

Ralf Bohnsack

# Qualitative Bild- und Videointerpretation

Die dokumentarische Methode

2., durchgesehene und aktualisierte Auflage

Verlag Barbara Budrich  
Opladen & Farmington Hills 2011

Online-Angebote oder elektronische Ausgaben sind erhältlich unter [www.utb-shop.de](http://www.utb-shop.de)

Bibliografische Informationen der Deutschen Nationalbibliothek  
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

GE  
2012  
1195

K

Universitäts-  
Bibliothek  
Freiburg i. Br.

Gedruckt auf säurefreiem und alterungsbeständigem Papier.

Alle Rechte vorbehalten.

© 2011 Verlag Barbara Budrich, Opladen & Farmington Hills  
[www.budrich-verlag.de](http://www.budrich-verlag.de)

UB Freiburg



291431

**UTB-Band-Nr.: 8407**  
**ISBN 978-3-8252-8482-4**

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Satz: Beate Glaubitz, Redaktion + Satz, Leverkusen  
Umschlaggestaltung: Atelier Reichert, Stuttgart  
Druck: Friedrich Pustet KG, Regensburg  
Printed in Germany

Für Sabine, Karla und Johanna

# Inhalt

<b>1. Einleitung</b> .....	11
<b>2. Die dokumentarische Methode: allgemeine Grundlagen</b> .....	15
2.1 Der methodische Zugang zum handlungsleitenden Wissen .....	15
2.2 Interpretatives Paradigma und praxeologische Wissenssoziologie .....	16
2.3 Konjunktion und Kommunikation: die Doppelstruktur alltäglicher Verständigung .....	17
2.4 Methodik und empirische Verfahrensweise .....	19
<b>3. Dokumentarische Bildinterpretation</b> .....	25
3.1 Die Marginalisierung des Bildes in der empirischen Sozialforschung ...	25
3.2 Implizites Wissen, Ikonologie und Habitus .....	28
3.3 Die Eigensinnigkeit des Bildes und die Suspendierung des textlichen Vorwissens .....	32
3.4 Die Differenzierung des ikonografischen Vor-Wissens: kommunikatives und konjunktives Wissen .....	34
3.5 Die essentielle Ambiguität des Bildes: die Sinnkomplexität des Übergegensätzlichen .....	36
3.6 Zur Rekonstruktion der Formalstruktur des Bildes .....	38
3.7 Sequenzanalyse versus Kompositionsvariation .....	42
3.8 Zum Problem der Polysemie: die Vieldeutigkeit des Bildes .....	45
3.9 Simultaneität, Synchronizität und Sequenzialität .....	47
<b>4. Zur Forschungspraxis dokumentarischer Bildinterpretation</b> .....	55
4.1 Die Arbeitsschritte dokumentarischer Bildinterpretation .....	56
4.2 Exemplarische Interpretation eines Werbefotos .....	58
4.2.1 Formulierende Interpretation .....	60
4.2.2 Reflektierende Interpretation .....	61
4.2.3 Bild-Text und Bild-Logo .....	68
4.2.4 Zusammenfassung .....	68

4.2.5	Komparative Analyse: der propagierte Lifestile im Werbefoto eines anderen nationalen Marktes .....	70
4.3	Exemplarische Interpretation von Familienfotos und Methodentriangulation .....	73
4.3.1	Zur Auswahl der Fotos .....	75
4.3.2	Familie Schiller: „Museum“ .....	78
4.3.3	Familie Telchow: „Gartenfest“ .....	85
4.3.4	Kommunion von Frau Schiller .....	94
4.3.5	Jugendweihe in der Familie Telchow .....	101
4.3.6	Kommunion der Kinder der Familie Schiller .....	106
4.3.7	Schluss: Triangulation, Validierung und Vertiefung .....	115
<b>5.</b>	<b>Dokumentarische Video- und Filminterpretation .....</b>	<b>117</b>
5.1	Video und Film als Erhebungsinstrument oder Alltagsdokument .....	117
5.2	Zum Verhältnis von Produkt- und Rezeptionsanalyse .....	120
5.2.1	Exkurs: Produkt-, Rezeptions- und Diskursanalyse in den Cultural Studies .....	120
5.2.2	Die Polysemie in der Produkt- und Rezeptionsanalyse .....	124
5.2.3	Zur Differenzierung des Rezeptionsbegriffs: Interpretation, Verstehen und Aneignung .....	129
5.3	Ansätze der Film- und Videointerpretation in der aktuellen qualitativen Forschung .....	133
5.3.1	Alltagsinterpretation und wissenschaftliche Interpretation .....	133
5.3.2	Grenzen des interpretativen Paradigmas .....	135
5.3.3	Videoanalyse in ergänzender Funktion zur Gesprächsanalyse .....	137
5.4	Die Bewegungen der abgebildeten Bildproduzent(inn)en .....	140
5.4.1	Der Zugang des Films zu elementaren Ausdrucksformen und Ebenen sozialer Realität .....	141
5.4.2	Gebärden, Operationen und institutionalisierte Handlungen .....	144
5.4.3	Motivkonstruktion versus Rekonstruktion des Habitus .....	148
5.4.4	Die AnalyseEinstellung auf das Performative .....	149
5.4.5	Die unterschiedlichen Ebenen dokumentarischer Interpretation und die Primordialität der vor-ikonografischen Ebene .....	150
5.5	Fotogramm und Simultaneität .....	151
5.5.1	Die Bedeutung des Fotogramms für die Analyse von Gebärden oder Kinemorphemen .....	151
5.5.2	Grenzen der Interpretation von Fotogrammen .....	154
5.5.3	Zur Rekonstruktion von Einstellung und Perspektivität im Fotogramm .....	156
5.6	Montage, Einstellung und Sequenzialität .....	158
5.6.1	Einstellungen, Szenen und Sequenzen .....	159
5.6.2	Montage und Einstellung als Produkt der abbildenden Bildproduzent(inn)en ....	162
5.6.3	Montage und Räumlichkeit .....	163
5.6.4	Relationierungen als Leistung dokumentarischer Interpretation .....	164

Exkurs: die Relationierung der Relationen am Beispiel .....	165	
5.6.5	Sequenzielle und simultane Relationen und das Kontext-Wissen .....	168
5.7	Zur Transkription .....	170
5.7.1	Interpretation, Transkription und Protokoll .....	170
5.7.2	Das Transkriptionssystem MoViQ .....	171
5.8	Arbeitsschritte dokumentarischer Film- und Videointerpretation .....	172
5.8.1	Zum Verhältnis der Interpretationen auf der Bild- und Textebene .....	173
5.8.2	Zur Auswahl der für die Interpretation relevanten Sequenzen und Fotogramme .....	174
5.8.3	Überblick über die Arbeitsschritte im Ablauf .....	176
<b>6.</b>	<b>Zur Forschungspraxis dokumentarischer Videointerpretation: Exemplarische Analyse einer Fernsehshow: „Istanbul Total“ .....</b>	<b>177</b>
6.1	Zur Auswahl des Genres und des Videos .....	177
6.2	Videotranskript .....	179
6.3	Interpretation in der Bilddimension .....	195
6.3.1	Zur Auswahl der Sequenzen .....	196
6.3.2	Formulierende Interpretation von Sequenzen und von Einstellungswechsel und Montage .....	198
6.3.3	Zur Auswahl der Fotogramme .....	201
6.3.4	Formulierende und Reflektierende Interpretation der Fotogramme .....	202
6.3.5	Reflektierende Interpretation von Einstellungswechsel und Montage .....	213
6.3.6	Fokussierungen aufgrund von Steigerungen, Verdichtungen und Diskontinuitäten .....	215
6.4	Interpretation in der Dimension von Text und Ton .....	222
6.4.1	Texttranskript .....	222
6.4.2	Formulierende Interpretation des Textes: Thematische Gliederung .....	225
6.4.3	Reflektierende Interpretation in der Dimension von Text und Ton .....	226
6.5	Reflektierende Gesamtinterpretation .....	233
6.6	Komparative Analyse und Gattungsanalyse .....	239
<b>7.</b>	<b>Anhang .....</b>	<b>241</b>
7.1	Transkriptionssystem MoViQ: Movies and Videos in Qualitative Research .....	241
7.2	Transkriptionssystem TiQ: Talk in Qualitative Social Research .....	241
7.3	Zur Rekonstruktion der Perspektivität .....	243
7.4	Einstellungsgrößen: Übersicht über Begriffe .....	247
Abbildungsverzeichnis .....	249	
Literaturverzeichnis .....	253	

## 1. Einleitung

Die entscheidenden Fortschritte im Bereich der Methodologie und Forschungspraxis qualitativer Methoden im 20. Jahrhundert sind – abgesehen von den Arbeiten der Chicagoer Schule in den 20er Jahren – ganz wesentlich in den 60er bis 70er Jahren entfaltet worden. In dieser Zeit richtete sich die Aufmerksamkeit sozialwissenschaftlicher Empirie zum ersten Mal in systematischer Weise auf die Strukturen und die Eigensinnigkeit der „ganz normalen kunstvollen Praktiken“ des „alltäglichen Lebens“, der „artful practices“ of „everyday life“, wie Harold Garfinkel (1967: vii) dies genannt hat. Dieser Erkenntnisfortschritt ist wesentlich in der Einsicht der Notwendigkeit einer *Rekonstruktion* der Alltags-Konstruktionen bzw. der Alltags-Praktiken in ihrer Eigenlogik begründet. Der Fortschritt war also eng mit dem *rekonstruktiven* Charakter sozialwissenschaftlicher Erkenntnis verbunden.

Die Alltagspraktiken in ihrer Geordnetheit und in ihrem kunstvollen Charakter ernst zu nehmen, bedeutet, sie einer detaillierten, einer mikroskopischen Betrachtung für wert zu erachten, um sie in ihrer Eigengesetzlichkeit erfassen zu können. Es ging und geht also darum, diese ‚profanen‘ Produkte mit einem Respekt zu behandeln, wie dies bis dahin nur den Werken der Literatur und Kunst zuteil wurde. Dieses Anliegen hat die in der praxeologischen Wissenssoziologie fundierte dokumentarische Methode sowohl mit der in der Phänomenologischen Soziologie verankerten Tradition qualitativer Forschung gemeinsam, als auch mit den Cultural Studies – auch wenn die dokumentarische Methode sich in der Realisierung dieses Anliegens von diesen Traditionen wiederum deutlich unterscheidet (siehe auch Kapitel 5).

Diese Entwicklung, also die Hinwendung zur Re-Konstruktion alltäglicher profaner Praktiken und Produkte, fiel dann in ihren Anfängen zeitgeschichtlich zunächst mit jener Entwicklung in Philosophie, Erkenntnistheorie sowie sozialwissenschaftlicher Handlungstheorie und Empirie zusammen, die als „linguistic turn“ bezeichnet wurde. Dies führte dazu, dass die ‚rekonstruktive Wende‘ in der empirischen Sozialforschung zuallererst im Bereich sprachlicher Verständigung und hier – zunächst in Form der ethnomethodologischen Konversationsanalyse von Harvey Sacks – in der Interpretation von Texten entfaltet worden ist. Von wesentlicher Bedeutung für die Gültigkeit dieser Methoden sind dabei einerseits die empirische Rekonstruktion der *Formalstruktur* umgangssprachlicher Kommunikation und Darstellung und andererseits die systematische Kontrolle des in die Interpretation einbezogenen Kontextwissens. Auf diesem Wege wurde den alltagssprachlich produzierten Texten der Status eigengesetzlicher, selbst-referentieller Systeme zuerkannt.

Im Bereich der qualitativen oder rekonstruktiven Textinterpretation erscheint es – zumindest unter denjenigen, die mit qualitativen Methoden genauer vertraut sind – inzwischen als selbstverständlich, die ‚profanen‘ Texte mit ähnlicher Sorgfalt und Systematik zu behandeln, wie sie die Literaturwissenschaft traditionellerweise künstlerisch anerkannten Produkten angedeihen lässt. Im Bereich der Interpretationen stehender und bewegter Bilder befinden wir uns demgegenüber in dieser Hinsicht noch ganz am Anfang. Nunmehr geht es darum, die bildhaften, die ikonischen Produkte alltäglicher Verständigung in ihrer Eigenlogik, ihrer Formalstruktur und Alltagsästhetik in analoger Weise wie die Texte einer genauen Rekonstruktion zu unterziehen. Erst wenn dies gelingt, haben wir es – zumindest in Ansätzen – auch in der qualitativen resp. rekonstruktiven Sozialforschung mit dem viel beschworenen ‚iconic turn‘ zu tun.

Diese konsequente Anerkennung der Eigengesetzlichkeit ikonischer Produkte erscheint allerdings nicht nur denjenigen keineswegs selbstverständlich, die sich der Bild- und Filminterpretation methodisch nähern, ohne mit den methodologischen Grundlagen im Bereich der qualitativen Textinterpretation vertraut zu sein. Auch manche Experten der qualitativen Sozialforschung teilen hier jene Vorbehalte, wie sie in analoger Weise vor allem in den 70er und 80er Jahren gegenüber der rekonstruktiven Textinterpretation und der mikroskopischen Analyse der formalen Strukturen alltäglicher Verständigung geltend gemacht worden sind: Elaborierte Verfahren, wie sie u.a. mit der dokumentarischen Methode und insbesondere in Anlehnung an die Kunstgeschichte entwickelt worden sind, erscheinen nicht so ohne Weiteres geeignet für die Analyse nicht-professioneller Produkte und spontaner Schnappschüsse (allzumal für diejenigen von Kindern und Jugendlichen). Ebenso wie im Bereich der Textinterpretation werden derartige Vorbehalte letztlich nur durch die empirische Analyse selbst und deren Ergebnisse überwunden werden können.

Einige Grundlagen für einen validen Zugang zur Eigenlogik und formalen Ästhetik alltäglicher, ‚profaner‘ Bildprodukte kann uns die Kunstgeschichte und in gewisser Weise auch die Filmwissenschaft resp. Filmtheorie und Filmästhetik liefern, wie dies in Kapitel 3 und auch Kapitel 5 zu zeigen sein wird. Dabei geht es wesentlich darum, theoretische Konzepte, Begrifflichkeiten und forschungspraktische Verfahrensweisen aus diesen disziplinspezifischen Traditionen auf ihre Relevanz für die sozialwissenschaftliche Analyse und ihre Kompatibilität mit sozialwissenschaftlichen Begriffen hin zu befragen.

Wenn wir die (in Kapitel 5 näher erläuterte) Unterscheidung treffen zwischen solchen Bild- und Film- resp. Videoprodukten, welche von den Erforschten insgesamt selbst und solchen, die von den Forschenden zu Erhebungszwecken produziert worden sind (wie dies etwa im Bereich der Unterrichtsforschung der Fall ist), so liegt in diesem Buch der Schwerpunkt der forschungspraktischen Demonstration in ersterem Bereich. Dies hat den Sinn, dass die zunächst in diesem Bereich erarbeiteten methodischen Anforderungen und Interpretationsschritte eher auf den anderen Bereich, also denjenigen der von den Forschenden mitproduzierten Bilder, Videos und Filme, übertragen werden können, als dies umgekehrt der Fall wäre.

Der Anspruch dieses Buches geht dahin, eine in sich geschlossene Methodologie und Methodik der Interpretation des stehenden und bewegten Bildes von sozi-

alwissenschaftlicher Relevanz vorzulegen, welche der Eigenlogik des Ikonischen und dessen unterschiedlichen Gattungen im Bereich des stehenden wie des bewegten Bildes gerecht zu werden vermag. Ein Begriff, welcher die bewegten Bilder insgesamt als ‚Gattung‘ umfassen würde, steht uns zumindest im Deutschen nicht zur Verfügung. Während der Begriff des ‚Films‘ auf ganz spezifische Gattungen bewegter Bilder zugeschnitten ist, erscheint derjenige des ‚Videos‘ hier offener und allgemeiner, da er weniger besondere Gattungen bewegter Bilder, als vielmehr deren für den alltäglichen Gebrauch und die Analyse nutzbare Speicherung bezeichnet. Aus diesem Grunde lautet der (Ober-) Titel des Bandes ‚Qualitative Bild- und Videointerpretation‘, obschon in Kapitel 5 auch die Filminterpretation im engeren Sinne und der hier relevante Diskurs mit einbezogen und in ihrer Relevanz für die sozialwissenschaftliche Analyse diskutiert werden.

Der vorliegende Band ist derart aufgebaut, dass zunächst jeweils die methodologisch-theoretische Vergewisserung und Selbstverortung der Bildinterpretation (Kapitel 3) sowie dann der Video- und Filminterpretation (Kapitel 5) ausgearbeitet wird, um schließlich jeweils im Anschluss daran, die forschungspraktische Durchführung (Kapitel 4 u. 6) exemplarisch zu entfalten. In Kapitel 4.2 wird zunächst die Forschungspraxis einer Interpretation von Fotos aus dem öffentlich-medialen Bereich am Beispiel von Werbefotos vorgestellt und dann in Kapitel 4.3 die exemplarische Interpretation von Fotos aus dem privaten Bereich am Beispiel von Familienfotos. In Kapitel 6 werde ich die dokumentarische Film- resp. Videointerpretation am exemplarischen Fall der Folge einer Fernseh-Show (TV-Total) demonstrieren, die durch die Häufigkeit ihrer Ausstrahlung und ihre Popularität den Alltag der deutschen Fernsehlandschaft wesentlich mitprägt.

Mit der exemplarischen Interpretation von Familienfotos in Kapitel 4.3. wird zugleich eine Kombination oder Integration (Methodentriangulation) der Fotointerpretation mit der Analyse von Gesprächen (Tischgesprächen und Gruppendiskussionen) vorgeführt.

Um eine Triangulation oder Integration von Fotointerpretationen mit Gesprächsanalysen leisten zu können, bedarf es einer umfassenden Methodologie und Grundlagentheorie (Metatheorie), welche den Bild- wie den Textinterpretationen zugrunde gelegt werden kann, gleichwohl aber deren entscheidenden Differenzen Rechnung zu tragen vermag. Die dokumentarische Methode ist schon von ihrer Geschichte her (vgl. Kapitel 2 u. 3) hierfür besonders geeignet. Die besonderen Anforderungen einer Integration von Bild- und Textinterpretation bei gleichzeitiger Berücksichtigung ihrer (kategorialen) Unterschiede stellen sich im Bereich der Film- und Videoanalyse selbstverständlich bereits innerhalb des jeweiligen Produkts selbst. Text- und Bilddimension müssen hier immer schon aufeinander bezogen werden. Im Bereich der Gesprächsanalyse und Textinterpretation verfügen wir inzwischen bereits über eine mehr als zwanzigjährige Erfahrung. Daran gemessen stehen wir im Bereich der Bild- und insbesondere der Film- und Videointerpretation eher am Anfang.

Die in diesem Band gewählte Abfolge, nach der zunächst die methodologisch-theoretische Vergewisserung und dann die forschungspraktische Durchführung entfaltet wird, ließe sich – im Sinne der Logik qualitativer, genauer: rekonstruktiver Sozialforschung – auch umkehren. Denn im Bereich der rekonstruktiven Sozialfor-

schung haben wir es mit einer für diese charakteristischen zirkelhaften (oder reflexiven) Bewegung zwischen Methodologie und Forschungspraxis, zwischen Theorie und Empirie, zu tun. Eine derartige Dynamik des Forschungsprozesses ist Voraussetzung dafür, die Beschränkungen sowohl einer empirielosen Theorie wie auch einer theorielosen Empirie zu überwinden.

## 2. Die dokumentarische Methode: allgemeine Grundlagen

Mit der von ihm begründeten „Wissenssoziologie“ hat Karl Mannheim (1952b) in den 20er Jahren des 20. Jahrhunderts sowohl einen theoretischen wie auch methodologisch-methodischen Zugang zu jenen Bereichen des Wissens eröffnet, welche unsere alltägliche *Handlungspraxis orientieren*. Mannheim (1964a) bezeichnet dieses in die Alltagspraxis eingelassene vorreflexive Wissen auch als *atheoretisches Wissen* – im Unterschied eben zum *theoretischen Wissen*, also zu den Alltags-Theorien oder Common Sense-Theorien. Das atheoretische Wissen lässt sich in mancher Hinsicht mit einem Begriff von Michael Polanyi (1985) auch als „tacit knowledge“, als stillschweigendes oder *implizites Wissen* bezeichnen.

Auch das *inkorporierte Wissen* im Sinne von Bourdieu (u.a. 1976) gehört zum Bereich des atheoretischen Wissens. Die „dokumentarische Methode der Interpretation“ (Mannheim 1964a) ermöglicht den methodisch kontrollierten Zugang zu diesen handlungsorientierenden Wissensbeständen. Sie ist dabei der Ikonologie von Panofsky (u.a. 1975) verwandt, die für Bourdieu (1970) die Grundlage eines methodologisch fundierten Zugangs zum Habitus darstellt. Panofsky seinerseits (1932: 115) hat sich in der Begründung seiner Ikonologie explizit auf Karl Mannheim und dessen dokumentarische Methode bezogen.

### 2.1 Der methodische Zugang zum handlungsleitenden Wissen

Mannheim (1980: 73) erläutert den Charakter dieses Wissens am Beispiel (der Herstellung) eines Knotens. Das handlungsleitende Wissen, welches mir ermöglicht, einen Knoten zu knüpfen, ist ein atheoretisches Wissen. Diese Handlungspraxis vollzieht sich intuitiv und vorreflexiv. Das, was ein Knoten ist, *versteh* ich, indem ich mir jenen Bewegungsablauf (von Fingerfertigkeiten) einschließlich der motorischen Empfindungen vergegenwärtige, „als dessen ‚Resultat‘ der Knoten vor uns liegt“ (Mannheim 1980: 73).

Es erscheint ausgesprochen kompliziert, wenn nicht sogar unmöglich, diesen Herstellungsprozess in adäquater Weise *begrifflich-theoretisch zu explizieren*. Wesentlich unkomplizierter ist es, den Knoten auf dem Wege der *Abbildung*, also der bildlichen Demonstration des Herstellungsprozesses zu vermitteln. Das Bild erscheint somit in besonderer Weise geeignet für eine Verständigung im Medium des atheoretischen Wissens. Die im Medium des Textes zu leistende begrifflich-theo-

retische Explikation dieses intuitiven Herstellungsprozesses, dieses impliziten Wissens, nennt Mannheim „Interpretieren“ (1980: 272).

Zugleich bietet uns das Beispiel des Knotens aber auch – über Mannheim hinausgehend – die Möglichkeit der Differenzierung zwischen dem *impliziten* und dem *inkorporierten* Wissen: Solange und soweit ich mir im Prozess des Knüpfens eines Knotens dessen Herstellungsprozess, also die Bewegungsabläufe des Knotens, bildhaft – also in Form von mentalen (inneren) Bildern – vergegenwärtigen muss, um in der Praxis erfolgreich zu sein, habe ich den Prozess des Knüpfens eines Knotens noch nicht vollständig *inkorporiert*. Der Habitus ist dann das Produkt eines *modus operandi*, welcher auf *impliziten* Wissensbeständen basiert. Er kann aber auch das Produkt inkorporierter – gleichsam automatisierter – Praktiken sein. Während der empirisch-methodische Zugang zum Habitus in ersterer Hinsicht über die empirische Rekonstruktion von Erzählungen und Beschreibungen der Handlungspraktiken durch die Akteure, also über die Interpretation von Texten, führt, ist der Habitus in letzterer Hinsicht vor allem als Bild (körperlicher Bewegungsabläufe) in methodisch kontrollierter Weise zugänglich.<sup>1</sup> In diesem Sinne ist die „Habitus-theorie (auch) als eine Wissenssoziologie des Körpers“ zu verstehen (Meuser 2007: 222).

Bei der Unterscheidung zwischen dem theoretischen und dem atheoretischen Wissen geht es grundlegend um diejenige zwischen einer *theoretischen* und einer *praktischen* Beziehung zur Welt, um das Verhältnis der „theoretischen Logik“ zur „praktischen Logik“, wie man mit Bourdieu (1976: 228) sagen könnte (dazu auch Bohnsack 2006c). Erst die genaue Kenntnis dieser praktischen Logik – der Logik des Handelns jenseits der *Theorien* und der begrifflichen Konstruktionen und Definitionen, welche die Akteure in Wissenschaft und Alltag *über* ihre eigene Praxis haben – schafft die Bedingungen der Möglichkeit für eine umfassende Erkenntnis des alltäglichen Handelns.<sup>2</sup> Da die Wissenssoziologie im Sinne von Mannheim uns den Zugang zur Praxis eröffnet, habe ich sie auch als praxeologische Wissenssoziologie bezeichnet.

