

»Bei allen beschreibenden Verfahren der Datenerhebung, z. B. den ethnographischen [...], lassen sich, weil die Beschreibungen von intelligenten Subjekten durchgeführt werden und immer schon eine gegenstandsspezifische Interpretation enthalten, die im Grunde eine Dateninterpretation ist, Datenerhebung und Dateninterpretation bzw. -auswertung nicht mehr auseinanderhalten. Daraus resultiert eine vermeidbare, die methodische Geltungsüberprüfung einschränkende Zirkelhaftigkeit« (Oevermann 2013, 69).

Dennoch sind teilnehmende Beobachtungen und deren Protokolle notwendig, um ein fallspezifisches Kontextwissen zu erhalten und damit die Verdichtung oder Modifikation sequenzanalytisch ermittelter Fallstrukturhypothesen zu ermöglichen. Teilnehmende Beobachtungen sind somit ein unabdingbares Korrektiv und notwendig für eine valide Auswertung der Daten.

Für textliche Protokolle, das heißt Interview-Transkripte und textliche Primärquellen, ist die sequenzanalytische Verfahrensweise gut dokumentiert und vielfach angewendet worden. Für fotografische Protokolle ist diese Methode jedoch erst wenig verwendet worden. Daher widme ich diesem Verfahren unter dem Titel »Bildsequenzanalyse« ein eigenes Kapitel (siehe unten). Selbst der transparente und schrittweise Umgang mit Fotografien als Quellen sozialwissenschaftlicher Forschung ist in der Methodenliteratur noch dürtig dargestellt, was teilweise daran liegt, dass einige Forscher:innen der Meinung sind, ein solcher systematisch-analytischer Umgang mit Fotografien sei kaum möglich. So schreibt Douglas Harper:

»Wie andere Aspekte qualitativer Forschung auch kann die visuelle Dimension am ehesten durch Praxis begriffen werden. Sie lässt sich systematisch nicht beschreiben und nicht lehren wie Survey-Methoden oder statistische Verfahren. Denn mit der Fotografie gelangt die Soziologie als Wissenschaft ganz in die Nähe der künstlerischen Tätigkeit« (Harper 2015, 402).

In deutlicher Abgrenzung von dieser Position schlage ich in dieser Studie vor, dass Fotografien zwar andere Arten von Daten im Vergleich zu textlichen Daten sind, aber prinzipiell ebenfalls einer kontrollierten und nachvollziehbaren Analyse zugänglich gemacht werden können (weitere methodologische Überlegungen finden sich in Kapitel 9.5).

## 9.4 Bildsequenzanalyse

Mit der Sequenzanalyse stelle ich einen methodologischen Vorschlag zur Debatte, der auf den Prämissen der Objektiven Hermeneutik basiert (grundlegend Oever-

mann et al. 1979). Daher umreiße ich zunächst einige Grundbegriffe dieser Methodologie (ausführlicher siehe bspw. Krech und Radermacher 2022) und gehen dann auf deren Anwendung im Hinblick auf fotografische Protokolle ein.

**Objektivität** Die Verwendung des Begriffs »objektiv« in der Bezeichnung der Objektiven Hermeneutik ist als Abgrenzung gegenüber einer »subjektiven« Hermeneutik zu verstehen: Es geht nicht um eine *Interpretation* vorliegender Daten, sondern um eine *Rekonstruktion* objektiver, das heißt innerhalb bestimmter soziokultureller Zusammenhänge allgemein gültiger Regeln, die ermöglichen und bedingen, dass die untersuchte Ausdrucksgestalt so ist, wie sie ist. Mit der Formulierung »objektiv« wird nicht der Anspruch erhoben, dass die Methode letztgültige Ergebnisse von einem ›neutralen‹ Standpunkt aus liefern könnte. Der Ansatz will aber objektivierbare Ergebnisse zu produzieren, insofern er die Untersuchung empirischer Daten (die man als Protokolle bezeichnet) in einem kontrollierbaren, transparenten und schrittweisen Verfahren vorschlägt (Wernet 2014, 236).

**Latente Sinnstrukturen** Das zentrale Auswertungsverfahren der Objektiven Hermeneutik wird als Sequenzanalyse bezeichnet. Das Ziel dieser Methode ist es, latente Sinnstrukturen zu rekonstruieren, die das Protokoll ermöglichen und strukturieren, aber nicht explizit thematisiert werden, also nicht manifest sind. Manifeste Sinngehalte werden von Sprecher:innen ausdrücklich thematisiert, latente Sinngehalte bleiben implizit (Wernet 2014, 235–236). Die Unterscheidung latenter und manifester Sinngehalte basiert auf der Annahme, dass jeder *konkrete* Fall zugleich Ausdruck *allgemeiner* Muster ist, die aber in den jeweils vorliegenden Daten nicht explizit von Akteur:innen formuliert werden. Diese Ansicht wird in der einen oder anderen Variante in den meisten qualitativen Methoden der Kultur- und Sozialwissenschaften vertreten: »Researchers routinely probe *beyond the explicit* and the known to try to understand the worlds of research participants of which the participants themselves may be unconscious« (O'Toole und Were 2008, 621; Hervorhebung MR). Für dieses Ziel werden jedoch, so Paddy O'Toole und Prisca Were, die physischen und räumlichen Bedingungen mit ihren materiellen Artefakten und Arrangements noch zu selten berücksichtigt – ein Umstand, der mit dem hier vorgeschlagenen Verfahren bearbeitet werden soll.

