

2. Forschungsstand: Soziologie und Visualität

Bevor ich mich spezifisch dem Gegenstand der Videoanalyse zuwende, soll hier ein Überblick über die Rolle des Visuellen in der Soziologie (2.1), die Geschichte der Analyseverfahren (2.2), sowie der Videotechnologie (2.3) und den aktuellen Forschungsstand zur Videointerpretation (2.4) gegeben werden. Die ersten zwei Unterkapitel beziehen sich noch nicht direkt auf mein Thema, sondern stellen die Entwicklung der soziologischen Diskussion zum Visuellen – also zum Sehen, Zeigen und zu Bildern – vor. Video ist nicht nur ein rein visuelles sondern ein audiovisuelles Medium und unterscheidet sich damit recht grundlegend von unbewegten Bildern. In den unten aufgeführten Debatten um Visualität wird Video zwar mitgedacht, jedoch grundsätzlich nur als Randphänomen verhandelt. Die Besonderheiten der Videotechnik werden dann im zweiten Teil dieses Kapitel stärker hervorgehoben, wo ich mich auch stärker dem Forschungsthema dieser Arbeit annähere.

2.1 Exkurs: Sinne und Visualität in der Soziologie und Anthropologie¹

Das Sehen, das *Sichtbar-sein* und *Sichtbar-machen* sind in den klassischen Ansätzen der Soziologie immer mitgedacht und mitthematisiert, ohne explizit im Mittelpunkt der Forschung zu stehen. Erst in den vergangenen zwei Jahrzehnten erlebte das Interesse an diesen Phänomenen einen Aufschwung und wurde verstärkt in verschiedenen benachbarten Disziplinen wie der Wissenschaftsforschung oder den Bildwissenschaften verhandelt.²

Die Sinne, und vor allem auch der Sehsinn, wurden von einigen Autoren explizit theoretisiert, so schreibt Simmel in seinem Exkurs *zur Soziologie*

-
- 1 Dieses Kapitel habe ich gemeinsam mit Lisa-Marian Schmidt verfasst, es stellt die Einleitung zum Sammelband „Visuelles Wissen und Bilder des Sozialen“ dar, der im Rahmen eines gemeinsamen mit Petra Lucht organisierten Workshops zum Thema entstand. Der Text ist bis auf kleine Änderungen und Anpassungen direkt übernommen.
 - 2 Vorarbeit hat Jürgen Raab (2008) geleistet, der einige der wichtigsten Theoretiker in der visuellen Wissenssoziologie identifiziert und diskutiert. Für die Wissenschaftssoziologie sind insbesondere Heintz und Huber (2001) eine zentrale Quelle. Die Zusammenfassung erhebt keinen Anspruch auf Vollständigkeit, sondern dient der Einführung der momentan verhandelten Konzepte.

der Sinne gerade dem Auge eine besondere Bedeutung für Interaktion zu:

Unter den einzelnen Sinnesorganen ist das Auge auf eine völlig einzigartige soziologische Leistung angelegt: auf die Verknüpfung und Wechselwirkung der Individuen, die in dem gegenseitigen Sich-Anblicken liegt. Vielleicht ist dies die unmittelbarste und reinste Wechselbeziehung, die überhaupt besteht. (Simmel, 1908, S. 724)

Das gegenseitige Anblicken ist für Simmel die direkteste Form der Verbindung zwischen Subjekten – denn das Blicken ist nicht nur eine reine Informationswahrnehmung, sondern: „In dem Blick, der den andern in sich aufnimmt, offenbart man sich selbst; mit demselben Akt, in dem das Subjekt sein Objekt zu erkennen sucht, gibt es sich hier dem Objekte preis.“ (a.a.O.) Zentral ist der Unterschied zum Hören, denn beide Sinnesorgane sind jeweils für unterschiedliche Wissensformen zuständig: Das Sehen für das Gleichzeitige, Vielfältige und das Hören für das zeitlich geordnete Nacheinander. Die Verbindung der beiden stellt für Simmel eine zentrale Wechselwirkung dar – er geht in seinen Beobachtungen jedoch noch weiter. Simmel diagnostiziert, dass im Verkehr der Großstadt ein „unermessliches Übergewicht des Sehens über das Hören“ sich verfestige (A.a.O: S. 725). Mead nahm, etwa zeitgleich, verschiedene Formen von Gesten im sozialen Handeln in den Blick, und betont ebenso wie Simmel explizit verschiedene Aspekt der Sinne: Dass das Gehörte von mehreren gleichzeitig auf gleiche Art und Weise gehört werden kann begründet die besondere Qualität der akustischen, der Lautgeste. Hierauf bauen sein Kommunikationskonzept und die Erklärung der Herausbildung eines Bewusstseins auf. Auch das Visuelle spielt eine Rolle – gerade im weniger rezipierten, durch Joas (1989) wieder in den Vordergrund gerückten, pragmatistischen Kern des Werkes. Hier wird besonders die Koordination zwischen Auge und Hand betont:

The vast importance of the human hand for perception becomes evident when we recognize how it answers to the eye, especially among the distance senses. The development of space perception follows in normal individuals upon the interaction of the eye and the hand, and this interaction works a continual meeting of the discriminations of the eye by those of the skin, mediated through the manipulating hand. (Mead, 1907, S. 385)

Für Mead bietet das Auge dem Handelnden eine Reihe visueller, unzusammenhängender Eindrücke, die jedoch durch die dauerhafte Kontaktführung der Hand zusammengeführt werden können. Mead analysiert mit Rückgriff auf Dewey (1896), dass die Wahrnehmung der visuellen Reize beileibe kein passiver Akt, sondern ein aktives Schauen und Suchen nach diesen ist: Sehen ist Teil einer Handlung in spezifischen Kontexten. Wahrnehmung, insbesondere auch des Visuellen, begreift Mead als *Situ-*

ation – in welcher sich Sehender und Gesehenes befinden (Mead, 1969, S. 111).

Für die heutige Debatte wichtige Gedanken, die Anknüpfungspunkte an die Soziologie darstellen, finden sich weiterhin in der philosophischen Anthropologie. Sowohl Plessner als auch Gehlen setzten sich mit den Sinnen auseinander. Neben seiner Analyse der visuellen Gesten des *Lachens und Weinens* (1941) wird in Plessners Anthropologie der Sinne das Auge zwar als zentraler Fernsinn hervorgehoben, der mit dem Tastsinn eine Verbindung eingeht, vor allem aber warnt er vor unreflektierter Hinnahme des Sichtbaren:

Sehen ist strukturell etwas – Sehen, direkt ohne Vermittlung. Wir sagen, es zeigt sich selber, originär, es erscheint, und geraten durch diese Umschreibung allzu leicht auf erkenntnistheoretische oder ontologische Abwege. Die beliebte Formel ‚es zeigt sich von Selbst‘ suggeriert die Vorstellung einer Aktivität vom Gesehenen aus, Erblicken, ins Auge Fassen, vom Sehenden aus. (Plessner, 1970, S. 201)

Das Visuelle spielt bei den bislang genannten Autoren vor allem im Sehen des Anderen – der Bewegungen, Gesten und Zeichen – eine Rolle. Diese Linie lässt sich weiterführen bis in die gegenwärtige, mit Interaktionen befasste Soziologie. Zentral ist hier sicherlich Erving Goffman – der das Soziale als Theater, als *Schaubühne* begreift (1959) und für den das sichtbare Gesicht und die Darstellung die zentrale Rolle spielt und beispielsweise in Werbebildern untersucht wurde (1979). Auch im symbolischen Interaktionismus, dem es um die Symbole geht, gibt es die Hinwendungen zur Bildlichkeit – insbesondere zur Kunst und Photographie, etwa bei Howard Becker (1974).

In der französischen Debatte thematisiert Foucault den Blick vor allem als zentrales Element in seiner Beschreibungen der Macht, so wie bereits im ärztlichen Blick (1976), später im disziplinierenden Blick im Panoptikum (1977) sowie in seiner feinen Analyse von Velázquez Gemälde *Las Meninas* (1974).

Aber nicht nur die Beschäftigung mit dem Blick und dem Sehen des Anderen in *Face to Face* Interaktionen, sondern auch die Verwendung (visueller) Symbole ist grundlegend für die Soziologie. So ist das Totem, das bei Durkheim die Einheit des Clans symbolisiert, ein visuelles Objekt. Bilder erfüllen so eine spezifische soziale Funktion. Explizit verweist Gehlen (1965) in seiner Kunstsoziologie darauf, dass Kunst machtvolle Symbole schafft und dadurch handlungsleitend wirkt:

Keine Macht, welche sich zur Herrschaft im entschiedenen Sinn berufen fühlt, kann darauf verzichten, das Bewusstsein des Menschen zu besetzen, und die Endgültigkeit ihres Anspruchs drückt sich darin aus, dass sie dieses Bewusstsein vollständig bestimmt: also bis in die Anschauungen hinein. Was daher, vom einzelnen Menschen her gesehen, als Außenhalt des Inneren erscheint, stellt sich von den Institutionen aus als Repräsentation dar; denn sie verkörpern sich in sichtbaren und daseinsmächtigen Symbolen, unter denen die Künste stets einen hohen Rang einnahmen. (Gehlen, 1965)

Er unterscheidet verschiedene Formen der Kunst: Überkunst archaischer Gesellschaften, religiös geprägte ideelle Kunst, ‚rationale‘ realistische Kunst und subjektbezogene abstrakte Kunst, die sich jeweils in bestimmten Gesellschaftsformen herausbilden. Besonders gegenüber der abstrakten Kunst bleibt er sehr skeptisch – ein Ausdruck seiner konservativen Haltung. Weiterhin entwickelt Gehlen, ausgehend von verhaltensbiologischer Forschung und den oben beschriebenen Gedanken Meads, den Begriff der motorischen Phantasmen (Gehlen, 2013, S. 210), in welchem Bewegungs- und Körperkonzepte als visuelle, handlungsanleitende Elemente, beschrieben werden. Für Gehlen sind optische Dinge „zunächst in Bewegungen und Handgriffen entwickelte Dinge“, die sich mit der Tastwahrnehmung verbinden, um das Handeln zu ermöglichen (Gehlen, 2013, S. 214).

Wenn Weber sich eher der Musik als dem Visuellen zuwendet, so hat aber auch er die Entwicklung des Sehens zumindest erwähnt. Für ihn spielt die Rolle der linearen und Luftperspektive (Raab, 2008; Weber, 2004) als Dokument der Rationalisierung der Ästhetik eine Rolle. Hier geht Mannheim noch einen Schritt weiter, wenn er die *Weltanschauung* einer Epoche oder Gruppe zu rekonstruieren sucht (1964). Für den Wissenssoziologen stehen die unterschiedlichen Wissensbestände im Fokus, die er als Weltanschauung beschreibt. Darunter versteht Mannheim gruppen- oder epochenspezifische, (vor allem) atheoretische Vorstellungen, die sich im Handeln, aber besonders in Objektivationen finden. Diese lassen sich nicht und nicht ohne weiteres in theoretische Konzepte übersetzen. Sein Vorschlag besteht in der Aufgliederung in drei Sinnschichten – er unterscheidet den *Dokumentsinn* einer Handlung oder Aussage von einem *objektiven Sinn* und einem *intendierten Ausdruckssinn*, also subjektiven Sinn des Akteurs (1964, S. 104). Auf dieser Unterscheidung baut Panowskys Ikonologie (1975) auf, die zunächst vor allem für die Kunstwissenschaft und spezifische Bildtypen genutzt wurde. Heute erheben daran anschließende Verfahren aber doch den Anspruch, auch breiter die sogenannten *schwachen Bilder*³ erfassen zu können.

Anschließend an diese beiden Autoren folgt Bourdieu, dessen Habitus-Begriff ja auch von Panofsky übernommen ist. Er selbst hat sich mit Visualität explizit auseinandergesetzt, in seinen eigenen Photographien in Algerien (2009) sowie in seiner Untersuchung zur *Photographie als Illegitimer Kunst* (2014), in welcher er den Gebrauch des Fotoapparates untersucht und sehr detailliert aufzeigt, wie diese scheinbar neutrale Praxis eingebunden ist in soziale Felder und ihre Wirkungen. Nicht jeder kann

3 Siehe für diesen gegenwärtigen Begriff Boehm (1994), das Konzept wird unten erläutert.

alles fotografieren, für die Bauern im Béarn ist dies etwas Außeralltägliches, für die Städter normal. An diesem Beispiel verbindet Bourdieu die jeweilige Ästhetik mit der Sozialstruktur und entwickelt somit das Konzept der Dispositionen.