## 2.2 Interpretatives Paradigma und praxeologische Wissenssoziologie

Auf beiden Ebenen, derjenigen eines theoretisierenden und definitorischen oder allgemeiner: *interpretativen* Verhältnisses zur Welt ebenso wie auf der Ebene des *handlungspraktischen* Verhältnisses zur Welt, haben wir es mit *Herstellungs-* und *Konstruktionsprozessen* zu tun. Das sog. *interpretative Paradigma*, also die Phäno-

1 Die (teilnehmende) Beobachtung ist – sofern sie ohne Foto oder Videografie arbeitet – hier nur die Methode zweiter Wahl, da sie das Bild erst in einen Text transformieren muss, ehe es der eigentlichen methodisch kontrollierten Interpretation zugänglich wird.

2 Wie Alexander Reckwitz (2003: 292) formuliert, gilt es zu verstehen, „dass das Wissen nicht als ein ‚theoretisches Denken‘ der Praxis zeitlich vorausgeht, sondern als Bestandteil der Praxis zu begreifen ist.“

menologie, der Symbolische Interaktionismus, die Ethnomethodologie<sup>3</sup> und der in diesen Traditionen stehende Konstruktivismus, erfasst allerdings lediglich die *interpretative* und *definitorische* Konstruktion oder Herstellung von Wirklichkeit (dazu auch: Bohnsack 2006b). Damit ist dieses Paradigma wesentlich eingegrenzt auf die Ebene des „*theoretischen Welt-Erkennens*“, wie man dies mit einem Begriff von Heidegger (1986: 67) nennen könnte.

Demgegenüber handelt es sich bei der praxeologischen Wissenssoziologie und auch der Praxeologie im Sinne von Bourdieu um einen Konstruktivismus im *erweiterten* Sinne. Dieser erweiterte Konstruktivismus erfasst nicht nur die interpretative, sondern auch die *handlungspraktische* Herstellung und Konstruktion von Welt (vgl. u.a. Bohnsack 2001, 2006a)<sup>4</sup> und vermag diese beiden Ebenen in ihrem Spannungsverhältnis zueinander zu analysieren.

Die Fixierung auf das theoretische Welt-Erkennen, auf die theoretische Vernunft der Akteure, wird ganz wesentlich dadurch befördert, dass der methodische Zugang zu dieser Ebene des Handelns unkomplizierter ist und somit die empirische Forschung vor größeren Anstrengungen bewahrt. Denn auf dieser Ebene kann die sozialwissenschaftliche Empirie sich darauf beschränken, die *Common Sense-Theorien* der Akteure zu rekonstruieren. Sie braucht lediglich das von diesen selbst bereits zur begrifflichen Explikation gebrachte Wissen zu systematisieren und neu zu formulieren. Die Aussagen der Akteure können ‚wörtlich‘ genommen werden.

## 2.3 Konjunktion und Kommunikation: die Doppelstruktur alltäglicher Verständigung

Das die Praxis orientierende, das handlungsleitende Wissen ist also nicht ein Wissen *über* etwas, sondern ein Wissen *um* und *innerhalb* von etwas. Letzteres wird in der selbst erlebten *Praxis*, also in einer Praxis, in welche die Akteure jeweils selbst eingebunden sind, erworben, eben er-lebt. Sofern den Akteuren dieses Erleben gemeinsam ist, sie somit über „Gemeinsamkeiten der Erlebnisschichtung“ verfügen, verstehen sie einander unmittelbar, ohne einander erst interpretieren zu müssen. Sie bilden einen gemeinsamen Erfahrungsraum, einen „*konjunktiven Erfahrungsraum*“ (Mannheim 1980: 230f. sowie Bohnsack/Schäffer 2002) und zeichnen sich aus durch Übereinstimmungen ihres Habitus.

Den beiden unterschiedlichen Arten des Wissens, dem theoretischen einerseits und dem atheoretischen oder impliziten andererseits, entsprechen somit auch zwei

3 Der Begründer der Ethnomethodologie, Harold Garfinkel (1961 u. 1967), hat die dokumentarische Methode der Interpretation von Mannheim in den 60er Jahren wieder entdeckt, ihren theoretischen Hintergrund allerdings in der skizzierten Weise eingeeengt.

4 Im Bereich der Evaluationsforschung hat Thomas A. Schwandt (1997 u. 2002: 47) in einer Argumentation, die der unsrigen, also derjenigen der praxeologischen Wissenssoziologie, verwandt ist, eine „praktische Hermeneutik“ gefordert. Er hat kritisiert, dass das gegenwärtig dominante Verständnis von Evaluation in weiten Bereichen an ein Konzept theoretischen Wissens und theoretischer Intelligenz gebunden sei, welches der Praxis unseres alltäglichen Handelns und der damit verbundenen praktischen Beziehung zur Welt nicht gerecht zu werden vermag. (s. genauer: Bohnsack 2006a).



unterschiedliche Arten der *Verständigung* und der *Sozialität*. Obschon Bourdieu einen entscheidenden Beitrag zur Praxeologie des Handelns geleistet hat, vermochte er die Konsequenzen für eine Differenzierung unterschiedlicher Modi der Verständigung und der Sozialität nicht bzw. lediglich an Ansätzen herauszuarbeiten.<sup>5</sup> Ein unmittelbares *Verstehen* ist unter denjenigen möglich, denen dieselben konjunktiven Erfahrungsräume gemeinsam sind, die also sozialisationsbedingt über gemeinsames atheoretisches Wissen verfügen. Die Mutter/Kind-Beziehung und die Familie stellen konjunktive Erfahrungsräume par excellence dar. Allerdings sind konjunktive Erfahrungsräume nicht allein an gruppenhafte Beziehungen oder an die direkte Interaktion gebunden. Gemeinsamkeiten der Erlebnisschichtung finden wir auch auf abstrakterer Ebene im Sinne von (beispielsweise) milieu-, generations- und geschlechtsspezifischen Erfahrungsräumen (s. dazu: Bohnsack/Schäffer 2002).

Diese Verständigung auf der Basis konjunktiver Wissensbestände, also dieses unmittelbare *Verstehen*, nennen wir auch *konjunktive* Verständigung. Eine Verständigung über die Grenzen unterschiedlicher (konjunktiver) Erfahrungsräume oder Milieus hinweg ist auf dem Wege des *Interpretierens* möglich. Dort, wo wir uns im Medium von Theorien, von theoretischen Wissensbeständen verständigen, *interpretieren* wir einander. In Anlehnung an Mannheim (1980: 285ff.) sprechen wir hier auch von „*kommunikativer*“ Verständigung, einer Verständigung auf der Basis kommunikativ-generalisierender Wissensbestände.

Um diesen beiden Arten des Wissens gerecht zu werden, ist es notwendig, die Doppelstruktur alltäglicher Erfahrungs- und Begriffsbildung zu beachten bzw. die „Doppeltheit der Verhaltensweisen in jedem einzelnen, sowohl gegenüber Begriffen als auch Realitäten“ (Mannheim 1980: 296). Denn Bezeichnungen und Äußerungen haben einerseits eine öffentliche oder gesellschaftliche und andererseits eine nicht-öffentliche oder milieuspezifische Bedeutung. So ist uns die öffentliche oder ‚wörtliche‘ Bedeutung des Begriffs ‚Familie‘ unproblematisch gegeben, da wir alle über ein Wissen um die Institution Familie verfügen.

Dieses kommunikative oder auch kommunikativ-generalisierende Wissen ermöglicht uns aber noch keinen Zugang zum Erfahrungsraum der je konkreten Familie in seinem milieuspezifischen oder auch individuell-fallspezifischen Eigensinn, der auf Gemeinsamkeiten der Sozialisationsgeschichte nach Art eines „kollektiven Gedächtnisses“ (Halbwachs 1985) basiert, durch welches sich das konjunktive Wissen auszeichnet.

Während für uns als Sozialforscher/innen der methodische Zugang zum kommunikativen Wissen relativ unproblematisch ist, da dieses direkt erfragt werden kann, erschließt sich uns das konjunktive Wissen nur dann, wenn wir uns (auf dem Wege von Erzählungen und Beschreibungen oder auch der direkten Beobachtung)

5 Derartige Ansätze finden sich dort, wo er das „unmittelbare ‚Verstehen‘“ definiert: Dieses „setzt ein unbewusstes Verfahren der Entschlüsselung voraus, dem nur dort voller Erfolg beschieden ist, wo die Kompetenzen beider: desjenigen, der sie in seiner Handlung oder seinem Werk verwirklicht, und desjenigen, der sie in seiner Wahrnehmung dieses Handelns oder dieses Werkes objektiv einsetzt, zur Deckung kommt“ (Bourdieu 1976: 152). Diese Einsichten hindern Bourdieu aber nicht daran, den Habitus fast ausschließlich unter dem Aspekt der Distinktion, nicht aber unter demjenigen der Konjunktion zu betrachten. Darüber hinaus fehlt der systematisch-methodische Zugang zur *Mehrdimensionalität* des Habitus: s. dazu die Ausführungen zur Typenbildung weiter unten.

mit der Handlungspraxis vertraut gemacht haben. Es handelt sich um ein Wissen, welches von den Erforschten selbst nicht so ohne weiteres auf den Begriff gebracht, also *begrifflich-theoretisch expliziert* werden kann.

Die begrifflich-theoretische Explikation ist die Aufgabe und Leistung der dokumentarischen Interpretation. Die dokumentarische Methode eröffnet mit der Kategorie des „atheoretischen Wissens“ den Blick auf eine Sinnstruktur, die bei den Akteuren selbst wissensmäßig repräsentiert ist, ohne aber Gegenstand ihrer Reflexion zu sein. Somit gehen die Beobachter – und dies ist entscheidend – nicht davon aus, dass sie *mehr* wissen als die Akteure oder Akteurinnen, sondern davon, dass letztere selbst nicht wissen, was sie da eigentlich alles wissen. Die dokumentarische Methode geht damit erkenntnistheoretisch auf Distanz zu einer „Hierarchisierung des Besserwissens“, wie Niklas Luhmann (1990: 510) dies kritisch genannt hat, und stellt sich ganz wesentlich die Aufgabe, *implizites Wissen zur Explikation zu bringen*.

Das implizite Wissen ist zugleich das handlungsleitende, das die Handlungspraxis orientierende Wissen. Zwar habe ich in der (teilnehmenden) Beobachtung einen unmittelbaren und direkten Zugang zur Handlungspraxis (s. dazu auch: Vogd 2005b), allerdings interessiert den sozialwissenschaftlichen Interpreten (wie dargelegt) nicht nur (und zumeist auch gar nicht im Kern) die Faktizität dessen, was dort passiert, sondern die von den Akteuren mit diesen Ereignissen verbundenen *Orientierungen*. Denn nur diese Orientierungsmuster, also die das Handeln leitenden und orientierenden (individuellen oder kollektiven) Wissens- und Erfahrungsbestände, sind es, die diesem Handeln Dauer und Kontinuität verleihen. Nur wenn ich diese handlungsleitenden Wissensbestände kenne, kann ich prognostizieren, ob und wie dieses Handeln auch in Zukunft verlaufen wird.

## 2.4 Methodik und empirische Verfahrensweise

Die Leitdifferenz von kommunikativem und konjunktivem Sinngehalt bestimmt auch die Arbeitsschritte der dokumentarischen Methode. Wir beginnen mit der Rekonstruktion der Ebene des kommunikativ-generalisierten Wissens, der *formulierenden Interpretation*, und entfalten auf dieser Basis dann den Zugang zum konjunktiven Erfahrungswissen mit Hilfe des Arbeitsschritts der *reflektierenden Interpretation*, auf den dann schließlich derjenige der *Typenbildung* folgt.

### *Formulierende Interpretation: die Frage nach dem Was*

Im Bereich der Textinterpretation ist der erste Schritt derjenige der (Re-) *Formulierung* dessen, was von den Erforschten selbst expliziert, also *wörtlich* mitgeteilt wurde. Wir sprechen deshalb auch von *formulierender* Interpretation. Im Falle der Bildinterpretation bewegt sich die formulierende Interpretation auf der *ikonografischen* Ebene (hier wird formuliert, was Thema oder Sujet des Bildes ist). Hinzu kommt hier aber die *vor-ikonografische* Ebene (hier wird formuliert, was auf dem Bild zu sehen ist; s. Kap. 4).

Im Bereich der Textinterpretation bildet die thematische Gliederung, die Entschlüsselung der thematischen Struktur der Texte, das Grundgerüst der formulierenden Interpretation. Es gilt das, was *thematisch* wird und als solches Gegenstand der formulierenden Interpretation ist, von dem zu unterscheiden, *wie* ein Thema, d.h. in welchem *Rahmen* oder in welchem *modus operandi*, es behandelt wird. Während also die formulierende Interpretation der Frage nachgeht, *was* mitgeteilt wird, geht die *reflektierende Interpretation* der Frage nach, *wie* das Mitgeteilte hergestellt wird, welcher *Orientierungsrahmen* oder welcher *Habitus* (zu den Begriffen s. Bohnsack 1998) sich in dem Gesagten oder bildhaft Dargestellten über eine Gruppe, ein Milieu, eine Generation oder auch ein Individuum *dokumentiert* (genauer dazu: Bohnsack 2001 u. 2003a).<sup>6</sup>

### *Reflektierende Interpretation: die Frage nach dem Wie*

Im Bereich der *Textauswertung* folgt die reflektierende Interpretation der *Sequenzanalyse*, in der Ausprägung, wie sie für die dokumentarische Methode spezifisch ist: Grundlegend konstituiert sich in der Relation von (empirisch beobachtbarer) erster Äußerung und (empirisch beobachtbarer) zweiter Äußerung, also der Anschlussäußerung (oder, wenn wir es im Sinne von G.H. Mead 1968 formulieren, in der Relation von Geste und Reaktion), eine Regelmäßigkeit, die es zu erschließen gilt, wenn die Bedeutung, die Signifikanz der ersten Äußerung herausgearbeitet werden soll.

Die Rekonstruktion dieser Regelmäßigkeit oder Signifikanz vollzieht sich – im Sinne der dokumentarischen Methode – nun derart, dass der Interpret nach (alternativen) *Anschlussäußerungen* sucht, die genauso gut als sinnvolle Reaktionen auf die vorherige Äußerung gelten können. Der Interpret sucht also nach funktionalen Äquivalenten zu der Art und Weise, *wie*, also in *welchem Rahmen*, das mit der ersten Äußerung gesetzte Thema durch die (empirisch gegebene) Anschlussäußerung bearbeitet wird. Der Interpret bildet auf diese Weise eine *Klasse* oder *Reihe* von Anschlussäußerungen, die *homolog* sind, die also derselben ‚Regel‘ oder Orientierung zuzuordnen sind. Dies ist der Weg, *abduktiv* (vgl. Bohnsack 2001: 336 u. 2008d: 197f.) eine Regel zu erschließen und zur Explikation zu bringen.

Im Bereich der Gesprächsanalyse der dokumentarischen Methode, wie sie vor allem im Zusammenhang mit dem Gruppendiskussionsverfahren angewandt wird, wird die Art und Weise der Bezugnahme von Äußerung und Anschlussäußerung(en) auch in ihrer formalen Struktur rekonstruiert, die ich *Diskursorganisation* genannt habe (vgl. Bohnsack 1989 sowie Bohnsack/Przyborski 2006).<sup>7</sup>

Wesentlich ist dabei allerdings, dass diese Suche nach funktionalen Äquivalenten immer auch einen *Vergleichshorizont*, eine Kontrastfolie nicht dazugehöriger, d.h. zu anderen ‚Klassen‘ von Orientierungen gehörender, Anschlussäußerungen

6 Der Übergang von der Frage nach dem Was zur Frage nach dem Wie entspricht dem Übergang von „Beobachtungen erster Ordnung“ zu „Beobachtungen zweiter Ordnung“ bei Niklas Luhmann (1990: 86).

7 Im Bereich der dokumentarischen Bildinterpretation geht es um die Rekonstruktion der formalen Komposition des Bildes: s. dazu Kap. 2 u. 3

voraussetzt. Dieser Vergleichshorizont bleibt implizit: „Alles Beobachten ist Benutzen einer Unterscheidung zur Bezeichnung der einen (und nicht der anderen) Seite. Die Unterscheidung fungiert dabei unbeobachtet“ (Luhmann 1990: 91). Dieser „blinde Fleck“ (a.a.O: 85) ist das, was wir in der praxeologischen Wissenssoziologie im Sinne von Mannheim (1952: 227) auch die Standortgebundenheit oder auch Seinsverbundenheit der Interpret(inn)en nennen.

Sie kann und muss im Sinne der dokumentarischen Methode derart einer methodischen Kontrolle zugeführt werden, dass *empirisch überprüfbare* Vergleichshorizonte in Form eines Fallvergleichs dagegehalten werden (Wie wird dasselbe Thema in anderen Fällen – Gruppendiskussionen oder Interviews – in einer anderen Art und Weise, d.h. innerhalb eines anderen Orientierungsrahmens, bearbeitet?). Diesen kontrollierten Fallvergleich bezeichnen wir in Anknüpfung an die Chicagoer Schule (vgl. Glaser/Strauss 1967) auch als *komparative Analyse* (vgl. dazu: Bohnsack 2007c u. 2008d sowie Nohl 2001). Deren Bedeutung für die dokumentarische Methode besteht u.a. darin, dass der (empirisch fundierte) Fallvergleich möglichst frühzeitig in die Analyse einbezogen werden sollte.

Im Sinne der dokumentarischen Methode lässt sich also die Sequenzanalyse, welche nicht nur in der dokumentarischen Methode, sondern auch in anderen Traditionen rekonstruktiver Sozialforschung als generelles Prinzip der Textinterpretation gilt<sup>8</sup>, auf ein noch generelleres Prinzip zurückführen: auf dasjenige der Operation mit Vergleichshorizonten, also auf das Prinzip der *komparativen Analyse*. Dieses Prinzip gilt, wie in Kap. 3 dargelegt wird, auch für die Bildinterpretation – allerdings nicht in Form der Sequenzanalyse, welche der Eigenlogik des Bildes keineswegs entspricht. Sobald wir das Prinzip der Sequenzanalyse direkt auf das Bild zu übertragen suchen, zielen wir an dessen Eigengesetzlichkeit vorbei.

### *Typenbildung und Generalisierung*

Die Typenbildung (s. dazu auch Bohnsack 2007c u. 2009a) vollzieht sich in aufeinander aufbauenden Stufen der Abstraktion bzw. der Abduktion auf der Grundlage der komparativen Analyse – nach Art der Rekonstruktion von Gemeinsamkeiten im Kontrast und von Kontrasten in der Gemeinsamkeit. Das den (Fall-) Vergleich strukturierende Dritte, das *tertium comparationis*, ist zunächst ein gemeinsames *Thema*.

### *Sinngenetische Typenbildung*

Die erste Stufe der Typenbildung ist dann erreicht, wenn durch die Kontraste zwischen den Fällen hindurch ein ihnen *allen* gemeinsamer Orientierungsrahmen bzw.

8 Die Sequenzanalyse wird u.a. von der Konversationsanalyse wie auch der objektiven Hermeneutik und der dokumentarischen Methode im Bereich der *Textinterpretation* in Anspruch genommen. Allerdings zeigen sich hier auch erhebliche Unterschiede zwischen diesen Traditionen. Zur dokumentarischen Methode der Textinterpretation und ihren Unterschieden zur objektiven Hermeneutik siehe Bohnsack 2001 u. 2003b sowie Bohnsack 2008a.

ein gemeinsames Orientierungsproblem identifizierbar ist. In unserer Untersuchung von Jugendlichen türkischer Herkunft, die als Beispiel dienen soll, ist dies das Orientierungsproblem der „Sphärendifferenz“, das Problem der Differenz zwischen der Sphäre der Familie und Verwandtschaft einerseits und der gesellschaftlichen Öffentlichkeit andererseits. Dieser allen Fällen gemeinsame Orientierungsrahmen kann als *Basistypik* bezeichnet werden. In unserem Beispiel ist die Basistypik die Migrationstypik, ein in den Gemeinsamkeiten der Migrationsgeschichte, des Erfahrungsraums der Migration fundiertes Orientierungsproblem. In der unterschiedlichen Art und Weise der Bewältigung dieses gemeinsamen Bezugsproblems der Basistypik zeigt sich die besondere Charakteristik der Fälle, der je fallspezifische *Erfahrungsraum*.

### *Soziogenetische Typenbildung und die Mehrdimensionalität dokumentarischer Interpretation*

Das konjunktive (Orientierungs-) Wissen als ein in die Handlungspraxis eingelassenes und diese Praxis orientierendes und somit vorreflexives oder implizites Erfahrungswissen ist dem Interpretieren nur zugänglich, wenn er sich den dazugehörigen Erfahrungsraum erschließt. Die Komplexität dieser Erfahrungswirklichkeit und somit die Komplexität der empirischen Analyse stellen uns allerdings vor das Problem, dass das Individuum bzw. die konkrete Gruppe, welche jeweils den zu untersuchenden Fall bilden, immer schon teilhaben an *unterschiedlichen* Erfahrungsräumen. Mit Bezug auf unser Beispiel bedeutet dies, dass uns der migrationspezifische Erfahrungsraum (die Migrations- oder Basistypik) immer schon in der Überlagerung bzw. wechselseitigen Durchdringung unterschiedlicher Erfahrungsräume bzw. Dimensionen begegnet – beispielsweise bildungs-, geschlechts- und generationstypischer, aber auch alterstypischer, d.h. lebenszyklischer Erfahrungsräume (vgl. Bohnsack 1989).

Somit lässt sich aber das rekonstruierte Orientierungsmuster erst dann als eines der Migrationstypik validieren und generalisieren, nachdem wir kontrolliert haben, ob es sich nicht etwa um Orientierungen handelt, die ganz allgemein typisch sind für die junge Generation (Generationstypisches) oder für die Adoleszenzphase (Alters- oder Entwicklungstypik) oder für männliche Jugendliche (Geschlechtstypik). Es ist also erst dann in valider Weise möglich, das beobachtete Orientierungsmuster dem ‚migrationstypischen Erfahrungsraum‘ zuzuordnen und es somit als eine *migrationstypische* Orientierung zu generalisieren, nachdem in *komparativer Analyse* kontrolliert wurde, ob diese Orientierung bei (Migranten-) Jugendlichen unterschiedlichen Alters, Geschlechts und unterschiedlicher Milieuzugehörigkeit, also durch milieu- und entwicklungsspezifische Variationen oder Modifikationen von Erfahrungsräumen hindurch bzw. in der Überlagerung durch andere Dimensionen oder Erfahrungsräume, auf einer abstrakten Ebene als Gemeinsamkeit identifizierbar bleibt.

Mit der dokumentarischen Methode eröffnet sich damit die Möglichkeit zur Bewältigung des Problems der *Generalisierung* in der qualitativen Sozialforschung (s. dazu auch Bohnsack 2005a u. 2009a). Das Niveau der Validität der einzelnen Typik und die Möglichkeit ihrer Generalisierung sind davon abhängig, inwieweit sie

von anderen, auf der Grundlage der fallspezifischen Beobachtungen ebenfalls möglichen, Typiken unterscheidbar ist, also davon, wie vielfältig, d.h. multidimensional, der einzelne Fall innerhalb einer ganzen Typologie verortet werden kann.

Im Bereich der Textinterpretation ist die Ausdifferenzierung mehrdimensionaler Typiken auf der Grundlage der dokumentarischen Methode inzwischen weit fortgeschritten (s. dazu als Forschungsbeispiele u.a.: Bohnsack 1989; Bohnsack et al. 1995; Nentwig-Gesemann 1999; Nohl 2001 u. 2006a; Schäffer 1996 u. 2003; Schittenhelm 2005; Vogd 2004; Weller 2003; Asbrand 2009). Im Bereich der Bild- und Filminterpretation stehen wir insbesondere in dieser Hinsicht noch ganz am Anfang.