Die Annahme einer latenten Sinnstruktur besagt, dass die Bedeutung einer beliebigen Aussage sich nicht darin erschöpft, was Akteur:innen damit ›meinen‹. Die Methode ist nicht primär an individuellen, persönlichen Intentionen und Einstellungen interessiert. Daher haben auch die Schlussfolgerungen der Analyse nichts einzelnen Personen, zum Beispiel Interviewpartnern oder Autor:innen der untersuchten Texte, zu tun. Die Methode ist im Gegenteil an Strukturen und Mustern

interessiert, die unabhängig von einzelnen Personen bestehen.<sup>4</sup> Diese Strukturen kann man als objektiven Sinn bezeichnen, der in der Regel nicht mit dem subjektiven Sinn deckungsgleich ist (Oevermann et al. 1979, 380), dadurch aber nicht als universell und jenseits aller Standpunkte gültiger Sinn zu betrachten ist.

Für diese Studie ist zu berücksichtigen, dass jedes sozialräumliche Arrangement auf verschiedene Weisen sinnvoll beschrieben werden kann. Expert:innen für katholische Liturgie können kirchliche Räume und darin stattfindende Rituale erklären; aber auch Laien und Anwesende können eine »sinnvolle« – und das heißt hier: verstehbare und mit Anschlussmöglichkeiten versehene – Adressierung eines Rituals oder eines Kirchenraumes vornehmen. Diese explizit kommunizierten Deutungen sind wichtige Bestandteile der empirischen Daten. Doch besteht der Anspruch des hier vorgeschlagenen Verfahrens nicht nur darin, diese manifesten Sinngehalte wiederzugeben und zu ordnen, sondern die latenten Strukturen jener ausdrücklich formulierten Zuschreibungen zu rekonstruieren, weil diese als konstitutive Regeln und Ermöglichungsbedingungen fungieren.

**Ausdrucksgestalten** Betrachtet man ein konkretes, räumlich und zeitlich definiertes sozialräumliches Arrangement als Ausdrucksgestalt im Sinne der Objektiven Hermeneutik, dann lässt sich die darin liegende Sinnstruktur rekonstruieren, indem ein Protokoll dieser Ausdrucksgestalt sequenzanalytisch untersucht wird (Oevermann 2013, 73). Methodisch bedeutet dies, dass die Analysierenden das semantische Potenzial sequenzanalytisch rekonstruieren und mit dem realisierten Fall in Verbindung bringen. Anders gesagt: Welche Lesarten sind alltagsweltlich erwartbar und was haben diese auf einer allgemeinen Ebene gemeinsam. Welche Anschlüsse sind alltagsweltlich möglich, welche sind tatsächlich realisiert und welche allgemeinen Muster lassen sich daraus ableiten?

Latente Sinnstrukturen aktualisieren sich in konkreten Ausdrucksgestalten, die verbaler, textlicher sowie nicht-sprachlicher Natur sein können: Dazu gehören Artefakte, Architektur, Musik und auch Bilder. In den empirischen Daten, die wir sequenzanalytisch als Protokolle untersuchen, sind die latenten Sinnstrukturen außerdem in der Regel überdeterminiert: Sie treten vielfach und in verschiedenen Varianten auf, damit sie – ohne ausdrückliche text- oder verbalsprachliche Kommunikation – soziale Wirksamkeit entfalten können.

---

4 Die Idee, dass es etwas Allgemeines ›hinter‹, ›über‹ oder ›in‹ den konkreten Begebenheiten gibt, ist kein Proprium der Objektiven Hermeneutik, sondern zieht sich durch die Geschichte der Geistes-, Kultur- und Sozialwissenschaften, angefangen bei de Saussures Rede von »langue« und »parole« (Saussure 1967 [1916], 11), über das strukturalistische Denken (z. B. Lévi-Strauss 1975) bis hin zu Diskurstheorien, in denen der Diskurs ein strukturierendes Wissen ist, das sprachliche Äußerungen unbewusst lenkt (z. B. Hall 2003, 6).

Solche nicht-sprachlichen Ausdrucksgestalten können zum Beispiel architektonische Arrangements sein (Schmidtke 2008). Auch in der Architektur gibt es eine Differenz zwischen dem subjektiven Sinn, den ein Bauwerk beispielsweise durch die Architekt:innen oder durch die Zuschreibungen der Besucher:innen erhält und dem objektiven Sinn. Letzterer ist für die objektive Hermeneutik rekonstruierbar, auch wenn er den Architekt:innen oder Rezipient:innen nicht zwangsläufig bewusst ist und auch nicht bewusst sein muss. Das heißt nicht, dass diese sequenzanalytisch ermittelte Sinnschicht ‚richtiger‘ oder ‚wahrer‘ ist. Sie ist lediglich von allgemeiner Art und verweist darauf, dass es neben dem konkreten psychischen Sinn einzelner Personen weitere Sinnstrukturen gibt, die prägend für die konkrete Ausdrucksge- stalt sind.

**Lesarten und gedankenexperimentelle Kontextvariationen** Das zentrale Mittel der Sequenzanalyse, um das genannte Ziel zu erreichen, ist die Formulierung von Lesarten, die ohne fallspezifisches Kontextwissen gedankenexperimentell entwickelt werden. Das bedeutet, dass eine Gruppe von Analysierenden kurze Passagen (wenige Worte eines Textes oder einen kleinen Ausschnitt eines Bildes) aus dem Protokoll hinsichtlich der Frage diskutiert: In welchen alltagsweltlichen Zusammenhängen könnte diese Sequenz erwartbarer Weise auftauchen und was würde sie dort üblicherweise bedeuten? Man bezeichnet diesen Arbeitsschritt auch als gedanken-experimentelle Kontextvariation (Pickel und Sammet 2014, 45).

