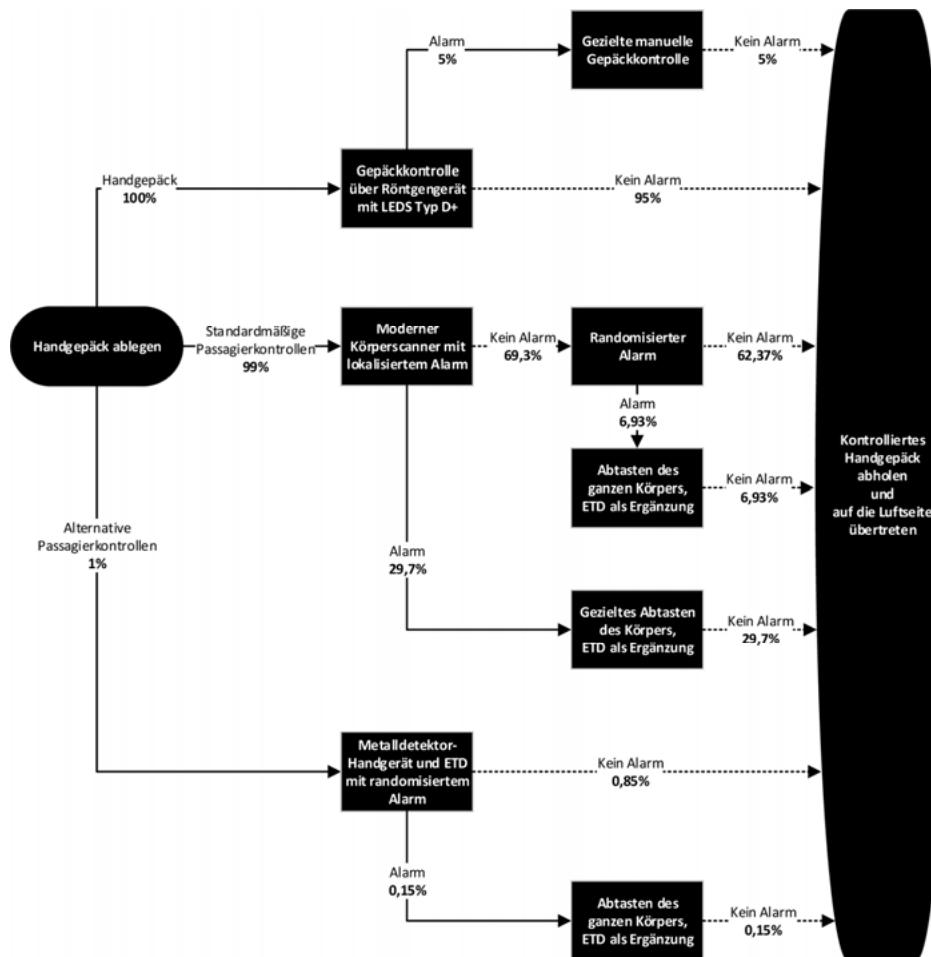
Wie oben bereits erwähnt ist der Einsatz von Sprengstoffspuren-Detektoren an europäischen Flughäfen mittlerweile recht verbreitet. Als Kontrolltechnik – gegebenenfalls in Verbindung mit den Metalldetektor-Handgeräten – kann dieses Verfahren auch eingesetzt werden, um jenes Problem der Schlechterstellung von religiösen Minderheiten zu entschärfen, das für den einfachen Beispielcheckpoint über das Framework identifiziert werden konnte. Statt religiöse Kopfbedeckungen beim manuellen Abtasten alternativlos mit einzubeziehen, können diese bei einem Alarm des Körperscanners im Kopfbereich auch über die Kombination aus Sprengstoffspuren-Detektion und Metalldetektor-Handgeräten kontrolliert werden – eine Möglichkeit, die in den europäischen Regularien explizit ermöglicht wurde, „wenn nach Ansicht des Kontrollpersonals die Durchsuchung ... von Hand ineffizient und/oder nicht geboten ist“ (EU 2015: 16).