### *Schluss*

Die dokumentarische Methode hat in den Sozial- und Erziehungswissenschaften bis hin zur Informatik und Medizin inzwischen ein weites und vielfältiges Anwendungsfeld gefunden (dazu u.a.: Bohnsack 2003a u. Bohnsack/Nentwig-Gesemann/Nohl 2007). Im Bereich der Sozial- und Erziehungswissenschaft reicht dies von der Kindheits-, Jugend-, Gender- und Migrationsforschung und der Erwachsenenbildung über die Medizinsoziologie, die Polizei- und die Organisationskulturforschung bis hin zur Mediennutzungsanalyse. Das methodische Spektrum umfasst u.a. die Gesprächsanalyse und das Gruppendiskussionsverfahren (s. dazu: Bohnsack/Przyborski/Schäffer 2006; Bohnsack 2004a; Przyborski 2004), die Auswertung von Interviews (s. u.a. Nohl 2006b), die teilnehmende Beobachtung (s. u.a. Bohnsack et al. 1995 u. Vogd 2005b), die Evaluationsforschung (s. u.a. Bohnsack 2006c u. Nentwig-Gesemann 2006c sowie Bohnsack/Nentwig-Gesemann 2009) und schließlich die Bild-, Film- und Videoanalyse, welche Gegenstand des vorliegenden Bandes ist.

Da die dokumentarische Methode ein derart weites Spektrum von methodischen Zugängen und Erhebungsverfahren umfasst, eröffnet sie die Möglichkeit einer validen Methodentriangulation, d.h. Möglichkeiten des koordinierten Bezugs unterschiedlicher Methoden aufeinander – u.a. eben auch der Triangulation text- und bildbasierter Verfahren (s. dazu am Forschungsbeispiel Kap. 4.3).<sup>9</sup>

<sup>9</sup> Für eine Kombination resp. Triangulation von dokumentarischer Methode und anspruchsvollen quantitativen Verfahren siehe: Pfaff 2006.

### 3. Dokumentarische Bildinterpretation

Der methodisch kontrollierte Zugang zum Bild stellt eine der größten Herausforderungen für die gegenwärtige sozialwissenschaftliche Forschung dar. Während im Bereich der Geisteswissenschaften, insbesondere der Philosophie und Kunstgeschichte, wesentliche Vorarbeiten in der Auseinandersetzung mit dem Bild geleistet worden sind, steht die sozialwissenschaftliche Analyse, die ja grundlegend eine empirisch fundierte zu sein hat, noch ganz am Anfang. Zugleich ist evident, dass der hier notwendige empirisch-methodische Zugang nicht – oder allenfalls am Rande – auf der Basis standardisierter Verfahren zu bewältigen ist. Es sind vor allem die qualitativen oder rekonstruktiven Verfahren, die sich den Herausforderungen des Bildes zu stellen haben.

Und es sind ebenso – wie auch im Bereich der Textinterpretation – allein die qualitativen Verfahren, welche den Anschluss der sozialwissenschaftlichen Empirie an die geisteswissenschaftlichen Traditionen der Philosophie und Kunstgeschichte sowie an deren Verstehensdebatten zu leisten vermögen. Es stellt überhaupt eine der wesentlichen Leistungen methodologisch und theoretisch anspruchsvoller qualitativer Forschung dar, dass sie zentrale theoretische Kategorien aus den geisteswissenschaftlichen Traditionen mit den Anforderungen der empirischen Sozialforschung zu verbinden vermag (s. dazu auch Bohnsack 2005a). Im Bereich der Bildinterpretation steht dieser Diskurs allerdings noch ganz am Anfang. Und er ist – im Vergleich zur Textinterpretation – immer noch von marginaler Bedeutung.

#### **3.1 Die Marginalisierung des Bildes in der empirischen Sozialforschung und in den qualitativen Methoden**

Es herrscht weitgehend Übereinstimmung unter den Vertreterinnen und Vertretern qualitativer Sozialforschung, dass – in deutlichem Kontrast zur gesellschaftlichen Bedeutung des Bildes – der Stellenwert, welcher ihm in der Praxis qualitativer Forschung zukommt, gegenüber dem des Textes ein marginaler ist. Allerdings hat diese Einsicht, welche in den letzten Jahren zunehmend artikuliert worden ist, bisher kaum Konsequenzen gehabt. Dies lässt vermuten, dass die Gründe für diese Marginalisierung tiefer liegender Art sind, dass sie im Kernbereich der methodologischen Grundlagen sozialwissenschaftlicher Empirie selbst zu suchen sind – und hier er-

staunlicherweise gerade auch in den methodologischen Grundlagen qualitativer Forschung.

Mindestens vier Gründe sind für diese Marginalisierung geltend zu machen:

Zunächst ist hier jene Entwicklung in Philosophie, Erkenntnistheorie und sozialwissenschaftlicher Handlungstheorie und Empirie zu nennen, die (zuerst von Richard Rorty 1967) als „linguistic turn“ bezeichnet wurde. Welche Bedeutung diese Entwicklung für die methodische Marginalisierung des Bildes hatte, ist spätestens dann deutlich geworden, als dieser sprachwissenschaftlich fundierten Wende die Antizipation eines „iconic“ oder „pictorial turn“ (Mitchell 1994 u. 1997) entgegengehalten wurde. Zu wenig beachtet wurde in der hier geführten Diskussion allerdings, dass diese sprachwissenschaftliche Wende nicht nur philosophisch-erkenntnistheoretisch (u.a. Ricoeur 1972) und sozialphilosophisch-handlungstheoretisch (u.a. Habermas 1971), sondern auch *empirisch-rekonstruktiv* – zuerst durch die ethnomethodologische Konversationsanalyse (vgl. Sacks 1995) – eingeläutet worden ist.

Die Renaissance der qualitativen Methoden in den 1970er Jahren wurde entscheidend durch den empirisch-rekonstruktiv fundierten linguistic turn mit getragen. Methodologische Reflexionen und methodische Verfeinerungen sind seitdem fast ausschließlich im Medium der Textinterpretation fortentwickelt worden. Allerdings ist diese Dominanz des Modells der Textinterpretation in den qualitativen Methoden nicht allein vom linguistic turn her zu plausibilisieren.

Vielmehr ist – *zweitens* – ein weiterer Grund geltend zu machen, welcher schon vorher bzw. unabhängig vom linguistic turn die sozialwissenschaftliche Methodologie geprägt hat und auch als Voraussetzung dafür anzusehen ist, dass der linguistic turn derart nachhaltige Konsequenzen im Bereich der sozialwissenschaftlichen Empirie haben konnte. Unabhängig von ihrer spezifischen Ausrichtung gilt für alle sozial- und naturwissenschaftlichen Methoden und Methodologien die (zuerst von Karl Popper 1971 umfassend ausformulierte) Prämisse, dass soziale Wirklichkeit, wenn sie wissenschaftliche Relevanz gewinnen will, in Form von Beobachtungssätzen oder „Protokollsätzen“, also in Form von Texten, vorliegen muss. Die qualitative oder rekonstruktive Sozialforschung ist dem gefolgt: Als letzte – nicht mehr hintergehbare – Grundlage des *wissenschaftlichen* Zugangs zur Wirklichkeit gelten auch hier Texte. Jegliche Beobachtung, die wissenschaftlich relevant werden soll, muss also durch das Nadelöhr des Textes hindurch.

Hieran anschließend ist dann – *drittens* – im Bereich der qualitativen Sozialforschung die Argumentation zugunsten des Textes in folgender Weise weitergeführt worden: Wenn alles, was wissenschaftlich relevant werden soll, durch das Nadelöhr des Textes muss, dann sollte man als Grunddaten der Analyse auch die *Originaltexte* heranziehen: also jene Texte, die von den Erforschten selbst produziert werden – im Unterschied zu jenen Texten, die als Beobachtungsprotokolle oder Interpretationstexte von den Forschern und Forscherinnen verfasst worden sind. Nur diese Art von Ursprungsdaten kann *unmittelbar* zu Grunddaten der Analyse werden.

Und es erscheint insbesondere dort unerlässlich, auf die Texte der Erforschten, also deren Originaläußerungen, zurückzugreifen, wo die Analyse über die oberflächliche oder manifeste Ebene der Bedeutungen von Äußerungen hinausgreifen will auf die eher impliziten oder latenten Bedeutungsgehalte, für deren Interpretation

Nuancen des Ausdrucks und die genaue Kon-Textuierung entscheidend sind. Vor dem Hintergrund derartiger Überlegungen ist dann eine Fixierung auf das Paradigma der Textinterpretation in den qualitativen Methoden entstanden. Bildinterpretationen mussten von vornherein weniger valide erscheinen. Denn Bilder sind, wenn sie wissenschaftlich relevant werden sollen, grundsätzlich erst einmal in Beobachtungssätze bzw. -texte umzuformulieren.

Im Fahrwasser dieser Argumentation ist (auf den ersten Blick oft unbemerkt) dann – *viertens* – noch eine ganz andere Bedeutung von „Textförmigkeit sozialer Wirklichkeit“ in die methodologische Begründung qualitativer Methoden eingebracht worden: Nicht mehr nur allein die *wissenschaftlich relevante* Wirklichkeit sollte textförmig sein. Vielmehr wurde – vor allem von Seiten der „objektiven Hermeneutik“ – schließlich die Sprach- und somit auch die Textförmigkeit jeglicher Verständigung behauptet, also auch der alltäglichen, der außerwissenschaftlichen Verständigung. Damit wird das Bild bzw. die Bildhaftigkeit, die Ikonizität als ein Medium der Verständigung in seiner Eigenlogik und Eigensinnigkeit gegenüber dem Text dann grundsätzlich in Frage gestellt.<sup>10</sup>

Die im Zusammenhang des linguistic turn entstehende Orientierung am Paradigma der Textinterpretation hat insbesondere seit Ende der 70er Jahre zu einer enormen Präzisierung der Standards und der forschungspraktischen Verfahrensweisen geführt. Einer der Gründe für diese Erfolge kann darin gesehen werden, dass – als methodologisches Prinzip – Texte im Zusammenhang mit dem linguistic turn als selbst-referentielle Systeme betrachtet worden sind. Eine Vorreiterrolle hat hier die sogenannte Konversationsanalyse gespielt. Harvey Sacks (1995: 536) hat dies schon in den 60er Jahren auf die Formel gebracht: „Wenn wir so etwas wie eine Soziologie des Gesprächs betreiben, dann geht es um die Frage, was das System selbst als Grundlagen, Motive, oder was auch immer, zur Verfügung stellt, um etwas Wesentliches für das System zu tun.“<sup>11</sup>

Dieses Prinzip oder diese AnalyseEinstellung, Gespräche und Texte ganz allgemein als selbst-referentielle Systeme zu betrachten, ist im Folgenden auch von anderen Methodologien der Textinterpretation im Bereich der qualitativen Methoden übernommen worden. Bis heute hat dies jedoch keinerlei Anwendung oder Relevanz im Bereich der Bildinterpretation gewonnen. Die Notwendigkeit der Orientie-

10 Vertreter anderer Strömungen qualitativer Forschung haben diese Prämisse kritisiert. Für den Bereich der Bildinterpretation bestreitet Jo Reichertz (1992: 142) als Verfechter einer „hermeneutischen Wissenssoziologie“ in Abgrenzung gegenüber der objektiven Hermeneutik, „dass die Welt nichts außer Text sei – auch wenn Oevermann die Welt gerne so sieht“. Allerdings begrenzt auch Reichertz die Bedeutung und die Reichweite der Ikonizität in entscheidender Weise, wenn er betont: „Will man sich über die Wahrnehmung und den Eindruck des Fotos mit anderen verständigen, muss man ein Protokoll des privaten Ereignisses herstellen, es in einen intersubjektiven Code übersetzen.“ (Reichertz 1992: 143) Wenn – in der Perspektive der hermeneutischen Wissenssoziologie – lediglich ein „privates“, also ein individuelles oder monologisches Verstehen, im Medium des Bildes möglich ist, impliziert dies aber den Ausschluss einer *intersubjektiven* Verständigung in diesem Medium. Letztere ist wiederum auf Sprache und Texte angewiesen.

11 Zitate aus nicht deutschsprachiger Literatur sind dort, wo der nicht deutschsprachige Titel zitiert wird, von mir selbst übersetzt worden – mit Ausnahme der Zitate in den Anmerkungen, die ich in der Originalsprache belassen habe.

nung an einer derartigen AnalyseEinstellung ist eines der zentralen Anliegen dieses Buches.<sup>12</sup>

Wenn wir Bildern den Status von selbst-referentiellen Systemen zuerkennen, so hat dies auch Konsequenzen für die Art und Weise des Verständnisses von Bildern als Medien der Kommunikation. Und zwar lassen sich zwei sehr unterschiedliche Arten der bildhaften Verständigung differenzieren. Es gilt, eine Verständigung *über* das Bild zu unterscheiden von einer Verständigung *durch* das Bild, wie ich dies formulieren möchte.

Eine intersubjektive Verständigung durch das Bild, also im Medium des Bildes und somit jenseits des Mediums von Sprache und Text, bleibt stillschweigend bzw. ohne weiter greifende Begründung aus der Methodologie und auch aus der Handlungstheorie ausgeschlossen. In den Blick gerät lediglich die im Medium von Sprache und Text sich vollziehende Verständigung *über* das Bild. Erst wenn wir der Verständigung im Medium des Bildes selbst theoretisch und methodisch gerecht werden, ermöglicht dies, wie Hans Belting (2001: 15) mit Bezug auf William J. T. Mitchell (1994) formuliert, „Bilder nicht mehr mit Texten zu erklären, sondern von Texten zu unterscheiden.“

### 3.2 Implizites Wissen, Ikonologie und Habitus

In dieser Unterscheidung zwischen einer Verständigung *durch* das Bild von jener *über* das Bild sind Annahmen über unser alltägliches Verstehen und Handeln impliziert, die weit in die Handlungs-, Zeichen-, Wissens- und Erkenntnistheorien hineingreifen. Dass wir uns im Alltag *durch* Bilder verständigen, bedeutet, dass unsere Welt, unsere gesellschaftliche Wirklichkeit „durch Bilder nicht nur repräsentiert, sondern durch die Herstellung von Bildern tatsächlich konstituiert und zur Existenz gebracht wird“ (Mitchell 1994: 41). Dabei kann eine derartige *Herstellung* der Welt durch Bilder allerdings wiederum in mindestens zweifacher Hinsicht verstanden werden. Das eine Verständnis geht dahin, dass lediglich die *Deutung* der Welt sich wesentlich im Medium der Ikonizität vollzieht. Ein darüber hinaus gehendes, weiter greifendes Verständnis einer Konstitution der Welt durch das Medium des Bildes umfasst auch die *handlungsleitende* Qualität der Bilder.

In der sozial- und erziehungswissenschaftlichen Handlungs-, Kommunikations- und Bildungstheorie bleibt letzterer Aspekt, nämlich derjenige, dass Bilder auf einer ganz fundamentalen Ebene der Verständigung und des Lernens, der Sozialisation und der Bildung (auch außerhalb der Massenmedien) Medium alltäglicher Verständigung und alltäglichen Handelns sind, systematisch unbeachtet. Dies gilt insbesondere in der Hinsicht, dass soziale Situationen oder Szenarien in Form von mentalen Bildern gelernt werden, dass sie u.a. im Medium des Bildes erinnert werden, in wesentlicher Hinsicht bildhaft im Gedächtnis sedimentiert sind.<sup>13</sup>

<sup>12</sup> Zum Verhältnis von Systemtheorie und dokumentarischer Methode siehe auch: Vogd 2005a.

<sup>13</sup> Dazu heißt es bei Assmann (1997: 37f.): „So abstrakt es beim Denken zugehen mag, so konkret verfährt die Erinnerung. Ideen müssen versinnlicht werden, bevor sie als Gegenstände im Ge-

So verfüge ich über typenhafte mentale Bilder, die mir Aufschluss über die Bedeutung von Mimik und Gestik meiner jeweiligen Interaktionsgegenüber vermitteln. Wir sind ständig gefordert, Gesichtsausdruck und Körperhaltung der relevanten anderen in adäquater Weise zu interpretieren, um Handlungssicherheit gewinnen zu können. Als Voraussetzung dafür müssen wir aber bereits über mentale Bilder, über bildhafte, also ikonische Schemata verfügen, durch die Mimik und Körperhaltung für uns erst zu sinnvollen Zeichen werden. Mehr noch sind in jeder Art von Zeichen oder Bedeutungssystemen Bilder impliziert. Das zu jedem Signifikant (einem Wort beispielsweise) gehörende Signifikat ist nicht ein Ding, sondern ein mentales oder „psychisches“ Bild: „das Signifikat des Wortes Ochs ist nicht das Tier Ochs, sondern sein psychisches Bild“ (Barthes 1983: 53).

Nach Alfred Schütz (1971: 4) basiert jegliche Symbol- bzw. genauer: Typenbildung auf einer „Imagination hypothetischer Sinnesvorstellungen“. Und Cornelius Castoriadis (1984: 217) nimmt in dieser Hinsicht bezug auf die „imaginäre Komponente jedes Symbols“. Eine derartige mentale Bildhaftigkeit, das „Imaginäre“ (Wulf 1999: 333f.), in dem auch die Bildhaftigkeit sozialer Orientierungen fundiert ist, stellt also nicht nur die Grundlagen für die Produktion von Ikonizität, von äußeren Bildern dar, sondern ist auch der Symbolik der Sprache vorgeordnet. Auch kreative Entwürfe sozialer Handlungen und Szenarien vollziehen sich im Medium mentaler Bilder und Imaginationen.

Die Verständigung im Medium des Bildes, d.h. im Medium von mentalen Bildern, ist weitgehend eine vorreflexive, eine implizite. Es handelt sich um eine Verständigung, die sich unterhalb der begrifflich-sprachlichen Explizierbarkeit vollzieht. Die bildhafte Verständigung ist eingelassen in die stillschweigenden oder „atheoretischen“ Wissensbestände, wie sie bei Karl Mannheim (1980) genannt werden. Diese Wissensbestände strukturieren vor allem das habituelle, das routinemäßige Handeln und werden ganz wesentlich erlernt im Modus der Verinnerlichung bzw. der „mimetischen“ Aneignung (vgl. Wulf 1998) von sozialen Szenarien, von Gebärden, Gestik und Mimik.<sup>14</sup>

Dieses Wissen wird einerseits in Form von Erzählungen und Beschreibungen vermittelt, d.h. in Form von Metaphern, von metaphorischen, also von bildhaften textlichen Darstellungen sozialer Szenarien, andererseits – und ganz wesentlich – im Medium des Bildes selbst, im Medium der Ikonizität. Das Medium der Vermittlung des atheoretischen Wissens ist also ganz allgemein dasjenige der „Bildlichkeit“, wenn wir den Begriff der Bildlichkeit mit Gottfried Boehm (1978: 447) so fassen, dass „Bild und Sprache an einer gemeinsamen Ebene der ‚Bildlichkeit‘ partizipieren“. Und diese Ebene der Bildlichkeit ist diejenige des impliziten oder atheoretischen Wissens. Vor diesem Hintergrund wird dann aber auch deutlich, dass nur solche sozialwissenschaftlichen Methodologien den empirisch-methodischen Zugang zur Sinnstruktur des Bildes zu eröffnen vermögen, die über einen

dächtnis Einlaß finden können. Dabei kommt es zu einer unauflösbaren Verschmelzung von Begriff und Bild.“

<sup>14</sup> Derartige „innere Bilder, bildhafte Vorstellungen, die in den äußeren Bildern in Erscheinung treten und eine deutlichere Gestalt gewinnen“ sind nach Schulze (1999: 63) Gegenstand bzw. Ziel einer „pädagogischen Ikonologie“ (a.a.O.: 62).

grundlagentheoretischen Zugang zu derartigen atheoretischen oder impliziten Wissensbeständen verfügen.

Der Wechsel von der Ebene des expliziten Wissens, der Ebene von Common Sense-Theorien, hin zum impliziten oder atheoretischen Wissen ist bei Panofsky derjenige von der Ikonografie zur Ikonologie. Panofsky ist ganz wesentlich durch die sozialwissenschaftliche Diskussion beeinflusst worden, insbesondere durch seinen Zeitgenossen Karl Mannheim und dessen dokumentarische Methode.<sup>15</sup> Dem Wechsel von der Ikonografie zur Ikonologie entspricht der Wechsel vom immanenten zum dokumentarischen Sinngehalt bei Karl Mannheim (1964a). Hieran anknüpfend sowie im Anschluss an Heidegger (1986) und Luhmann (1990) lässt sich dieser Wechsel der AnalyseEinstellung als derjenige vom *Was* zum *Wie* bezeichnen. Es geht um den Wechsel von der Frage, *was* kulturelle oder gesellschaftliche Phänomene oder Tatsachen *sind*, zur Frage, *wie* diese *hergestellt* werden.

Bei Panofsky umfasst die Frage nach dem *Was* aber nicht nur die Ebene der Ikonografie, sondern auch die *vor*-ikonografische Ebene. Die sozialwissenschaftliche Relevanz dieser Unterscheidung wird vor allem dort deutlich, wo Panofsky (1975: 38) die von ihm entworfenen Interpretationsschritte nicht im Bereich der Kunst, sondern des „Alltagslebens“ und dort am Beispiel der Gebärde eines Bekannten erläutert.<sup>16</sup> Diese Gebärde, die auf der vor-ikonografischen Ebene zunächst als „Hutziehen“ identifizierbar ist, kann im Sinne von Panofsky erst auf der ikonografischen Ebene als ein „Grüßen“ analysiert werden. In sozialwissenschaftlicher Fortentwicklung der Argumentation von Panofsky lässt sich dieser Schritt der Interpretation als derjenige der Unterstellung von „Um-zu-Motiven“ charakterisieren: Der Bekannte zieht seinen Hut *um zu* grüßen. Wir begeben uns – wie wir dies auch im Common Sense tun – auf die Suche nach den subjektiven Intentionen. Die Ikonografie bleibt methodisch ungesichert, da sie auf Introspektion und Unterstellungen basiert.

Im Unterschied zur ikonografischen vollzieht die ikonologische AnalyseEinstellung den „Bruch mit den Vorannahmen des *common sense*“, wie man mit Bourdieu (1996: 278) sagen könnte. Sie unterscheidet sich radikal von der Frage nach dem *Was* und fragt nach dem *Wie*, nach dem modus operandi der *Herstellung* bzw. *Entstehung* einer Gebärde. Nach Panofsky erschließt sich auf diese Weise „die eigentliche Bedeutung“ oder der „Gehalt“ einer Gebärde (Panofsky 1975: 40), der „Wesensinn“ oder eben „Dokumentsinn“ (Panofsky 1932: 115 u. 118) – ein Begriff, mit dem Panofsky sich explizit auf Karl Mannheim (1964a) und dessen dokumenta-

15 Zum Verhältnis von Mannheim und Panofsky siehe auch Barboza 2005 sowie Hart 1993.

16 Der Paradigmawechsel, wie er von Panofsky von der Ikonografie zur Ikonologie (und von Mannheim mit der immanenten zu dokumentarischen Interpretation) vollzogen wurde, ist nicht allein durch den damit verbundenen Wechsel der AnalyseEinstellung bestimmt, sondern zugleich auch durch den Übergang zu einem gänzlich unterschiedlichen Verhältnis zum Alltag. Nunmehr erscheint nicht allein und weniger das explizierbare ikonografische Vor-Wissen des (kunsthistorischen) Gelehrten als die wesentliche Voraussetzung für eine wissenschaftliche Interpretation, sondern die intuitiven interpretativen Kompetenzen der Alltagsinterpretation, die „Alltagsmethoden“ und deren Systematisierung, werden zur wesentlichen Grundlage. Damit ist ein wesentlicher Schritt getan hin zu einer Methodologie, die – wie 40 Jahre später dann in der Ethnomethodologie explizit gefordert – die Rekonstruktion der Alltagsmethoden und deren Systematisierung zur Grundlage wissenschaftlicher Interpretation macht, also ein Schritt zu einer rekonstruktiven Sozialforschung.

rische Methode bezieht. Auf dem Wege der ikonologischen Interpretation werden wir an der Gebärde den „Eindruck einer ganz bestimmten Wesensart erhalten können (...), die sich in der Grußhandlung ebenso klar und ebenso unabhängig vom Willen und Wissen des Grüßenden ‚dokumentiert‘, wie sie sich in jeder anderen Lebensäußerung des betreffenden Menschen dokumentieren würde“ (Panofsky 1975: 115f.).

Es ist eben dieser sich hier dokumentierende „Wesensinn“, der ikonologische Sinngehalt, welchen Panofsky (1989) auch als „Habitus“ bezeichnet hat. Der ikonologische Sinngehalt „wird erfaßt, indem man jene zugrundeliegenden Prinzipien ermittelt, die die Grundeinstellung einer Nation, einer Epoche, einer Klasse, einer religiösen oder philosophischen Überzeugung enthüllen“ (Panofsky 1975: 40).