Die Analysierenden sollen dafür nicht auf *fallspezifisches* Kontextwissen zurückgreifen. Das heißt, sie sollen den untersuchten Fall nicht aus seinen konkreten Entstehungsbedingungen wie persönlichen Intentionen der Beteiligten oder deren Handlungen und ebenfalls nicht aus einem konkreten Wissen über thematisierte Dinge und Menschen heraus erklären, sondern *fallunspezifisches* Kontextwissen verwenden: Jede Sequenz ist in gewisser Hinsicht verständlich, weil sie auf fallun-abhängigen Regeln basiert, die häufig unbewusst eingesetzt werden, um alltägliche Kommunikation zu betreiben. Diese latenten Strukturen will die Analysegruppe rekonstruieren, indem sie verschiedene Lesarten formuliert und miteinander vergleicht (Maiwald 2005, 25).

Die sequenzanalytische Forderung nach dem Ausschluss von Kontextwissen wird oft missverstanden: Es geht nicht darum, dass die Mitglieder der Analysegruppe sämtliches Wissen über die soziale Wirklichkeit ausschließen sollen. Reichertz bringt es auf den Punkt, wenn er – etwas überspitzt – formuliert: »Wer [...] dumm in die Analyse geht, kommt in der Regel auch dumm heraus« (Reichertz 2011, 21). Im Gegenteil sollen die Analysierenden ihre allgemeinen lebensweltlichen Kompetenzen und Wissensbestände einfließen lassen. Nur so sind sie in der Lage, Lesarten – also mögliche und erwartbare Zusammenhänge – für eine Sequenz zu formulieren. Das erfordert auch, dass sie mit dem sozio-kulturellen Umfeld des untersuchten Gegenstandes vertraut sein müssen. Nur so können sie erwartbare

und übliche Kontexte formulieren, aus denen dann allgemeine Regeln und Muster abgeleitet werden können. Was sie jedoch nicht in die Analyse einbeziehen sollen, ist das konkrete Wissen um den untersuchten Einzelfall und um die abgebildeten Menschen, Dinge und Orte (Oevermann 2014, 35–36). Erst in einem späteren Schritt der Analyse soll in kontrollierter Weise fallspezifisches Kontextwissen eingeführt werden, um die entwickelten Fallstrukturhypothesen mit dem konkreten Fall in Verbindung zu bringen (Oevermann 2014, 50).

Man formuliert auf die gleiche Weise mögliche Anschlüsse an die zuvor diskutierte Sequenz und vergleicht dabei die möglichen mit den im untersuchten Protokoll realisierten Anschlüssen. Durch diese Vergleiche werden Rückschlüsse auf die impliziten Strukturen des Falls gezogen, die man übersehen würde, wenn ein Fall aus seinem direkten Kontext (also z. B. den Intentionen der beteiligten Akteur:innen) heraus erklärt würde (Wernet 2014, 240).

**Analysegruppe** Um ausgewogen zu arbeiten und eine subjektive Verzerrung zu vermeiden, empfiehlt die Literatur zur Sequenzanalyse (z. B. Pickel und Sammet 2014, 47), dass Forscher:innen zu einer Analysegruppe eingeladen werden, die einerseits mit der soziokulturellen Wirklichkeit vertraut sind, aus der das Protokoll stammt – ohne notwendigerweise mit dem konkreten Fall selbst vertraut zu sein –, und die andererseits möglichst »streitlustig« sind, damit eine möglichst breite Vielfalt an Lesarten entwickelt werden kann.

**Protokoll** Eine notwendige methodologische Voraussetzung dieses Vorgehens ist, wie erwähnt, ein Protokoll. Als Protokoll dient in der Objektiven Hermeneutik üblicherweise ein Text, beispielsweise das Transkript eines Interviews oder einer anderen verbalsprachlichen Interaktion. Da sich sozialräumliche Arrangements aber nicht primär verbalsprachlich artikulieren, sind in dieser Studie fotografische Protokolle eine zentrale und unverzichtbare Datenquelle. Ein Foto erreicht methodologisch das gleiche wie ein Interview-Transkript: Das intuitive Verstehen wird sequenzanalytisch regelgeleitet verlangsamt, um zu rekonstruieren, welche soziokulturellen Dispositionen sich in welcher Weise realisieren.

Der methodologische Unterschied zwischen Textprotokollen und Bildprotokollen ist, dass Bilder nicht im engeren Sinne sprachlich verfasst sind. Es gibt eine breite Debatte dazu, inwiefern Bilder überhaupt innerhalb des sprachlogisch fundierten Paradigmas gedeutet werden können. Die Vertreter eines *iconic turn* im Sinne von Boehm (2006), Mitchell (1994) und Beltig (2001) sprechen von einer »Eigenlogik des Bildes«, die grundsätzlich nicht in eine sprachliche Logik überführt werden könne (Boehm 2004, 28–29; siehe auch Bohnsack 2014, 867).