Kontrollprozesse eines verbesserten Beispielcheckpoints

Auf Basis der oben erarbeiteten Chancen zur Entschärfung der für die Bewertung zentralen sozialethischen Konflikte soll in diesem Abschnitt nun ein Verbesserungsvorschlag erarbeitet werden. Ziel ist es dabei *nicht*, eine Art Idealcheckpoint vorzustellen, sondern vielmehr zu zeigen, dass über das vorgestellte Framework zur operationalisierten Bewertung von Systemen der Fluggastkontrolle auch konkrete Handlungsvorschläge entwickelt werden können. Im Zentrum meiner Ausführungen steht also weniger der hier vorgestellte Checkpoint, als vielmehr der Nachweis darüber, dass mit diesem methodischen Ansatz in Entscheidungssituationen zur Ausgestaltung derartiger Kontrollprozesse auch eine konkrete moralische *Beratungsfunktion* eingelöst werden kann.

Abb. 14 stellt hierzu die Kontrollprozesse eines entsprechend veränderten Checkpoints schematisch dar. Auch hier sind wieder die Anteile der jeweils betroffenen Passagiere mit angegeben.¹⁸⁰ Der Kontrollprozess beginnt wieder damit, dass alle Passagiere zunächst Jacken und Handgepäck ablegen. Allerdings müssen durch die Nutzung neuerer Röntgengeräte diesmal keine Gegenstände aus den Gepäckstücken herausgenommen werden – Laptops, Tablets und Flüssigkeiten können also in Taschen und Koffern verbleiben. Auch Gürtel mit me-

¹⁸⁰ Wie schon für den einfachen Beispielcheckpoint habe ich auch hier die zugrunde gelegten Alarmraten recht beliebig festgelegt, da die tatsächlichen Alarmraten der verschiedenen Technologien als vertraulich eingestuft und entsprechend nicht öffentlich zugänglich sind.



LEDS Typ D+: Flüssigsprengstoff-Detektion, bei der weder Flüssigkeiten noch elektronische Geräte separat abgelegt werden müssen.

ETD: Sprengstoffspuren-Detektion auf Basis von Wischtests ohne direkte Berührung durch das Kontrollpersonal

Abb. 14: Logik der Kontrollprozesse eines verbesserten Checkpoints mit Alarmraten.

tallischer Schnalle müssen nun nicht mehr abgelegt werden, da zur Primärkontrolle keine Metalldetektoren mehr eingesetzt werden.

Ich gehe von der gleichen Ausgangssituation aus wie für den obigen Beispielcheckpoint, bei der die allermeisten Passagiere (99%) im Standard-Kontrollverfahren überprüft werden. Im Unterschied zum obigen einfachen Beispielcheckpoint werden die Primärkontrollen dabei über einen modernen Körperscanner mit automatischer Gefahrenerkennung und lokalisierter Alarm durchgeführt. Der Körperscanner, so nehmen wir an, erkennt bei einem Großteil der

so kontrollierten Passagiere (70%) keine Bedrohung. Dennoch sollen für einen kleinen Anteil (10%) von diesen Passagieren stichprobenartig Nachkontrollen stattfinden, für die der Körperscanner dann einen nicht-lokalisierten Alarm ausgibt. Lokalisierte Alarne werden durch gezieltes Abtasten, nicht-lokalisierte Kontrollen durch Abtasten des ganzen Körpers aufgelöst. In beiden Fällen werden in Ausnahmefällen bestimmte Bereiche wie religiöse Kopfbedeckungen dabei lediglich über Sprengstoffspuren-Detektion kontrolliert, nicht aber berührt.

Hieraus ergeben sich somit drei Wege für die standardmäßige Personenkontrolle: Der größte Teil der Passagiere durchläuft den Körperscanner, ohne dass ein Alarm ausgelöst wird ($99\% \times 70\% \times 90\% = 62,37\%$); für einen recht kleinen Anteil löst der Körperscanner einen nicht-lokalisierten Alarm aus, was zum manuellen Abtasten des ganzen Körpers führt ($99\% \times 70\% \times 10\% = 6,93\%$); für einen weiteren Teil erkennt der Körperscanner hingegen eine Bedrohung und erzeugt einen lokalisierten Alarm. Dies führt dann zu gezielten Nachkontrollen durch manuelles Abtasten der entsprechenden Körperregion ($99\% \times 30\% = 29,7\%$).