Die Frage nach dem ikonologischen Sinngehalt zielt im Sinne von Panofsky also auf den *Habitus* der Bildproduzent(inn)en. Allerdings erscheint es hier notwendig, grundsätzlich zwei Dimensionen oder Arten von Bildproduzent(inn)en zu unterscheiden: Auf der einen Seite haben wir die (wie ich es nennen möchte) *abbildenden* Bildproduzent(inn)en, also u.a. Fotografen oder Künstler sowie alle diejenigen, die als Akteure, als Produzenten *hinter* der Kamera und noch nach der fotografischen Aufzeichnung an der Bildproduktion beteiligt sind. Auf der anderen Seite haben wir die *abgebildeten* Bildproduzent(inn)en, also die Personen, Wesen oder sozialen Szenarien, die zum Sujet des Bildes gehören bzw. *vor* der Kamera agieren.

Die sich aus der komplexen Relation dieser beiden unterschiedlichen Arten von Bildproduzent(inn)en ergebenden methodischen Probleme sind dann leicht zu bewältigen, wenn beide zu *demselben* ‚Erfahrungsraum‘, d.h. zum selben Milieu, gehören. Dies ist beispielsweise dort der Fall, wo ein Angehöriger der Familie ein Familienfoto produziert oder wenn (wie im Falle historischer Gemälde, die mir Aufschluss über eine historische Epoche zu geben vermögen) der Maler ebenso wie die Modelle oder die abgebildeten Szenarien zur selben Epoche gehören. Denn bei der ikonologischen Interpretation geht es darum, einen Zugang zum Erfahrungsraum der Bildproduzent(inn)en zu finden, dessen zentrales Element der individuelle oder kollektive Habitus darstellt. Methodisch komplexer wird das Problem dort, wo der Habitus der *abgebildeten* Bildproduzent(inn)en mit demjenigen der *abbildenden* Bildproduzent(inn)en, also der Fotografen oder Maler, nicht so ohne weiteres in Übereinstimmung zu bringen ist (als Beispiel siehe: Bohnsack 2008c: 249-257).

Die besondere Leistung der Ikonologie von Panofsky, die abgesehen von den Arbeiten Bourdieus und einigen anderen Ausnahmen (s. dazu: Pilarczyk/Mietzner 2000 u. 2005 sowie Michel 2006a u. 2007) in den Sozial- und Erziehungswissenschaften bisher kaum Anwendung gefunden hat, ist darin zu sehen, dass er den Habitus bzw. den Wesensinn oder Dokumentsinn (beispielsweise einer Epoche wie der Renaissance) aus den Analogien oder Homologien *unterschiedlicher* Medien, unterschiedlicher Darstellungsgattungen oder Kunstgattungen (von der Literatur über die Malerei und Architektur bis zur Musik) dieser Epoche hervortreten lässt. Für Max Imdahl (1994 u. 1996a) ist aber gerade diese besondere Leistung Panofskys Ausgangspunkt für die zentrale Frage danach, wo dann noch das Besondere des Mediums Bild in den Interpretationen von Panofsky zu finden sei. Panofsky ist nicht primär an jenen Sinngehalten interessiert, die *nur* durch das Bild, sondern an jenen, die unter anderem *auch* durch das Bild zu vermitteln sind.

### 3.3 Die Eigensinnigkeit des Bildes und die Suspendierung des textlichen Vorwissens

In diesem Zusammenhang kritisiert Imdahl auch die reduzierte Bedeutung von „Formen“ und „Kompositionen“ bei Panofsky. Formen und Kompositionen würden auf die Funktion reduziert, die (natürlichen) Gegenständlichkeiten des Bildes und die ikonografischen Narrationen (z.B. der biblischen oder heilsgeschichtlichen Texte) wiedererkennbar zu gestalten. Dieser Reduktion auf das „wiedererkennende Sehen“ stellt Imdahl das „sehende Sehen“ gegenüber, welches nicht von den einzelnen Gegenständlichkeiten ausgeht, sondern von der Gesamtkomposition und der Ganzheitlichkeit des Bildes.<sup>17</sup>

Im Unterschied zur Ikonologie von Panofsky, die auf der Ikonografie und damit auf dem textlich-narrativen Vorwissen aufbaut, setzt Imdahl mit seinem methodischen Zugang, den er als *ikonische* Interpretation oder *Ikonik* bezeichnet, bereits auf der vor-ikonografischen Ebene an – vor allem aber bei der *formalen Komposition* des Bildes. „Der ikonischen Betrachtungsweise oder eben der Ikonik wird das Bild zugänglich als ein Phänomen, in welchem gegenständliches, wiedererkennendes Sehen und formales, sehendes Sehen sich ineinander vermitteln zur Anschauung einer höheren, die praktische Seherfahrung sowohl einschließenden als auch prinzipiell überbietenden Ordnung und Sinntotalität.“ (Imdahl 1996a: 92).<sup>18</sup>

Die ikonische Interpretation kann sich nach Imdahl vom ikonografischen Vorwissen oder von ikonografischen Sinnzuschreibungen weitgehend freihalten. Sie kann – wie es bei Imdahl heißt – „von der Wahrnehmung des literarischen oder szenischen Bildinhalts absehen, ja sie ist oft besonders erfolgreich gerade dann, wenn die Kenntnis des dargestellten Sujets sozusagen methodisch verdrängt wird“ (1996b: 435).

Eine derartige Suspendierung erscheint methodisch geboten, wenn wir das Bild im Sinne von Imdahl (1979: 190) als „ein nach immanenten Gesetzen konstruiertes und in seiner Eigengesetzlichkeit evidentes System“ erfassen wollen.<sup>19</sup> Hinsichtlich

17 Der Differenzierung von wiedererkennendem, d.h. auf die Umwelt des Bildes und deren Gegenständlichkeit bezogenem Sehen und dem auf die Gesetzmäßigkeit des Bildes selbst bezogenen Sehen entspricht im Bereich der Textinterpretation die Differenzierung von propositionaler und performativer Dimension der Darstellung (vgl. Bohnsack 2007f.).

18 Für den Kunsthistoriker und Kunstwissenschaftler Steffen Bogen (2005: 57) stellt die Ikonik von Max Imdahl „einen kaum zu überbietenden Höhepunkt dar, künstlerische Ausdrucksmöglichkeiten im Medium des Einzelbildes anschaulich und analytisch nachzuvollziehen.“

19 Dieses methodische Postulat hat in der Rezeption der Methode dokumentarischer Bildinterpretation zu Missverständnissen geführt. So heißt es bei Fuhs (2007, S. 211): „Ein solches Vorgehen der asketischen Interpretation formaler Bildkompositionen geht unausgesprochen davon aus, dass der Alltagssinn und das intuitive Verstehen von Bildern nicht nur keinen Beitrag zum ikonischen Verständnis der Weltsymbolik eines Bildes liefert, sondern dass die Alltagsinterpretation sogar den eigentlichen Sinn des Bildes verstellt.“ (S. 211) Dies trifft insofern nicht zu, als die dokumentarische Methode sich eben gerade als Methode der Explikation intuitiven, d.h. *atheoretischen*, also impliziten oder inkorporierten Alltagswissens versteht, welches allerdings vom *theoretischen* Wissen, den Common Sense-Theorien unterschieden wird. Dies habe ich immer wieder betont (siehe u.a. Bohnsack 2008d: 194ff.), ebenso wie auch die erkenntnistheoretische Distanz der dokumentarischen Methode von einer „Hierarchisierung des Besserwissens“, wie ich dies mit Bezug auf Luhmann genannt habe (s. dazu u.a. Kap. 2 in diesem Band), also die Distanz gegenüber dem Anspruch eines sozialwissenschaftlichen Zugangs, der das Privileg einer *höheren* Rationalität in An-

dieser Suspendierung des ikonografischen Vor-Wissens zeigen sich gewisse Parallelen zur Semiotik mit ihren beiden prominenten Vertretern Umberto Eco und Roland Barthes, die (unabhängig von den Unterschieden zwischen ihnen) dahingehend übereinstimmen, dass wir – um zur Besonderheit und Eigensinnigkeit des Bildes vorzudringen – *unterhalb* der konnotativen Ebene anzusetzen haben. Die konnotative Ebene lässt sich aber – und dies wird bei Eco (1994: 243) auch explizit hervorgehoben – in mancher Hinsicht analog zur ikonografischen Ebene bei Panofsky verstehen.<sup>20</sup>

Die Besonderheit und Eigensinnigkeit des Bildes im Unterschied zum Text, d.h. die besondere Botschaft der bildhaften, der ikonischen Zeichen, entscheidet sich also auf der *vor-ikonografischen* oder denotativen Ebene.<sup>21</sup> Bei der Entschlüsselung dieser Botschaft müssen wir allerdings immer zuerst durch die darüber gelagerte Ebene des ikonografischen oder konnotativen Code hindurch, der sich uns gleichsam aufdrängt. So neigen wir im Common Sense beispielsweise immer dazu, nicht-abstrakte Bilder zunächst in der Weise zu interpretieren, dass wir Handlungen und Geschichten gedanklich entwerfen, die sich auf dem Bild abspielen könnten. Die Entschlüsselung der Botschaft des Bildes „entledigt“ sich aber, wie es bei Roland Barthes (1990: 37) heißt, der Konnotationen und ist somit „eine Restbotschaft, die aus dem besteht, was vom Bild übrig bleibt, wenn man (geistig) die Konnotationszeichen ausgelöscht hat.“

Hier zeigen sich auch Parallelen zu Foucaults Bildinterpretation am Beispiel des Gemäldes „Las Meninas“ von Velásquez (s. Abb.1). Foucault (1971: 38) betont dort: „Man muß also so tun, als wisse man nicht.“ Dabei geht es im Sinne von Foucault nicht so sehr darum, das *institutionalisierte* Wissen auszuklammern (also das Wissen darum, dass es sich hier im Rahmen der Institution des Hofes um Hofedamen und -fräulein, um Höflinge und Zwerge handelt). Vielmehr geht es darum, dass wir die „Eigennamen auslöschen“ müssen (wie er sagt), also das Wissen um die *je fall- oder auch milieuspezifische Besonderheit* des Dargestellten und seiner konkreten Geschichte, wenn man „die Beziehung der Sprache und des Sichtbaren offenhalten will, wenn man nicht gegen, sondern ausgehend von ihrer Unvereinbarkeit sprechen will“ (Foucault 1971: 38).

spruch nimmt. Die dokumentarische Methode beansprucht lediglich einen *anderen* Zugang, einen Wechsel der AnalyseEinstellung. Demgegenüber schreibt Fuhs (2007: 210): „es wird auch eine Hierarchie des Verstehens und Interpretierens postuliert.“

20 Siehe zu den Korrespondenzen von denotativer und vor-ikonografischer Ebene auch van Leeuwen 2001.

21 Übereinstimmungen zeigen sich in dieser Hinsicht in einigen Punkten auch zu der von Erving Goffman (1979: 24) vorgelegten Fotointerpretation. Denn auch Goffman setzt ganz wesentlich auf der vor-ikonografischen Sinnebene an. Zentraler Gegenstand seiner Bildinterpretation sind die Gebärden, die er – unterhalb der Ebene einer Interpretation von *Handlungen* – als „*small behaviors*“ bezeichnet.



Abbildung 1: Diego Velázquez, Las Meninas, 1656. Madrid, Museo del Prado.  
Entnommen aus: Greub 2001: 295



### 3.4 Die Differenzierung des ikonografischen Vor-Wissens: kommunikatives und konjunktives Wissen

Foucault betont also, dass nicht alle sprachlichen Vorbegrifflichkeiten, nicht alle Namen, „ausgelöscht“ werden sollten, sondern lediglich die „Eigennamen“. Am Beispiel eines Familienfotos bedeutet dieses Auslöschen von Eigennamen, dass wir zwar aufgrund gesicherter Informationen oder aufgrund von Vermutungen davon ausgehen können oder müssen, dass es sich bei den abgebildeten Personen um eine

*Familie* handelt und wir somit unser entsprechendes Wissen um die Institution Familie aktualisieren. Sofern wir aber auch wissen oder vermuten, dass es sich um die Familie *Meier* handelt, sollten wir das, was wir über diese Familie mit ihrer konkreten Familienbiografie wissen, weitest möglich suspendieren. Diese beiden Arten des Wissens lassen sich im Rahmen der dokumentarischen Methode als *kommunikatives* Wissen einerseits und *konjunktives* Wissen andererseits bezeichnen.

Beim kommunikativen Wissen handelt es sich um generalisierte und weitgehend stereotypisierte, genauer: um institutionalisierte Wissensbestände. Institutionen finden wir im Sinne von Berger/Luckmann (1969: 58) dort, wo: „habitualisierte Handlungen durch Typen von Handelnden reziprok typisiert werden.“ Dieses Wissen betrifft das Rollengefüge der Gesellschaft, bspw. die Institution Familie. Davon zu unterscheiden ist das *konjunktive* Wissen als eines, welches sich hinter den „Eigennamen“ verbirgt, ein Wissen um die Familie „Meier“ in ihrer je individuellen, fallspezifischen Besonderheit einerseits und in ihrem milieutypischen Charakter andererseits.

Auch dort, wo wir über ein valides familienbiografisches Vorwissen in sprachlich-textlicher Form verfügen (etwa auf der Basis von Interviews oder Gesprächsanalysen), sollten wir dieses im Zuge der Bildinterpretation suspendieren.

Es zeichnen sich also gewisse Übereinstimmungen zwischen prominenten Ansätzen und Traditionen in der Bildinterpretation ab. Diese gehen dahin, dass spezifische Sinngelänge auf der konnotativen oder ikonografischen Ebene, welche in besonderer Weise durch (sprachliche) Narrationen, also durch unser textförmiges Wissen, geprägt sind, gleichsam suspendiert werden müssen, um auf diese Weise das (Spannungs-) Verhältnis von Bild und Sprache bzw. Text „offen halten“ zu können, wie es bei Foucault (1971: 38) heißt,<sup>22</sup> und somit das Bild der sprachlich-textlichen Logik nicht von vornherein unterzuordnen.<sup>23</sup>

Dem ist bisher im Bereich der qualitativen Methoden nicht Rechnung getragen worden.

Im Bereich der Semiotik hat Roland Barthes einige beispielhafte Interpretationen vorgelegt, die der skizzierten Suspendierung folgen, die also einsetzen, nachdem „man (geistig) die Konnotationszeichen ausgelöscht hat“ (Barthes 1990: 37). Er hat jene Sinnebene, die das Resultat dieser Interpretationen darstellt, als „*stumpfen Sinn*“ („sens obtue“) bezeichnet. Die Signifikanz dieser Sinnebene lässt sich in sprachlich-textlicher Form allenfalls in Gegensätzlichkeiten und Ambiguitäten fassen. So zeigt Barthes (1990:53f.) an Fotos aus dem Film „Der Panzerkreuzer Potemkin“ von Eisenstein, dass beispielsweise die Mimik einer weinenden alten Frau weder im schlichten Sinne eine ‚tragische Mimik‘ ist noch in eine ‚Komik‘ umkippt.

<sup>22</sup> s. dazu auch Renggli 2007: [10].

<sup>23</sup> In ähnlicher Weise argumentiert Mollenhauer (1997: 254), indem er Kant dahingehend referiert, „daß Kunstobjekte sich erst dann dem an Erkenntnis interessierten Urteil erschließen, wenn solche ‚bestimmenden Verstandesurteile‘ suspendiert würden, um die vielleicht passenden Begriffe erst zu suchen“. Wobei Mollenhauer diese Suchbewegung den „phänomenologischen Verfahren“ und die Operation mit „Verstandesbegriffen“ den „ikonografischen Verfahren“ zuordnet.

### 3.5 Die essentielle Ambiguität des Bildes: die Sinnkomplexität des Übergegensätzlichen

Roland Barthes unterscheidet den „stumpfen Sinn“ von dem wesentlich weniger komplexen „entgegenkommenden Sinn“, welcher – ähnlich dem ikonografischen Sinngehalt – an stereotypisierenden Sinnzuschreibungen haftet. Den stumpfen Sinn beschreibt Barthes (1990: 54) in Kategorien der Widersprüchlichkeit, als „eine Schichtung von Sinn, die den vorhergehenden Sinn immer bestehen läßt, wie in einer geologischen Konstruktion; das Gegenteil sagen, ohne auf das Widersprochene zu verzichten“. Ein evidentes Merkmal des stumpfen Sinnes ist also dessen Widersprüchlichkeit. In ähnlicher Weise charakterisiert Eco (1994: 145ff.) jene Botschaft, welche die Ebene der Konnotationen überschreitet und die er als „ästhetische Botschaft“ bezeichnet, durch ihre „produktive Ambiguität“ (a.a.O.: 146). Aufgrund dieser Ambiguität „beginne ich zu beobachten, um zu sehen, wie sie gemacht ist.“

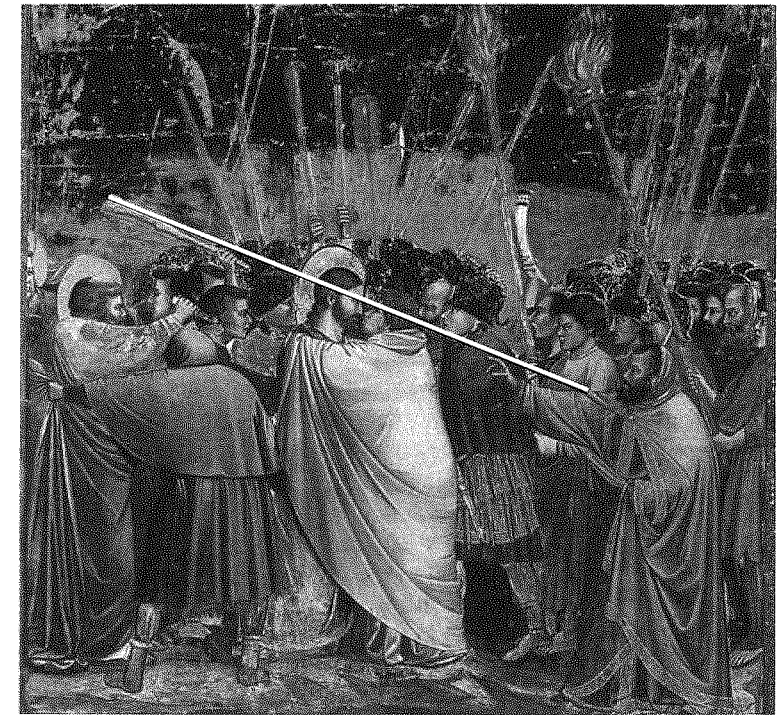
Es stellt sich hier also – analog zur ikonologischen Sinnebene bei Panofsky – im Zuge einer tiefer gehenden semantischen Analyse die Frage nach der *Herstellung* dieser Botschaft, nach dem *modus operandi* des Herstellungsprozesses, die Frage nach dem *Wie*. Vielleicht kann man auch sagen: Während die konnotative Botschaft oder (im Sinne von Barthes) der entgegenkommende Sinn stereotypisiert, bricht die ästhetische Botschaft bzw. der stumpfe Sinn diese Stereotypisierungen durch Ambiguitäten, Widersprüchlichkeiten oder Gegensätzlichkeiten auf.

In ähnlicher Weise sieht Imdahl (1996a: 107) das Spezifische des ikonischen Sinnes in einer (wie er es nennt) „Sinnkomplexität des Übergegensätzlichen“. Imdahl erläutert (1994: 312) am Beispiel des Fresko „Die Gefangennahme“ von Giotto (*siehe Abbildung 2*), dass „vermöge besonderer Bildkomposition Jesus sowohl als der Unterlegene wie auch als der Überlegene erscheint“. Diese Sinnkomplexität basiert ganz wesentlich auf der planimetrischen Komposition, der Komposition des Bildes in der Fläche (genauer dazu weiter unten). In diesem Fall ist es die Schräge, die nach Imdahl die Komposition des Bildes entscheidend bestimmt. Diese Sinnkomplexität des Übergegensätzlichen ist sprachlich nur schwer fassbar, und ihre sprachliche, intersubjektiv verständliche Vermittlung gelingt nur im direkten Verweis auf das Bild.<sup>24</sup>

In einem Forschungsprojekt zum familialen Habitus haben wir Familienfotos sowie Tischgespräche und Gruppendiskussionen ausgewertet (dazu: Kap. 4.3). Dabei zeigte sich, dass wesentliche Elemente des Familienmilieus, des familialen Erfahrungsraums, ihren unmittelbaren Ausdruck in adäquater Weise lediglich in *Übergegensätzlichkeiten*, in *Ambiguitäten* finden konnten. Und zwar beispielsweise in der Relation bzw. im Spannungsverhältnis zwischen dem Charakter des *Provisorischen und Ungesicherten* und der (*sozialen*) *Isolation* der abgebildeten Gruppe *einerseits* und der *Rigidität und Strenge* (in der planimetrischen Gesamtkomposition, aber auch in Mimik, Gestik und Körperhaltung) *andererseits*, welche die zentrale Stimmung des Bildes ausmacht.

24 Im Bereich der erziehungswissenschaftlichen Ikonologie stellt die von Dieter Lenzen (1993) vorgelegte Interpretation zweier Säuglingsbilder von Otto Dix ein treffendes Beispiel für eine derartige ‚Sinnkomplexität des Übergegensätzlichen‘ dar. Lenzen (1993: 62ff.) verweist auf die ‚Koinzidenz der Gegensätze‘ der Darstellung des Neugeborenen zwischen liebenswertem Wesen und komisch-hässlicher Kreatur.

Abbildung 2: Giotto, Gefangennahme, um 1305. Padua, Arena-Kapelle.  
Entnommen aus: Imdahl 1996a: Abbildungsverzeichnis S.45



Der Habitus dieser Familie ließe sich somit formulieren als derjenige *der Rigidität und Strenge im Kontext des Provisorisch-Ungesicherten*. Der Zugang zu dieser Übergegensätzlichkeit hat sich uns erst über das Bild erschlossen. Die eigengesetzliche Sinnstruktur des Bildes eröffnet uns einen systematischen Zugang zur Eigengesetzlichkeit des Erfahrungsraums der Bildproduzent(inn)en, hier also zu demjenigen der Familie.

Ein anderes Beispiel für eine derartige Sinnkomplexität des Übergegensätzlichen finden wir bei Werbefotos. So zeigt die genaue Interpretation eines Werbefotos der (Bekleidungs-) Firma Burberry (dazu: Kap. 4.2), dass der Burberry-Style in einer Weise präsentiert wird, die nahe legt, dass es mit ihm gelingt, zugleich Zugehörigkeit und Zusammengehörigkeit und somit Gemeinschaftlichkeit zu erfahren und dennoch die Individualität zu wahren. Dabei wird in diesem Foto aber auch eine andere – für das Medium Werbefoto generell gültige – Übergegensätzlichkeit deutlich: Da die Werbung nur die Pose zur Verfügung hat (vgl. dazu Bohnsack 2007b; Imdahl 1996d, steht sie vor dem (paradoxen) Problem, Individualität auf dem Wege des Posierens, also durch Stereotype darzustellen.

In diesem Werbefoto geschieht dies – bezogen auf die bildinternen sozialen Relationen – in der Widersprüchlichkeit von (körperlicher) Nähe einerseits und Abwendung des Blicks andererseits. Genauer bzw. von der formalen Komposition her betrachtet, handelt es sich darum, dass durch die planimetrische Komposition (also

durch die für die ikonische Interpretation wesentlichste Dimension formaler Komposition) eine Zugehörigkeit der Personen untereinander hergestellt wird, die dann durch das Fehlen bzw. die Verweigerung des Blickkontakts wieder negiert oder eingeschränkt wird.

In der genauen Analyse eines Fotos der Zigarettenwerbung der Firma „West“ (siehe Bohnsack 2007b) zeigt sich, dass in der auf dem Foto abgebildeten Person der Sennerin („Heidi“) eine ausgesprochen traditionelle Rolle und Lebenswelt sich mit einem unkonventionellen Habitus in einer Sinnkomplexität nach Art einer Hybridisierung von Stilelementen verbindet.