Es wird zutreffen, dass der bildliche Modus sinnhafter Operationen sich vom sprachlichen Modus in mancher Hinsicht unterscheidet, dennoch sind auch Bilder sinnvoll, insofern sie aus bildlichen Elementen durch wechselseitige Bezüge (syn-

taktisch, semantisch und pragmatisch) Sinn konstituieren, also semiotisch operieren – genau wie dies Texte tun. Insofern scheint es legitim, beide Ausdrucksformen (Bilder und Texte) als semiotische Kommunikationsweisen zu verstehen (siehe hierzu ausführlicher den methodologischen Exkurs in Kapitel 9.5).

Für gesprochene Sprache werden Protokolle in der Regel durch Tonaufnahmen und Transkription angefertigt. So wird eine konkrete und flüchtige Ausdrucksgestalt (gesprochene Sprache) fixiert und in Form einer Transkription der Analyse zugänglich gemacht. In Bezug auf räumliche Arrangements lässt sich kein Transkript anfertigen; hier kann das Foto die Funktion eines Protokolls übernehmen, weil es die konkrete Ausdrucksgestalt – ein sozial-räumliches Arrangement zu einem gegebenen Zeitpunkt – der zeitlichen Flüchtigkeit enthebt und somit der Analyse zugänglich macht. Für die Sequenzanalyse wird diskutiert, ob eine Übertragung auf Fotos und Bilder ohne Weiteres möglich sei. Oevermann hält dies, wie auch andere Autor:innen für vertretbar (Oevermann 2014, 35; Betz und Kirchner 2016). Er betont sogar: »Im Hinblick auf die konkrete Realisierung von Bedeutungs- oder Sinnstrukturen dagegen ist die Sprache eine Ausdrucksmaterialität unter vielen anderen und häufig nicht einmal die wirksamste« (Oevermann 2013, 73).

Dass ein Bild und das, was es abbildet, nicht dasselbe sind, darüber herrscht spätestens seit der viel diskutierten »Krise der Repräsentation« Einigkeit (z. B. Hasse 2012, 37). Denn ein Foto protokolliert zum einen nur einen bestimmten, zeitlich definierten Ausschnitt der Wirklichkeit und es protokolliert neben dem abgebildeten Arrangement zugleich auch den Blick der Person, die das Foto aufgenommen hat. In dieser Studie steht jedoch nicht die *Produktion* eines Fotos im Mittelpunkt, sondern das, was es abbildet, und die Strukturen, die in diesem abgebildeten sozialräumlichen Arrangement zum Ausdruck kommen. Der Umstand, dass ein Foto nicht unverändert die fotografierte Situation repräsentiert, ändert aber nichts daran, dass es auch eine Aussagekraft bezüglich des abgebildeten Raumes behält. Auch Interview-Transkripte repräsentieren nur einen Bruchteil der ursprünglichen verbalen und non-verbalen Interaktion unter Anwesenden. Protokolle bilden nie eins zu eins die erlebte ›Wirklichkeit‹ ab:

»Methodologisch haben wir es [...] nie mit dieser Unmittelbarkeit des Hier und Jetzt zu tun, sondern immer nur mit Protokollen davon. Alle unsere Daten in den Erfahrungswissenschaften, welcher Disziplin auch immer, sind Protokolle der im Hier und Jetzt unmittelbar gegebenen Wirklichkeit und nicht diese selbst« (Oevermann 2009, 130).

Weder textliche Dokumente noch Transkriptionen von verbalen Interaktionen sind demzufolge ›wirklichkeitsgetreu‹. Dennoch wird gerade Fotografien oft unterstellt, dass sie die dokumentierten Daten besonders stark verzerrten, in erster Linie durch Selektion und Perspektive. Diese Problematik gilt es zwar zu beachten, doch haben

Fotos den Vorteil, dass sie ein räumliches Arrangement in reduzierter und verdichteter Weise konservieren und die körperliche Anwesenheit vor Ort für die Analyse verzichtbar machen. Das Foto ist eine zweidimensionale Verdichtung der Empirie (Oevermann 2014, 45–46).

Ein weiterer methodologischer Diskussionspunkt betrifft die Anwendbarkeit einer primär sequenziell operierenden Methode auf einen vermeintlich simultan vorliegenden Gegenstand, die Fotografie. Auch hierzu gibt es eine breite Auseinandersetzung (siehe z. B. Breckner 2003; Bohnsack 2003; Peez 2006; Raab 2012), in der oft vermutet wird, sequenzanalytische Verfahren würden der Simultaneität des Bildes nicht gerecht (z. B. Bohnsack 2003, 251; Raab 2014, 107; zur Zusammenfassung dieser Diskussion siehe auch Betz und Kirchner 2016, 264–265). Andere Autor:innen gehen jedoch davon aus, dass sowohl Texte als auch Bilder – wenn auch mit unterschiedlicher Gewichtung – prinzipiell simultan *und* sequenziell operieren (siehe den methodologischen Exkurs in Kapitel 9.5) und dieser methodologischen Prämissen folge ich.