Auch diese Standard-Kontrollverfahren sind nicht für alle Passagiere geeignet und entsprechend muss ein kleiner Teil der Passagiere (1%) über alternative Prozesse kontrolliert werden. Dies betrifft beispielsweise Personen mit Rollstuhl, Personen, die im Körperscanner nicht alleine stehen oder die Arme nicht heben können, wie auch Personen mit größeren Prothesen oder medizinischen Geräten. Diese Passagiere werden über eine Kombination aus Metalldetektor-Handgeräten und Sprengstoffspuren-Detektoren kontrolliert, wobei der Sprengstoffspuren-Detektor auch zufällige Alarne ausgibt. Ein Großteil der so kontrollierten Passagiere durchläuft die Kontrollen ohne einen Alarm (85%); wird hingegen ein Alarm ausgelöst, finden für einen kleineren Anteil der Passagiere (15%) Nachkontrollen in Form eines manuellen Abtastens am ganzen Körper statt. Prothesen, religiöse Kopfbedeckungen etc. werden hierbei nicht berührt, sondern ergänzend gezielt nach Sprengstoffspuren kontrolliert. Bei den alternativen Personenkontrollen ergeben sich somit noch zwei weitere Wege durch den Checkpoint: Zum einen, wenn bei den Primärkontrollen über die Kombination aus Sprengstoffspuren-Detektion und Metalldetektor-Handgeräten kein Alarm ausgegeben wird ($1\% \times 85\% = 0,85\%$), sowie zum anderen wenn, Nachkontrollen stattfinden ($1\% \times 15\% = 0,15\%$).

Auch hier finden wieder parallel zur Personenkontrolle auf diesen fünf Wegen noch die Kontrollen des abgelegten Handgepäcks statt. Eine Sichtkontrolle findet hier nicht mehr statt, die Gepäckstücke werden statt dessen über ein modernes Röntgengerät kontrolliert, das auch Flüssigsprengstoffe detektieren kann und das es zudem nicht mehr erforderlich macht, elektronische Geräte aus dem Gepäck herauszunehmen. Wir nehmen an, dass bei den neuartigen Geräten zudem auch seltener Alarne ausgelöst werden, weshalb ein noch größerer Teil der Fluggäste (95%) das kontrollierte Gepäck abholen und auf die Luftseite übertreten kann. Für einen Teil der Passagiere (5%) wird vom Kontrollpersonal bei den

Tabelle 6: Gesamt- und Teilwertungen zu den sozialethischen Risiken für den verbesserten Beispielpoint

Weg	Kombination der Kontrolltechniken	Anteil	Teilwertung zur Brisanz der Problemkategorien (und -typen)
i	Körperscanner; Röntgengerät	59,25 %	Eingriff in Privatheit: A (AABB) Schlechterstellung: B (ACAA) Restriktivität: A (AAAB)
ii	Körperscanner; Röntgengerät, gezielte manuelle Gepäckkontrolle	3,12 %	Eingriff in Privatheit: B (CABB) Schlechterstellung: C (CCAA) Restriktivität: A (BAAB)
iii	Körperscanner, Abtasten; Röntgengerät	6,58 %	Eingriff in Privatheit: B (ACBB) Schlechterstellung: A (AAAA) Restriktivität: A (BAAB)
iv	Körperscanner, Abtasten; Röntgengerät, gezielte manuelle Gepäckkontrolle	0,35 %	Eingriff in Privatheit: C (CCBB) Schlechterstellung: B (CAAA) Restriktivität: B (CAAB)
v	Körperscanner, gezieltes Abtasten; Röntgengerät	28,21 %	Eingriff in Privatheit: A (ABBB) Schlechterstellung: A (ABAA) Restriktivität: A (BAAB)
vi	Körperscanner, gezieltes Abtasten; Röntgengerät, gezielte manuelle Gepäckkontrolle	1,49 %	Eingriff in Privatheit: C (CBBB) Schlechterstellung: B (CBA) Restriktivität: B (CAAB)
viii	Metalldetektor-Handgerät und Sprengstoffspuren-Detektion; Röntgengerät	0,81 %	Eingriff in Privatheit: A (AAAB) Schlechterstellung: B (ACAA) Restriktivität: A (AAAB)
viii	Metalldetektor-Handgerät und Sprengstoffspuren-Detektion; Röntgengerät, gezielte manuelle Gepäckkontrolle	0,04 %	Eingriff in Privatheit: B (CAAB) Schlechterstellung: C (CCAA) Restriktivität: A (BAAB)
ix	Metalldetektor-Handgerät und Sprengstoffspuren-Detektion, Abtasten; Röntgengerät	0,14 %	Eingriff in Privatheit: B (ACAB) Schlechterstellung: A (AAAA) Restriktivität: A (BAAB)
x	Metalldetektor-Handgerät und Sprengstoffspuren-Detektion, Abtasten; Röntgengerät, gezielte manuelle Gepäckkontrolle	0,01 %	Eingriff in Privatheit: C (CCAB) Schlechterstellung: B (CAAA) Restriktivität: B (CAAB)
Gesamter Checkpoint:		100 %	Eingriff in Privatheit: A Schlechterstellung: A Restriktivität: A