Theorien des Visuellen haben in der Soziologie in den vergangenen Jahrzehnten eine untergeordnete Rolle gespielt. Dominant war der Linguistic Turn, ein Paradigmenwechsel in der Soziologie, durch den vor allem das gesprochene und geschriebene Wort im Vordergrund stand und von den Sozialwissenschaften aufs Feinste untersucht wurde (wozu auch die Entwicklung des Audiorekorders beitrug). Impulse, das Visuelle zu thematisieren, stammten in den letzten Jahrzehnten vor allem aus anderen Disziplinen:

2.1.1 Bildwissenschaft

Unter anderen haben sich Wissenschafts- und Kunsthistoriker, Medienwissenschaftler und Philosophen in den vergangenen Jahren intensiv mit den Themen bzw. dem Themenkomplex Bild, Visualität und Bildpraktiken beschäftigt. Diese in Deutschland recht junge interdisziplinäre Debatte mündete in der Forderung nach und ersten Versuchen der Etablierung eines neuen Forschungsfeldes, der Bildwissenschaft, die diese Fächer und ihre jeweiligen Perspektiven auf das Thema bündelt. Eine längere Tradition besteht hingegen im angloamerikanischen Raum der *Visual Studies* und *Visual Culture Studies*, die explizit die Hinwendung zu visuellen Kulturen des Alltags fordern:

Visual Culture directs our attention away from structured, formal viewing settings like the cinema and art gallery to the centrality of visual experience in everyday life. At present different notions of viewing and spectatorship are current both within and between all the various visual subdisciplines. (Mirzoeff, 1998, S. 7)

Die zentralen Bezugspunkte und paradigmatischen Ansätze in den Bildwissenschaften beziehen sich zumeist auf die zwei programmatischen Ausrichtungen des *Pictorial Turn* und des *Iconic Turn*. Mit dem *Pictorial Turn* beschreibt Mitchell (Mitchell, 1997, S. 15–18) zwei Beobachtungen. Zum einen, dass verbunden mit technischen Entwicklungen vor allem der Informations- und Kommunikationstechniken, Bilder in einem hohen Maße das kulturelle Leben bestimmen, aber gleichzeitig auch eine ‚Furcht vor dem Bild‘ und seiner manipulierenden, mächtigen Wirkung besteht. Zugespitzt formuliert Bredekamp (2007, S. 102) die These, dass sich in der Moderne ein radikaler Wandel vollzogen hat, in der Bilder durch eine Art *Kopernikanischer Wende* das kulturelle Leben dominieren. Zum anderen sieht Mitchell aber auch deutliche Anzeichen für einen Wandel in der

akademischen Debatte, in der sprachanalytische Modelle, die bisher paradigmatisch für die Analyse kultureller Phänomene waren, abgelöst werden, durch eine stärkere Beschäftigung mit Bildern und Versuchen, diese anhand nicht-sprachorientierter Modelle zu analysieren. Auch Boehm (1994) proklamiert mit dem *Iconic Turn* die ‚Wende zum Bild‘ in ähnlicher Weise wie Mitchell. Boehm diagnostiziert, dass Bilder allgegenwärtig in der Gegenwartskultur sind, dagegen als Forschungsgegenstand in der akademischen Debatte lange marginalisiert wurden (a.a.O., S. 11-17). Er fordert eine Abwendung von sprachanalytischen Modellen in der Auseinandersetzung mit Bildphänomenen, um der *Ikonischen Differenz*, die Bilder auszeichnet, gerecht zu werden. Anders als im *Pictorial Turn* geht es Boehm um eine ontologische Bestimmung der Eigenlogik des Bildes und nicht explizit um eine historische und soziokulturelle Berücksichtigung von konkreten Bildpraktiken sowie Bild- und Sehtraditionen.

Diese Ausschnitte der Debatten in den Bildwissenschaften lassen es erkennbar werden, dass die These der Hegemonie von Bildern soziologisch äußerst relevant ist, auch wenn sie weiterer empirischer Analysen bedarf, um einen tatsächlichen Wandel feststellen zu können.⁴ Es lässt sich fragen, ob und wie sich Wissen, Institutionen und soziale Beziehungen durch den zunehmenden Einsatz bildlicher Kommunikation, bzw. dem Einsatz von Bildern bei der Wissensproduktion, verändern. Generell ist festzustellen, dass Bilder als Kommunikationsform oder kulturelles Artefakt, als „Konstante menschlicher Kulturen“ (Schnettler & Pötzsch, 2007, S. 476), immer eine Rolle gespielt haben.⁵ Die These der Dominanz der Bilder in der Moderne übersieht, dass Bilder selten allein auftreten. Vielmehr lässt sich die Verbreitung von ‚hybriden‘ Formen, also Text-Bild-Zahlen-Kombinationen, beobachten (a.a.O., S. 480), beispielsweise im Internet oder im Fernsehen. Zu betonen ist auch, dass der Zugang und die Nutzung digitaler Informations- und Kommunikationstechniken, die laut der oben dargestellten Thesen die Bilderflut vorantreiben, ungleich unter verschiedenen gesellschaftlichen Gruppen und Gesellschaften verteilt ist und dringend weiterer Forschung bedarf.⁶ Boehms Forderung nach der Analyse der Eigenlogik des Bildes bzw. des Visuellen ist aus soziologischer Perspektive auch deshalb nicht fruchtbar, da er Bilder entkoppelt von den sozialen Praktiken der Herstellung, Nutzung und Rezeption betrachtet

4 Für ein exemplarisches Beispiel zum Anstieg des Fernsehkonsums in Deutschland siehe Raab (2008).

5 Zur Bedeutung von Bildern in der mittelalterlichen Kommunikation siehe Knoblauch (Knoblauch, 2010, S. 331). Exemplarisch für die Wissenschaftskommunikation siehe Shapin/Schaffer (Shapin & Schaffer, 2011, S. 60–69).

6 Exemplarisch hierzu die Debatte um die „digitale Spaltung“ (Zillien, 2006, S. 82 ff).

(Burri, 2008, S. 72; Schnettler & Pötzsch, 2007, S. 475). Aus soziologischer Perspektive haben Bilder zunächst keine feststehende Bedeutung oder Funktion, da diese sich erst in den sozialen Verwendungszusammenhängen konstituiert. Diese sozialen Konstruktionsprozesse der Bedeutungen, der Formen und sozialen Funktionen von Bildern wurden insbesondere in der Wissenschaftssoziologie analysiert.

2.1.2 *Sehen und Visualisieren in den Wissenschaften*

Wissenschaftsbilder – die Praktiken der Herstellung, Deutung, Verbreitung und Wirkmächtigkeit von Bildern in der Wissenschaft – sind Gegenstand der Wissenschafts- und Techniksoziologie seit Ende der 1970er Jahre. Schwerpunkt der Forschung sind zumeist naturwissenschaftliche Bildpraktiken, während kaum Arbeiten zu Bildern in den Geistes- oder Sozialwissenschaften bestehen.

Insgesamt wird eine ‚zunehmende Piktoralisierung der Wissenschaft‘ verzeichnet, die sich nicht nur in der erhöhten Quantität von Bildern, sondern auch als eine bessere Qualität der Bilder ausdrückt. So stellen diese zunehmend selbst das *Epistemische Objekt* der Forschung dar (Heintz & Huber, 2001). Im Zuge der Digitalisierung der Wissenschaft und damit der Verbreitungs- und Bearbeitungsmöglichkeiten der Bilder, wird somit auch die Gewissheit, was diese Bilder zeigen und welchen Status diese haben, zunehmend unklar (Heßler, 2009, S. 11 ff).

Als gemeinsame Einsicht der verschiedenen Analysen zu der Arbeit an und mit Bildern in der Wissenschaft lässt sich erstens festhalten, dass die Wissenschaftsbilder, insbesondere der Naturwissenschaften, weniger als Abbilder der Forschungsobjekte der Wissenschaften verstanden werden, sondern vielmehr hochgradig artifizielle und durch viele Transformations- und Bearbeitungsschritte im Labor hergestellte soziale Produkte, eben „Visualisierungen, Sichtbarmachungen oder Inskriptionen“ darstellen. Zweitens wird in der Zusammenschau dieser Arbeiten deutlich, dass das Sehen und die Verwendung von Bildern als *Epistemische Objekte* zentrale Erkenntnispraktiken in den Wissenschaften darstellen.

Das Labor stellt dabei den lokalen, *artifiziellen Handlungskontext* dar, der spezialisierte Erzeugungsvorgänge der Wissensproduktion ermöglicht. Der *soziale Fabrikationsprozess* von Bildern in der Wissensproduktion involviert dabei eine Reihe von Selektionen, Entscheidungen und Verhandlungen, die wiederum die weitere Arbeit strukturieren (Knorr-Cetina, 2002, S. 27). Diese Entscheidungen sind abhängig von u.a. situationsspezifischen Bedingungen im Labor. Der lokale Kontext der Bildherstellung und Bilddeutung schreibt sich dabei durch die Arbeit am Bild in die Bilder ein. Dies ermöglicht und beschränkt zugleich alle weiteren Schritte

der Wissensproduktion. Die Wissensproduktion an und mit Bildern wird aus dieser Perspektive als eine Serie von sozialen Verhandlungsprozessen verstanden, in deren Verlauf Deutungen, beispielsweise was auf einer astronomischen Aufnahme zu sehen ist, letztendlich zu anerkanntem Wissen stabilisiert und häufig wiederum in Bildern objektiviert werden.

Dabei wird auch der Doppelstatus von Bildern als einerseits visuelles Medium und andererseits materielles Artefakt hervorgehoben (Burri, 2008, S. 346). Die visuelle Ebene des Bildes kann dabei zu Gunsten der materiellen Seite des Bildes in den Hintergrund rücken. Im Hinblick auf die zunehmende Computerisierung und Digitalisierung von Wissenschaften wird es meines Erachtens daneben wichtig, die verschiedenen Ebenen der Bilder noch feingliedriger zu unterscheiden. Es gilt unter anderem die Handlungen auf der numerischen Ebene von digitalen Bildern – und inwiefern diese Relevanz hat in der konkreten Praxis bspw. der Bearbeitung von Bildern – zu berücksichtigen.

Ein zentraler Forschungsgegenstand in wissenschaftssoziologischen Arbeiten ist die Dominanz des Sehsinnes gegenüber anderen Sinnen in den Wissenschaften (Burri, Schubert, & Strübing, 2011, S. 3). In der „Hierarchie der Sinne“ hat besonders Ludwik Fleck (1983) sich dem Sehen als zentrale Erkenntnispraktik in den Wissenschaften gewidmet. Durch die Sozialisation in einen spezifischen Denkstil einer Wissenschaftlergemeinschaft – des Denkkollektives – erlernt der ‚Novize‘ typische visuelle Formen zu unterscheiden und zu erkennen. Sein anfängliches *Schauen* wandelt sich in diesem Prozess zu einer denkstilgemäßen sinnhaften visuellen Wahrnehmung: dem *Sehen*. Fleck bringt es wie folgt auf den Punkt: „Wir schauen mit den eigenen Augen, aber wir sehen mit den Augen des Kollektivs“ (A.a.O., S. 154). Auch die Bildgestaltung und Ästhetik im Prozess der Wissensproduktion folgt dabei den Stilformen des Denkkollektives.

Spezifische Bildstile und ästhetische Gestaltungen der Bilder dienen in den Wissenschaften, so die Analysen, beispielsweise der Durchsetzung von Deutungsansprüchen, der Verdichtung von Argumenten oder aber der Vermittlung komplexer Wissensbestände in die Öffentlichkeit: „Das Bild der Wissenschaft ist ganz wesentlich durch Bilder geprägt“ (Adelmann, Frercks, Heßler, & Hennig, 2009, S. 41). Besonders deutlich wird dies sicherlich an der Hirnforschung und Nanotechnologie, deren Bilder Einzug gehalten haben in eine Reihe an populären Magazinen, Fernsehsendungen usw. Deutlich wird in der prominente Studie von Lynch & Edgerton (1988) zur Astronomie, dass eine Reihe an Gestaltungen und ästhetischen Entscheidungen in der wissenschaftlichen Arbeit an Bildern vorgenommen wird, je nach Adressaten der Bilder. So lassen sich

Bildformen und Bildgestaltungen, die an die Fachgemeinschaft adressiert sind (Zeitschriften, Lehrbücher usw.), von denen unterscheiden, die in der Öffentlichkeit zu finden sind. Dies verweist auf die Frage nach der Vermittlung spezialisierter Wissensbestände im Medium Bild, aber auch auf die Frage nach der Überzeugungsmacht von Bildern. Die besondere Überzeugungswirkung von Bildern, beispielsweise in wissenschaftlichen Kontroversen, wird in einer Reihe von Studien hervorgehoben. Diese Bildermacht kann dabei auf unterschiedlichen sozialen und medialen Mechanismen beruhen.

So argumentiert zum Beispiel Heintz (2007), dass die spezifischen semiotischen Logiken von Sprache, Zahlen und spezifischen Bildtypen jeweils unterschiedlich operieren und dass sich damit auch die Überzeugungswirkungen, die erzielt werden können, unterscheiden. Naturwissenschaftliche Bilder etwa blockieren im Gegensatz zu künstlerischen Bildern jede Kontingenz, da sie ihre eigene Bildhaftigkeit nicht zum Thema machen, und suggerieren, dass sie lediglich auf etwas Bildexternes verweisen. Ihr hergestellter, artifizieller Status wird also negiert und dadurch wird eine gewisse Eindringlichkeit erzielt (a.a.O., S. 78). Zu betonen ist hier, dass die besondere Eindringlichkeit und Wirkmacht der einzelnen Zeichensysteme soziale Zurechnung ist, die auf bestimmten Traditionen und Konventionen beruhen. Der Macht von Bildern und Bildpraktiken in den Wissenschaften hat sich auch eine Reihe an Arbeiten aus der feministischen Wissenschaftsforschung (bspw. Lammer, 2001) gewidmet sowie Studien, die an Foucaults Diskursanalyse anschließen (Maasen, Mayerhauser, & Renggli, 2006). Sie zeigen an und mit Bildern deutlich die Einbettung und Verweisungszusammenhänge der Bilder und Handlungen an und mit Bildern in spezifischen Machtkonstellationen und Herrschaftsverhältnissen beispielsweise der Geschlechterverhältnisse in den Wissenschaften, wie auch in weiteren Diskursen auf. Die Frage, was wie in den Bildern sichtbar gemacht wird (und von wem) und welche Standpunkte hierbei marginalisiert werden, muss dies mit berücksichtigen.