**Gebaute Räume und Architektur** Im Hinblick auf gebaute Räume und Architektur wird die Sequenzanalyse bislang noch selten verwendet. Ausnahmen sind etwa die Arbeiten von Oliver Schmidtke (2008), Jan Egger (2017) und Veronika Eufinger (2019), wobei nur letztere sich auch auf religionsbezogene Räume bezieht, nämlich urbane Arrangements von Citykirchen-Projekten (Eufinger 2018; Eufinger 2019, 175–176). Andererseits ist die Anwendung einer sequenziellen Analyse auf Protokolle räumlicher Arrangements unmittelbar einsichtig, wenn man architekturhistorische Arbeiten konsultiert:

»Architektur wird ›durchwandert, durchschritten‹, betont Le Corbusier. Während der *promenade architecturale* soll der Mensch innere Bewegung empfinden, ›das Ergebnis einander folgender Erschütterungen‹ (1962, 29). Das Hauptmerkmal der Sequenz ist die lineare Abfolge, das eine folgt auf das andere« (Janson und Tigges 2013, 284–285; ähnlich auch Janson und Bürklin 2002).

Wenn man hier die leibphänomenologische Grundierung außer Acht lässt, ist der Hinweis auf die Sequenzialität des Erlebens von Architektur ein wichtiges methodologisches Argument. Werner Hager schreibt ähnlich über den mittelalterlichen Kirchenraum, dass dieser »sukzessiv erlebt« wird und architektonisch unterteilt wird (Hager 1957, 641). Damit wird, wenn auch aus anderer fachlicher Perspektive, ange deutet, dass sozialräumliche Arrangements genau wie verbal- und textsprachliche Ausdrucksgestalten einen sequenziellen Aspekt haben (auch wenn das nicht bedeutet, dass sie keinen simultanen oder rekursiven Aspekt hätten).

Die sequenzanalytische Herangehensweise an ein fotografisch dokumentiertes räumliches Arrangement hat – im Rahmen der übergeordneten Forschungsfrage

– das Ziel, das intuitive ›Verstehen‹ des Bildes methodisch zu verlangsamen und damit aus der Intuition herauszuheben. Was zu sehen ist und was die einzelnen Elemente bedeuten, mag (mehr oder weniger) offensichtlich sein – wie genau aber diese Bedeutung ausgehend vom sozial-räumlichen Arrangement entsteht, wird im Handeln, beispielsweise beim Betreten eines solchen Raumes, in der Regel nicht reflektiert. Dies ist im sozialen Vollzug auch nicht nötig; es wäre sogar hinderlich, wenn Menschen sich immer wieder aufs Neue kommunikativ versichern müssten, dass sie sich beispielsweise in einem Raum der Stille befinden und welche Verhaltensweisen dies nahelegt.<sup>5</sup> Sie verhalten sich schlicht gemäß (oder auch im Widerspruch zu) den Dispositionen eines Raumes, aber der soziale Vollzug kann und muss seine Ermöglichungsbedingungen nicht mitreflektieren. Dies wird erst in der handlungsentlasteten Situation der Analyse eines Protokolls möglich. Den »Handlungsdruck aufzulösen«, ist in der Sequenzanalyse eine zentrale Voraussetzung (z. B. Reichertz 2012, 225).

**Arbeitsschritte** Nach diesen Ausführungen kann ich das angewendete Verfahren und die konkreten Arbeitsschritte erläutern: Methodisch folge ich einem von GREGOR J. BETZ und BABETTE KIRCHNER vorgeschlagenen Vorgehen, bei dem das Foto zunächst in Segmente unterteilt wird, die dann nacheinander aufgedeckt und sequenziell ausgewertet werden (Betz und Kirchner 2016, 263).<sup>6</sup>

- a) Der erste Arbeitsschritt besteht demzufolge darin, ein fotografisches Protokoll in mehrere Segmente einzuteilen. Die Einteilung in Segmente orientiert sich an den bildlichen Einheiten, bleibt aber (vorbehaltlich künftiger methodologischer Entwicklungen) contingent. Auch die Zahl und Größe der Segmente kann unterschiedlich ausfallen. Darin unterscheidet sich die Einteilung von bildlichen Protokollen nicht von der Einteilung textlicher Protokolle in kurze Sequenzen. Die methodologische Hypothese ist jedoch, dass die Zahl und Größe der Sequenzen keinen Einfluss auf die rekonstruierte Sinnstruktur des Protokolls hat, solange hinreichend viele Sequenzen ausführlich genug diskutiert werden, um alle mög-

---

5 In einer anderen Rahmung, z. B. Kunst, kann diese Debatte selbstverständlich Teil der Atmosphäre sein, wenn ein Kunstwerk beispielsweise mit den Differenzen von Kunst und Religion spielt und Betrachter/innen dies ausdrücklich diskutieren.

6 Neben der Bildsequenzanalyse gibt es weitere, teilweise ähnlich angelegte Methoden zur Analyse von Bildern, z. B. das von Roswitha Breckner entwickelte Verfahren der Segmentanalyse (Breckner 2010; Breckner 2012). Die ebenfalls einflussreiche dokumentarische Methode (Bohnsack 2003, Bohnsack 2013, Bohnsack 2014, Bohnsack 2017) wirft der Sequenzanalyse eine unzulässige Übertragung des Textparadigmas auf Bilder vor und geht von einer eigenen Bildlogik aus. Mit der Konstellationsanalyse (Raab 2012) macht Jürgen Raab einen weiteren Vorschlag zur methodisch geleiteten Untersuchung von Bildern.

lichen Lesarten formulieren zu können (zur Durchführung einer solchen Vorgehensweise am konkreten Material siehe bspw. S. 216 ff).