Handgepäckkontrollen aber ein Gegenstand als verdächtig deklariert und die entsprechenden Gepäckstücke manuell gezielt durchsucht. Insgesamt ergeben sich somit zehn mögliche Wege durch den Checkpoint, denn an jeden der fünf Wege für die Personenkontrollen schließen sich wieder zwei mögliche Wege für

die Gepäckkontrollen an.¹⁸¹ Tabelle 6 zeigt eine Übersicht über die Gesamt- und Teilwertungen zu den zehn Wegen dieses verbesserten Beispielcheckpoints.

Die letzte Spalte gibt wie oben für jeden Weg auch die Wertungen hinsichtlich der Brisanz der drei sozialethischen Problemkategorien an. In Klammern sind auch die Wertungen zu den ausdifferenzierten Problemtypen beigegeben, auf denen die Bewertungen der Kategorien jeweils beruhen und die im Detail im Anhang C nachvollzogen werden können. Die letzte Zeile zeigt schließlich das Ergebnis der Gesamtwertung für den Beispielcheckpoint für jede der drei Problemkategorien. Dabei wird deutlich, dass die entsprechenden Wertungen für diesen Checkpoint deutlich besser ausfallen. In allen drei Problemkategorien wurde die Bestwertung ‚A‘ vergeben, was auf einen Checkpoint nahe dem Bestfall hinweist. Insbesondere mit Blick auf die Schlechterstellung verletzlicher Gruppen konnte hier eine drastische Verbesserung erzielt werden: Angehörige von Minderheiten, für die eine bestimmte Kopfbedeckung Ausdruck ihrer religiösen Überzeugungen ist, die sie weder in der Öffentlichkeit abnehmen noch vom Kontrollpersonal berühren lassen wollen, werden nun aktiv berücksichtigt und ein Mindeststandard des Schutzes vor Schlechterstellung entsprechend realisiert.

Die hier vorgeschlagenen Verbesserungen bei den Kontrolltechniken werden so oder in ähnlicher Form durchaus bereits an europäischen Flughäfen eingesetzt oder getestet. Der verbesserte Beispielcheckpoint wäre also durchaus kein gänzlich unrealistisches System zur Fluggastkontrolle, allerdings berücksichtigen die Verbesserungsvorschläge aktuell ausschließlich die sozialethischen Problemkategorien. Eine konkrete Analyse der Kostenfaktoren – von Personal- und Betriebskosten über Kundenzufriedenheit bis hin zum Platzbedarf – blieb für die Vorschläge ebenso unberücksichtigt wie eine umfassende Analyse des durch den Checkpoint gewährleisteten Schutzes vor Bombenanschlägen oder Entführungen von Flugzeugen aus der Gruppe der Passagiere heraus. Für eine umfassend informierte Entscheidungsfindung müssen also, wie bereits mehrfach erwähnt, noch eine Reihe weiterer wertungsrelevanter Kriterien berücksichtigt werden, die allerdings nicht Gegenstand des in dieser Arbeit entwickelten Frameworks sind. Dennoch lässt sich bereits klar erkennen, dass die so aufbereiteten, sozialethischen Risiken von Fluggastkontrollen strukturiert in die Entscheidung einbezogen und hermeneutisch über den kulturell-geschichtlichen Wertehorizont europäischer Gesellschaften erschlossen werden können.

¹⁸¹ Wie schon für den obigen einfachen Beispielcheckpoint nehmen wir auch hier der Einfachheit halber den durchaus realistischen Fall an, dass intensive Nachkontrollen und Übergaben an Polizeikräfte so selten auftreten, dass sie für die Wertungen über das Framework nicht weiter ins Gewicht fallen.