2.1.3 *Gegenwärtige soziologische Theorie und Empirie des Visuellen*

Einerseits durch die Forschung in den Laborstudien und der Wissenschaftssoziologie beeinflusst und andererseits durch die sich formierenden Bildwissenschaften aufgerüttelt, gibt es auch innerhalb der Soziologie eine verstärkte Zuwendung zu Visualität. Eine Vielzahl von Studien hat sich mit den verschiedensten Aspekten auseinandergesetzt – eine Formierung des Feldes steht meiner Ansicht nach jedoch noch aus.

Vertiefende theoretische Analysen und Aufarbeitungen wurden in den

letzten Jahren durchaus vorgelegt, dennoch muss man ihnen innerhalb der Soziologie noch einen Pionierstatus zugestehen. Insbesondere Raab (2008), den ich oben bereits erwähnt habe, hat das Feld aufgearbeitet und sein hermeneutisches Verfahren mit der begrifflichen Konzeption der *Sehgemeinschaften* verknüpft. Er beschreibt damit bestimmte visuelle Stile, die typisch sind für jeweils bestimmte Gruppen. Er untersucht dies an den verschiedenen *Schnittmustern* der Amateur-Videoproduktion, also von Hochzeitsvideos, Musikvideos und Dokumentarfilmen. Hierbei verbindet er die Untersuchung der jeweiligen (ästhetischen) Sehkonventionen der verschiedenen Gruppen mit einer feinen Analyse ihrer jeweiligen Produkte.

Die neuen visuellen Kommunikationsformen, die Einzug in verschiedenste Bereiche nehmen, wurden am Beispiel von Powerpoint Präsentationen untersucht (Knoblauch, 2012; Schnettler & Knoblauch, 2007). Powerpoint und seine Verbreitung lässt sich nicht auf ein technisches Programm oder die gezeigten Bilder reduzieren, sondern stellt eine Kommunikationsform dar, wobei besonders die verschiedensten Verknüpfungen zwischen den lokalen Handlungen, Körperanordnungen, Bildern und dem räumlichen und sozialen Kontext im Mittelpunkt stehen.

Das Interesse an der Visualität geht Verbindungen mit dem Interesse an der Verwendung von Bildern in den Sozialwissenschaften einher. So stellt Gerald Beck (2013) sich die Fragen, wie Visualisierungen in der Soziologie genutzt werden und welche Rolle sie in der Produktion wissenschaftlichen Wissens spielen. Er untersucht dies semiotisch-bildhermeneutisch an einer Reihe von Fallbeispielen, wie etwa der *Bolte-Zwiebel*, *Photographischen Verdichtungen* und *Risikokartierungen* aus verschiedenen Forschungsprojekten. Er wendet sich hierbei vor allem der Wissenschaftskommunikation im Rahmen der reflexiven Moderne zu und verbindet dies mit den Konzepten der autoritativ-selbstreferentiellen Wissensproduktion im *Modus 1* und der dialogischen, praxisorientierten im *Modus 2* (Gibbons u. a., 1994). Wenn seine Studie auch nicht die einzige ist, die offensiv den Ikonoklasmus in der Soziologie angeht, verweist er dennoch auf ein wichtiges Thema, das in Zukunft breiter diskutiert werden muss.

Der Umgang mit Visualität wird in der Soziologie einem anderen Kontext bereits reflektiert: In den Diskussionen über den Umgang mit empirischen Daten in visueller Form, die ich im folgenden Kapitel genauer besprechen werde. Zunächst gehe ich aber auf die oft vergessene Entwicklung der Nutzung audiovisueller Daten im Methodenkanon der Sozialwissenschaften ein.

2.2 Geschichte und Entwicklung visueller Analysen⁷

Obgleich es sich bei der sozialwissenschaftlichen Analyse audiovisueller Daten um eine relativ junge Methode handelt, kann sie sich auf eine erstaunlich lange Reihe von Vorläufern stützen. Dazu zählen zweifellos der ethnographische Film und die sozialwissenschaftlichen Foto- und Filmanalysen, wie sie in den letzten Jahren in der Anthropologie, den Visual Studies und der visuellen Soziologie entwickelt wurden. Allerdings kann hier keine unilineare Geschichte rekonstruiert werden, sondern eher verschiedene Entwicklungslinien, die in der Videographie münden. Es sollen einige der relevantesten Linien nachgezeichnet werden, in denen visuelle Aufzeichnungen zur Analyse kommunikativer Handlungen verwendet wurden. Eine historische Darstellung der Nutzung von filmischem Material findet sich in dem Buch von Ramón Reichert, ‚Kino der Humanwissenschaften‘ (2007).

2.2.1 Vorläufer visueller Analysen

Bevor ich mich mit der Bedeutung von Videos beschäftige, muss ich deren Vorgänger – Filme und Bilder – kurz in den Blick nehmen. So wurden beispielsweise Photographien bereits kurz nach ihrer Erfindung in den späten 1830er-Jahren für die Analyse von Gesichtsausdrücken eingesetzt. Einige Dekaden später nutzte Charles Darwin in seinem Buch über *Die Gemütsbewegungen bei Menschen und Tieren* (2000) Fotos, um den Ausdruck von Gefühlen bei Menschen und Tieren miteinander zu vergleichen. Etwa um dieselbe Zeit, in den 1870er-Jahren, entwickelte Eadweard Muybridge mithilfe einer speziellen Aufnahmetechnik durch Aneinanderreihung von Einzelbildern die ersten filmähnlichen Sequenzen. Zweifellos stand die Entwicklung des Films seit seiner Anfangsphase in direktem Bezug zu verhaltenswissenschaftlichen Studien. Dies kommt vor allem in Muybridges berühmten Analysen der Bewegungsabläufe von Tieren und Menschen zum Ausdruck. Dabei gelang es ihm, die bis dahin ungeklärt gebliebene Frage zu beantworten, ob Pferde während des Galopps einen Huf am Boden halten oder tatsächlich in einem bestimmten Augenblick alle vier Hufe abheben. Seine Aufnahmen bewiesen, dass Pferde tatsächlich einen Moment lang ‚fliegen‘. In den 1880er-Jahren wandte sich Muybridge menschlichen Verhaltensabläufen zu, was wiederum in der Ethnologie aufgenommen wurde. Regnault drehte 1890 den ersten ethnographischen Film und vor allem Haddon propagierte die Verwendung des Films in der Ethnographie.

7 Diese Rekonstruktion der Geschichte ist ursprünglich erschienen in der methodischen Einführung in die Videographie (Tuma, Knoblauch, & Schnettler, 2013). Sie basiert auf den Vorarbeiten meiner Co-Autoren und wird hier in leicht angepasster Fassung übernommen.

Die frühe Anerkennung von Bildern wird deutlich, wenn man bedenkt, dass zwischen 1896 und 1916 insgesamt 31 Artikel im *American Journal of Sociology* veröffentlicht wurden, in denen visuelles Material zu finden ist. Jedoch ebte schon kurz danach das Interesse in der Soziologie am Einbezug visueller Materialien in die Forschung deutlich ab und beschränkte sich auf einige wenige Ausnahmen. Dies ist vermutlich dem damals wachsenden Einfluss statistischer Methoden geschuldet, in dessen Folge Photographien abrupt von Tabellen, Formeln und Graphen ersetzt wurden, die nun als die einzig angemessenen wissenschaftlichen Illustrationen Geltung errangen (Stasz, 1979). Es ist deshalb wenig verwunderlich, dass es in den folgenden Jahren vor allem Initiativen von außerhalb der Wissenschaft waren, die entscheidende Impulse für die Entwicklung einer erst spät wieder aufkeimenden ‚visuellen Soziologie‘ lieferten.

Zu den wenigen Nutzern visuellen Materials zählt Henry Ford, der – aufbauend auf den Prinzipien des wissenschaftlichen Managements von Frederick W. Taylor – Filmaufnahmen einsetzte, um die Organisation der betrieblichen Arbeit zu studieren und die Teilung der Arbeit zwischen Mensch und Maschine effizienter zu gestalten (Bryan, 2003). In Deutschland untersuchte Kurt Lewin um 1923 die ersten Filmsequenzen menschlichen Konfliktverhaltens. Sein Schüler Gsell veröffentlichte 1935 das Buch *Filmanalyse als Methode der Erforschung des Verhaltens* (Thiel, 2003).

2.2.2 Proxemik, Kontextanalyse und Kinesik

Eine besondere Bedeutung spielt hier die genaue Beobachtung menschlichen Handelns anhand von visuellen Daten. Die wissenschaftliche Verhaltensbeobachtung hat vor allem in der Ethologie⁸ und der sozialpsychologischen Forschung eine starke Verankerung. Entsprechend wurden zum Studium von emotionalem Ausdruckverhalten und zur Interaktionsanalyse schon sehr früh visuelle, filmische und audiovisuelle Techniken eingesetzt. Ein eindrückliches Beispiel ist der von Edward T. Hall (1990) in den 1950er-Jahren entwickelte Ansatz der *proxemischen Analyse*, der sich mit der Rolle der Nutzung von Räumen und des spatialen Verhaltens in der Interaktion beschäftigt und diese für den interkulturellen Vergleich einsetzt. Die Proxemik beruht auf Alltagsbeobachtungen sozialer Formen räumlichen Verhaltens, ihre Entdeckung geht aber auf die sorgfältige Analyse einer Filmsequenz zurück. Nach zahlreichen Wiederholungen einer Aufzeichnung konnte Hall dabei den Grund für die Störung der Interaktion zwischen einer amerikanischen Touristin und einer Indiofrau

8 Ethologie (Verhaltensbiologie) ist zu unterscheiden von Ethnologie (Völkerkunde) und beide wiederum nicht zu verwechseln mit Ethnographie (Methode der Feldforschung) und der Ethnomethodologie (eine spezifische Forschungsrichtung der Soziologie).

auf einem Markt in den Anden entdecken: Es handelte sich um eine Überschreitung kulturell unterschiedlicher körperlicher Zonen. Solche Zonen, meinte Hall, schichten sich gewissermaßen in einer Abfolge von intimer, sozialer und öffentlicher Distanz um das Individuum auf und sind kulturell je unterschiedlich zugerichtet. Diese „stumme Sprache des Raumes“ kann dann insbesondere in der interkulturellen Interaktion zu nachhaltigen Störungen und Missverständnissen führen, weil sie sich als eingelebte Selbstverständlichkeit der Kontrolle der Handelnden oft entzieht.

So anregend die proxemischen Analysen sind, so sehr sind sie mit einer wenig produktiven Separierung einer einzelnen kommunikativen Modalität verbunden – der des Raumverhaltens. Die Proxemik sensibilisiert für eine oft vernachlässigte Dimension, steht aber in der Gefahr, diese aus ihrem Zusammenhang zu reißen und überzubetonen.

Diese Zergliederung wird in einem anderen Ansatz überwunden. Eine der sozialwissenschaftlich bedeutsamsten Richtungen nennt sich bezeichnenderweise *Kontextanalyse*, die von Argyle auch als ‚strukturelle Analyse‘ bezeichnet wird. Sie wird heute u. a. von Adam Kendon vertreten, der Ray Birdwhistell, Albert Scheflen (der den Begriff 1963 prägte) und Erving Goffman als deren Begründer nennt. Genau genommen gingen Einflüsse aus der interpersonellen Psychiatrie, der Anthropologie, der Informationstheorie und der strukturellen Linguistik in diesen Ansatz ein. Die Kontextanalyse betrachtet Kommunikation in der Interaktion als einen fortwährenden Multikanal-Prozess und versucht die strukturellen Merkmale des in der Interaktion ablaufenden Kommunikationssystems zu beschreiben.

Historisch sehr bedeutsam für die Entwicklung audiovisueller Analysen war die sogenannte *Palo Alto-Gruppe*, die sich aus Psychiatern (Frieda Fromm-Reichmann), Anthropologen (Gregory Bateson), Kybernetikern und Linguisten zusammensetzte. Sie zeichneten Interaktionen in Familien mit schizophrenen Kindern auf und untersuchten sie hinsichtlich der Frage: Was unterscheidet Interaktionen hier von denen in anderen Familien. Die Art der Analyse folgt einem ähnlichen Muster wie die, die Margaret Mead mit ihren kulturvergleichenden Filmen vorgenommen hat. Es wird also auf kleinste Unterschiede in der Behandlungsweise geachtet, auf die sogenannten *micropatterns*. Berühmt wurde die Gruppe auch für die Analyse eines Filmausschnittes („Doris-Film“) aus verschiedenen disziplinären Blickwinkeln.