Die ikonologische oder ikonische Interpretation hat es also u.a. zu leisten, diese Sinnkomplexität des Übergegensätzlichen auf den Begriff zu bringen. Während es bei Imdahl keineswegs aussichtslos erscheint, diese zu verbalisieren, beharrt Roland Barthes darauf, dass man den stumpfen Sinn „theoretisch situieren, aber nicht beschreiben kann“ (1990: 63). „Der stumpfe Sinn ist nicht in der Sprache“ (a.a.O.: 58). Die Semiologie von Barthes wie die Semiotik von Eco tun sich gleichermaßen schwer, die Transzendenz der konnotativen Botschaft in systematischer Weise mit den ihnen zur Verfügung stehenden Kategorien zu erfassen.

Eines der Probleme besteht sicherlich darin, dass der stumpfe Sinn (bei Barthes) bzw. die ästhetische Botschaft (bei Eco) nicht – wie der entgegenkommende oder konnotative Sinngehalt – an *einzelnen* Bildgegenständlichkeiten festgemacht werden kann, wie Roland Barthes dies versucht (was insbesondere am Beispiel des „punctum“ evident wird; vgl. Barthes 1985). Vielmehr erschließt die tiefer gehende Semantik sich erst auf dem Weg über die Rekonstruktion der *Gesamt*-Komposition, die – wie ich weiter unten darlegen werde – wiederum an die Rekonstruktion der *Formalkomposition* des Bildes gebunden ist.

### 3.6 Zur Rekonstruktion der Formalstruktur des Bildes

Mit Max Imdahl können wir drei Dimensionen der Formalstruktur, des formalen kompositionalen Aufbaus des Bildes unterscheiden: die „*planimetrische Ganzheitsstruktur*“, die „*szenische Choreographie*“ und die „*perspektivische Projektion*“.

Durch die *perspektivische Projektion*, also durch die je unterschiedliche Herstellung von Perspektivität werden in der bildspezifischen Art und Weise Räumlichkeit und Körperlichkeit konstruiert. Es wird damit eine Gesetzmäßigkeit in das Bild hineingetragen, die im wahrsten Sinne Einblicke in die Perspektive der abbildenden Bildproduzent(inn)en und ihre Weltanschauung, ihren Habitus eröffnet. So hat Panofsky (1964a u. 2001a) den für die jeweilige Epoche typischen Modus von Perspektivität – insbesondere die Einführung zunächst der Achsen- und dann der Zentralperspektive in der Früh- und Hochrenaissance – als *das* zentrale Dokument für die Unterschiede zwischen der mittelalterlichen Weltanschauung und derjenigen der Renaissance herausgearbeitet (ausführlicher dazu siehe Anhang: 7.4 sowie Edgerton 2002).

Indem die Perspektivität dazu dient, Gegenstände in ihrer Räumlichkeit und Körperlichkeit identifizierbar zu machen, ist sie an den Gesetzmäßigkeiten der im Bild dargestellten Außen- oder Umwelt des Bildes orientiert. Im Rahmen empiri-

scher sozial- und erziehungswissenschaftlicher Forschung ist die Rekonstruktion der Perspektivität dort sinnvoll, wo nicht (nur) das Produkt, sondern (auch) der Bildproduzent (der Maler oder Fotograf) Gegenstand der Untersuchung ist.<sup>25</sup>

Im Bereich der sozialwissenschaftlichen *Foto*-Interpretation haben wir es grundsätzlich mit der Zentralperspektive zu tun, und auf dieser Grundlage gewinnt die Rekonstruktion der Perspektivität ihre Bedeutung vor allem durch die Beantwortung der Frage, welche Personen und sozialen Szenarien durch den abbildenden Bildproduzenten, durch das ‚Kameraauge‘ in Form des Fluchtpunktes und der Horizontlinie fokussiert und somit ins Zentrum des (sozialen) Geschehens gerückt werden und welche Ausprägung, welcher *Modus* der Zentralperspektive (bspw.: Frontal-, Übereck-, Froschperspektive) gewählt wird (dazu auch im Anhang Kap. 7.4).

Die *szenische Choreografie* des Bildes ist bei Imdahl (1996a: 19) definiert als „die szenische Konstellation der in bestimmter Weise handelnden oder sich verhaltenden Figuren in ihrem Verhältnis zueinander“, also als deren soziale Bezogenheit. Dies betrifft die Bewegungen der abgebildeten Bildproduzent(inn)en (dazu auch: Kap. 5.4), ihre räumliche Positionierung zueinander ebenso wie den Bezug ihrer Gebärden, aber auch Blicke, aufeinander. Die sozialwissenschaftliche Bedeutung dieser Ebene der Formalstruktur des Bildes ist offensichtlich.

Allerdings stellt eine begriffliche Charakterisierung der körperlichen Positionierung der abgebildeten Bildproduzent(inn)en und ihrer Blicke sowie Gebärden auf einer Ebene unterhalb von Ikonografie und Ikonologie, also auf der vor-ikonografischen Ebene, hohe Anforderungen an eine Beschreibungssprache (siehe dazu auch Kap. 5.4.2). Die Grundbegrifflichkeit vorliegender sozialwissenschaftlicher Handlungstheorien ist diesen Anforderungen kaum gewachsen. Zugleich wird auch deutlich, dass die Rekonstruktion der szenischen Choreografie im Sinne von Imdahl hinsichtlich ihrer Möglichkeiten der Formalisierung hinter die beiden anderen Dimensionen – Perspektivität und auch Planimetrie – weit zurücktritt. Sie lässt sich weniger als die anderen Dimensionen geometrisch und mathematisierbar formalisieren.

Die Rekonstruktion der *planimetrischen Komposition* bezieht sich auf die formale Konstruktion des Bildes in der Fläche. Da das Bild selbst eine Fläche ist, ist es die planimetrische Position, welches das Bild in der ihm eigentümlichen Gesetzmäßigkeit bestimmt.

Die Rekonstruktion der planimetrischen Komposition oder „planimetrischen Ganzheitsstruktur“ führt uns das Bild als „ein nach immanenten Gesetzen konstruiertes und in seiner Eigengesetzlichkeit evidentes System“ (Imdahl 1979: 190) vor Augen. Diese Dimension der Bildkomposition ist von besonderer Bedeutung für die ikonische Interpretation, da sie die entscheidende Grundlage für das *sehende Sehen* darstellt, welches Imdahl, wie dargelegt, vom *wiedererkennenden* (d.h. Gegenstände identifizierenden) Sehen unterscheidet.

Die beiden anderen Dimensionen, die „perspektivische Projektion und die szenische Choreographie erfordern ein wiedererkennendes, auf die gegenständliche Au-

25 Im Fall laienhafter Malerei (z.B. Kinderzeichnungen) ist aufschlussreich, a) ob und welcher *Modus* von Perspektivität (z.B. Achsenperspektivität) und b) welcher Fluchtpunkt, welche Perspektive gewählt wird. So hat Mollenhauer (1996, 126ff.) nach ähnlichen Kriterien Stile kindlicher Malerei in empirischer Analyse ausdifferenziert und Analogien zu kunstgeschichtlich ausdifferenzierten epochalen (Mal-) Stilen herausgearbeitet.

Benwelt bezogenes Sehen (...). Dagegen geht die planimetrische Komposition, insofern sie bildbezogen ist, nicht von der vorgegebenen Außenwelt, sondern vom Bildfeld aus, welches sie selbst setzt“ (Imdahl 1996a: 26). Während also die perspektivische Projektion sich auf die Körperlichkeiten und Räumlichkeiten der Außenwelt oder *Umwelt* des Bildes bezieht, dort ihren Maßstab hat und in deren Wiedererkennen fundiert ist und für die szenische Choreographie das Gleiche in Bezug auf soziale Beziehungen und Konstellationen der Umwelt gilt, schafft die planimetrische Komposition ihre eigenen Gesetzmäßigkeiten, ihre eigene formale Ganzheitsstruktur im Sinne einer Totalität, eines selbst-referentiellen Systems sozusagen: „Das Ganze ist von vornherein in Totalpräsenz gegeben und als das sinnfällige Bezugssystem in jedem einzelnen kopräsent, wann immer jedes einzelne in den Blick genommen wird“ (a.a.O: 23).

Dies hat zweierlei Konsequenzen hinsichtlich der Bedeutung der Rekonstruktion der planimetrischen Komposition. Diese ist somit wesentliche Grundlage für eine dem Medium des Bildes bzw. des Bildhaften überhaupt angemessene Interpretation, da diese nicht durch die Interpretationen anderer Darstellungsgattungen (vor allem ein sprachlich-textliches Vorwissen um die abgebildeten Gegenständlichkeiten) vorgeprägt ist. Zum anderen lenkt die Rekonstruktion der planimetrischen Komposition in ihrer systemischen Eigengesetzlichkeit die Analyseinstellung auf die Totalität des im Bild Dargestellten.

Der planimetrischen Kompositionsvariation kommt also für die *ikonische Interpretation* die entscheidende Bedeutung zu. Die Rekonstruktion der planimetrischen Komposition oder Koordination sollte somit im Zuge einer Bildinterpretation den ersten Schritt darstellen und unabhängig von der Ikonografie geleistet werden.

In der abstrakten Malerei ist – wenn wir Gehlen (1986, S. 64) folgen – beginnend mit dem Impressionismus eine zunehmende Hinwendung zur Planimetrie, zur „Verflächung des Darstellungsraumes und der dargestellten Gegenstände“ zu beobachten: „Die entscheidende Erfindung bestand in der *Zweischichtigkeit* des Bildes, nämlich in seiner Verselbständigung zu einer *Reizfläche eigenen Rechts* bei *dennoch festgehaltener* Gegenständlichkeit“ (S. 64). Die in der modernen Malerei zu beobachtende „Verflächung“ des Bildes kann man im Sinne von Imdahl als Hinwendung zum Primat der planimetrischen Komposition des Bildes verstehen und somit als eine zunehmende Hinwendung zum „sehenden Sehen“.

Allerdings ist – so möchte ich hier einwenden – nicht die Zweischichtigkeit des Bildes als solche als eine „Erfindung“ des Impressionismus und der modernen Malerei anzusehen, sondern die *Betonung des Primats* der Fläche und ihres kompositionalen Aufbaus gegenüber der Reproduktion von Gegenständlichkeit und Räumlichkeit. Auch in der Fotografie finden wir bereits diese Zweischichtigkeit. Wobei – aufgrund der vorgegebenen Zentralperspektive – die Orientierung an Gegenständlichkeit und Räumlichkeit nicht grundlegend zu hintergehen ist. Gleichwohl oder gerade deswegen ist die Ausdruckskraft und Ästhetik des Fotos im Sinne der dem Bild eigentümlichen systemischen Eigengesetzlichkeit von der Gestaltungsleistung (der abbildenden wie abgebildeten Bildproduzenten und -produzentinnen) im Bereich der planimetrischen Komposition abhängig.<sup>26</sup>

26 Analog zur Zweischichtigkeit des Bildes kann man von einer Zweischichtigkeit des Textes und der Sprache bzw. des Diskurses insofern sprechen, als auch der Diskurs in der einen Dimension –

Was die *Interpretation* des Bildes anbetrifft, so „zwingt“ uns die Rekonstruktion der formalen, insbesondere der planimetrischen, Komposition gewissermaßen dazu, die Einzelelemente des Bildes nicht isoliert, sondern grundsätzlich immer im Ensemble der anderen Elemente zu interpretieren, wohingegen wir im Common Sense dazu neigen, einzelne Elemente des Bildes herauszugreifen.

Wenn es uns – u.a. mit Hilfe der Formalstruktur und ggf. auch der Farbkomposition (vgl. Imdahl 2003) – gelingt, einen Zugang zum Bild als eigengesetzlichem oder selbst-referentiellen System zu erschließen, dann eröffnet sich uns auf diese Weise auch ein systematischer Zugang zur Eigengesetzlichkeit des Erfahrungsraums der Bildproduzent(inn)en, beispielsweise zum Erfahrungsraum einer Familie mit ihrem spezifischen familialen Habitus.

Die Vorschläge zur Rekonstruktion der Formalstruktur des Bildes von Imdahl bieten erste Hinweise für die Entfaltung formaler Analysen im Bereich qualitativer Bildinterpretation, die es nun aufzugreifen gilt. Ansätze dazu sind im Bereich der dokumentarischen Bildinterpretation forschungspraktisch entwickelt worden (s. Bohnsack 2003, 2005b, 2007a u. b, 2008c, 2009b; Nohl 2002; Nentwig-Gesemann 2006b u. 2007b; Dörner 2007; Schäffer 2009 sowie Selke 2005a und andere Beiträge in: Fetzner/Selke 2005: Teil I).

Auch im Bereich der Textinterpretation hat sich die qualitative Methodik auf die Rekonstruktion der formalen Strukturen gestützt, um auf diese Weise den von den Textproduzent(inn)en selbst *hergestellten* (Gesamt-) *Kontexten* Rechnung tragen zu können. So kommen wir z.B. der Semantik offener Interviews dann auf die Spur, wenn wir rekonstruieren, wie die Interviews in sich durch unterschiedliche „Textsorten“ strukturiert sind (s. Schütze 1987 u. Nohl 2009). Analog stellt im Bereich der Gesprächsanalyse und des Gruppendiskussionsverfahrens u.a. die genaue Charakterisierung der Art der interaktiven Bezugnahme der Beteiligten aufeinander im Sinne der Analyse der „Diskursorganisation“ (vgl. Bohnsack/Przyborski 2006) die Voraussetzung für tiefer greifende semantische Interpretationen dar.

Für die sozialwissenschaftliche Methodik der Bildinterpretation erscheint es in analoger Weise sinnvoll, sich der umfangreichen Vorarbeiten zur formalen Ästhetik im Bereich der Kunstgeschichte zu vergewissern. So hat Klaus Mollenhauer (1983: 179) betont, „daß bereits die *nur* formalästhetischen Charakteristika inhaltliche Hinweise enthalten. In linguistischer Metapher gesprochen: Die Bild-Syntax zeigt schon der Bildsemantik ihren Weg.“ Die von Mollenhauer (1983: 173) in „bildungstheoretischer Absicht“ vorgelegten Bildinterpretationen auf der Basis der Ikonologie von Panofsky und der Ikonik von Imdahl oder bspw. die von Parmentier (2001) in ähnlicher Absicht vorgelegten Bildinterpretationen lassen Zusammenhänge von Bildungsprozessen im Sinne der geisteswissenschaftlichen Tradition und der Vermittlung qualitativer Methoden deutlich werden.

der *propositionalen* – die Funktion hat, Erfahrungen mit Gegenständen, Handlungen und Regeln der *Umwelt* des Diskurses zu reproduzieren bzw. ‚abzubilden‘, und er in der anderen Dimension – der *performativen* – als ein sich selbst steuerndes Handlungs- und Regelsystem operiert. Der Zweischichtigkeit des Bildes entspricht also die Doppelstruktur des Diskurses, die Doppelstruktur von Propositionalität und Performativität. Und ebenso wie beim Diskurs wird es beim Bild darum gehen, die *Homologien* zwischen diesen beiden Dimensionen oder Schichten herauszuarbeiten.

Denn die in bildungstheoretischer Absicht vorgelegten Arbeiten, die auf eine „ästhetische Bildung durch Bilder“ oder ein „Sich bilden durch Bilder“ zielen (Schäffer 2005: 220f.), lassen sich zugleich zu einem Teil auch lesen als Anregungen für die Entwicklung von qualitativen Methoden der Bildinterpretation. Umgekehrt betrachtet ist die Systematisierung und Verfeinerung von intuitiven Kompetenzen der Rekonstruktion formaler Strukturen der Darstellung und der Interpretation nicht nur Grundlage für die Methoden-Entwicklung und die Ausbildung in Methoden. Vielmehr lässt sich die Vermittlung dieser intuitiven Kompetenzen (im Bereich der Fotointerpretation bis hin zur Gesprächsanalyse) auch als Beitrag zur ästhetischen Bildung verstehen – indem der Blick geschärft wird für die ästhetischen Grundlagen profaner, alltäglicher Darstellung und Kommunikation.

### 3.7 Sequenzanalyse versus Kompositionsvariation

Die Berücksichtigung der formalen Bildkomposition ist also entscheidende Voraussetzung, um einen Zugang zu jener Sinnebene zu gewinnen, die durch den Text und unser textförmiges Vor-Wissen, also das ikonografische Wissen, nicht zu vermitteln ist.<sup>27</sup> Ein textförmiges Wissen und somit auch ein textförmiger Zugang zum Bild ist vor allem durch seine *Sequenzialität*, durch eine am zeitlichen Nacheinander orientierte Ordnung, charakterisiert. Diese findet u.a. auch in der Narrativität sprachlicher Darstellungen ihren Ausdruck. Eine solche Bindung an Sequenzialität und Narrativität, an das (narrative) Nacheinander, kritisiert Imdahl (1996a: 137) an den konventionellen Interpretationsverfahren der Kunstgeschichte. Demgegenüber sei für das Bild die „*Simultanstruktur*“ konstitutiv, ein „kompositionsbedingtes, selbst sinnstiftendes Zugleich“, bei der das „Ganze (...) von vornherein in Totalpräsenz gegeben ist“ (Imdahl 1996a: 23).<sup>28</sup> Demzufolge ist ein Interpretationsverfahren, welches der Eigensinnigkeit der Bildhaftigkeit gerecht werden will, mit einer sequenziellen oder sequenzanalytischen Vorgehensweise unvereinbar, steht ihr diametral entgegen.

Indem (wie dargelegt) die erheblichen Erkenntnisfortschritte der qualitativen Methoden in den letzten zwanzig Jahren im Bereich der Textinterpretation gewonnen wurden, stellt für alle neueren qualitativen Methoden die *Sequenzanalyse* ein zentrales, bisweilen sogar *das* zentrale methodische Prinzip dar. Sobald wir allerdings – wie dies beispielsweise in der objektiven Hermeneutik versucht worden

27 Eine derartige Fundierung der ikonologischen Interpretation in der Formalstruktur findet sich auch bei Parmentier, dessen Arbeiten als kunsthistorische Analysen von erziehungswissenschaftlicher bzw. pädagogischer Relevanz zu verstehen sind. Parmentier stützt sich in seinen Interpretationen auf elaborierte Rekonstruktionen der Perspektivität (bspw. der „doppelten Perspektive“ in Bruegheles „Kinderspielen“; Parmentier 2001) und der Planimetrie in ihrem Zusammenspiel mit der Perspektivität (bspw. in Chardins ‚L'enfant au totot‘; Parmentier 1993).

28 Vgl. dazu auch die Unterscheidung von diskursiven und visuellen Formen bei Susanne Langer (1984, S. 99): Visuelle Formen „bieten ihre Bestandteile nicht nacheinander, sondern gleichzeitig dar, weshalb die Beziehungen, die eine visuelle Struktur bestimmen, in einem Akt des Sehens erfasst werden.“

ist<sup>29</sup> – dieses Prinzip direkt auf das Bild zu übertragen suchen, zielen wir an der Eigensinnigkeit des Bildes vorbei.

Erfolgversprechender erscheint es, prinzipieller anzusetzen und die Frage zu stellen, in welchem generelleren methodischen Prinzip die Sequenzanalyse ihrerseits fundiert ist, um dann von dieser prinzipielleren Ebene her Gemeinsamkeiten und Unterschiede der Bild- und Textinterpretation zu begründen. Dieses generellere Prinzip ist dasjenige der Operation mit Vergleichshorizonten: das Prinzip der *komparativen Analyse*. Die Sequenzanalyse stellt nur eine der möglichen Ausprägungen der erkenntnisgenerierenden Methodik einer Operation mit Vergleichshorizonten dar (vgl. auch: Bohnsack 2008d: 198ff.). Die Bedeutung der komparativen Analyse für das Feld der Textinterpretation zeigt sich beispielsweise darin, dass sich mir das, was den Sinngehalt eines spezifischen Diskurses ausmacht, dadurch erschließt, dass ich dagegenhalte, wie dasselbe (oder ein vergleichbares) Thema auch in anderer Weise, in einem anderen Diskurs hätte behandelt werden können oder (besser noch) bereits behandelt worden ist (siehe dazu Bohnsack 2001 u. Bohnsack/Nohl 2007). Diese Vergleichshorizonte, die ich im Zuge der Interpretation des Diskurses an ihn herantrage, können also imaginativer Art oder in empirischen Vergleichsfällen fundiert sein.

Auch im Medium der Bildinterpretation ist der Interpret als Beobachter in unterschiedlicher Weise und auf unterschiedlichen Ebenen auf Vergleichshorizonte angewiesen, die zunächst implizit bleiben. Dies gilt auch bereits auf der Ebene der Rekonstruktion der formalen Komposition eines Bildes. So vollzieht sich bereits deren Wahrnehmung vor dem Vergleichshorizont (intuitiv vollzogener) anderer, kontingenter Kompositionen. Imdahl hat die je spezifische Komposition eines Bildes – beispielsweise diejenige der mittelalterlichen Miniatur „Der Hauptmann von Kapernaum“ – in experimenteller Weise verändert und konnte auf diese Weise zeigen, dass der Sinn einer verbildlichten Szene direkt mit der formalen Komposition variiert (vgl. u.a. Imdahl 1994: 302ff.). Diese experimentelle Veränderung der Komposition wie auch das Heranziehen von empirischen Vergleichsfällen, die sich durch systematische Variationen der Komposition voneinander unterscheiden, habe ich als *Kompositionsvariation* bezeichnet (vgl. Bohnsack 2007a).

Auch an den (ikonologischen bzw. dokumentarischen) Interpretationen Panofskys (2001) lassen sich (unbeschadet der Einsicht, dass das Besondere der Ikonizität bei Panofsky in systematischer Weise unberücksichtigt bleibt) die Prinzipien der komparativen Analyse und der Kompositionsvariation im Rahmen ausgewählter Fallanalysen in komplexer Weise rekonstruieren. Beispielhaft sind hier seine Untersuchungen zur Entstehung innovativer Stilelemente im italienischen Trecento, wie sie etwa in dem um 1305 entstandenen „Abendmahl“ von Giotto in der Cappella dell’Arena ihren Ausdruck finden.

29 So geht Thomas Loer (1994), der sich selbst der Methodologie der objektiven Hermeneutik verpflichtet sieht, *zum einen* von der Unterstellung aus, dass der Malprozess, der Prozess der Herstellung des Bildes, einem sog. „ikonischen Pfad“ folgt, welcher sequenziell strukturiert ist. Und *zum anderen* muss dann unterstellt werden, dass dieser Pfad bzw. mehrere mögliche Pfade für den Bildinterpretieren sequenziell rekonstruierbar sind (siehe auch Loer 1996). Für ähnliche Versuche siehe auch Englisch 1991.

Indem er die formale Komposition dieses Werkes (hier insbesondere die bahnbrechenden Innovationen im Bereich der Räumlichkeit und Perspektivität betreffend) einer komparativen Analyse mit Werken der Vorläuferperiode (also des 13. Jahrhunderts) unterzieht, gelingt es ihm aufzuweisen, dass einzelne Elemente oder Merkmale dieser Komposition in diesen früheren Werken bereits entfaltet worden waren, diese Merkmale aber isoliert voneinander blieben, da sie in unabhängigen Strömungen dieser Vorläuferperiode anzutreffen waren: in der nordeuropäischen Hochgotik einerseits und der südeuropäischen byzantinischen Tradition andererseits. Erst im 14. Jahrhundert, im Trecento, gelangten wesentliche Elemente dieser beiden Strömungen zu einer Synthese, so dass „der ‚moderne‘ Raum in dem Moment geboren wurde, als der von Skulptur und Architektur genährte hochgotische Sinn für Volumen und Kohärenz auf jene illusionistische Tradition zu wirken begann, die in der byzantinischen und byzantinsierenden Malerei fortgelebt hatte – mit anderen Worten: im italienischen Trecento.“ (Panofsky 2001: 24).

Ein weiteres eindrucksvolles Beispiel einer komparativen Analyse findet sich dort, wo sich Panofsky (2001: 12) der europäischen Malerei des 15. Jahrhunderts auf dem Wege des Vergleichs zweier ihrer zentralen „beinahe gleichzeitigen und relativ vergleichbaren“ Werke der italienischen Malerei einerseits und der niederländischen Malerei andererseits nähert: Jan van Eycks Doppelbildnis des Giovanni Arnolfini und seiner Frau von 1434 und Masolino da Panicone „Tod des Heiligen Ambrosius“ um 1430. Indem er die Unterschiede und zugleich Gemeinsamkeiten der italienischen und flämischen Malerei an diesen beiden Fällen herausarbeitet und seine Interpretation somit einem Prinzip folgt, welches ich (ursprünglich: Bohnsack 1989: Kap. 4.5 sowie u.a. Bohnsack 2009a) als dasjenige des „Kontrasts in der Gemeinsamkeit“ bezeichnet habe, gelingt es ihm, allein an diesen beiden ausgewählten Fällen den Stil der Malerei des Quattrocento und die für ihn konstitutiven Polarisierungen zu rekonstruieren und zu belegen.