- b) Der zweite Schritt des Verfahrens beinhaltet, dass für die erste aufgedeckte Sequenz, das heißt für den ersten aufgedeckten Bildausschnitt, gedankenexperimentelle Kontextvariationen durchgeführt werden. Die Teilnehmer:innen der Analysegruppe formulieren alltagsweltliche Szenarien, in denen diese Sequenz erwartbarerweise vorkommen könnte und sinnvolle Anschlüsse ermöglichen würde. Sie greifen dabei nicht auf fallspezifisches Kontextwissen zurück, weil dies die Formulierung der möglichen Lesarten massiv einschränken würde. Sie versuchen auch nicht, zu ermitteln, was einzelne Personen (z. B. Architekt:innen, Gestalter:innen, Eventmanager etc.) »sich gedacht« haben könnten (Peez 2006, 122). Es werden möglichst alle typischen erwartbaren Situationen gesammelt, um die Bandbreite der Möglichkeiten abzudecken (Betz und Kirchner 2016, 268).
- c) Im dritten Arbeitsschritt werden die möglichen Anschlüsse formuliert, die sich in der Regel aus den bereits formulierten Lesarten ergeben: Was könnte in der nächsten Sequenz folgen? Was ermöglicht die bereits diskutierte Sequenz, was erschwert sie? Hier wird erkennbar, warum es notwendig ist, sequenziell vorzugehen: Jede Sequenz ermöglicht eine Reihe von Anschlussmöglichkeiten, von denen das Protokoll aber immer nur eine realisiert, was wiederum dazu führt, dass man einige der zuvor formulierten Lesarten ausschließen kann. Aber nur, wenn man die ganze Bandbreite möglicher Anschlüsse öffnet, werden deren gemeinsame Merkmale sichtbar und der tatsächlich realisierte Anschluss kann als singulärer Ausdruck einer allgemeinen Struktur erkannt werden.
- d) Im vierten Arbeitsschritt kann dann die Folgesequenz aufgedeckt werden und man beginnt erneut mit der Formulierung von Lesarten und der gedankenexperimentellen Kontextvariation, so dass die Schritte zwei und drei alternierend wiederholt werden, bis das gesamte Protokoll oder ein gewählter Ausschnitt bearbeitet wurden. Das gesamte Verfahren kann für ein fotografisches Protokoll mehrere Analysesitzungen mit jeweils mehreren Stunden in Anspruch nehmen. Das verlangsame Bearbeiten eines Protokolls ermöglicht die Rekonstruktion der latenten Strukturen.

**Regeln** Bei allen Arbeitsschritten gelten einige Regeln (in Anlehnung an Wernet 2009, 21; Oevermann 2013, 78–79; Pickel und Sammet 2014, 45–46):

- a) Kontextfreiheit: Wie bereits geschildert werden Sequenzen ohne Berücksichtigung ihres unmittelbaren Kontexts diskutiert. Möglicherweise vorhandenes Wissen um den Entstehungszusammenhang von Protokollen oder um die Absichten und Beweggründe der beteiligten Akteur:innen wird ausgeblendet.

- b) Sequenzialität: Analysierende greifen nicht vor, um eine bestimmte Stelle im Protokoll zu erklären. Sie halten sich an den sequenziellen Ablauf des Protokolls und überspringen in den zu untersuchenden Passagen keine Elemente. Dies hat mit der Grundannahme zu tun, dass implizite Strukturen sich sequenziell entfalten. Nach jeder Sequenz eröffnen sich zahlreiche Möglichkeiten, von denen durch das Fortschreiten der Kommunikation genau eine gewählt wird. Um diesen Prozess der Optionen und der sequenziellen Schließung nachzuvollziehen, muss die Analyse ebenfalls sequenziell vorgehen. Auch wenn man häufig mit dem Beginn eines Protokolls anfängt, kann die Sequenzanalyse prinzipiell an jedem Punkt des Protokolls beginnen, muss dann aber sequenziell voranschreiten (Wernet 2014, 242).
- c) Sparsamkeit: Die Analysegruppe soll sich auf Lesarten beschränken, die ohne Zusatzannahmen und direkt aus dem Protokoll ableitbar sind. Die Teilnehmenden sollen nicht ausufernde oder unwahrscheinliche Lesarten entwickeln. Es geht um alltagsweltlich erwartbare Zusammenhänge.
- d) Extensivität: Im Gegensatz zur Sparsamkeitsregel besagt die Regel der Extensivität, dass für jede Sequenz alle Lesarten erschöpfend diskutiert werden müssen, um sicherzugehen, dass das gesamte semantische Potenzial einer Sequenz erschlossen wird (Wernet 2009, 33). Dies bedeutet aber nicht, dass zwingenderweise immer das gesamte vorliegende Protokoll sequenzanalytisch untersucht werden müsste, denn: »the particular structure of a case can be reconstructed in every part of a protocol« (Wernet 2014, 243). Für fotografische Protokolle hat es sich bislang bewährt, diese immer vollständig zu untersuchen; für textliche Protokolle untersuche ich jeweils ausgewählte Passagen.
- e) Wörtlichkeit: Das Protokoll muss so gelesen werden, wie es vorliegt. Es werden keine Lücken gefüllt oder vermeintliche Fehler korrigiert. Es werden keine Zusatzannahmen eingeführt, um vermeintlich unklare Elemente des Protokolls verständlich zu machen. Jede Lesart muss sich durch das Protokoll transparent begründen lassen und mit fallunspezifischem Kontextwissen plausibilisiert werden.