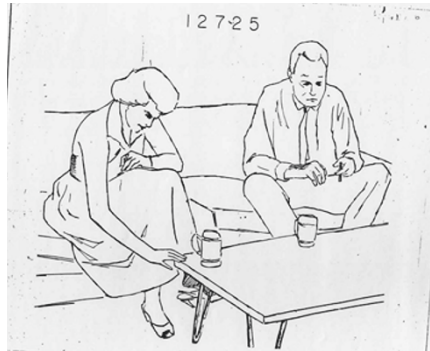


Abbildung 1: Abbildung aus den auf Mikrofilm verfügbaren Materialien zum Doris Interview (McQuown, 1971)

Obwohl diese Analysen bedauerlicherweise nie veröffentlicht wurden und nur die Berichte als seltene Mikrofilme vorliegen, zeigte sich ihr Einfluss doch in den Arbeiten von Birdwhistell (1970) sowie von Pittenger, Hockett und Danehy (1960). Man bezeichnete diesen Ansatz auch als natural history approach, weil es sich um detaillierte Beschreibungen von Beobachtungen der Abläufe bei Interaktionen handelt. Es war vor allem Bateson, der das visuelle Element einführte, denn er hatte schon zuvor mit Filmen über Interaktionen gearbeitet. Von Bateson stammen die folgenden methodologischen Empfehlungen (Birdwhistell, 1970, S. 183):

1. Wie bei allen anderen Naturereignissen muss auch das körperliche Verhalten als etwas angesehen werden, das in seinem Kontext Sinn ergibt.
2. Wie alle anderen Aspekte menschlichen Verhaltens sind auch Körperhaltung, Bewegung und Gesichtsausdruck strukturiert und deshalb Gegenstand systematischer Analyse.
3. Auch wenn man biologische Beschränkungen einräumt, sollte man bis zum Beweis des Gegenteils die menschlichen Körperbewegungen als soziale Phänomene betrachten.
4. Sichtbare körperliche Bewegungen beeinflussen ebenso wie sprachliche das Verhalten anderer Gesellschaftsmitglieder.
5. Bis zum Beweis des Gegenteils sollte davon ausgegangen werden, dass diese Bewegungen eine erforschbare kommunikative Funktion haben.
6. Die damit verbundenen Bedeutungen hängen mit dem Verhalten und mit den untersuchten ‚operations‘ zusammen.
7. Die besondere biologische Situation und die individuellen Lebenserfahrungen des Individuums führen zu kinetischen Idiosynkrasien, doch können diese nur festgestellt werden, wenn man längere Datenausschnitte zur Verfügung hat.

Auf dieser Grundlage interessierte sich Birdwhistell selbst von Beginn an für Körperbewegungen und „entwickelte eine Forschungsstrategie, um vor allem jene universellen Gefühlszeichen herauszudestillieren, die für die Gattung Mensch spezifisch sein sollten [...] Als die Forschungsarbeiten vorangingen, wurde noch vor der Entwicklung der Kinesik deutlich, dass diese Suche nach Universalien kulturgebunden sei [...] es gibt wahrscheinlich keine universellen Symbole für emotionale Befindlichkeiten“ (nach Ekman, 1982, S. 17). Er ging also von der Annahme aus, dass das Verhalten von Gesicht und Körper eine Sprache sei, die die gleichen Bestandteile und Organisationsebenen wie die gesprochene Sprache aufweise und daher am angemessensten mit Hilfe linguistischer Methoden zu untersuchen sei. Kommunikation ist für ihn – im Sinne der Strukturalisten – ein strukturelles System signifikanter Symbole aller Sinnesmodalitäten, die eine geordnete menschliche Interaktion ermöglichen (Birdwhistell, 1970, S. 95). Im Unterschied zum linguistischen Wissen können wir das kinetische indessen schlechter lehren und lernen.

Birdwhistell ist deswegen von Interesse, weil er schon in den Fünfzigern mit sehr teuren Filmaufnahmen arbeitete und hunderttausende Meter Film analysierte. Dabei ist zu beachten, dass damals Zeitlupe gerade erst möglich geworden und Rückwärtsspulen offenbar geradezu unmöglich war. Filmlabors waren noch sehr selten. Deshalb konnte er zu den Schwierigkeiten im Umgang mit diesem Datenmaterial noch bemerken: „Das schnelle Vorüberbrauschen der Mengen von Daten und die Ausrichtung und Fokussierung der Aufnahme machen es sehr schwer den Ton, Film oder das Band zu analysieren“⁹ (Birdwhistell, 1970, S. 152). Methodisch warnt er vor der vorgängigen Festlegung von Deutungen, die auf „ähnliches Training und daher auch der gleichen Reaktion“ zurückgehe. Es sei, so betont er, nicht nötig, die beobachtete äußere Umgebung zu messen (S. 153).

Kinesik konzipiert er analog zur Linguistik als ein geordnetes System isolierbarer Elemente. Birdwhistell neigte dabei zu natürlichen Kontexten – und gibt zu bedenken, man könne ja das soziale Verhalten der Fische auch nicht studieren, wenn man sie aus dem Wasser nehme. Birdwhistell reduzierte die Bewegungen in einzelne Elemente, die sogenannten *Kineme*. Kineme sind die kleinsten, noch sinnhaft beschreibbaren Elemente. Von den vielen tausend Bewegungen des Gesichtes werden 32 Kineme gebildet. So wird etwa die Haltung der Augenlider als offen, weit aufgerissen, geschlossen und eng beschrieben. Kine(m) bezeichnet diejenige Einheit, die von einer Beobachterin aus der sozialen Gruppe der sich Be-

9 Die Übersetzung dieses und aller weiteren fremdsprachlichen Originalzitate stammen, soweit nicht anders ausgewiesen, vom Autor.

wegenden als eigene, von anderen unterschiedene Bewegung beobachtbar ist bzw. von verschiedenen als ‚dieselbe Bewegung‘. Sie werden also durch Abstraktion und durch Kontrast gewonnen.

Kineme fasst er dann in *Kinemorphe* zusammen, die in kinemorphischen Klassen beschrieben werden. Kinemorph bezeichnet eine Reihe verschiedener Bewegungen im selben körperlichen Bereich (Kopf, Schulter, Oberleib, Unterleib, rechter/linker Arm, Beine). Im Bewegungsstrom werden sie als komplexe Kinemorphe erfasst und, zusammen mit anderen Bewegungsweisen, in vielschichtigen kinemorphischen Konstruktionen zusammengefasst. Dazu kommen noch kinemorphische Markierungen, wie etwa die kinetische Betonung (leichtes Nicken mit dem Kopf, Schulterzucken usw.).

Eine in der Psychologie sehr einflussreiche Version der Videoanalyse wurde von Ekman (1982) geprägt. Zusammen mit unterschiedlichen Mitarbeitern hat er seine Aufmerksamkeit auf die filmisch aufgezeichneten Bewegungen beim menschlichen und tierischen Ausdruck gerichtet. Vor allem sein System zur Erfassung mimischer Bewegungen des Gesichts und der Gesichtsmuskeln (*FACS*) wurde prägend für die Tendenz einer Standardisierung und Systematisierung von Beobachtungen anhand visueller Daten, die sich in der Psychologie, aber auch in anderen Disziplinen, durchgesetzt hat. Zugleich bilden Ekmans, noch mehr aber Birdwhistells und Halls Untersuchungen den Ausgangspunkt einer interdisziplinären Bewegung, die sich auf das Zusammenspiel verschiedener *Modalitäten* des menschlichen Ausdrucks richtet, also Gesten, Mimik, Prosodie, Körperhaltung usw..

2.2.3 *Ethnologischer Film*

Eine weitere Strömung, die vor allem für die Darstellung visuellen Materials relevant ist, bildet der ethnologische Film. In der auch teilweise *ethnographisch* genannten Filmtradition liegt das Hauptaugenmerk auf der Dokumentation von Verhaltensweisen und Ausdrucksformen fremder Kulturen. Häufig wurden diese Filme angefertigt, um diese Kulturen anderen (westlichen) Beobachtern vor Augen zu führen oder sie vor dem Vergessen zu retten. Die Forschenden gingen also insbesondere in ihnen fremde Felder und nutzten die zunächst sehr aufwändig zu transportierende und bedienende Filmtechnik, um dort das Gesehene zu konservieren und mit nach Hause zu nehmen. Die Analyse spielt in dieser Disziplin eine untergeordnete Rolle, denn es gilt immer noch das Primat der Feldnotizen. Jedoch ist zu betonen, dass schon der Filmschnitt eine Form der Interpretation darstellt.

Der ethnographische Film ist vermutlich so alt wie der Film überhaupt.

Sogar schon vor der ersten Projektion der Gebrüder Lumière¹⁰ fertigte Félix-Louis Regnault 1895 in Paris für die Ausstellung *Ethnographie de l'Afrique Occidentale* vier Filmsequenzen über eine senegalesische Wolof-Frau beim Töpfern an, die vermutlich jedoch nicht gezeigt wurden (vgl. Marks, 1995). Besonders A. C. Haddon propagierte die Verwendung des Films in der Ethnographie. Der frühe ethnographische Film zeichnete sich durch den starken Einfluss der wissenschaftlichen Ethnographie Malinowskis aus. Dazu gehörte, dass die Beforschten selbst aufgesucht wurden. Dies gilt mustergültig für Robert Flahertys *Nanook of the North* von 1922 – ein Film über die Inuit in Nordkanada. Flaherty verbrachte Monate bei den Eskimos, bevor er diesen Film drehte.¹¹ Wie von Malinowski gefordert, machte der Forscher sich mit der Sprache und der Kultur der Inuit vertraut und drehte mit ihnen zusammen die Szenen. Wie Malinowski ging der gesamte frühe ethnographische Film davon aus, dass menschliche Handlungen als Kombinationen von Bewegungen aufgezeichnet und gezeigt werden können. Überdies unterstellen die Filmausschnitte, die Wirklichkeit der Inuit wiederzugeben (und wurden dafür kritisiert, dass sie theoretisch unterbelichtet wären).

Vor allem in der Zeit nach beiden Weltkriegen festigte sich der ethnographische Film. Durch die immer bessere Aufnahmetechnik, leichtere Geräte sowie verbesserte Ton- und Bildaufnahmetechnik bei geringerem Energieverbrauch nahm die Qualität der Aufnahmen stetig zu und ihre Anfertigung wurde immer einfacher. Mit der Ausbreitung des Fernsehens gewann er schließlich enorm an Breitenwirkung, wie etwa der Erfolg der BBC-Serie ‚Disappearing World‘ zeigt. Eine deutliche Veränderung nahm der ethnographische Film schließlich mit dem Film ‚The Ax Fight‘ von Timothy Asch und Napoleon Chagnon. Der Film zeigt eine authentische kurze, gewalttätige Begebenheit bei den Yanomami im Amazonas, als in einem Dorf plötzlich ein blutiger Streit beginnt. Dieser Streit zwischen einem Mann und einer Frau um eine Kochbanane führt zu tätlichen Angriffen mit Unterstüt-

10 Die Gebrüder Lumière führten am 28. Dezember 1895 im Grand Café auf dem Boulevard des Capucines in Paris mit ihrem Kinematographen der Öffentlichkeit einen selbst gedrehten Kurzfilm vor, der Arbeiter beim Verlassen ihrer Fabrik zeigt und als erster Film überhaupt gilt. Allerdings hatten die Brüder Skladanowsky schon vorher kleine Szenen von turnelnden Menschen gedreht, die sie am 1. November 1895 im Berliner Varieté Wintergarten präsentierte.

11 Flaherty schildert das Leben Nanooks und seiner Familie und beobachtete sie mit der Kamera bei alltäglichen Verrichtungen wie Jagd, Fischfang, Iglubau, Fellhandel, Kinderpflege, Betreuung der Schlittenhunde, beim Kampf gegen die Kälte und der Nahrungsbeschaffung, aber auch auf der gefährlichen und seltenen Walrossjagd. Flahertys Film war bereits in den 20er-Jahren international erfolgreich und revolutionierte durch seine einfühlsamen Aufnahmen den Dokumentarfilm.

zung durch Verwandte. Der Film dokumentiert den Ausbruch der Feindseligkeiten, ihren Verlauf, das Auftreten konfligierender Gruppen, den Schlag mit der Axt, den Rückzug und die Beschimpfung, die den Abschluss der körperlichen Feindseligkeiten bildete. Asch filmte den Kampf aus einer gewissen Distanz, während Chagnon neben der Kamera stand. Beide unterhielten sich während der Szene über die Handlung, kommentierten das Geschehen und versuchten, ihm einen Sinn abzugewinnen. Die Kommentare von Asch und Chagnon sind Ad-hoc-Erklärungen, die eher alltäglich als wissenschaftlich ausfallen. Sie wurden im Original ohne Hinzufügung erhalten und der Film wurde nicht ediert. Damit folgten sie der konstruktivistischen Devise: „We do not discover order, we create it“.¹²

Eine über diese Entwicklung hinausgehende Fassung filmischer Reflexivität stellt der Film „Die weißen Geister“ von Bob Connolly und Robin Anderson dar, der überhaupt keine Kamera nutzt, sondern die Kameraführung der Imperialisten einspannt und sie mit den Kommentaren der Kolonialisierten unterlegt.¹³ Dies spiegelt einen Trend zur Reflexivität der Feldforschung wider, der vor allem in der Ethnologie unter dem Begriff der *Krise der Repräsentation* heftige erkenntnistheoretische und methodologische Debatten ausgelöst hat. Unter Reflexivität versteht Ruby (2000, S. 156), dass „die Hersteller ihrem Publikum ausdrücklich und absichtlich die epistemologischen Annahmen erläutern, die sie dazu führten, bestimmte Fragen zu stellen, bestimmte Antworten zu suchen und schließlich ihre Ergebnisse auf eine bestimmte Art darzustellen“.¹⁴

Werfen wir einen Blick in die Gegenwart der ethnologischen Forschung, so führt dieser Trend zur Reflexivität zuweilen sogar zu einer ausdrücklichen

12 Der Film wurde 1971 gedreht und 1975 veröffentlicht. Etwa 10 Minuten des halbstündigen Disputs wurden mit der Kamera aufgezeichnet. Der Film zeigt und analysiert den Streit in vier Abschnitten: 1. Ungeschnittene Fassung des Rohmaterials; 2. Kommentierung des Gesehenen durch die filmenden Ethnologen (auf Schwarzfilm); 3. Analyse des Streits durch die filmenden Ethnologen mit Hilfe von Diagrammen, Standbildern und Zeitlupe; 4. Geschnittene Fassung.