Während Panofsky von den drei Dimensionen der formalen Komposition des Bildes – Perspektivität, szenische Choreographie und planimetrische Komposition – sich in seinen komparativen Analysen bzw. Kompositionsvariationen vor allem auf die formale Konstruktion von Perspektivität und Räumlichkeit stützt, stellt Max Imdahl, dem wir aus den dargelegten Gründen in der dokumentarischen Bildinterpretation folgen, die planimetrische Komposition, also die rein flächenhafte Komposition, ins Zentrum der Kompositionsvariation.

Die Methodisierung der Bildinterpretation bedarf also der Rekonstruktion der Formalstruktur des Bildes und – wie dargelegt – der Suspendierung des *ikonografischen* bzw. *konnotativen* Sinngehalts, also des sprach- und textförmigen Vorwissens. Im Bereich sozialwissenschaftlicher Bildinterpretationen erscheint dies besonders notwendig, weil hier die Ikonografien, also die kommunikativen Wissensbestände, nicht in kodifizierter Form vorliegen – wie wir dies in der Kunstgeschichte beispielsweise in Form von Bibeltexten finden.

### 3.8 Zum Problem der Polysemie: die Vieldeutigkeit des Bildes

Genauer betrachtet, ist es, wie dargelegt, nur die eine Dimension innerhalb der ikonografischen Sinnebene, die der Suspendierung bedarf, nämlich diejenige des *konjunktiven* Wissens im Unterschied zum *kommunikativen* Wissen. Bei letzterem handelt es sich um gesamtgesellschaftlich geteiltes, um institutionalisiertes Wissen (bspw.: ‚auf dem Bild ist eine Familie – Mutter, Vater, drei Kinder – abgebildet‘). Die Interpretationen oder Sinnzuschreibungen sind hier also relativ eindeutig, eben nicht vieldeutig oder „polysem“, wie es bei Roland Barthes (1990: 34) heißt.

Demgegenüber ist auf der Ebene des *konjunktiven* Wissens, also im Hinblick auf die je besondere Geschichte, welche ‚das Bild erzählt‘, eine nahezu unendliche Polysemie möglich, wenn der Interpret sein eigenes standortgebundenes Wissen in naiver Weise an das Bild heranträgt.<sup>30</sup> Das Problem der Polysemie ist somit eine Variante jener Problematik des Verstehens und Interpretierens, auf die ich an anderer Stelle mit dem Problem der Standort- oder Seinsver- bzw. -gebundenheit der Interpretation eingegangen bin (vgl. u.a. Bohnsack 2008a: Kap. 10 sowie 2009a).<sup>31</sup>

Solange der Interpret diese Standortgebundenheit nicht reflektiert, nimmt er das Bild in den eigenen (milieu-, generations-, geschlechtsspezifischen) Erfahrungsraum hinein. Wir sprechen hier mit einem Begriff von Matthes (1992) auch von „Nostrifizierung“. Diese interpretative Beliebigkeit könnte durch den Rückgriff auf textgebundenes Kontextwissen bewältigt werden. Dann aber würden wir uns den Zugang zu den aus der Eigenlogik des Bildes zu erschließenden Sinnstrukturen durch das textliche Vorwissen verstellen. Hierin liegt einer der Gründe dafür, das konjunktive Vor-Wissen weitest möglich zu suspendieren.

Die Differenzierung des ikonografischen Vor-Wissens in ein *kommunikatives* und ein *konjunktives* Wissen eröffnet also neue Perspektiven auf das Problem der Polysemie und dessen Bewältigung. Weiterhin ist nun aber auch die Polysemie auf der Ebene des ikonografischen (resp. konjunktiven) Vor-Wissens von derjenigen auf der *ikonologischen* Ebene zu differenzieren (die Roland Barthes allerdings nicht systematisch herausgearbeitet hat, allenfalls ansatzweise mit dem „stumpfen Sinn“). Der Unterschied zwischen der ikonografischen resp. konnotativen und der ikonolo-

30 Das bestätigt sich in den Analysen und auch methodischen Ausführungen derjenigen, die daran festhalten, Bildanalyse als Narrationsanalyse zu betreiben. So heißt es bei Marotzki/Stoetzer (2007), deren Forschungsprogramm auf „die Geschichten hinter den Bildern“ (a.a.O., S. 15) zielt: „dass das Bild selbst nur in wenigen Fällen eine eindeutige Sinnstruktur entfaltet. Es kommt gerade darauf an, die Vielfältigkeit zur Geltung zu bringen“ (a.a.O., S. 23). Und Fuhs (2006) betont: „Das Foto ist also in einer Botschaft mehrdeutig und wird in verschiedenen historischen Kontexten unterschiedlich ‚gelesen‘ und verstanden.“ Wobei dies im Widerspruch steht zu dem, was Fuhs ebenso evident erscheint: „Die Fotografie (...) formt sich zu einer Bild-Sprache, die eine weltweite Kommunikation über staatliche, sprachliche und kulturelle Grenzen hinweg ermöglicht.“

31 Im Bereich der Kunstgeschichte wird diese Problematik – wenn wir Gottfried Boehm (1985: 124) folgen – traditionellerweise kaum in angemessener Weise methodologisch reflektiert. Es dominieren hier objektivistische Zugänge: „Das Pathos historischer Objektivität lebt von der – nicht selten undurchschauenden – Überzeugung, daß sich zwar alles unter der Sonne historisch betrachten lasse, daß alles einen historischen Index trägt, außer dem einen: der wissenschaftlichen Einstellung des Forschers selbst. Nur sie vermag auch jenen objektiven, der Zeit enthobenen Anspruch auf Richtigkeit und Wahrheit zu garantieren.“

gischen Ebene besteht, wie dargelegt, in einem fundamentalen Wechsel der Analyse-einstellung.

Auf der ikonografischen Ebene frage ich danach, *was* dargestellt wird im Sinne der Thematik, des Sujet des Bildes. Auf der ikonologischen Ebene frage ich nach dem *Wie*, nach dem *modus operandi* der Herstellung der Darstellung. Hier kann ich auf ein Kontextwissen vollständig verzichten. Allerdings ist die Beobachtung des *modus operandi* abhängig von meinen *Vergleichshorizonten*: Wenn ich den auf einem Familienfoto sich dokumentierenden Habitus als durch „Starrheit“ und „Rigidität“ gekennzeichnet identifiziere (vgl. das Beispiel in 4.3.1), rekurriere ich zunächst auf Vergleichshorizonte, welche sich auf der Grundlage meiner mentalen Bilder konstituieren. Diese sind das Produkt meiner je individuellen oder kollektiven Sozialisationsgeschichte und Element meines je spezifischen Erfahrungsraums, der sich in unterschiedliche kollektive oder konjunktive (u.a. milieu-, geschlechts- und generationsspezifische) Erfahrungsräume differenzieren lässt.

Jene derart fundierten Vergleichshorizonte, welche ich in der alltäglichen Interpretation intuitiv an das Bild herantrage, lassen sich allerdings methodisch-empirisch kontrollieren. Dies gelingt insoweit, als an die Stelle imaginativer Vergleichshorizonte solche treten, die empirisch fundiert sind, uns also in Form jener Art von Bildern vorliegen, die in die empirische Analyse einbezogen werden, die selbst Gegenstand empirischer Analyse sind. Wir sprechen in diesem Fall von *komparativer Analyse* (s. u.a. Bohnsack 1989 u. 2008d: 198ff.; Nohl 2001; als Forschungsbeispiel: Kap. 4.3).

Auf dem Wege der komparativen Analyse wird die Mehrdeutigkeit nicht eliminiert, sondern methodisch kontrolliert. Auf diese Weise wird der Blick auf die Mehrdimensionalität des Bildes eröffnet. Je nach *Vergleichshorizont* geraten unterschiedliche – einander aber nicht ausschließende – Dimensionen oder Erfahrungsräume des Bildes in den Blick. In der dokumentarischen Methode wird dem (auch im Bereich des Textes) mit der *Mehrdimensionalität* der Sinn- und Typenbildung Rechnung getragen (vgl. Bohnsack 2007c u. 2009a).

Um es an einem Beispiel zu erläutern: Wenn wir das Foto einer Kommunionfeier Anfang der 80er Jahre in der DDR (vgl. Kap. 4.3) vor dem Vergleichshorizont eines Fotos einer Kommunionfeier aus der BRD der gleichen zeitgeschichtlichen Phase interpretieren, geraten uns *kulturtypische* Merkmale in den Blick, also Unterschiede des kulturtypischen Habitus der DDR zu demjenigen der BRD. Interpretieren wir dieses Fotos einer Kommunionfeier vor dem Vergleichshorizont eines anderen Kommunion-Fotos oder eines Fotos einer Jugendweihe, welches ebenfalls aus der DDR und aus derselben zeitgeschichtlichen Phase stammt (vgl. Kap. 4.3), so geraten *Milieudifferenzen* in den Blick: bspw. kleinbürgerlicher vs. proletarischer Habitus. Vergleichen wir innerhalb desselben Milieus und innerhalb derselben Kultur Fotos ähnlicher Thematik aus unterschiedlichen zeitgeschichtlichen Phasen, geraten uns Differenzen generationsspezifischer Habitus in den Blick.

Die methodisch kontrollierte Bewältigung des Problems der Polysemie, des Problems der (beliebigen) Vieldeutigkeit des Bildes, ist also an die Operation mit empirisch fundierten Vergleichshorizonten, an die *komparative Analyse*, gebunden. Im Bereich der Bildinterpretation sind wir – ebenso wie auch in demjenigen der Textinterpretation – nicht erst auf der Stufe einer mehrdimensionalen Typenbildung

auf Vergleichshorizonte angewiesen. Vielmehr erschließt sich uns, wie Imdahl (1994) gezeigt hat, bereits die formale Komposition eines Bildes in ihrer Besonderheit erst vor dem Vergleichshorizont anderer, kontingenter Kompositionsmöglichkeiten. Im Unterschied zur *Sequenzanalyse* im Bereich der Textinterpretation gelingt der Zugang zur Eigenlogik des Bildes allerdings nur dann, wenn wir, wie Imdahl (1996a: 23) betont, dessen „*Simultanstruktur*“ auf der Grundlage der formalen Komposition erfassen, als ein „kompositionsbedingtes, selbst sinnstiftendes Zugleich“, bei der das „Ganze (...) von vornherein in Totalpräsenz gegeben ist.“ Sobald wir das Prinzip der Sequenzanalyse direkt auf das Bild zu übertragen suchen, zielen wir an dessen Eigengesetzlichkeit im Sinne von Imdahl vorbei.

### 3.9 Simultaneität, Synchronizität und Sequenzialität

Um die Differenz zwischen einer Sequenzanalyse und einer Analyse auf der Basis der Simultaneität zu präzisieren, ist es zunächst notwendig, den Begriff der Simultaneität von demjenigen der Synchronizität zu unterscheiden. Bei letzterem handelt es sich (entsprechend dem griechischen Wortstamm) um das *gleichzeitige* Zusammentreffen nicht zusammenhängender Ereignisabläufen, also unterschiedlicher Sequenzen. Der Begriff der Synchronizität bezieht sich also ausschließlich auf die *zeitliche* Dimension. Der Begriff der *Simultaneität* umfasst (entsprechend dem lateinischen Wortstamm) demgegenüber nicht allein die zeitliche Dimension, also die „Gleichzeitigkeit“, sondern auch weitergehende „Gemeinsamkeiten“ (Duden 1982: 705). In einer spezifischeren Bedeutung meint Simultaneität laut Duden (ebenda): „die Darstellung von zeitlich od. räumlich auseinander liegenden Ereignissen auf einem Bild.“

Es ist eben genau diese Bedeutung, die bei Imdahl im Zentrum steht. Es soll deshalb der Begriff der Simultaneität als *Obergriff* zu dem – spezifischeren, weil rein zeitlichen – der Synchronizität verwendet werden. Die Simultaneität ist nach Imdahl (1994: 300) wesentliches Element der Eigenlogik des Bildes „als eine solche Vermittlung von Sinn, die durch nichts anderes zu ersetzen ist.“ In dem Potential zu einer (in diesem Sinne verstandenen) Simultaneität des Bildes ist dann auch ganz wesentlich die dem Bild eigentümliche „Sinnkomplexität des Übergegensätzlichen“ (Imdahl 1996a: 107) verankert.

#### *Simultaneität in der Fotografie*

Auf den ersten Blick scheint es so, als sei die Simultaneität in diesem Sinne eine Besonderheit der Malerei, auf die der Begriff bei Imdahl zunächst ja auch bezogen ist, und wäre somit auf die (nicht-montierte) Fotografie nicht zu übertragen. Denn nach Imdahl (1994: 308) ist es nicht nur so, dass das, was wir auf dem Gemälde oder der Zeichnung sehen, nicht „durch sprachliche Narration ersetzt werden kann.“ Vielmehr hat es auch nicht „in der Empirie eines Geschehens ein Vorkommen“. Mit anderen Worten: Es ist nicht schlicht die Abbildung des empirisch Beobachtbaren.

Würden wir die Fotografie als die ‚reine Abbildung‘ des empirisch Beobachtbaren verstehen, so träfen die von Imdahl geltend gemachten spezifischen Gestaltungs- und Kompositionsleistungen des Bildes, an denen er besonders die Simultanstruktur hervorhebt, auf sie nicht zu. Wir würden die Fotografie dann in der Sprache von Roland Barthes als ‚reine Denotation‘ begreifen, was dieser als ‚Mythos‘ bezeichnet, also als eine in den Selbstverständlichkeiten des Common Sense verwurzelte Unterstellung oder Vorstellung (Barthes 1990: 14). Weder im Sinne von Barthes noch im Sinne von Eco, auf deren unterschiedliche Argumentationen ich hier nicht eingehen kann, ist das Foto als reine Abbildung zu verstehen. Im Sinne von Eco (1994: 244) haben wir es auch auf der denotativen Ebene mit einem ‚Code‘ zu tun. Dieser beruht zwar nicht auf verbalem Vor-Wissen, aber auf ‚Konventionen‘ und zwar ‚viel mehr als in der verbalen Sprache‘.

Das Spiel mit derartigen Konventionen und die daraus resultierenden Gestaltungs- und Kompositionsmöglichkeiten, in denen sich die Habitus der abbildenden und abgebildeten Bildproduzent(inn)en dokumentieren, sind aber nicht allein konstitutiv für die Malerei, sondern auch bereits für die Fotografie – und auch dort, wo diese nicht montiert ist. Somit ist auch eine auf der Simultaneität basierende Sinnkomplexität des Bildes nicht allein für die Malerei, sondern auch für die Fotografie konstitutiv.

Auch Bourdieu können wir so interpretieren, dass derartige Gestaltungsmöglichkeiten der Fotografie – und auch der Amateurfotografie – offen stehen bzw. untrennbar mit ihr verbunden sind. Nach Bourdieu (1983a: 17) ‚wählt, jedenfalls innerhalb der theoretischen Unendlichkeit aller Photographien, die ihr technisch möglich sind, jede Gruppe praktisch ein endliches und bestimmtes Sortiment möglicher Gegenstände, Genres und Kompositionen aus (...) so daß noch die unbedeutendste Photographie neben den expliziten Intentionen ihres Produzenten das System der Schemata des Denkens, der Wahrnehmung und der Vorlieben zum Ausdruck bringt, die einer Gruppe gemeinsam sind.‘

Die hier getroffene ‚Wahl‘ ist also keineswegs auf die expliziten Intentionen der ProduzentInnen zu reduzieren – und zwar ebenso wenig wie dies im Bereich der Kunst der Fall ist.<sup>32</sup> Ebenso wie der ‚Habitus‘ (...) den Künstler mit der Kollektivität und seinem Zeitalter verbindet und, ohne daß dieser es merkte, seinen anscheinend noch so einzigartigen Projekten Richtung und Ziel weist“ (Bourdieu 1970: 132), gilt dies auch für die Amateurfotograf(inn)en – auch wenn deren Projekte nicht unbedingt als ‚einzigartig‘ erscheinen. „Diese objektive Intention, die sich niemals auf die bewusste Absicht des Künstlers beschränkt, ist eine Funktion der Denk-, Wahrnehmungs- und Handlungsmuster, die der Künstler seiner Zugehörigkeit einer bestimmten Gesellschaft oder Klasse verdankt“ (a.a.O: 154), also einem spezifischen Milieu oder konjunktiven Erfahrungsraum.

32 Panofsky erläutert das Verhältnis zur Intentionalität mit Bezug auf den von Alois Riegl stammenden Begriff des ‚Kunstwollens‘, welcher ‚die Summe oder Einheit der in demselben (dem Kunstwerk; R.B.) zum Ausdruck gelangenden, es formal wie inhaltlich von innen heraus organisierenden schöpferischen Kräfte bezeichnen sollte“ (Panofsky 1964b: 30). Er grenzt diese Bestimmung des Begriffes ‚Kunstwollen‘ von der ‚missverständlichen Deutung‘ des Begriffes ab, welche ‚die künstlerische Absicht, das künstlerische Wollen, als den psychologischen Akt des historisch greifbaren Subjektes ‚Künstler‘ betrachtet“ (a.a.O: 31).

Wenn wir (wie dargelegt) vom Foto als gemeinsamem Produkt der abbildenden und abgebildeten Bildproduzent(inn)en ausgehen, so ist der Schnappschuss nicht lediglich als Akt des Registrierens der vorgefundenen, d.h. durch die *abgebildeten* Bildproduzent(inn)en hergestellten, Simultaneität zu verstehen. Vielmehr wird aus der Fülle von Möglichkeiten gegebener Simultaneitäten durch die *abbildenden* Bildproduzent(inn)en – gemäß ihrer Schemata von Präferenzen, gemäß ihrem Habitus – eine spezifische ausgewählt. Anders formuliert: Die mit der Fotografie verbundene code- oder habitusspezifische Transformation oder Konstruktion umfasst – so ist weiterführend zu argumentieren – nicht allein die Selektivität der *Herstellung* des Arrangements oder der Komposition, sondern auch die Selektivität in der Situationsauswahl der Ablichtung *gegebener* Arrangements, vorgefundener Kompositionen.

Es geht also nicht allein darum, *wie* die Bildgegenständlichkeiten durch die abbildenden und ggf. auch durch die abgebildeten Bildproduzent(inn)en zueinander komponiert und somit in ein Verhältnis der (zeitlichen und sachlichen) Simultaneität zueinander gebracht werden. Vielmehr geht es auch darum, *dass* (aufgrund der ‚Entscheidung‘ im Sinne der intentionlosen Intention der abbildenden, aber durchaus auch der abgebildeten Bildproduzent(inn)en) gerade diese Situation abgelichtet wurde, in der die Bildgegenständlichkeiten sich in einem so und nicht anders gearbeteten Verhältnis der Simultaneität zueinander befanden. Bei Bourdieu heißt es dazu: „In Wirklichkeit hält die Photographie einen Aspekt der Realität fest, d.h. das Ergebnis einer willkürlichen Wahl und somit einer Bearbeitung: Von den Eigenschaften eines Gegenstandes werden nur jene erfasst, die in einem besonderen Augenblick und unter einem besonderen Blickwinkel hervortreten.“ (Bourdieu 1983b: 85).

Diese Selektivität, welche die Bildgegenständlichkeiten in ein je spezifisches Verhältnis der Simultaneität zueinander bringt, ist Ausdruck der Alltagsästhetik und deren stilistischer Präferenzen. Deren Rekonstruktion bietet die Möglichkeit, Ausführungen der Kunstgeschichte, die auf die ästhetischen Grundlagen der Kunst zielen, sowie auch die Standards und die Methodik kunstgeschichtlicher Interpretation sozialwissenschaftlich auch für die profane Fotografie relevant werden zu lassen. So betont auch Jürgen Wittpoth (2003: 78) in Bezug auf die Unterscheidung in ‚einfache‘ Fotos (‚Schnappschüsse‘), die Ausschnitte der Wirklichkeit so aufnehmen/abbilden, wie wir sie vorfinden“ und den ‚professionellen“ und ‚künstlerischen“ Fotografien: „Mittlerweile denke ich, dass diese Unterscheidung weniger von der Sache als vielmehr von der besonderen Art der wissenschaftlichen Arbeiten nahegelegt wird, die wir zur Fotografie vorfinden.“

Die Frage, inwieweit die Methoden der Kunstgeschichte auf die profane Fotografie übertragbar sind, ist bei Bourdieu schon im Titel seines Buches zur Fotografie beantwortet. Denn dieser lautet: „Eine illegitime Kunst“ oder – im Original (Bourdieu et al. 1965) – „Un art moyen“. Zum anderen ist diese Frage bei Bourdieu (implizit) bereits dort beantwortet worden, wo er die Ikonologie von Panofsky in ihrer Relevanz für die sozialwissenschaftliche Methodik herausgearbeitet hat (vgl. Bourdieu 1970: 127ff.).

Im Bereich der qualitativen oder rekonstruktiven Textinterpretation sind wir es inzwischen gewohnt, die Texte mit ähnlicher Sorgfalt und Systematik zu behan-



deln, welche die Literaturwissenschaft künstlerisch anerkannten Produkten ange-  
deihen lässt. Wir machen dabei immer wieder die – für manche immer noch er-  
staunliche – Entdeckung, dass auch diese alltäglichen Produkte in einer Weise  
durchkomponiert und ‚organisiert‘ sind, die man ihnen zunächst nicht ansieht. Das,  
was uns im Bereich der Textinterpretation selbstverständlich erscheint, stößt im Be-  
reich der Bildinterpretation nun wiederum auf eben jene Vorbehalte, wie wir sie im  
Bereich der Textinterpretation insbesondere aus den 1970er Jahren kennen: Die  
elaborierten Verfahren, wie sie in Anlehnung an jene der Kunstgeschichte entwik-  
kelt werden, seien für nicht-professionelle Produkte und spontane Schnappschüsse  
nicht geeignet.

Aber gerade beim spontanen Schnappschuss dokumentieren sich in der Selekti-  
vität der Auswahl des Bildausschnittes, also der Kadrierung, der perspektivischen  
Fokussierung und des Arrangements der Bildgegenständlichkeiten in ihrer je spezi-  
fischen Simultaneität die Stile, die stilistischen Präferenzen der abbildenden und  
(teilweise auch) der abgebildeten Bildproduzent(inn)en.<sup>33</sup> Durch die intuitive Wahl  
des Bildausschnittes, der Kadrierung, wird spontan u.a. auch über die planimetrische  
Komposition des Schnappschusses und damit in besonderer Weise über die eigen-  
logische Struktur des Bildes als selbst-referentielles System entschieden. „Kadrie-  
rung ist die Kunst, Teile aller Art für ein Ensemble auszuwählen. Dieses Ensemble  
ist ein relativ und künstlich geschlossenes System“ (Deleuze 1997a: 35).

Somit finden beim Schnappschuss, insbesondere auch in der Kadrierung, Stil  
und Habitus der Bildproduzent(inn)en ihren Ausdruck. In der Praxis des For-  
schungsprozesses können wir dann – im Zuge der notwendigen Auswahl der Fotos,  
die dann schließlich interpretiert werden – den stilistischen Präferenzen der Bild-  
produzent(inn)en in ihrer eigentümlichen Selektivität noch einmal in verdichteter  
Form zum Ausdruck verhelfen, wenn wir ihnen (also beispielsweise den Familien-  
angehörigen) die Auswahl jener ‚Schnappschüsse‘ (aus ihrem ‚Fotoalbum‘) über-  
lassen, die sie uns für eine Interpretation zur Verfügung stellen.