**Sequenzanalyse: Vor- und Nachteile** Wie alle Methoden so hat auch die Sequenzanalyse ihre Stärken und Schwächen. Man ist oft verwundert, wie viel Aufhebens um scheinbar Selbstverständliches gemacht wird. In den für diese Studie untersuchten Fotografien ist beispielsweise meistens ›auf den ersten Blick‹ erkennbar, dass es sich um kirchliche Räume oder Veranstaltungen handelt. Es könnte also mäßig erscheinen, diesen Umstand erst ausführlich zu diskutieren. Andererseits sollte gerade der Umstand, dass es sich ›offensichtlich‹ um einen Kirchenbau handelt, aus sozial- und kulturwissenschaftlicher Sicht Interesse auslösen. Dass wir ›auf den ersten Blick‹ wissen, was Sache ist, geschieht nicht von ungefähr und ist bei Weitem nicht selbst-

verständlich.<sup>7</sup> Dafür sind, so die Vertreter:innen der Sequenzanalyse, soziokulturelle Muster und Dispositionen verantwortlich, die diskursiv und bildlich in vielfältigen Varianten realisiert werden und auf Grund von Sozialisierung individuell – aber meistens ohne ausdrückliche Reflektion und Kommunikation – abrufbar sind. Diese soziokulturellen Muster zu rekonstruieren und dabei etwas über die Atmosphäre, also das realisierte Potenzial des sozialräumlichen Arrangements auszusagen, bleibt eine methodische und methodologische Herausforderung. Die Erhebung und Auswertung von Interviews oder teilnehmenden Beobachtungen ist dabei ein zentraler Bestandteil, doch liegt der Vorteil des hier vorgeschlagenen Ansatzes darin, zusätzlich einen Bereich der Empirie zu berücksichtigen, der andernfalls größtenteils brachliegen würde: Materielle Kultur, Architektur, Artefakte und deren Arrangements in sozialen Situationen. Damit kann man eine wesentliche Forderung des *material turn* – nämlich Materialität in ihrer kulturellen Bedeutsamkeit ernst zu nehmen – auch methodisch erfüllen. Wenn die Ergebnisse einer sequenzanalytischen Rekonstruktion sich wesentlich von denen unterscheiden, die durch Befragungen gewonnen werden, so wäre dies ein neuer Anlass zur Hypothesenbildung und weiteren empirischen Forschung.

Bei der Untersuchung von fotografischen Protokollen stoßen Analysierende häufig auf »Irritationen«: Der methodologische Einwand beim Stichwort »Irritation« (oder »Paradoxie«) ist oft, dass die Artikulation einer Irritation mehr über die Erwartungen der Analysierenden als über das Material aussagt. Wenn ich beispielsweise einen mit bunten Lichtern erleuchteten Kircheninnenraum zwar ästhetisch ansprechend aber zugleich irritierend finde, kann es sein, dass das nur daran liegt, dass mir persönlich protestantisch-schlichte Kirchen, aus welchen Gründen auch immer, eher vertraut sind.

Es ist aber unvermeidlich, dass die Bildung von Hypothesen über die Sinnstruktur eines Protokolls bei konventionellen Ansichten über das, was in einem bestimmten sozialen und kulturellen Kontext erwartbar ist, beginnt. Und genau hier liegt auch die Stärke einer sequenzanalytischen Auswertung: Da diese in einer Gruppe von Personen stattfindet, die einerseits mit den Regeln des Verfahrens vertraut sind und sich andererseits mit dem sozio-kulturellen Umfeld, aus dem das Protokoll stammt, auskennen, darf alles, was in der Gruppe als Lesart prinzipiell teilbar ist (und den Kriterien der Sparsamkeit und Extensivität entspricht; z. B. Wernet 2009, 21) als Lesart in die Hypothesenbildung aufgenommen werden. Es lässt sich dann im Laufe der Sequenzanalyse im Detail aufschlüsseln, wie und warum bestimmte Elemente und insbesondere deren Kombination im vorliegenden Protokoll als irritierend wahrgenommen werden. In anderen Worten: Nur wenn man die Struktur der

---

<sup>7</sup> Im Prolog zur zweiten Auflage seiner Studie zum Jonestown-Massaker von 1978 erinnert David Chidester: »One of the responsibilities of the academic study of religion is to make the strange familiar and the familiar strange« (Chidester 2003, xxv).

Irritation explizieren kann, können Irritationen etwas über die Struktur des Falls aussagen.<sup>8</sup>

Ein weiterer methodologischer Einwand, der gerade bei natürlichen Protokollen – also solchen Fotografien, die im Feld gefunden und nicht von Forschenden selbst aufgenommen wurde – auftritt, ist der Verdacht, das Foto könne ›gestellt‹ oder ›nachbearbeitet‹ sein. Damit sei seine dokumentarische Funktion eingeschränkt und es zeige nicht, was ›eigentlich‹ geschehen sei. Beispielsweise könnte der:die Fotograf:in nachträglich die Farben verstärkt haben oder durch die Einnahme einer bestimmten Position die Perspektive verzerrt haben. Inwiefern sind diese Einwände relevant, wenn es darum geht, die Sinnstruktur eines (fotografisch) protokollierten sozial-räumlichen Arrangements zu erschließen?