13 „Als die weißen Geister kamen“ zeigt historische Aufnahmen der ersten Ankunft europäischer »Entdecker« in einem bislang unberührten Teil Papua-Neuguineas Anfang des 20. Jahrhunderts. Diese werden mit deren Kommentierung durch die Nachfahren 50 Jahre später in einem filmischen Dialog über verschiedene Perspektiven montiert, welche neben der okzidentalen auch die der Einheimischen in Wort und Bild fassen will.

14 Es sei darauf hingewiesen, dass der Begriff der Reflexivität verschiedene Bedeutungen hat. Wenn ich von Reflexiver Videoanalyse spreche, meine ich damit lediglich, dass die Videoanalyse auf Videoanalyse angewendet wird, verwende ich den Begriff im theoretischen Zusammenhang mit der Ethnomethodologie der Reflexivität des Handelns, ist damit der Umstand gemeint, dass Handelnde beständig im Tun auch kommunizieren was sie tun; schlussendlich ist an dieser Stelle die Reflexivität der Forschenden in der Ethnographie gemeint, die Ihre Perspektive (selbstkritisch) hinterfragen müssen.

Ablehnung von Videodaten. Emmison und Smith (2000, S. 4) vertreten etwa die Meinung, dass man in der Soziologie überhaupt keine Kameras benutzen sollte, wenn sie fordern: „Visuelle Daten sollten nicht als das, was die Kamera aufzeichnen kann, betrachtet werden, sondern als das, was man mit dem Auge beobachten kann. Fotos sind zwar manchmal hilfreich um die sichtbare Dimension des sozialen Lebens einzufangen. Normalerweise sind sie dazu aber nicht notwendig“. Stattdessen empfehlen sie verdeckte Beobachtungen (S. 110). Andererseits hat die reflexive Methodologie auch zu *partizipativen* Konzepten geführt. So forderte etwa Holliday (2000) ihre ‚Untersuchungssubjekte‘ auf, die Videokamera selbst in die Hand zu nehmen und eigene Videotagebücher zu führen. Damit solle die Inszeniertheit des eigenen Alltags bloßgestellt und die Vorstellung einer ‚wirklichen‘ unveränderlichen Wirklichkeit überwunden werden. Ein ähnliches Konzept wurde von Michaels verfolgt, der bei seiner Untersuchung australischen Ureinwohnern selbst die Kamera in die Hand gab und sie damit zu Akteuren machte, um nicht seine ‚zentralisierte‘ Sicht aufzuoktroieren (Michaels, Kelly, & andere, 1984).

Die zweite Reaktion auf die Reflexivität ist die Rückbesinnung auf eine klassische Forderung der Ethnographie: Das Überdenken der eigenen Methoden als konstitutives und integrales Merkmal der Forschung selbst zu betrachten. Schon von klassischen Ethnographen wird erwartet, dass sie ihre eigene Feldrolle und ihre Lage gleichsam wissenssoziologisch reflektieren und die daraus folgende Perspektivität offen zu legen versuchen. Hier ergeben sich Anschlüsse an aktuelle Debatten in der Videographie, wie sie in dieser Arbeit behandelt wird, denn ähnliche Fragen werde auch ich stellen. Die soziologische Videographie baut auf ethnographischer Arbeit in der eigenen Gesellschaft auf, und daher muss auch hier beständig die Feldrolle reflektiert werden. Tatsächlich liegen schon entsprechende methodologische Arbeiten vor, die hier als Orientierung dienen können. Ein Beispiel dafür stellen etwa die Arbeiten von Elisabeth Mohn dar, die sich mit den Fragen der Selektion der Ausschnitte, der Kamerapositionen von Forschenden und schließlich auch der ‚Haltung‘ des ethnographischen Filmens beschäftigt hat (Mohn, 2002). Ebenso nehmen diese methodologischen, epistemologischen und kulturkonstruktiven Rahmungsaspekte der visuellen Forschung in den Arbeiten von Raab eine prominente Stellung ein (Raab 2008). Bei Pink (2006) werden Videos im Rahmen einer breiter angelegten und sehr stark visuelle und audiovisuelle Daten berücksichtigenden postmodernen Feldforschung eingesetzt.

2.2.4 *Video in der Soziologie*

Bis in die 1970er-Jahre stützten sich audiovisuelle Methoden weitgehend auf das Medium Film, das eine gewisse Schwerfälligkeit für die Analyse aufwies und keine nachhaltigen Spuren in der Soziologie hinterließ.¹⁵ Mit der Entwicklung der Videotechnologie verbesserten sich die Voraussetzungen jedoch allmählich. Videoaufzeichnungen ermöglichen eine detaillierte Betrachtung von Interaktionsabläufen. Allerdings waren die ersten Geräte teuer, groß und unhandlich, weshalb die Aufnahmen häufig in besonderen Räumlichkeiten vorgenommen wurden. So untersuchte eine Forschergruppe um Luckmann Ende der 1970er-Jahre die Interaktionsabläufe von Interviews, die man in einem Medienraum mit vier Kameras aufgezeichnet hatte. Luckmann beschreibt diese frühen Vorläufer der heutigen Videointeraktionsanalyse folgendermaßen:

Wir haben eine eigene Untersuchung durchgeführt. Grundlage hierfür war die Aufnahme von vier kurzen, eigentlich nur einigen Minuten dauernden Episoden eines Gespräches. Wir haben uns entschlossen, trotz eines gewissen Realitätsverlustes die Leute (...) nicht in einer normalen natürlichen Interaktion, sondern mit fixierten Sitzen mit einer Kamera in der Totalen und einer Kamera, die jeweils aus dem Blickpunkt des einen den anderen erfasste und umgekehrt zu filmen. Weiterhin haben wir noch Zuseher und Zuhörer mit in die Situation geholt. Diese wurden dann jeweils nachträglich befragt. Im Anschluss haben wir das ganze Geschehen transkribiert.

Es gab bei der Auswertung eine Reihe von Problemen (...) – vor Allem dabei, von den Detailbetrachtungen auf die nächsthöhere Ebene zu gelangen. Die phonetische Analyse, zum Beispiel, wurde maschinengestützt durchgeführt. Das war sehr einfach zu bewerkstelligen, denn wir konnten dadurch Frequenzen ermitteln. Aber damit funktionale Eindrücke des Hörens zu verbinden, ist uns nicht richtig gelungen. Ebenso wenig glückte es uns, bei der Videoanalyse eine quasi objektivierende Verbindung (im Sinn von Ekman & Friesen) zwischen dem Sehen des einen und den Reaktionen des anderen zu finden. Immerhin haben wir verhältnismäßig viele Detail-Analysen, aber eben nicht integrierend. Das sind die Kosten des Wechsels auf das nächste Niveau. (Interview mit Thomas Luckmann im August 2011)

In diesem Projekt wurde ein höchst differenziertes Transkriptionssystem entwickelt, das die verschiedensten Modalitäten der Kommunikation in Form einer Partitur wiedergibt (Luckmann & Gross, 1977). Während die Ergebnisse dieser Analysen allgemein wenig rezipiert wurden, fand die Videoanalyse seit den 1980er-Jahren bei Vertretern der anthropologischen Linguistik und der Konversationsanalyse stärkeren Anklang. Dies mag paradox erscheinen, weil sich die Konversationsanalyse durch eine äußerst streng eingehaltene Beschränkung auf auditive Daten auszeich-

15 Soziologisch relevante Filmaufnahmen wurden allerdings im Rahmen des Dokumentarfilms erstellt, die jedoch noch einmal einen eigenen Bereich darstellen, siehe Blanc (2013) und Heine (2012, 2013).

nete (Bergmann, 1981). Zugleich schien aber gerade diese Beschränkung den Blick für das Visuelle zu schärfen, das aus dieser Betrachtungsweise ausgeschlossen war. Dies gilt zum einen für Charles Goodwin, der sich zunächst auf das Zusammenspiel von sprachlichen und gestischen Handlungsmustern konzentrierte (1981, 1986); einen etwas breiteren Ansatz entwickelte Christian Heath, der umfassende soziale Interaktionssituationen mit Video analysierte, wie etwa die ärztliche Untersuchung von Patienten (Heath 1986). Auch Erickson und Schultz (1982) hatten die soziale Interaktionssituation des schulischen Bewerbungsgesprächs untersucht. So beschreibt Erickson (2011, S. 181) die mühsame Handhabung der Videokamera in einem Projekt, das dem von Luckmann sehr ähnelte, folgendermaßen:

The first videotape I made was done in the fall of 1967 in a study of small-group discussions of young people of middle school age. The camera I used weighed at least 25 pounds and was mounted on a heavy, wheeled tripod, with recording being done on reels of tape that were an inch thick and about 16 inches in diameter. This setup could only be used in a studio, so I brought a discussion group to that space, seated them in front of the camera, turned on extra lights around the room, and recorded continuously using the widest-angle camera lens. Having previously recorded such discussions using audiotape alone, one single videotape seemed marvelously illuminating; I could see who the speakers were addressing as they spoke – a particular individual, a subset of the group, or the whole group. A multimodal and multi-party analysis of the locally situated ecological processes of interaction and meaning making became possible using such an audiovisual record.

By 1970, a camera weighing about four pounds was available. I used it in my next study, which focused on interaction processes in job interviews and academic advising interviews (Erickson & Schultz, 1982). The video camera was mounted on top of a soundproof box containing a 16-mm cinema camera for simultaneous recording on cinema film. The box sat on a tripod and sound was recorded using a directional 'shotgun' microphone about a meter long, which was mounted on another tripod, facing the interlocutors. Clumsy as this rig was, it could be moved by car to field sites and be setup in a corner of the interviewers' offices. The video recording deck and a monitor were set up in an adjoining room and the whole system was operated by remote control. Both cameras were 'locked on' – no panning or zooming was done while recording – and the interviewee and the interviewer were fixed in space as they interacted. They sat facing one another across the interviewer's desk.

Mit der technischen Entwicklung wurde auch die Videoaufzeichnung und -analyse allmählich erleichtert. Eine breitere Anerkennung fand die Videotechnologie durch ihren Einsatz in den (konversationsanalytisch ausgerichteten) *Workplace Studies*, die das Zusammenspiel menschlicher Interaktion und technischer Artefakte zum Gegenstand videographischer Untersuchungen erhoben (Heath & Knoblauch, 1999). In diesem Forschungszusammenhang machten Suchman und Trigg (1991),

Jordan und Henderson (1995) sowie Heath (1997) und Lacoste (1997) die ersten Vorschläge zur Methodologie der Videoanalyse, auf die ich in Kapitel 3 näher eingehen werde.

Video wird in der Soziologie also zur Erweiterung des empirischen Materials, sozusagen als Erweiterung der Perspektive verwendet. Die technikhistorischen und medialen Besonderheiten werden im kommenden Abschnitt behandelt.

2.3 Zeitmaschine und Mikroskop – zur Rolle der Videotechnologie

Unter Video können wir mit Zielinski (2010) ein konkretes technisches Sachsystem verstehen, einen „Apparat, mit dessen Hilfe man Ton-Bilder elektromagnetisch aufzeichnen kann“ (S. 11). In seinem Rückblick zur Neuauflage des Klassikers der Medienwissenschaften von 1986 stellt der Autor fest, dass sich dieses Medium im Umbruch zwischen Massenmedien (wie dem Kino) und netzwerkartigen neuen Formen (wie dem Internet) entwickelte, da es elektronisch funktioniert und die Produktionsmittel der privaten Aneignung zugänglich macht. Diese Eigenschaften sind sicherlich relevant für die weite Verbreitung und vor allem für die Vielfalt der Nutzungsweisen, die Video – als Camcorder oder in CCTV Systemen – nach sich zog. Video ermöglicht es nahezu jedem, Ton und Bilder, also lokale Ereignisse, aufzuzeichnen und damit zu konservieren. Video, wie auch andere Medien, erlaubt es also, bestimmte Aspekte von Geschehnissen zu objektivieren, in den materiellen Träger einzuschreiben und sie so immer wieder verfügbar zu machen. Bei solchen Objektivationen sprechen wir zumeist davon, dass Informationen abgespeichert wurden, sollten aber nicht übersehen, dass dies kein neutraler Vorgang sondern eine spezifische, technische und auswählende Handlung ist.