Wenn wir also ein Familienfoto interpretieren, so hat dies mehrere Etappen der  
Selektivität im Rahmen des familienspezifischen Erfahrungsraums durchlaufen: die  
Selektivität im Moment der Ablichtung und durch die Wahl dieses Moments, ggf.  
die technische Bearbeitung durch die familialen Amateurfotograf(inn)en nach der  
Ablichtung sowie die Auswahl der Fotos durch Angehörige der Familie für die For-  
scher. Die in der Selektivität im Moment der Ablichtung in ihrer je spezifischen  
Simultaneität hergestellte Komposition und die sich darin dokumentierenden stilis-  
tischen Präferenzen der abbildenden und (teilweise auch) der abgebildeten Bildpro-  
duzent(inn)en werden dadurch sozusagen *authentisiert* oder auch verworfen oder  
korrigiert. Und eine derartige „Authentisierung“ als letzter Selektionsfilter ist auch  
im Bereich der Kunst von zentraler Bedeutung, wie Ernst Gombrich (1978: 390)  
betont: „Es nicht wichtig, ob man vorhersehen konnte, wo das Papier, aus dem Pi-  
cassos unheimliche Maske entstand, reißen würde. Wichtig ist nur, warum er es  
aufbewahrt hat.“

33 Dies weist Analogien auf zur Bedeutung der „Stegreiferzählung“, also der spontanen erzähleri-  
schen Darstellung in der Methodologie des narrativen Interviews (vgl. Schütze 1987), oder der  
Bedeutung der Generierung von Selbstläufigkeit in der Methodologie von Gruppendiskussionen  
(vgl. Bohnsack 2008a: Kap. 12.1).

## Simultaneität und Synchronizität

Verhältnisse der Gleichzeitigkeit finden wir auch in anderen Bereichen der Dar-  
stellung als derjenigen des Bildes: so bspw. im synchronen Verhältnis von sprach-  
lich-textlichem Ausdruck zu den sprachbegleitenden Äußerungen vokaler Art wie  
Intonation einerseits oder körpergebundenen Ausdrucksformen wie Gestik und Mi-  
mik andererseits. In allen diesen Fällen handelt es sich aber um (synchroner) Rela-  
tionen zwischen *unterschiedlichen* Darstellungsmedien. Im Sprachgebrauch der  
qualitativen Methoden kann man auch sagen, dass das *eine* Medium (bspw. der kör-  
perliche Ausdruck) den *Kontext* für das *andere* Medium (bspw. den sprachlichen  
Ausdruck) darstellt bzw., dass die Medien wechselseitig füreinander Kontexte bil-  
den.<sup>34</sup>

Die Besonderheit des Mediums der Ikonizität, des ikonischen Codes, liegt aber  
offensichtlich darin, dass hier die Gleichzeitigkeit, die ich mit Imdahl als Simultan-  
struktur oder Simultaneität bezeichne, innerhalb *desselben* Mediums möglich und  
konstitutiv für die Eigenlogik und Eigensinnigkeit dieses Mediums ist, wohingegen  
dies im Medium der Sprache und des Textes eher die Ausnahme bzw. – wie dies in  
der Konversationsanalyse (vgl. Sacks 1995) verstanden wird – die Abweichung  
darstellt.

## Simultaneität und Rezeptionsprozess

Die bisherigen Ausführungen zur Simultaneität betrafen den Bereich der Bildinter-  
pretation als *Produktinterpretation*. Davon zu unterscheiden ist die Simultaneität im  
Bereich der *Rezeption* (welche der derart produzierten Simultaneität allerdings ge-  
recht zu werden hat). Susanne Langer (1984: 99) betont hier, „dass visuelle Formen  
nicht diskursiv sind. Sie bieten ihre Bestandteile nicht nacheinander, sondern  
gleichzeitig dar, weshalb die Beziehungen, die eine visuelle Struktur bestimmen, in  
einem Akt des Sehens erfaßt werden.“

Unter „Sehen“ wird hier jedoch sowohl die *sinnliche Wahrnehmung* als auch  
die *Interpretation, die interpretative Sinnbildung*, also die *Semantik* des Visuellen,  
verstanden. Die Frage der Simultaneität im Bereich der sinnlichen Wahrnehmung  
ist als ein wahrnehmungspsychologisches Problem für unsere Betrachtung von sek-  
undärer Bedeutung. Uns geht es um die Prozesse der interpretativen Sinnbildung,  
der Semantik. Der Zugang zur Simultaneität ist hier Voraussetzung für die Erfas-  
sung von Ganzheitlichkeit. Die Erfassung von Ganzheitlichkeit auf dem Wege der  
Simultaneität findet ihren Ausdruck auch im hermeneutischen Zirkel, in dem zirkel-  
haften Oszillieren zwischen Teil und Ganzem.

Den Zugang zur Ganzheitlichkeit auf dem Wege der Simultaneität muss man  
sich aber in der Regel erst erarbeiten. Dabei kann der Weg hin zu einer die Totalität

34 In diesem Sinne betont etwa Charles Goodwin (2001: 166): “that images in interaction are lodged  
within endogenous activity systems constitute through the ongoing, changing deployment of mul-  
tiple ‘semiotic fields’ which mutually elaborate each other (...) Bringing signs lodged within dif-  
ferent fields into a relationship of mutual elaboration produces locally relevant meaning and action  
that could not be accomplished by one sign system alone.”

des Bildes erfassenden Simultaneität durchaus sukzessive vonstatten gehen. Der Unterschied von Bild- und Textinterpretation besteht aber wohl darin, dass es – aufgrund der essentiellen Simultaneität der Semantik des Bildes – nicht gelingt, diese Semantik an *Teilen* oder *Abschnitten* des Bildes festzumachen, so wie wir dies beispielsweise bei der Auswahl von Passagen aus Gesprächen tun. Das Bild ist eine „in seiner Ganzheitlichkeit invariable und notwendige, das heißt alles auf alles und alles aufs Ganze beziehende Simultanstruktur. Diese erzwingt keineswegs ein starres Hinsehen auf das Bildganze, wohl aber eröffnet sie ein Sehen simul und singulariter: Das Ganze ist von vornherein in Totalpräsenz gegeben und als das sinnfällige Bezugssystem in jedem Einzelnen kopräsent, wann immer jedes Einzelne in den Blick genommen wird.“ (Imdahl 1996a: 23).

Sofern wir es mit gegenständlichen Bildern zu tun haben, kommt den abgebildeten Objekten zwar ihre je zeichenhafte Bedeutung zu. Für die Identifikation des ikonischen (oder ikonologischen) Sinngehalts, für den ein „sehendes Sehen“ (Imdahl 2003: 23) konstitutiv ist, stellen diese Objekt- oder Gegenstandsbedeutungen, wie ein „wiedererkennendes Sehen“ (ebenda) sie identifiziert, jedoch nicht die entscheidenden Elemente dar. Sie führen den Betrachter, der an jenen Sinngehalten interessiert ist, wie sie nur durch das Bild zu vermitteln sind, eher auf Ab- oder Nebenwege. Vielmehr geht es im Bereich der ikonologischen oder ikonischen Bildinterpretation darum, dass „das wiedererkennende Sehen und das sehende Sehen zu den ungeahnten oder unvordenklichen Erfahrungen eines erkennenden Sehens zusammenwirken“ (Imdahl 1996a: 92).

### Schluss

Die Kunstgeschichte bietet wesentliche Vorarbeiten, die uns Zugänge zum Bild in seiner Eigenlogik und Eigensinnigkeit eröffnen – und dies vor allem auch in forschungspraktischer Hinsicht, im Hinblick auf die Fortentwicklung von qualitativen Methoden der Bildinterpretation. So kommt denn auch William James T. Mitchell (1997: 17; s. auch im Original: 1994: 15), der den Begriff des „pictorial turn“ geprägt hat, zu dem Ergebnis: „Wenn sich tatsächlich ein pictorial turn in den Humanwissenschaften ereignet, so könnte sich die theoretische Marginalität der Kunstgeschichte durchaus in eine Position des intellektuellen Zentrums wandeln, indem nämlich die Humanwissenschaften von ihr eine Erklärung ihres grundlegenden theoretischen Gegenstandes erwarten.“ Zugleich betont Mitchell die hervorragende Stellung der Ikonologie Panofskys in diesem Zusammenhang, wie auch Bredekamp (2004: 15) feststellt, der die Arbeiten von Mitchell als Versuch interpretiert, „Erwin Panofskys Ikonologie zu aktualisieren“.<sup>35</sup> Und Panofsky selbst hat, wie erwähnt, explizit gemacht, in welcher grundlegenden Weise seine Ikonologie

35 Mitchell (1994: 16) bezeichnet die Ikonologie von Panofsky als „inevitable model and starting point for any general account of what is called ‘visual culture’“. Seine gleichwohl artikulierte Kritik an Panofsky zeigt gewisse Parallelen zu Imdahl, indem Mitchell bezweifelt, ob Panofsky den Zugang zur Eigensinnigkeit der Ikonik gefunden habe: „Panofsky’s is an iconology in which the ‚icon‘ is thoroughly absorbed by the ‚logos‘, understood as a rhetorical, literary, or even (less convincingly) a scientific discourse“ (Mitchell 1994, S. 28).

durch die dokumentarische Methode von Karl Mannheim beeinflusst wurde. Damit hat er in seiner Ikonologie auch von vornherein die Brücke zur sozialwissenschaftlichen Methodologie und Methodik geschlagen. Das Bild wird zum Dokument für die Weltanschauung, genauer: den Habitus einer Epoche und eines Milieus.

Allerdings konnte, wenn wir u.a. der Kritik von Max Imdahl folgen, Panofsky dem Bild in seiner Eigenlogik und Eigensinnigkeit letztlich nicht umfassend gerecht werden. Perspektiven zur Überwindung der Textgebundenheit des methodischen Fortschritts in der qualitativen Sozialforschung (vgl. auch: Krüger 2000: 338) vermag uns das Werk von Panofsky weniger zu eröffnen. Erst wenn auch Bilder als selbst-referentielle Systeme zum Gegenstand der Analyse werden, wie sich dies im Bereich der qualitativen Methoden im Zusammenhang mit dem linguistic turn durchzusetzen vermochte, werden wir hier ähnliche Fortschritte zu verzeichnen haben wie im Bereich der Textinterpretation, so dass wir dem viel beschworenen „iconic turn“ auch in der qualitativen resp. rekonstruktiven Sozialforschung zumindest in Ansätzen näher kommen werden.

Dabei geht es insbesondere auch um einen Wechsel in der AnalyseEinstellung, der eine vom sprachlich-textlichen Vorwissen in kontrollierter Weise weitgehend unabhängige Interpretation des Bildes ermöglicht. In diesem Sinne habe ich mit Bezug auf Semiotik, Philosophie und Kunstgeschichte zu zeigen versucht, dass es bei der Entwicklung von Methoden der Bildinterpretation darum gehen muss, „Bilder nicht mehr mit Texten zu erklären, sondern von Texten zu unterscheiden“ (Belting 2001: 15).

Beinahe gleich wichtig ist es jedoch, gemeinsame Standards oder methodologische Prinzipien zu entwickeln, die gleichermaßen für die Text- wie für die Bildinterpretation ihre Gültigkeit haben. Als Beispiele für derartige gemeinsame Standards lassen sich – neben dem der Prinzip der Behandlung von Texten ebenso wie Bildern als selbstreferentiellen Systemen – der Wechsel der AnalyseEinstellung vom *Was* zum *Wie* nennen sowie die Rekonstruktion der *Formalstruktur* von Bildern wie auch von Texten, um auf diese Weise die Einzelelemente eines Bildes oder Textes in den übergreifenden Kontext integrieren zu können, und – nicht zuletzt – die komparative Analyse.

Die Anwendung oder Realisierung dieser gemeinsamen Standards und dieser methodologischen AnalyseEinstellung im Bereich der Interpretation von Bildern ist allerdings nachhaltig zu unterscheiden von deren Anwendung in der Textinterpretation, wenn es darum gehen soll, der Ikonizität, dem Bild, in seiner ihm eigentümlichen Sinnstruktur gerecht zu werden.

#### 4. Zur Forschungspraxis dokumentarischer Bildinterpretation

Mit der dokumentarischen Methode der Bildinterpretation ist auf der Grundlage der Wissenssoziologie Mannheims eine an die Kunstinterpretation von Panofsky anschließende methodische Verfahrensweise von sozialwissenschaftlicher Relevanz entwickelt worden. Diese Methode räumt, über Panofsky hinausgehend, aber auch – im Sinne von Imdahl – den (ästhetischen) Formalstrukturen des Bildes (insbesondere der Planimetrie) einen zentralen Stellenwert ein und sucht jenen von Seiten der Kunstgeschichte, der Semiotik und der Philosophie formulierten Ansprüchen gerecht zu werden, welche einen konsequenten Zugang zur Eigenlogik des Bildes im Unterschied zu derjenigen des Textes fordern. Dabei leistet die dokumentarische Methode u.a. die Bewältigung folgender Aufgaben:

- Der erkenntnistheoretisch-philosophisch geforderten Suspendierung oder auch „methodischen Verdrängung“ (Imdahl 1996b: 435) des textlich-narrativen Vorwissens kann eine sozialwissenschaftliche Relevanz verliehen und im Rahmen sozialwissenschaftlicher Empirie methodisch-forschungspraktisch umgesetzt werden, um der Eigensinnigkeit des Bildes, der Ikonizität gerecht werden zu können.
- Die in diesem Zusammenhang ebenfalls notwendige Rekonstruktion der formalen Komposition des Bildes kann in einer über die rein ästhetische Analyse hinausgehenden und für die sozialwissenschaftliche Empirie relevanten Weise umgesetzt werden.
- Auf der Grundlage gemeinsamer methodologischer und erkenntnistheoretischer Grundlagen können generelle forschungspraktische Verfahrensweisen und Arbeitsschritte entwickelt werden, die es ermöglichen, Bild- und Textinterpretationen in einander ergänzender Weise (im Sinne einer Methodentriangulation; vgl. Kap. 4.3) aufeinander zu beziehen.
- Auf dem Wege der Typenbildung kann den für die sozialwissenschaftliche Empirie wesentlichen Qualitätskriterien der Generalisierung und ‚Erklärung‘ auch im Bereich der Bildinterpretation Rechnung getragen werden (siehe dazu: Bohnsack 2009a). Während im Bereich der Textinterpretation die empirische Generierung von Typologien und Generalisierungen inzwischen weit fortgeschritten ist (vgl. auch Kap. 2.4), stehen wir im Bereich der Bildinterpretation in dieser Hinsicht noch am Anfang.
- Insgesamt wird es auf diese Weise möglich, (allgemeine) gemeinsame Standards von Bild- und Textinterpretation zu entwickeln (vgl. dazu Bohnsack

2005a) und dabei zugleich aber auch den Besonderheiten und Eigengesetzlichkeiten des Bildes im Unterschied zum Text Rechnung zu tragen, d.h. die Semantik des Bildes nicht dem textlichen (Vor-) Wissen zu subsumieren.

#### 4.1 Die Arbeitsschritte dokumentarischer Bildinterpretation

Die Arbeitsschritte der Interpretation von Bildern folgen ebenso wie diejenigen der Interpretation von Texten der Leitdifferenz von *immanentem* und *dokumentarischem Sinngehalt* und der daraus resultierenden Differenzierung von *formulierender* und *reflektierender Interpretation*. Die formulierende Interpretation fragt danach, *was* auf dem Bild bzw. im Text dargestellt wird. Die reflektierende Interpretation fragt nach dem *Wie* der Herstellung der Darstellung, nach dem *modus operandi*.

##### *Formulierende Interpretation*

Panofsky hat mit seiner Ikonologie an die Unterscheidung von immanentem und dokumentarischem Sinngehalt bei Mannheim angeknüpft, um dann aber innerhalb des immanenten Sinngehalts, also innerhalb der Frage nach dem *Was* (dargestellt ist), noch einmal zu differenzieren zwischen der *vor-ikonografischen* Ebene, als dem Bereich der auf einem Bild sichtbaren Gegenstände, Phänomene und Bewegungsabläufe, und der *ikonografischen* Ebene, als dem Bereich der auf dem Bild identifizierbaren Handlungen. Um Handlungen zu identifizieren, muss ich Motive unterstellen, genauer: „Um-zu-Motive“. Die auf der vor-ikonografischen Ebene beschreibbare Bewegung des „Hutziehens“ wird auf der ikonografischen Ebene als ein „Grüßen“ interpretiert (so das Beispiel von Panofsky 1975: 38): Der Bekannte zieht seinen Hut *um zu* grüßen.

Mit dieser Interpretation oder Konstruktion eines subjektiv gemeinten Sinns sind aber schwerwiegende Probleme der Zuschreibung von Motiven, von Intentionen verbunden. Diese (auf der Grundlage des narrativ-textlichen Vorwissens vollzogenen) Attribuierungen und Unterstellungen gilt es – soweit wie möglich – zu suspendieren. Motivunterstellungen sind nur dort unproblematisch, wo wir es mit institutionalisierten oder (wie es in der dokumentarischen Methode genannt wird) „kommunikativ-generalisierten“ Bedeutungen zu tun haben, wie beispielsweise im Falle der Interpretation des „Hutziehens“ als „Grüßen“ (vgl. auch Kap. 3.3). Hierzu gehört auch das Wissen um gesellschaftliche Institutionen und Rollenbeziehungen, also beispielsweise das allgemeine Wissen darum, was eine ‚Familie‘ ist. Davon zu unterscheiden ist im Sinne der dokumentarischen Methode ein Wissen um die *fall- oder auch milieuspezifische* Besonderheit des Dargestellten und seiner konkreten Geschichte, das „*konjunktive*“ Wissen, also beispielsweise unser Wissen um die (auf dem Bild dargestellte) Familie.

Dieses konjunktive Wissen, das Wissen um die „Eigennamen“, wie es bei Foucault (1971: 38) heißt, gilt es – auch dann, wenn es in empirisch valider Form vor-

liegt – einzuklammern, zu suspendieren. Im Sinne von Foucault (ebenda) „muß man die Eigennamen auslöschen“.

Im Bereich der *ikonografischen* Beschreibung oder Interpretation findet unser sprachlich-textliches Vorwissen also lediglich insoweit Berücksichtigung, als es sich um die kommunikativ-generalisierten Wissensbestände handelt. Dieser bedienen wir uns weitgehend in stereotypisierender Form, wodurch dann auch die ikonografische Interpretation charakterisiert ist. Ein derartiges stereotypes Wissen umfaßt im Sinne von Panofsky (1975) auch die „Typengeschichte“ (wir wissen bspw., dass das Sujet des Bildes das „Letzte Abendmahl“ ist) und die „Stilgeschichte“ (wir wissen bspw., dass der Kinderwagen auf dem Foto aus den 50er Jahren stammt, dem Stil der 50er Jahre entspricht).

##### *Reflektierende Interpretation*

Das Grundgerüst der reflektierenden Interpretation des Bildes bildet – wie auch im Bereich der Textinterpretation (vgl. Bohnsack 2001) – die Rekonstruktion der Formalstruktur, der formalen Komposition. Imdahl (1996a: Kap. II) unterscheidet drei Dimensionen des formalen kompositionalen Aufbaus des Bildes: die „*perspektivische Projektion*“, die „*szenische Choreografie*“ und die „*planimetrische Ganzheitsstruktur*“. Die planimetrische Ganzheitsstruktur, also die formale Konstruktion des Bildes in der Fläche, erscheint nach Imdahl von entscheidender Bedeutung für das „*sehende Sehen*“, welches den Zugang zum Eigensinn des Bildes eröffnet. Die *planimetrische Komposition* schafft ihre eigenen bildinternen, systemimmanenten Gesetzmäßigkeiten, ihre eigene formale Ganzheitsstruktur im Sinne einer Totalität.

Insbesondere die Planimetrie führt uns das Bild als „ein nach immanenten Gesetzen konstruiertes und in seiner Eigengesetzlichkeit evident System“ (Imdahl 1979: 190) vor Augen. Demgegenüber dient die *Perspektivität* primär dazu, Gegenstände und Personen in ihrer Räumlichkeit und Körperlichkeit identifizierbar zu machen. Sie ist somit an den Gesetzmäßigkeiten der im Bild dargestellten *Außen- oder Umwelt* des Bildes orientiert.

Im Bereich der sozialwissenschaftlichen Bildinterpretation ermöglicht uns die Rekonstruktion der Perspektivität im wahrsten Sinne des Wortes Einblicke in die *Perspektive* des abbildenden Bildproduzenten und in seine *Weltanschauung*.

Während es im Sinne von Imdahl und auch Panofsky bei der Rekonstruktion der Perspektivität im Wesentlichen darum geht, die „Perspektive als ‚symbolische Form‘“ (Panofsky 1964a) im Kontext epochaler Wandlungsprozesse (insbesondere die uns heute selbstverständliche Zentralperspektive vor dem Hintergrund ihrer historischen ‚Vorläufer‘, z.B. der Achsenperspektive) genauer zu beleuchten,<sup>36</sup> haben wir es im Bereich der Fotointerpretation in der Regel mit der Zentralperspektive zu tun. Es geht vor allem um die Frage, welche Personen und sozialen Szenerien durch den abbildenden Bildproduzenten, durch das Kameraauge sozusagen, in Form des Fluchtpunktes fokussiert und somit ins Zentrum des sozialen Geschehens gerückt

36 Zur Entwicklung und Geschichte der Zentralperspektive siehe auch: Panofsky 2001 sowie Edgerton 2002 und Kap. 7.4 (Anhang).

werden. Innerhalb der Zentralperspektive lassen sich u.a. die Frontalperspektive (mit einem Fluchtpunkt), die Schräg- oder Übereckperspektive (mit zwei Fluchtpunkten) sowie die Untersicht und die Aufsicht unterscheiden (siehe Anhang: 7.4).

Indem die dokumentarische Bildinterpretation die Rekonstruktion der formalen – allen voran der planimetrischen – Komposition zum Ausgangspunkt und Grundgerüst der reflektierenden Interpretation nimmt, nähert sie sich der Ikonik von Im Dahl, der hierin wesentlich den Unterschied seiner *ikonischen* Interpretation zur *ikonologischen* Interpretation von Panofsky gesehen hat. Da die dokumentarische Bildinterpretation aber auch der starken sozialwissenschaftlichen Relevanz der ikonologischen Interpretation von Panofsky verbunden bleibt, bezeichne ich sie auch als *ikonologisch-ikonische* Interpretation.

Im Folgenden wird die dokumentarische Bildinterpretation an exemplarischen Fällen von Fotos erläutert: einmal aus dem öffentlich-medialen Bereich am Beispiel von Werbefotos (Kap. 4.2) und zum anderen aus dem privaten Bereich am Beispiel von Familienfotos (Kap. 4.3) in Kombination (Methodentriangulation) mit der Interpretation von Gesprächen (Tischgesprächen und Gruppendiskussionen). Für eine dokumentarische Interpretation von Zeichnungen, die hier nicht exemplarisch dargestellt wird, sei auf die Interpretation derjenigen von Kindern am Übergang zum Jugendalter (in Kombination mit der Interpretation von Gruppendiskussionen) von Gabriele Wopfner (2008) verwiesen. Eine dokumentarische Interpretation des komplexen Zusammenhangs von bildlichen und schriftlichen Ausdrucksformen in Tagebüchern findet sich bei Andrea Sabisch (2007 u. 2009) und die dokumentarische Interpretation fotografischer Selbstportraits bei Klika/Kleynen (2007).

## 4.2 Exemplarische Interpretation eines Werbefotos

Die zunehmende Bedeutung von Bildern in der öffentlichen, genauer: der medialen Kommunikation und damit auch im Bereich der Werbung ist unbestritten. Dies gilt nicht nur für Bildmedien im engeren Sinne, also Film und Fernsehen, sondern auch für die Zeitschriftenwerbung.<sup>37</sup> Trotz des enormen Bedeutungszuwachses des Bildhaften in der kommunikativen Verständigung ist die Bildinterpretation in der modernen sozialwissenschaftlichen Empirie eine Marginalie geblieben. Der „iconic“ oder „pictorial turn“ (Mitchell 1994 u. 1997) ist im sozialwissenschaftlichen Forschungsalltag noch nicht angekommen (genauer dazu: Kap. 3). Und auch in der Werbewissenschaft wird „Bildforschung“, wie Schierl (2005: 309) konstatiert, „weiterhin kaum in einem der Bedeutung des Gegenstandes angemessenen Umfang betrieben. Dies mag damit zusammenhängen, dass Bilder in der Gesellschaft vorwiegend wegen ihres Unterhaltungswertes geschätzt, aber bezüglich ihres Beeinflussungspotentials absolut unterschätzt werden.“

37 So konnte Wehner (1996: 76ff.) am Beispiel der Berliner Illustrierten Zeitung (1900-1944) und des Stern (1949-1992) zeigen, dass sowohl der Anteil von bebilderten Inseraten als auch der Anteil der Bilder an der Gesamtfläche der Inserate deutlich zugenommen hat.