Ziel des hier entwickelten ethischen Frameworks war es, die in Kapitel 2 und 3 entwickelte moralische Landkarte typischer sozialethischer Wertungskonflikte bei der Herstellung öffentlicher Sicherheit durch Fluggastkontrollen zu operationalisieren. Dabei wollte ich an die hermeneutisch ausgerichteten Überlegungen so anschließen, dass die entwickelten Typiken auch im Sinne von Planungsgrößen für kollektives zukunftsbezogenes Handeln fungieren können. Hierin besteht, wie oben unter anderem mit Verweis auf Grundwald (1999: 75) entwickelt, in der aktuellen deutschsprachigen Technikfolgenforschung ein Desiderat für eine gelungene philosophische TA. Den Nachweis darüber, dass die in der moralischen Landkarte identifizierten typischen Problemkonstellationen über das in diesem Kapitel entwickelte Framework tatsächlich in einer abwägenden Entscheidung strukturiert und in produktiver Weise einbezogen werden können, habe ich im Laufe dieses Abschnitts ebenfalls erbracht. Entsprechend verbleibt nun noch zu explizieren, welche Rolle dieses Framework, das als Beitrag zur philosophischen Technikfolgenforschung begriﬀen werden soll, für eine Politik- und Gesellschaftsberatung spielen kann.

Zunächst kann das Framework insbesondere für jene Personen eine Orientierungs-, Strukturierungs- und Leitungsfunktion übernehmen, die direkt mit den Entscheidungen zu den Einzelheiten der Prozesse der Fluggastkontrolle betraut sind. Wie ich in Abschnitt 2.1.3 ausgeführt habe, ist dabei innerhalb der Europäischen Union nicht einheitlich geregelt, wer konkret zu diesem Personenkreis gehört. So betrifft dies in manchen Ländern Angestellte der jeweiligen Flughäfen, in anderen Ländern – wie in Deutschland – hingegen Beamte der entsprechenden Polizeibehörden. Auch die involvierten Personen im Gesetzgebungsprozess (Parlamentsabgeordnete, Referenten etc.) zählen zu diesem Personenkreis und können über das Framework und die moralische Landkarte einen strukturierten Zugang zu den relevanten Wertungskonflikten, wie auch zu den konkreten Folgen bestimmter Entscheidungen gewinnen. Wie ich oben an verschiedenen Stellen dargelegt habe, zeigt sich anhand der in den letzten Jahren bisweilen heftig geführten öffentlichen Debatten um Privatheits-, Diskriminierungs- und Restriktionsprobleme von Fluggastkontrollen, dass eine solche Beratungsleistung durchaus einen wichtigen Beitrag im Sinne einer Politikberatung leisten kann.

Die Planungsgrößen, die sich über das Framework für konkrete Prozesse der Fluggastkontrolle entwickeln lassen, sollen dabei keine konkreten Zukünfte prognostizieren oder empirisch erheben, ob eine bestimmte politische Position gesellschaftlich populär oder mehrheitsfähig ist. Vielmehr zielt das Framework und die entwickelte moralische Landkarte auf eine vorausgreifende Bearbeitung gesellschaftspolitischer Wertungskonflikte und damit auf eine moralische Deliberation ab, die in produktiver Weise eine Reihe relevanter Wertungen mit in die Entscheidung einbeziehen kann. Sie soll somit im moralpragmatischen Sinne

eine Transformation des Erlebens einer Entscheidungssituation für jene Personen erschließen, die mit Entscheidungen zur Ausgestaltung von Fluggastkontrollen betraut sind – was im hermeneutischen Sinne letztlich auf ein vertieftes Verstehen der handlungsrelevanten Konsequenzen durch eine historisch-kulturelle Kritik und Kontextualisierung verweist.

Statt die hierbei identifizierten Wertungskonflikte aufzulösen, eine Handlungsoption als ‚rational geboten‘ vorzugeben und sich damit zu Recht einem Technokratie-Vorwurf auszusetzen, zielt der hier entwickelte Ansatz zur Operationalisierung darauf, diese handlungsrelevanten Konsequenzen so aufzubereiten, dass auch Kompromisslösungen entwickelt werden können, die aus rein sozialethischer Sicht eher zweit- oder drittbeste Lösungen sein mögen. Darauf, dass eine solche Kompromissfähigkeit in der angewandten Ethik zentral ist, um wirklich auch in der Praxis anwendbar zu sein (BAYERTZ 2008: 167; MESSNER 2012), habe ich in Abschnitt 1.1.2 bereits hingewiesen. Hierin zeigt sich die Stärke des gewählten moralpragmatischen Ansatzes, über den sich Begriffe und Theorien als intellektuelle Werkzeuge für den Umgang mit moralischen Problemsituationen und Dilemmata konzipieren lassen.