2.3.1 Die Geschichte des Videorecorders

Zielinski untersucht den Videorekorder als *soziokulturelles* bzw. *sozio-technisches* Sachsystem, das er jedoch, unter Einbezug verschiedener medientheoretischer und auch techniksoziologischer Konzepte (er bezieht sich unter anderem auf einen frühen Text von Rammert, 1983) theoretisch fasst. Er verhandelt den Videorekorder, der ja auch den Kern des *konservierenden* Aspekts der Videokamera ausmacht, ausführlich aus einer innovationstheoretischen Perspektive und rekonstruiert verschiedene Schritte oder *Niveaus der Entwicklung des Videorekorders* (Zielinski 2010: 73):

„Entwicklungs-Niveaus des Artefaktes Videorecorder als ‚Innovation im weiteren Sinne‘

- Erfindung (im engeren Sinne 1948 – 1956)
- Technische Konzipierung von Magnetbandmaschinen für die Aufzeichnung von Fernsehimages: Innovation (1956 – 1958)
- Erste wirtschaftliche Umsetzung des Videorecorders als Gerät für die Programmdistribution bei den US-Networks: Erste Verbreitung (1958 – 1961)
- Etablierung als Gerät für die Produktion von Programmen im Fernseh-rundfunk außerhalb der US-Networks, Anfänge seiner Integration im bundesdeutschen TV-Rundfunk: Zweite Verbreitung (ab 1962)
- Diversifikation des Videorecorders für Praxisfelder von Fernsehkommuni-kation außerhalb des Rundfunks: Dritte Verbreitung (BRD: ab 1977)

Durchsetzung des Videorecorders auf dem Massenmarkt der Unterhaltungs-elektronik.“

Zielinski nimmt in seiner Rekonstruktion vor allem die empirisch sehr klare Verbindung der Technikentwicklung zum Fernsehsystem in den Blick, die für ihn auch die Grundlage der Durchsetzung des Videorekorders darstellt. Es muss an dieser Stelle noch einmal betont werden, dass die elektrotechnische Konservierung visueller Abbildungen weitreichende Folgen für die Massenkommunikationsmedien hatte, da die Empfangsgeräte vorher nur zur Verteilung, nicht zum Speichern in der Lage waren (abgesehen vom aufwendigeren und mit vielen Nachteilen behafteten Medium Film oder mittels Zwischenfilm-Verfahren). Der Darstellung Zielinskis folgend stelle ich im Folgenden einige wichtige historische Punkte dar; die Entwicklungsgeschichte ist jedoch deutlich komplizierter und breiter, wie ein Blick in eine technikhistorisch detaillierte Darstellung zeigt (Polzer, Engel, Kuper, & Bell, 2013).

Die ‚apparative Keimzelle‘ der für die Videoaufzeichnung maßgeblichen Magnetbandtechnik identifiziert Zielinski die Arbeit des dänischen Physikers *Waldemar Pulsens*, der 1898 ein Patent auf das *Telegraphon* anmeldete, das auf Klavierdraht und Stahlbändern magnetisch elektrische Impulse konservieren konnte. Das erste Gerät für eine Aufzeichnung von elektrischen Bildsignalen besteht im *Phonoscope* des schottischen Ingenieurs *John Logie Baird* im Jahr 1927, jedoch noch auf dem Medium einer Wachsplatte. Weitere Fortschritte der Kombination der Speicherung von Bildern mit dem Medium des Magnetbandes wurden in Deutschland bei AEG entwickelt, wurden jedoch vom zweiten Weltkrieg unterbrochen. Dort spielten Weiterentwicklungen in der Aufzeichnungstechnik allerdings bei den alli-

ierten Militärs (zur Aufzeichnung der Flugbahnen von Bomben usw.) eine große Rolle. Nach dem Ende des Krieges ging die Verbreitung und Verbesserung vor allem in den USA weiter, wo der Radio- und auch Fernsehmarkt boomte. *John Mullin*, ein Mitarbeiter der US -Militär Nachrichtentruppe soll einige bei *AEF* weit entwickelten *Magnetophone* nach dem Krieg aufgefundenen, in die USA verschifft, dort untersucht und weiterentwickelt haben. Dort wurde das Magnetband zunächst für die Aufnahme von Ton entwickelt (im Jahr 1948 erschien der berühmte *Ampex 200*). Nach einigen Jahren der Forschung und Ausreifung gelang es 1951 dem im Unterhaltungsfernsehen tätigen *Crosby Laboren*, die Laufgeschwindigkeit der Videobänder so zu steigern, dass diese auch die Bewegtbilder mit hinreichender Qualität aufzeichnen und wiedergeben konnten. Die kommerziellen Fernsehsender waren sehr an solchen Aufnahmetechniken interessiert, um die Kosten der aufwendigen Live-Produktionen in den Griff zu bekommen. Jedoch waren die damaligen Videogeräte noch nicht für Privatanwender nutzbar, sondern fanden sich in den folgenden Jahren als große, komplizierte Apparaturen vor allem in Fernsehstudios und Distributionszentren. Das Fernsehen wurde durch diese Innovation revolutioniert, denn nun war es möglich, Konserven zu produzieren, Sendungen ohne größeren Qualitätsverlust (wie beim *Zwischenfilm*) erneut auszustrahlen, zwischen verschiedenen Sendern einfach auszutauschen und Zeitzonen zu überbrücken.

Ich möchte mich hier aber nicht weiter der Verwendung des Videos im Fernsehen oder für Unterhaltungsfilme widmen, sondern einen zweiten Gedanke aufnehmen, den Zielinski darlegt: Neun Prozent, also ein zunächst kleiner Teil der Anlagen ging an Abnehmer außerhalb des Felds des Fernsehens. Vor allem das Militär war schon sehr früh interessiert. Mit der *Verkleinerung*, *Vereinfachung* und *günstigeren Preisen* wurde die Videotechnik aber auch zunehmend interessanter für andere Anwendergruppen (vgl. Zielinski 219 ff). Eine wichtige technische Verbesserung war die sogenannte *Schräg-Spur-Aufzeichnung* (die einer *Faltung der Zeit*, im materiellen Sinne gleichkommt), die zuerst von *Toshiba* in Japan entwickelt wurde. Solche Weiterentwicklungen lösten einen Wettkampf zwischen den Anbietern in den USA, Japan und Europa aus, der 1967 mit dem von Sony entwickelten *Portapak* den Höhepunkt erreichte.



Der Video Rover des Herstellers Sony, der aufgrund seiner Kompaktheit als *Portapak* bezeichnet wurde: der transportable, auch mit Batterie zu betreibende Videorecorder, hier in der etwas späteren Version des AV-3420CE und die mobile Kamera, die nur 1,9 kg wog. Dazu gehörten selbstverständlich Netzadapter und Batterieladegerät.

Archiv Zielinski

Abbildung 2: Werbung für den Sony Portapak (Zielinski, 2010, S. 287)

Hier eröffneten sich mit einem ‚nur‘ 11kg schweren Videorekorder mit Camera, der mit Batterien betrieben werden konnte, neue Möglichkeiten und der Zugang für breitere Akteursgruppen zur Aufzeichnung von bewegten Bildern. Dieser Prozess wurde durch die Kassettierung weitergeführt, mit denen auch die Video-Datenträger unempfindlicher und einfacher zu handhaben wurden. Hierbei sind die als *format-wars* bekannten Auseinandersetzungen zwischen den Formaten (VHS vs. Betamax vs. Video2000) beliebter Gegenstand der öffentlichen Debatte wie auch wissenschaftstheoretischen Untersuchung, etwa in der Pfadabhängig-

keits-Debatte (Liebowitz & Margolis, 1995). Das Video erfuhr im Bereich der Unterhaltung zur Vermarktung von Spielfilmen und Ähnlichem eine weite Verbreitung (vgl. Greenberg, 2010); es wurde auch in der Kunst aufgenommen und in seiner Medialität reflektiert (Adorf, 2008; Spielmann, 2005).

Inhaltlich interessant für die vorliegende Arbeit sind jedoch vor allem weitere Nutzungsvisionen, die schon früh in die Entwicklung der Geräte eingebracht wurden und zunächst zumeist als *institutionelles* oder *industrielles Fernsehen* bezeichnet wurden. Die Entwickler planten bereits früh die Verbreitung über die relativ engen Grenzen der Verwendung im Fernseh-Sendebetrieb ein, und entwickelten Visionen für die verschiedensten Nutzungskontexte: Für schulische Zwecke fanden erste Fernsehübertragungen bereits um 1932 statt, die erste medizinische Operation wurde 1947 im Fernsehen übertragen. Solche Verwendungen des Fernsehens weckten ein Interesse und so wurde 1962 in der Fachzeitschrift „Electronic Markets“ eine auf Basis einer Studie der Universität Harvard eine Übersicht von 107 Praxisbereichen veröffentlicht, in denen Video Nutzungspotenzial besäße. Zielinski diagnostiziert mit Bezug darauf (in den 80ern), dass die Bereiche an sich nicht überraschend seien. Vielmehr überrasche die systematische Planung, die in rückblickender Betrachtung in Vergessenheit geraten sei; man habe eher die Vorstellung eines archaischen Verbreitungsprozesses gehabt (S. 256). Zielinski stellt insbesondere drei Bereiche ausführlich dar, die auch in dieser Arbeit ihren Niederschlag in meinen drei untersuchten Forschungsfeldern finden: Erstens *Überwachung, Kontrolle, Beobachtung* (Abb. 3), zweitens *Forschung, Aus-/Weiterbildung und Training* (Abb. 4), sowie drittens *Informationsdistribution, Promotion und Unterhaltung* (Abb. 5).

Erster Verwendungszusammenhang des Videorecorders im
institutionellen/industriellen Bereich und ausgewählte Praxisfelder⁸⁶

ÜBERWACHUNG, KONTROLLE, BEOBACHTUNG



Abbildung 3: Dritter Verwendungszusammenhang (Zielinski 2010, S. 261)

Zweiter Verwendungszusammenhang des Videorecorders im
institutionellen/industriellen Bereich und ausgewählte Praxisfelder⁸⁷



Abbildung 4: Dritter Verwendungszusammenhang (Zielinski 2010, S. 262)

**Dritter Verwendungszusammenhang des Videorecorders im
institutionellen/industriellen Bereich und ausgewählte Praxisfelder⁸⁸**



Abbildung 5: Dritter Verwendungszusammenhang (Zielinski 2010, S. 263)

Diese Darstellungen veranschaulichen eindrücklich die damaligen Visionen für den Umgang mit Video. Zielinskis Kategorisierung hebt darüber hinaus bereits die funktionalen Aspekte hervor (S.260 ff), mittels derer Video in den verschiedensten Bereichen seine transformative – hier vor allem zu Rationalisierungszwecken gedachte – Kraft entfalten konnte.

2.3.2 Besonderheiten von Videodaten

Diese transformative Kraft der Videodaten liegt in ihren Besonderheiten, die ich im Folgenden anhand der methodologischen Diskussion innerhalb der Sozialwissenschaften vorstelle. Denn die Videotechnologie erlaubt es, *mimetisch* auditiv und visuell erfahrbares Geschehen in seiner *Sequenzialität* festzuhalten, um es in anderen Kontexten wieder abzuspielen. Video fungiert somit als zeitverschiebendes Medium, da es die raumzeitlichen Kontexte zu überbrücken hilft: Es wird daher auch als *Zeitmaschine* bezeichnet (Zielinski, 2010, S. 273). Grimshaw (1982, S. 122) betont dies mit dem Begriff der *Permanenz*, der bei Videodaten mit einer besonderen Dichte einhergeht: Videodaten erlauben eine viel feinere und genauere Betrachtung eines Ablaufes, als es mit einfacher Beobachtung möglich wäre. Durch Veränderung der Abspielgeschwindigkeit können Bewegungsfolgen z.B. sehr fein zerlegt werden, sodass auch vom Video als *Mikroskop* gesprochen wird (Büscher, 2005; Heath, 1986, S. 4). Im Gegensatz zu anderen Forschungsdaten, wie etwa Interviews, die bestimmte Ereignisse oder Praktiken *rekonstruierend konservieren*, wird Video daher als *registrierend-konservierendes* Medium beschrieben (Bergmann, 1985). Video bezeichnet heute im alltäglichen wie auch im wissenschaftlichen Sprachgebrauch eine Reihe verschiedener Datensorten (Tuma, Knoblauch, & Schnettler, 2013), welche von ‚natürlichen‘ Forschungsaufnahmen bis zu aufwendig produzierten Medienprodukten reichen. In der vorliegenden Arbeit – und das ist auch durch die Auswahl des behandelten empirischen Materials dieser Arbeit vorgegeben – geht es um Erstere: Die Aufnahme wurde von Akteuren erstellt, um Vorgänge, die ‚vor der Kamera‘ ablaufen, in einer möglichst unbeeinflussten Art und Weise abzubilden¹⁶ und dann analysierend zu betrachten. Historisch sind solche Betrachtungsweisen (denken wir an Bateson oder die frühe Konversationsanalyse) gemeinsam mit der Verfügbarkeit der neuen Technologien – wie Film und Tonbandgerät – aufgetreten. Sie eröffneten eine neue Perspektive auf das

16 Dies trifft auf die Marktforschungsexperimente nur mit Einschränkungen zu; aber auch hier zielt die Analyse auf habitualisierte, ‚natürliche‘ Praktiken und diese Vorgänge.