Die (angebliche oder tatsächliche) Gestelltheit des Fotos wäre dann relevant, wenn es um die Rekonstruktion der Sinnstruktur des Kunstwerkes (des hergestellten Fotos) ginge, also die Erzeugungsprinzipien, die ermöglicht haben, dass das Bild so publiziert wurde, wie es als Protokoll im Feld aufgefunden wurde. Das ist hier aber nicht das Ziel der Analyse, denn es geht um die Sinnstruktur des fotografisch protokollierten sozial-räumlichen Arrangements und nicht um die des produzierten Bildes. Das räumlich-soziale Arrangement ermöglicht, dass diese Aufnahme gemacht werden konnte. Vielleicht sind die Farben tatsächlich nachträglich verstärkt worden, dann aber nur, weil es in diesem Raum farbiges Licht gibt, das offenbar eine Rolle für den Fall spielt.

Ein ebenfalls häufig thematisiertes methodologisches Problem sind die mögliche Symbolik der Farben und die Assoziationen, die individuelle Analysierende mit Farben haben. Blau wird mit Kälte assoziiert, Rot mag für Liebe, Opfer oder Blut stehen. Diese spontanen Assoziationen können als Lesarten, sofern sie argumentativ teilbar sind, mitgeführt werden. Vergleichbar kann es in textlichen Protokollen vorkommen, dass eine Sequenz das Wort »Rose« oder »Taube« enthält: Auch hier wäre einer der möglichen sinnhaften Anschlüsse, dass es um Symbole für »Liebe« oder »Frieden« geht. Aber erst im Zusammenhang mit den folgenden oder vorangegangenen Sequenzen müssten diese Möglichkeiten ausgeschlossen werden, wenn sie sich als hypothetische Anschlussmöglichkeiten nicht erfüllen. Das gleiche gilt auch in fotografischen Protokollen: Wenn sich beispielsweise die semantische Verknüpfung von Blau und Kälte im Laufe der Analyse mit weiteren Elementen des Protokolls wechselseitig verstärken würde, wäre dies Teil der Sinnstruktur.

Wie bei jeder Methode entscheidet das erzielte Ergebnis darüber, ob die Methode in der Lage ist, eine Forschungsfrage adäquat zu bearbeiten. Die Sequenzanalyse scheint mit allen Vor- und Nachteilen ein geeigneter Kandidat zu sein, um die semantischen Potenziale sozial-räumlicher Arrangements zu rekonstruieren.

---

8 Ich danke Silke Müller für diese Formulierung.

In Kombination mit den klassischen und etablierten Verfahren (Interviews, Beobachtungen, Textanalysen) kann sich ein möglichst vollständiges und gegenstandadäquates Bild der untersuchten Fälle ergeben.

## 9.5 Methodologischer Exkurs: Bild und Text

Bilder und Texte sind seit langem – und, was die Bilder betrifft, spätestens mit dem sogenannten *visual* oder *iconic turn* wieder verstärkt – Gegenstand der Methoden- und Theoriedebatte in den Geistes-, Kultur- und Sozialwissenschaften (z. B. Boehm 2004; Bräunlein 2004; Hartinger 2007; Breckner 2010). Wenn es um Bilder und Texte als Daten zu einem bestimmten Forschungsgebiet geht oder wenn der Status ihrer Existenzweisen bestimmt werden soll (»Was sind Bilder?«, »Was sind Texte?«), wird oft betont, dass Bilder und Texte grundverschieden seien.

In diesem methodologischen Exkurs möchte ich nicht bestreiten, dass Bilder und Texte verschiedene Dinge – genauer: verschiedene Datengattungen – sind. Aber sie sind, so meine Hypothese, auch nicht so verschieden, dass man sie nicht prinzipiell mit den gleichen Methoden untersuchen könnte. Dieses Argument werde ich im Folgenden ausführen und dabei weniger vorschlagen, dass »Bilder wie Texte sind« – weil dies reflexhaft (und manchmal begründet) die Kritik eines ›Textualismus‹ auslöst –, sondern umgekehrt zu zeigen versuchen, dass »Texte wie Bilder sind« bzw. Bildern ähnlicher als oft angenommen.

Um die angesprochene Debatte produktiv weiterführen zu können, gilt es eine zentrale, aber meines Wissens oft vernachlässigte Unterscheidung deutlich zu markieren: Die Unterscheidung zwischen *Texten und Bildern als Daten* in sozial-, kultur- und geisteswissenschaftlicher Forschung auf der einen Seite und »Text« und »Bild« als theoretischen Konzepten bzw. als epistemologischen Metaphern auf der anderen Seite. Dieser Unterschied wird in den betreffenden Diskussionen nicht immer deutlich markiert: Wenn man beispielsweise formuliert, dass »Bilder nicht wie Texte gelesen« werden sollen, dann sind »Bilder« in dieser Formulierung das untersuchte Datenmaterial. Der Begriff »Texte« in dieser Formulierung kann aber sowohl das textliche Datenmaterial meinen – dann vergleicht man zwei Datengattungen miteinander und kommt zu dem (wenig überraschenden) Schluss, dass sie verschieden sind –, als auch auf das theoretische Konzept »Text« (in welcher Variante auch immer) verweisen. Umgekehrt ist der Begriff »Bild« beispielsweise im Kompositum »Sprachbild« kein Verweis auf bildliches Datenmaterial, sondern ein theoretisches (metaphorisches) Konzept, um auf eine bestimmte Funktionsweise von Sprache hinzuweisen.

Es ist auffällig, dass Begriffe im semantischen Feld »Text« (z. B. »Bedeutung«, »Gewebe«, »Metapher«, »Semiotik«) als theoretische Konzepte deutlich stärker ausgearbeitet worden sind als Begriffe im semantischen Feld »Bild« (z. B. »Ikonogra-