Die Tauglichkeit des hier entwickelten Ansatzes für die Praxis wird sich dabei in zweierlei Hinsicht erweisen müssen. Zum einen ist dies, wie bereits erwähnt, mit Blick auf die drei entwickelten Problemtypiken zur Identifizierung der relevanten, wiederkehrenden sozialethischen Wertungskonflikte der Fall: Tauchen in der öffentlichen Diskussion um Fluggastkontrollen in der Zukunft immer wieder Problemkonstellationen auf, die von den Typiken nicht abgedeckt oder adäquat hermeneutisch zugänglich gemacht werden können, so erweisen sich diese als Verbesserungsbedürftig – was freilich nicht bedeutet, dass deshalb die Typiken gänzlich ihre Aussagekraft verlieren würden. Zum anderen wird sich aber auch die Tauglichkeit des Frameworks zur operationalisierten Bewertung insgesamt erweisen müssen, indem sich in der Praxis zeigt, dass entsprechende Wertungskonflikte hierüber tatsächlich bereits im Vorgriff besser bearbeitet werden können.

Dass sich eine solche Tauglichkeit nicht einfach im Ausbleiben eines medial vermittelten öffentlichen Empörtsein erweisen kann, lässt sich schon unter Verweis auf Deweys (1984a: 304–324) Skepsis gegenüber der Rolle von Massenmedien für eine öffentlich-demokratische Deliberation feststellen. Vielmehr muss sich die Tauglichkeit des Ansatzes dort erweisen, wo tatsächlich ein Austausch von Meinungen auch zu einer Transformation des situativen Problemerlebens führen kann – etwa unter Einbezug zivilgesellschaftlicher Akteure oder über partizipative Initiativen. In jedem Fall wird sich die Tauglichkeit des Frameworks dann aber dahingehend erweisen, dass es die Möglichkeit eröffnet, die jeweiligen Wertungen und Entscheidungsgründe bei Abwägungen strukturiert, transparent und kohärent zu kommunizieren, bzw. aus Sicht einer interessierten politischen Öffentlichkeit zu kritisieren.

Der in dieser Arbeit entwickelte Ansatz wird es dabei freilich nicht schaffen, die in Abschnitt 1.3 dargelegten institutionellen Schwächen moderner Massendemokratien zu überwinden. Es bleiben dementsprechend eine Reihe an Fragen hinsichtlich des Potentials politischer Öffentlichkeiten zur Multiperspektivität und zur Entfaltung gesellschaftlichen Lernens durch gemeinschaftlich kommuniziertes Erleben (DEWEY 1980: 93, 1982a: 186). Doch liegt zumindest eine der zentralen Herausforderungen für eine öffentliche Deliberation zu technischem Handeln darin, passende intellektuelle Werkzeuge für das Verstehen der jeweiligen Konsequenzen bereitzustellen. Ohne solche Werkzeuge bleibt die Erhebung von Zustimmung und Akzeptanz letztlich ohne wirkliche politische Aussagekraft: "Our Babel is not one of tongues but of the signs and symbols without which shared experience is impossible" (DEWEY 1984a: 324).

In diesem Sinne gehen die Darstellungen dieser Arbeit, die durch ihre Publikation auch einer wissenschaftlichen sowie politisch interessierten Öffentlichkeit zugänglich gemacht wurden, deutlich über eine reine Politikberatung im engen Sinne hinaus und sollen vielmehr als Beitrag zu einer *Politik- und Gesellschaftsberatung* verstanden werden. In jedem Fall bieten sie, das ist meine These und Hoffnung, ausreichend intellektuelle Werkzeuge an, um reflexartige und eigendynamische Denkmuster sowie pauschale Aussagen zu Prozessen der Fluggastkontrolle kritisch zu hinterfragen, Problemkonstellationen kulturell-historisch zu verorten und implizite Wertungen sowie Wertungskonflikte anhand konkreter handlungsrelevanter Folgen aufzudecken.