Sprechen, das Interagieren, das Handeln im Alltag. Durch solch eine, nun mögliche, Fokussierung auf die Handlungsvollzüge wird in vielen Bereichen spezifisches Wissen über etwas generiert, das Akteuren ansonsten lediglich als Routinewissen zur Verfügung stand: Wie genau sehen die Bewegungen beim Rauchen einer Zigarette aus, wie streifen Blicke von Anwesenden im Raum umher und treffen sich, welche feinen Zeichen können einen Lügner verraten. Mittels der Aufzeichnung wird eine neue Form der ‚Objektivität‘ hergestellt und in Analysen wird nach neuen Strukturen gesucht, die identifiziert und beschrieben werden. Dies ist Aufgabe meiner „Reflexiven Forschung“. Im Folgenden werden die Studien vorgestellt, die selbst bereits die Nutzung von Video zum Gegenstand gemacht haben, und auf denen meine Arbeit aufbaut, bzw. die sie fortsetzt.

2.4 Forschungsstand zur reflexiven Zuwendung zur Videoanalyse

Wie bereits erläutert, gibt es seit Kurzem ein neues Interesse an der Produktion von Erkenntnis in den Sozialwissenschaften. Dieses bezieht sich allgemein sowohl auf die Konstruktionsweisen und Wirkungen von Theorien und von standardisierten Forschungswerkzeugen, als auch auf die konkreten Handlungsformen der Interpretation. Die Interpretationstätigkeit findet in den Sozialwissenschaften anhand der unterschiedlichsten Datensorten statt, sie ist in den Sozialwissenschaften eingebettet in eine ganze Reihe von unterschiedlichen Forschungstätigkeiten, die von der Erhebung oder Sammlung von Daten über Recherchetätigkeiten, konzeptionelle Arbeit bis hin zum Schreiben von Artikeln und Büchern reichen. In der methodischen Literatur finden sich häufig theoretische Reflexionen über den Gegenstand, die verwendeten Technologien und die Vorgehensweisen, die für uns hilfreich für die Reflexion der *Vernacular Video Analysis* sein können. Dabei soll jedoch nicht der *scholastische Fehlschluss* begangen werden, an diese in anderen Feldern heimischen und mit den dortigen Handlungsproblemen befassten Verfahren nach Maßstäben, die für die Wissenschaft gelten, zu messen oder zu bewerten. Dieses Kapitel betrachtet die Reflexionen zur Interpretation von Daten aus der sozialwissenschaftlichen Methodenliteratur, aus der medienarchäologischen Betrachtung, den Science and Technology Studies wie auch aus unterschiedlichen Zweigen der Ethnomethodologie. Ziel der Ausführungen ist ein allgemeineres Verständnis dessen, was den Umgang mit audiovisuellen Daten auszeichnet.

2.4.1 Methodenliteratur

Die in der (interpretativen) Sozialwissenschaft üblichen Datensitzungen¹⁷ und Interpretationsrunden können als Interaktionsform¹⁸ beschrieben werden, in welcher ein spezifisches kommunikatives Problem bearbeitet wird. Ziel solcher Datensitzungen ist es sowohl, erste Interpretationen des Materials zu erarbeiten, als auch oft, andere Analysen, die von einzelnen Beteiligten vorbereitet wurden, zur Überprüfung vorzustellen. Reichertz (2013) hat sich jüngst diesem Thema zugewandt und Gruppeninterpretationsprozesse vor allem in der Hermeneutik (aber auch in anderen Traditionen) in den Blick genommen. Dabei hat er die Geschichte der verschiedenen Verfahren in den jeweiligen Teilströmungen der interpretativen Sozialforschung nachverfolgt und die Unterschiede und Gemeinsamkeiten aus der Geschichte der Hermeneutik, der Grounded Theory und der konversationsanalytischen Tradition rekonstruiert. Im Rahmen des (auch hier verwendeten) Konzeptes des kommunikativen Konstruktivismus stellt er mit methodischem wie theoretischem Interesse die Frage nach den Kommunikationsprozessen in diesen Sitzungen. Dabei legt er besonderen Wert auf die Frage, wessen Interpretationen sich durchsetzen und welche Rolle die Datensitzungen als Ort der Bewährung bedeuten. Den Interpretationsprozess charakterisiert er durch folgende Prozesse (Reichertz, 2013, S. 65)

- Abstrahieren heißt, ein Datum unter einer meist (theoretischen) Perspektive zu betrachten, und vor dieser theoretischen Perspektive aus dem konkreten Kontext zu lösen und auf andere Kontexte zu übertragen.
- Deuten selbst ist ein sukzessives Erarbeiten einer Lesart von Textstellen oder ganzen Texten (Abstraktion). Gemeinsam deuten heißt:
- Die Beteiligten äußern abwechselnd (ohne feste Ordnung durch Selbstauswahl, angeleitet durch einen Moderator mit besonderen Rechten und Pflichten) sukzessiv Lesarten;
- Sie geben Gründe an, weshalb etwas für sie eine Deutung sind, weshalb die Lesart zutreffend ist;
- zudem die gegenseitige kommunikative Validierung, ob die Gründe angemessen und hinreichend sind
- und zum Schluss eine implizite oder explizite (endgültige) Ratifizierung durch die Gruppe (oder Teile der Gruppe)

17 Datensitzungen, so nennen wir eine häufig vorzufindende soziale Situation, in welcher „eine Handvoll“ Teilnehmer gemeinsam ein Video betrachten, gemeinsam Interpretationen entwickeln, Bedeutungen diskutieren oder auch sich gegenseitig etwas aufzeigen. Diese Form findet sich nicht nur, aber vor allem, in der Wissenschaft.

18 Dieser Begriff kommt dem Konzept der Gattung, die in empirischen Studien zur sozialen Konstruktion entwickelt wurde sehr nahe (s. Kapitel 4.6). Der von Schindler & Liegl (2013) verwendete Terminus „Sprechformat“ weist hier Parallelen auf.

Eine genauere Analyse dieser Prozesse steht jedoch in seiner Schrift noch aus. Er bietet dem Leser die Möglichkeit, an einem Fallbeispiel eine Datensitzung nachzuverfolgen, ein Vorgehen, das auch einen methodendidaktischen Zug trägt. Bei dem Fallbeispiel geht es um eine hermeneutische Bildanalyse, also um ein visuelles Medium, jedoch noch nicht spezifisch um Videodaten.

Innerhalb der Soziologie¹⁹ wurde eine rekonstruierende Interpretation von Videos nur vereinzelt in den Blick genommen. Sicherlich gibt es eine ganze Bandbreite von methodischen und methodologischen Texten, die sich mit der Nutzung von Video zur soziologischen Analyse von Interaktion (Erickson, 2011; Heath, Hindmarsh, & Luff, 2010; Knoblauch, 2006; Tuma, Knoblauch, & Schnettler, 2013 usw.) oder medialer Produkte (Bohnsack, 2009; Raab, 2008; Reichertz & Englert, 2010 usw.) beschäftigen und durchaus teilweise reflexiv die Verwendung von Video als Handlungsform einordnen. Früh hat dies der oben bereits zitierte Bergmann (Bergmann, 1985) formuliert: Die registrierende Konservation ist beileibe kein passiver Akt. Nicht nur die Soziologie, auch die (visual) anthropology hat unter dem Stichwort der *Reaktanz* breit diskutiert, dass ein Video keine neutrale Aufzeichnung, sondern immer eine ‚Produktion‘ ist. Die Forscher und ihre Apparate sind durch ihre Anwesenheit Teil lokaler Interaktionen. Die Beobachter wählen einen Bildausschnitt, sie schließen andere mögliche Sichtweisen dadurch aus und produzieren also *einen* Blick auf das Geschehen. Sind solche Produktionsleistungen im Bild erkennbar, sprechen Reichertz und Englert (2010) vom Kamerahandeln.

Nicht nur die Aufzeichnung, sondern auch die Analyse und Rekonstruktion der Videodaten sind jedoch aktive, nicht ganz voraussetzungslose Prozesse. Bergmann merkt in seiner Erläuterung des Datentypus an, dass auch die audiovisuellen Daten „vom Betrachter deutend erschlossen werden [müssen]“ (Bergmann, 1985, S. 305). Auch Knoblauch und Schnettler haben bereits früh eine reflexive Anwendung der Methodendiskussion auf die Sozialwissenschaften selbst gefordert und auch später in einem methodischen Beitrag (2012) die *performativ körperliche Verortung* und die *Reenactments* der an Datensitzungen Beteiligten dargestellt (zum Forschungsergebnis im Folgenden mehr). Das Vorgehen in einer Datensitzung der Video-Interaktionsanalyse findet sich auch in unserem Methodenbuch (Tuma, Knoblauch & Schnettler 2013).

Es zeigte sich, dass Videoanalyse eben nicht nur eine technisch abstrakte Methode in der Wissenschaft darstellt, sondern genau im Zentrum

19 Zu verweisen ist noch auf die philosophische Behandlung z.B. bei Lazzarato (2002) und die ästhetische Dimension, die bei Spielmann (2005) ausführlich aufgearbeitet ist.

der Alltagspraxis der Forscher verortet ist. Der Umgang damit ist zu einem großen Teil Kommunikation in einem weiteren Sinne: Die Visualität des Wissens steht im Mittelpunkt. Dieses Forschungsprojekt entsprach der Forderung, die Schnettler (2007) für eine visuelle Wissenssoziologie sowie Burri & Dumit (2007) für die Science & Technology Studies eingebracht haben: *Visuelle Praktiken in der Herstellung, alltäglichen Verwendung und Verbreitung von ‚Bildern‘ zu untersuchen.*

2.4.2 Medienarchäologische Betrachtung

Medientheoretische und –historische Forschung hat sich mit Video vor allem im Bereich der Unterhaltung oder der Nutzung als politisches Werkzeug und auch der Aneignung durch Videoaktivisten als Produzenten auseinandergesetzt. Oft stehen auch abstrakte medientheoretische Modelle oder ästhetische Überlegungen im Mittelpunkt, bisher jedoch noch nie der konkrete Umgang mit Video zur Rekonstruktion in den *Vernacular Fields*. Die Verwendung von Video zur Produktion von Wissen nimmt Ramon Reichert (2007) aus einer wissenschaftshistorischen und medienarchäologischen Perspektive in den Blick. Er untersucht, post-strukturalistisch und stark von Foucault geprägt, Video als *Dispositiv*, in dem sich das *Kino* und die *Humanwissenschaften* miteinander verbinden und ein spezifisches Wissen über den Menschen herstellen. Das Kino wird bei ihm als Apparat beschrieben, welcher als „Transformationsmaschine zur Organisation psychischer Dispositionen und Blickstrukturen“ (a.a.O.: 11) fungiert. Der Autor analysiert, wie Forschungsvorhaben das Subjekt mittels bestimmter visueller Strategien vermessen, disziplinieren und normalisieren, indem sie Film, später auch Video, zur *Beobachtung*, *Aufzeichnung*, *Demonstration*, *Instruktion* und *Optimierung* menschlicher Tätigkeit nutzen. Es werden, so Reichert, in diesem Kino der Humanwissenschaften nicht nur Wissensinhalte hervorgebracht, sondern deren Form werde konstituiert. All dies zeigt Reichert an verschiedenen historischen Beispielen auf, wie etwa dem *Stanford Prison Experiment* und seiner für Filmaufnahmen geplanten Laborarchitektur, aber auch an Darstellungsformaten wie etwa dem *ethnologischen Film* oder *Lehr- und Propagandafilmen* in verschiedenen Epochen. Er kann aus der Analyse der Ergebnisse und Dokumente jedoch nicht die konkreten praktischen Formen des Umgangs mit dem visuellen Material rekonstruieren. Diese Perspektive steht vor allem in den Science und Technology Studies im Mittelpunkt.

2.4.3 Science and Technology Studies

Mit der Frage, wie Erkenntnis produziert bzw. sozial konstruiert wird, setzen sich bereits seit einigen Jahrzehnten die Science & Technology Studies

auseinander. Klassische Laborstudien, wie Latours und Woolgars Arbeiten über biologische Labore (Latour & Woolgar, 1979) oder Knorr-Cetinas Studien der Physik und Chemie (1981, 2002) stellen die Frage nach dem Prozess, in dem in den Naturwissenschaften Erkenntnis hergestellt wird. Hierbei kommen die Organisation, die praktische Tätigkeit und die Apparate und Techniken der Erkenntnis im Forschungsfeld in den Blick. Die sozialwissenschaftliche Forschung über die Naturwissenschaften beschäftigt sich häufig mit den Visualisierungen, die den Gegenstand der naturwissenschaftlichen Forschung, der häufig nicht sicht- oder greifbar ist, für die Forschenden visuell zugänglich machen (vgl. hierzu den Überblick bei Burri und Dumit (2007)).

In jüngster Zeit wenden sich die Science und Technology Studies wie auch die Wissens- und Wissenschaftssoziologie verstärkt den Sozialwissenschaften zu (Camic, Gross, & Lamont, 2012) und nehmen so ihre eigenen Erkenntniswerkzeuge reflexiv in den Blick. Zentral ist hierbei einerseits die Frage nach den praktischen Formen der Anwendung und andererseits die Frage nach den Wirkungen der Forschungswerkzeuge. Diese werden hier vor allem als performativ begriffen – in dem Sinne, dass sie bestimmte Wirkungen entfalten (für die Ökonomie MacKenzie, 2006).

2.4.4 *Ethnomethodologie, Konversationsanalyse und Workplace Studies*

Einige der Studien der Science & Technology Studies sind ethnomethodologisch informiert, denn sie entwickelten sich unter anderem aus den *Studies of Work and Science* (Garfinkel, 1986). Im Theoriekapitel (4) werde ich diesen theoretisch methodologischen Zugang genauer darstellen. Hier sollen zunächst einige Beispielstudien kurz dargestellt werden, die größtenteils im Kontext der *Workplace Studies* verortet werden können. Aufbauend auf den ethnomethodologischen Grundlagen nutzen die *Workplace Studies* die Verfahren der *Konversationsanalyse* und ergänzen sie um die Methoden der Videoaufzeichnung und –analyse, um die Interaktionsprozesse in den *Centers of Work and Coordination* ‚multimodal‘ zu untersuchen. Diese Forschung wird nicht nur von SoziologInnen betrieben, sondern reicht auch weit in die konversationsanalytische Linguistik und in die Untersuchung der Zusammenarbeit mit Hilfe von Computern (Computer Supported Collaborative Work, CSCW). Da die Workplace Studies in hoch technisierten Bereichen tätig sind, ist es nicht überraschend, dass sie schon sehr früh mit CCTV und Video konfrontiert waren. Diese Technologien spielten sowohl bei den Projekten der Goodwins im Auftrag der *Xerox Cooperation* eine große Rolle, als auch bei der Forschung von Heath, Luff und seiner Gruppe, deren groß angelegte Workplace Projekte sich zum Beispiel mit der Londoner U-Bahn beschäftigten; sie (Heath,

Luff, vom Lehn und Sanchez-Svensson 2002; 2008) haben auch die Rolle von CCTV Bildern, die in Kontrollräumen überwacht werden, betrachtet. Insbesondere zeigen sie auf, wie diese in die konkreten organisierten Arbeitsprozesse eingebunden sind und wie die Übertragung zur feinsten Koordination der Arbeitstätigkeiten genutzt wird. Das Betrachten der Videobildschirme ist also nicht einfach abtrennbar von der Arbeitspraxis und den zu lösenden Problemen. Zentrales Ergebnis dieser und einer Reihe weiterer Studien ist die genaue Identifikation von Sehpraktiken als Teil der Arbeitstätigkeit, die mit dem Begriff der *awareness* markiert werden.

Diese Studien setzten sich jedoch nicht explizit mit Video als Forschungsgegenstand und seiner aufzeichnenden Qualität auseinander, sondern vielmehr mit Liveübertragungen, in denen die Akteure unter Zeitdruck bestimmte Aufgaben lösen mussten. Auch aus dem Kontext der Workplace Studies kommend, haben Charles und Marjorie Harness Goodwin wegweisende Studien vorgelegt. Ihre konversationsanalytisch-ethnomethodologischen geprägten Veröffentlichungen zum *seeing as situated activity* (1996) und zur *professional vision* (1994), auf die vor allem im Kapitel zur Polizei weiter eingegangen wird (s. Kapitel 5), haben sich unter anderem mit der Betrachtung und genauer mit den diskursiven Praktiken der Analyse von Videos konkret auseinandergesetzt. Charles Goodwin illustriert die Sehpraktiken neben einigen anderen Beispielen vor allem am Rodney King Prozess: Einige Polizisten wurden in den 90ern angeklagt in Kalifornien einen afroamerikanischen Autofahrer (aus rassistischen Motiven) verprügelt und schwer verletzt zu haben. Mittels verschiedener (kommunikativer) Praktiken (*Coding Scheme, Highlighting und Graphical Representation*) gelingt es den Verteidigern vor Gericht, das Ereignis anhand von Videomaterial, das den Übergriff dokumentierte, auf eine bestimmte Art und Weise zu rekonstruieren und eine alternative Lesart einzuführen – und damit eine Wirklichkeit, die in weiterer Fortsetzung bis zu Aufständen in Los Angeles reichte, zu erzeugen. Das *Sehen* wird hier systematisch und folgenreich neu definiert:

Seeing is investigated as a socially situated, historically constituted body of practices through which the objects of knowledge which animate the discourse of a profession are constructed and shaped. (Goodwin 1994: S.606)

Auf diese diskursiven Praktiken, die ich in dieser Arbeit als *kleine performative Formen der Sichtbarmachung* fasse und auf die theoretische Grundlegung gehe ich im folgenden Abschnitt und in Kapitel 4.4 genauer ein.

Die Auseinandersetzung mit der Betrachtung und der Analyse von Video'daten' ist ein recht neues Forschungsfeld. So nimmt Hietzge (2008) im Rahmen der Erziehungswissenschaften die Nutzung von Video in der Lehrpraxis im Sportunterricht in den Blick und untersucht, wie Vi-

deoaufnahmen zur Selbstreflexion genutzt werden. Laurier, Strelow und Brown (2008) beschäftigen sich mit den Praktiken im professionellen Filmschnitt und zeigen, wie dort ausgewählt, verglichen und das Material von den Cuttern immer weiter reduziert wird, bis der Film als Produkt besteht. Mondada (2003, 2007, 2009) hat sich weiterhin mit der Nutzung der Videokamera als Aufzeichnungsinstrument und Interaktionsressource in konkreten Zusammenhängen (z.B. im Operationssaal) beschäftigt. All diese Untersuchungen widmen sich dem Video eben genau nicht als Forschungsmedium, sondern als Untersuchungsgegenstand. Eine Gruppe von am Thema interessierten Forschenden hat als Ergebnis eines Workshops zu ‚Video Practices‘ jüngst einen Sammelband herausgegeben (Broth, Laurier, & Mondada, 2014), der eine beeindruckende Sammlung ethnomethodologisch-konversationsanalytischer Studien zu Videopraktiken in verschiedensten Feldern, die ich als vernacular bezeichnen würde, enthält. Die Themen reichen unter anderem von einer Reflexion der Filmpraktiken im Forschungsfeld über die Arbeit in TV-Studios (Broth, 2014), die Verwendung von Videos zur Zahnpflegeerziehung (Lindwall, Johansson, & Rystedt, 2014), die Perspektivität in Computerspielen bis hin zum Live-Videoschnitt (Perry, Juhlin, & Engström, 2014). In allen Studien wird die Verwendung von Video ‚reflexiv‘ mit videoanalytischen Verfahren ausgewertet; dabei werden die Ethnomethoden im Umgang mit dieser audiovisuellen Technologie genau in den Blick genommen. Keine der Studien in diesem Band setzt sich jedoch dezidiert mit der Analysetätigkeit mittels Videos auseinander. Dieser haben sich in einigen Aufsätzen Dylan Tutt und v.a. Jon Hindmarsh (Hindmarsh & Tutt, 2012; Tutt, Hindmarsh, Shaikat, & Fraser, 2007; Tutt & Hindmarsh, 2011) gewidmet. Ähnlich wie die oben bereits angeführte Studie von Reichertz (2013) (Kapitel Methodenliteratur) filmten die Autoren Datensitzungen von Sozialwissenschaftlern und analysieren, wie die Beteiligten miteinander interagieren und eine gemeinsame Interpretation dessen entwickeln, was sich auf dem Video vor Ihnen abspielt. Um die Interaktionsorganisation des beobachteten Ablaufes zu durchblicken, werden die Datenfragmente in Datensitzungen analysiert, die Aufzeichnungen werden kommentiert und diskutiert. Für die Autoren bezeichnet dieses Analysieren und Rekonstruieren des Videos zunächst also einen Vorgang des Benennens der Teilschritte und ein sich gegenseitiges Anzeigen des Verständnisses:

As they [two Researchers; RT] are collaborating on the analysis of visible phenomena – namely, the embodied conduct of the father and daughter on-screen – much of their work is about debating and discussing its organisation. This demands that they make it visible to one another, but of course they cannot ‘see’ through the other’s eyes. So the commentary is key in this regard. They announce the order of events along with the video playback and thereby discriminate between analytically relevant and non-relevant conduct. This, in

turn, allows colleagues to assess the validity of the candidate order as it corresponds to the video playback. Establishing a 'shared seeing' entails not simply noticing (and having others notice) phenomena, but collaboratively piecing together and agreeing on (?) the order of analytically relevant conduct. (Hindmarsh & Tutt, 2012, S. 64)

Dabei handelt es sich aber nicht nur um ein sprachliches oder kognitives Zuordnen, sondern vielmehr spielt, wie Tutt und Hindmarsh betonen auch das körperliche Zeigen eine zentrale Rolle. In weiteren Studien nehmen die Autoren (2007) räumlich verteilte Analyseteams in den Blick. Diese wiederum hatten zuvor ein Video an verschiedenen Orten synchron betrachtet und bei den Analyseversuchen das Problem, zu verstehen, auf welche Elemente die andere Seite sich gerade bezieht. Aufgrund dieser Schwierigkeiten kamen besonders die lokalen *körperlichen Praktiken des Zeigens* und des *Nachspielens* in den Blick (wie bereits erwähnt, fasse ich diese in dieser Arbeit als *kleine performative Form der Sichtbarmachung*). Tutt und Hindmarsh zeigen, wie Sprache und Gestik von den Beteiligten miteinander verknüpft werden und somit die Interpretationsarbeit konstituieren: die Person, die ein Element auf dem Bildschirm hervorhebt, erschafft einen gestischen Raum, zieht die Aufmerksamkeit der anderen Beteiligten darauf (was wir an den Blicken rekonstruieren können) und verbindet die auf dem Bildschirm ablaufende Bewegung mit einer kodifizierten Fassung auf einem Transkript. Tutt und Hindmarsh betonen die Rolle dieses *gesture space*, denn die Akteure können in späteren Momenten der Interaktion wieder auf diese zurückverweisen. Ausgehend von einem Modell des *Re-Enactment*, das sie von Sidnell (2006) übernehmen, zeigen die Autoren weiterhin, wie nachgespielte Gesten von den beobachteten Forschern genutzt werden, um zu verstehen und sich anzuzeigen, was auf dem Videoschirm zu sehen ist. Die *Re-Enactments* sind zwar nicht unbedingt die einzige Methode, dieses Kommunikationsproblem zu lösen. Die an der Situation Beteiligten können durchaus andere Ressourcen heranziehen um ihr *visuelles Wissen* zu kommunizieren, Aber *Re-Enactments* zeigen sich auch in Beschreibungen und Hinweisen auf Forschungspraxis bei Herbrink (2011, S. 83 ff), die es mit Bezug auf Merleau Pontys Phänomenologie der Wahrnehmung (Merleau-Ponty, 1966) reflektiert.

Hier schließt sich der Kreis zu der Thematisierung dieses Aspekts in dem oben bereits zitierten methodologischen Arbeit von Knoblauch und Schnettler (2012). Dort zeigen die Autoren, wie eine Forschergruppe die räumliche Anordnung als körperliche Übersetzung in den anderen Kontext nachvollziehen und mittels diesem *Re-enactment* ein „*Hineinversetzen*“ in ihr Forschungsdatum erreichen. Eine weitere Studie mit ähnlichen Ergebnissen liegt von Meyer und Meier zu Verl. vor (2013), die sich

mit hermeneutischen Datensitzungen genauer auseinandersetzen hier diese Praxis als *Reinszenierung* bezeichnen. Arbeit mit Videos und Visualisierungen, ob in der Forschung oder in anderen Arbeitskontexten, ist immer eine rekonstruktive, interpretative und kommunikative Tätigkeit. Die Darstellungen müssen in aller Regel aufbereitet oder kommentiert werden, um ein gemeinsames Verständnis dessen, was zu sehen ist, zu erreichen.

Die verschiedenen Tätigkeiten im Umgang mit den Videodaten in ihrer eigenen Forschung und die Einbettung in den breiteren Arbeitsprozess des ethnographischen Forschens nimmt Schindler (2012) in den Blick, und liefert somit das Gegenstück zu den isolierten Betrachtungen der Datensitzungen als Situationen. Mit dem videographischen Ansatz, den ich im folgenden Kapitel genauer ausführe, verfolge auch ich ein Vorgehen, das beide Perspektiven kombiniert.

Videoprofis im Alltag

Die kommunikative Vielfalt der Videoanalyse

Tuma, R.

2017, XV, 308 S. 39 Abb., Softcover

ISBN: 978-3-658-15165-2