Das „Beeinflussungspotential“ des Bildes kann aber überhaupt erst adäquat eingeschätzt werden, wenn es gelingt, einen Zugang zu seiner tiefer liegenden Semantik zu gewinnen, also zu einer Sinnebene, die über eine bloße Beschreibung des Bildes hinausgeht. Und es ist häufig das Werbefoto, an dessen Beispiel derartige tiefer gehende Interpretationen exemplarisch entfaltet worden sind – so u.a. von den beiden prominenten Vertretern der Semiotik: Umberto Eco (1994) und Roland Barthes (1990).<sup>38</sup> Gernot Böhme (1999: 7) betont: „Es ist insbesondere der Bereich der Werbung, in dem Trivialität und Raffinement der Bildproduktion sich durchkreuzen.“

Das im Folgenden interpretierte Foto (Abbildung 3), welches auch im Original schwarz-weiß ist, entstammt der Werbung der Bekleidungsfirma Burberry für den US-amerikanischen und den russischen Markt. Es wurde veröffentlicht in der russischen und amerikanischen Ausgabe der Zeitschrift Vogue 2005.<sup>39</sup> Aufgrund der Heftbindung bleibt im Bereich der Mittelsenkrechten erkennbar ein schmaler Streifen verdeckt. Abgedruckt ist hier die russische Version.<sup>40</sup>

Abbildung 3: Burberry London. Entnommen aus: Vogue 2005 (russische Ausgabe)



38 Roland Barthes (1990: 29ff.) erläutert seine Differenzierung von unterschiedlichen Botschaften des Bildes am Beispiel der Pasta Panzani-Werbung, und Umberto Eco (1994: 267ff.) gibt uns Einblicke in unterschiedliche Ebenen der Botschaft, u.a. am Beispiel der Camay-, der VW- und der Knorr-Werbung.

39 Eine andere – weitgehend identische – Veröffentlichung der Interpretation unseres Fotos findet sich in Bohnsack 2007d. Ausgewählt wurde dieses Foto von Dr. Aglaja Przyborski vom Institut für Publizistik und Kommunikationswissenschaft und Prof. Dr. Thomas Slunecko von der Fakultät für Psychologie der Universität Wien im Rahmen eines von ihnen geleiteten Seminars zur Bildinterpretation. Dort war ich im Sommersemester 2005 zu einem Workshop geladen. Ich danke Aglaja Przyborski und Thomas Slunecko sowie den Teilnehmer(inne)n des Seminars, insbesondere Stefan Hampl, für viele wertvolle Anregungen und die außergewöhnlich inspirierende Atmosphäre.

40 Mit Ausnahme des Bildtextes (vgl. III. 2.2.3) sind beide Versionen identisch.

#### 4.2.1 Formulierende Interpretation

##### *Vor-ikonografische Interpretation*

Da aufgrund der vielen Details die vor-ikonografische Interpretation hier sehr umfangreich ausfällt, möchte ich darauf verzichten, sie hier abzudrucken (siehe dazu aber die Beispiele in 4.3 sowie in: Bohnsack 2006a u. 2007b). Die vor-ikonografische Interpretation gliedern wir üblicherweise in die Schritte: Bildvordergrund, Bildmittelgrund, Bildhintergrund. Wobei je nach Aufbau des Bildes diese noch einmal in sich differenziert werden können. In diesem Bild erscheint es bspw. sinnvoll, noch einmal den vorderen, mittleren und hinteren Bildvordergrund zu unterscheiden. Wenn Personen dargestellt sind (hier im Vordergrund), differenzieren wir deren Beschreibung jeweils personenbezogen nach (geschätztem) Alter, Kleidung, Frisur, Körperhaltung, Gestik und Mimik.

##### *Ikonografische Interpretation*

Von den auf dem Rasen ausgebreiteten Utensilien und der Situierung der gesamten Szene in einem Park her gesehen, lässt sich die Situation in stereotyper Weise als eine des Picknicks identifizieren. An ihm sind mehrere Generationen beteiligt, bei denen es sich um die Großeltern-, Eltern- und Kindergeneration und somit um eine Art Großfamilie handeln könnte.

Es ist charakteristisch für dieses Foto, dass sich hinsichtlich der Einschätzung der Verwandtschaftsverhältnisse, die wir bei Familienfotos auf der ikonografischen Ebene vornehmen, ebenso schwer eine intersubjektive Übereinstimmung herstellen lässt wie hinsichtlich der Altersschätzungen der einzelnen Personen, die bereits beim (hier nicht abgedruckten) Schritt der vor-ikonografischen Interpretation vorgenommen worden ist.

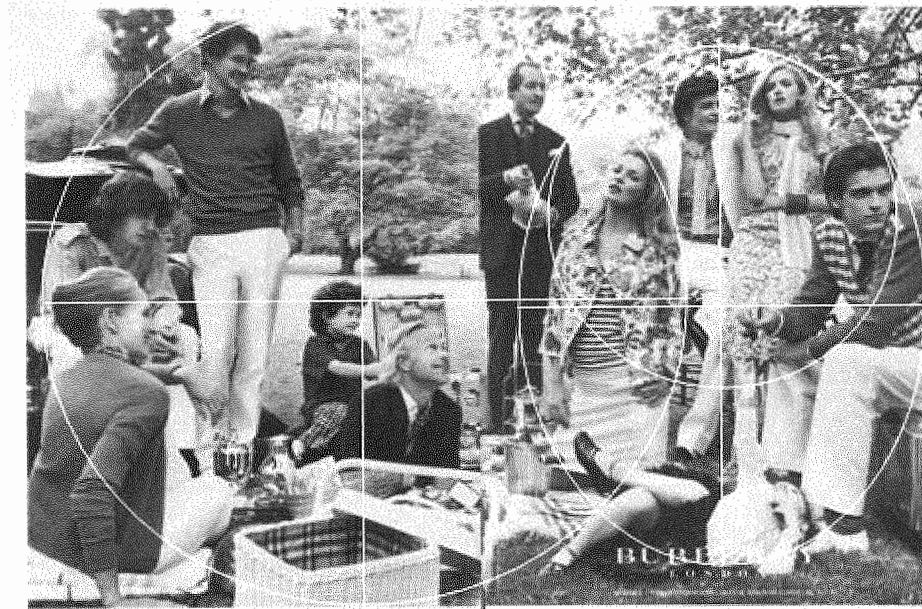
Wenn wir ein Wissen um Stilelemente, um die „Stilgeschichte“ einbeziehen, die nach Panofsky (1975) zur ikonografischen Ebene gehört, so vermitteln uns Kleidung und Accessoires den Eindruck eines gehobenen, mit einem gewissen Luxus versehenen (großbürgerlichen oder aristokratischen) Stils.

Eine der abgebildeten Personen, die vierte von rechts, lässt sich (sofern das entsprechende ikonografische Vorwissen verfügbar ist) als Kate Moss, ein bekanntes Model, identifizieren. – Es handelt sich um ein Werbefoto der Firma Burberry (zur genaueren Interpretation des Bild-Textes siehe weiter unten: 4.2.3).

#### 4.2.2 Reflektierende Interpretation

##### *Rekonstruktion der formalen Komposition*

Abbildung 4: Burberry London. Entnommen aus: Vogue 2005 (russische Version): Planimetrie I (Linien von R.B.)



##### *Planimetrie*

Bei der Rekonstruktion der planimetrischen Komposition (siehe Abb. 4) geht es darum, mit möglichst wenigen Linien die Gesamtkomposition des Bildes in der Fläche zu markieren. Dies erscheint hier am überzeugendsten durch zwei Kreise und eine Ellipse möglich. Die Ellipse auf der rechten Seite umschließt die relativ altershomogene Gruppe im Alter von ca. 25 bis 35 Jahren und das zwischen ihnen auf dem Rücken am Boden liegende Kind im Alter von ca. 6 Jahren. Ein in den oberen Teil der Ellipse integrierter Kreis rechts oben umschließt die Köpfe der altershomogenen Gruppe, welche durch diese Komposition noch einmal hervorgehoben wird.

Abbildung 5: Burberry London. Entnommen aus: Vogue 2005 (russische Ausgabe): Planimetrie II (Linien von R.B.)



Der große Kreis auf der linken Seite umschließt jene altersheterogene Gruppe, welche die rechte homogene Gruppe betrachtet bzw. in deren Richtung schaut. Zugleich wird die kniende Frau der rechten homogenen Gruppe aber auch in diesen linken Kreis einbezogen, verbindet sozusagen den linken mit dem rechten Kreis. Ihr Kopf wird zugleich durch beide Kreise und die Ellipse umschlossen und auf diese Weise planimetrisch fokussiert. Durch die beiden Schnittpunkte des (großen) linken Kreises und der rechten Ellipse verläuft die senkrechte Linie, welche den „Goldenen Schnitt“ des Bildes markiert (siehe Abb. 5).<sup>41</sup> Diese Linie stimmt darüber hinaus mit der Körpermittelachse der knienden Frau überein, deren zentrale Stellung dadurch noch einmal bestätigt wird.

41 Der Goldene Schnitt teilt eine Strecke derart in zwei ungleiche Teile, dass die gesamte Strecke sich zum größeren Teil verhält wie dieser zum kleineren (1: 0,6181). Er kann auch geometrisch ermittelt werden (vgl. Lexikon der Kunst 2004: 784 sowie Der Brockhaus Kunst: 316). Der Goldene Schnitt spielt in der Rekonstruktion der formalen Komposition bei Max Imdahl keine und in der kunstgeschichtlichen Betrachtung insgesamt eine eher untergeordnete Rolle spielt. Es ist jedoch interessant, dass er das Proportionsempfinden im Alltag und in der Gebrauchskunst offensichtlich intuitiv strukturiert. Entsprechend heißt es im Lexikon der Kunst (2004: 784): „Der Mensch empfindet den G.S. deshalb als bes. harmonisch, da er vermutlich als die geistig sinnl. Umsetzung von Modellvorstellungen entstanden ist, die sich beim Menschen bei der prakt. Aneignung der Welt gebildet haben“.

Abbildung 6: Burberry London. Entnommen aus: Vogue 2005 (russische Ausgabe): Perspektivität (Linien von R.B.)



– *Perspektivität*

Wir haben es mit der Zentralperspektive zu tun – und zwar in ihrer auf Fotos wohl häufigsten Variante: der Frontal-Perspektive (vgl. dazu: 7.4). Diese hat nur einen Fluchtpunkt, welcher hier am Dekolleté der durch die Planimetrie bereits fokussierten Frau positioniert ist (siehe Abb. 6). Die Frau wird also sowohl durch die *Planimetrie* wie auch zugleich durch die *Perspektivität* ins Zentrum gerückt. Der Baum und die nur schwach erkennbare Bank im Hintergrund sind auf der *Horizontlinie* positioniert, haben diese als Standfläche, sodass der Horizont hierdurch markiert wird. Die Horizontlinie verläuft somit etwas oberhalb der waagerechten Bildmittelachse und der waagerechten Mittelachsen der beiden die Komposition bestimmenden Kreise.

– *Szenische Choreografie*

Durch die drei Kreise, welche die Planimetrie bestimmen (siehe Abb. 4), werden zugleich zwei bzw. drei soziale Szenarien bzw. Gruppen ausdifferenziert: Durch den *rechten* oberen (kleinen) *Kreis* wird eine altershomogene Gruppe markiert. Wenn wir diesen Kreis in die (umfassendere) Ellipse integrieren, wird das auf dem Boden liegende Kind einbezogen und die Gruppe erhält Merkmale der Elterngeneration bzw. Elternschaft. Die rechte Gruppe erscheint somit von der in der Planimetrie fundierten szenischen Choreografie wahlweise mit und ohne Kind.

Der *linke* (große) *Kreis* umschließt und markiert eine altersheterogene Gruppe, deren Angehörige auf die rechte Gruppe schauen. Die linke Gruppe wird dadurch zur *betrachtenden*, die rechte zur *betrachteten*, sich vor Betrachtenden inszenieren-

den Gruppe. Letztere rückt damit ins Zentrum. Wir haben es hier also mit einer sehr elaborierten Integration von planimetrischer Komposition und szenischer Choreografie zu tun bzw. mit einer Elaboration der szenischen Choreografie auf der Basis der Planimetrie.

Die planimetrische Komposition und die darin fundierte szenische Choreografie weisen uns deutlich den Weg dahin, dass wir es mit zwei unterschiedlichen Gruppen zu tun haben. Auf der rechten Seite ist eine enger beieinander stehende und in sich geschlossene Gruppe von zwei Frauen, zwei Männern und einem Kind positioniert, die schon dadurch *homogener* wirkt als die andere Gruppe, dass sie (abgesehen von dem Kind) eine geringe Altersspanne aufweist: von ca. 25 bis 35 Jahren. Demgegenüber umfasst die Gruppe der Erwachsenen auf der linken Seite alle Alterstufen zwischen ca. 20 bis 65 Jahren.

Während die Angehörigen der linken Gruppe in Richtung der rechten schauen, blicken die Frauen der rechten Gruppe in Richtung der Bildbetrachter/innen, also in die Kamera. Die Männer der rechten Gruppe schauen nach rechts (vom Bildbetrachter aus gesehen) aus dem Bild heraus, das Kind blickt auf die Weinbeeren, welche der ganz rechts sitzende Mann zwischen Daumen und Zeigefinger hält. Im Gegensatz zu den beiden Männern öffnen die beiden Frauen sich also durch Körperhaltung und Blickrichtung zum Bildbetrachter hin. Dies dokumentiert sich auch in ihren (leicht) geöffneten Mündern.

Eine choreografische Sonder- und zugleich Randstellung nimmt der nahezu in der Bildmitte positionierte 60 bis 65-jährige Mann im dunklen Anzug mit Weste, Krawatte und weißem Stecktuch ein, da er sehr weit hinten steht und somit beinahe nach hinten aus der linken Gruppe heraustritt, in die er durch die Planimetrie allerdings wieder hereingeholt wird.

### *Ikonologisch-ikonische Interpretation*

Ich möchte zunächst die rechte – also die betrachtete und somit ins Zentrum gerückte – Gruppe einer ikonologisch-ikonischen Interpretation unterziehen und dann die linke, die betrachtende Gruppe.

#### *– Die rechte Gruppe*

Die rechte Gruppe wirkt stärker durchstilisiert und (damit zusammenhängend) auch stereotyper als die linke Gruppe. Dies findet in Frisur und Kleidung seinen Ausdruck: zum einen im Vergleich der Frisuren der beiden rechten jungen Männer mit den Frisuren der beiden linken, aber auch darin, dass die Krawatten und Bügelfalten in der Bekleidung der Männer der rechten Gruppe in der linken Gruppe (mit Ausnahme des stehenden Mannes mit der Sektflasche) ebenso fehlen wie die ausgeprägtere Musterung mit Blümchen, Karo und Streifen in der rechten Gruppe.

Der Eindruck des Posierens:

Die rechte Gruppe ist auch in ihrer Mimik, ihrem Gesichtsausdruck stilisierter bzw. stereotypisierter. Dieses Durchstilisieren führt zu Ähnlichkeiten. So ähneln vor allem die beiden Frauen aufgrund der Übereinstimmungen der Mimik und der Frisuren einander. (Die Mimik mit den geöffneten Mündern schafft den Eindruck von

‚Puppenhaftigkeit‘). In etwas weniger ausgeprägtem Maße lässt sich die Ähnlichkeit auch bei den jungen Männern beobachten: hinsichtlich Haarschnitt, Haarfarbe und Profil. Das Durchstilisieren und die damit verbundenen Ähnlichkeiten auf Seiten der rechten Gruppe erscheint also als ein Posieren, als eine Ent-Persönlichung, wie Imdahl (1996d: 575) die Pose definiert. Zugleich wird dadurch allerdings auch der Eindruck einer Zusammengehörigkeit der Gruppe bzw. einer Zugehörigkeit der Einzelnen zur Gruppe vermittelt. Dies wird, wie dargelegt, entscheidend durch die formale Komposition unterstützt.

Dieser Eindruck des Posierens wird dadurch verstärkt, dass diese Gruppe durch ihre stärkere Durchstilisierung im Kontext der Situation des Picknicks deplaziert wirkt. Ein wenig deplaziert wirkt auch, dass die beiden jungen Leute auf ein Fahrrad gelehnt sind, ohne dass man sich vorstellen kann, dass sie in dieser Bekleidung radeln würden. Auch das ca. 6-jährige Mädchen ganz unten im Bild fügt sich hinsichtlich seiner Kleidung und der übereinander geschlagenen Beine in dieses Bild des Posierens – im Kontrast zu dem kleinen Jungen links im Bild.

Die fehlende soziale Bezogenheit bzw. die ‚Pose der Individualität‘:

Im Widerspruch zu dem durch die Pose und die stilistischen Übereinstimmungen sowie auch durch die Planimetrie vermittelten Eindruck von Zusammengehörigkeit und Zugehörigkeit steht die fehlende soziale Bezogenheit: Insgesamt nehmen die Personen der rechten Gruppe nicht nur keinen Blickkontakt zueinander auf, sondern sie nehmen einander nicht einmal wahr. Sie schauen nicht zueinander hin, sondern aneinander vorbei. Und dies gilt auch für das Paar rechts im Hintergrund, obschon die beiden körperlich nahe zueinander positioniert sind.

Auch das ca. 6-jährige Mädchen, welches auf dem Boden liegt, ist zwar auf die Weinbeeren in der Hand des rechten jungen Mannes, aber nicht auf ihn bezogen. Auch letzterer wirkt, indem er die Weintrauben in der Hand hält, äußerst unbeteiligt, ohne Bezug zum Kind.

Die Paradoxie von *Zusammengehörigkeit* und *Zugehörigkeit* einerseits und fehlender *sozialer Bezogenheit* andererseits können wir dann auflösen, wenn wir in Rechnung stellen, dass die fehlende soziale Bezogenheit hier für etwas anderes stehen soll: nämlich für *Autonomie* und *Individualität*, welche aber im Kontext der Werbung lediglich auf Umwegen bzw. ex negativo dargestellt werden kann, da das für die Werbung konstitutive Posieren eine Ent-Individualisierung impliziert. An dieser Stelle erscheint ein theoretischer Exkurs zur Pose in der Werbung notwendig.

### *Theoretischer Exkurs: Werbung, Pose und Individualität*

Eine ent-individualisierende bzw. ent-persönlichende und somit stereotypisierende Stilisierung ist für die Werbung unumgänglich. Sie ist notwendigerweise auf das Posieren angewiesen (vgl. dazu auch Imdahl 1996d) – zumindest dann, wenn ein Lifestyle transportiert werden soll. Erst indem die Modelle bzw. die Akteure vor der Kamera ent-persönlicht werden, vermögen sie einen verallgemeinerbaren sozialen Lifestyle zu transportieren, der durch individuelle oder persönliche Stilelemente nicht ‚getrübt‘ wird. Dies ermöglicht es, ohne Umwege jene Zielgruppe mit ihrer spezifischen sozialen Identität anzusprechen, die mit diesem Konzept erreicht



werden soll, und bietet dieser die Möglichkeit zur Identifikation jenseits der je individuellen und persönlichen Stile.

Damit steht die Werbung aber vor dem Problem, Individualität mit Mitteln der Pose, also durch Stereotype, darzustellen. Realisiert wird dies mit dem vorliegenden Foto auf dem Wege der Demonstration fehlender sozialer Bezogenheit, der wechselseitigen Nicht-Wahrnehmung sowie der Vermeidung des Blickkontakts. Somit kommt es zu der Paradoxie, dass durch die Übereinstimmung der Stilisierungen und der Posen sowie durch die planimetrische Komposition (also durch die für die Bildhaftigkeit wesentlichste Dimension formaler Komposition) eine Zugehörigkeit der Personen untereinander hergestellt wird, die dann durch Nichtwahrnehmung und Verweigerung des Blickkontakts wieder negiert oder eingeschränkt wird.<sup>42</sup>

Wenn wir bereits separat aus der Interpretation der rechten Gruppe, also des rechten Teils des Fotos, so etwas wie eine Botschaft herausarbeiten wollen, so lässt sich diese folgendermaßen formulieren: Die stilistische Orientierung an Burberry schafft *Zugehörigkeit und Zusammengehörigkeit bei gleichzeitiger Wahrung der Individualität* – einer Individualität allerdings, die hier nur ex negativo, also auf dem Wege einer Demonstration fehlender sozialer Bezogenheit, ihren Ausdruck finden kann.

#### *Die linke Gruppe und das Verhältnis der Gruppen zueinander*

Die linke Gruppe wird, wie erwähnt, insgesamt zum Betrachter der rechten Gruppe. Ihre Blicke sind auf die rechte Gruppe gerichtet. Letztere wird dadurch fokussiert und somit zur eigentlichen Ziel- und Identifikationsgruppe dieses Werbefotos.

#### *Die Adressaten und Adressatinnen der Werbung und ihre lebenszyklische Stellung*

Über die spezifische Stellung dieser Zielgruppe im Lebenszyklus und im Familienzusammenhang gibt uns auch die unterschiedliche Positionierung der beiden Kinder auf dem Bild Aufschluss. Zu beiden Gruppen gehört jeweils ein Kind. Die Einbindung der Kinder in die Gruppe ist jedoch jeweils sehr unterschiedlich. Das Kind der *linken* Gruppe ist von der formalen Komposition her fokussiert. Sein Kopf bildet planimetrisch das Zentrum des linken Kreises und ist nicht weit von der Horizontlinie entfernt. Es nimmt den auf dem Boden positionierten Mann gleichsam als Spielzeug voll in Anspruch, traktiert ihn sozusagen, indem es auf ihm reitet und seine Haare ergreift. Das Kind der *linken* Gruppe ist also in diese integriert und nimmt offensichtlich einen zentralen Stellenwert ein.

Demgegenüber befindet sich das Kind der *rechten* Gruppe am unteren Rand des rechten Kreises und gehört nicht zum innersten Zirkel, der durch den rechten oberen Kreis markiert wird. Es greift auch im wahrsten Sinne des Wortes nur sehr vorsichtig in die Gruppe der Erwachsenen ein, indem es seine Hand spielerisch nach oben streckt und niemanden belästigt. Der junge Mann an der rechten Seite übernimmt, indem er die Weinbeeren über das Kind hält, zwar bereits in gewisser Wei-

42 Bestätigt wird dies auch in der spezifischen Art der Beziehung der beiden Gruppen – der rechten und linken – zueinander: Die beiden Gruppen sind nicht eigentlich kommunikativ aufeinander bezogen, ihre Beziehung ist nicht durch eine Reziprozität, eine Wechselseitigkeit des nonverbalen Austauschs charakterisiert. Die rechte Gruppe ist durch eine Selbstpräsentation, durch ein *Sich-vor-einem-Beobachter-Präsentieren* charakterisiert und die linke Gruppe durch ein *Betrachten* und *Beobachten*.

se eine ‚Ernährerfunktion‘, wirkt dabei aber gleichwohl unbeteiligt und keinesfalls in irgendeiner Weise durch das Kind in Anspruch genommen. Er distanziert sich damit zugleich wieder von dieser Funktion.

Die rechte Gruppe kann somit *wahlweise* bereits mit Elternfunktionen assoziiert oder noch als kinderlos wahrgenommen werden. Und dies ermöglicht es den Bildbetrachter(inne)n auch, sich wahlweise mit diesen unterschiedlichen Rollen zu identifizieren. Dies bestätigt, dass es die 20 bis 30-Jährigen sind, welche hier primär als Zielgruppe adressiert werden. Denn diese Alterskohorte befindet sich lebenszyklisch am Übergang von einer vor-familialen zur familialen Phase.

Dadurch, dass die linke Gruppe als Betrachterin der rechten konstruiert wird, erhalten ihre Angehörigen die Funktion der abgebildeten Bildbetrachter/innen. Auf diese Weise *vermittelt* die linke Gruppe sozusagen zwischen den Bildbetrachter/innen und der rechten Gruppe, erhält also eine Vermittlungsfunktion. Diese hat, genauer betrachtet, unterschiedliche Komponenten. Zunächst können wir die Komponente der Vermittlung zwischen den *Generationen* identifizieren.

#### *Die Vermittlung zwischen den Generationen*

Indem die rechte Gruppe der 20-30-Jährigen ins Zentrum der Aufmerksamkeit gerückt wird, ist es insbesondere diese Generation, die durch dieses Werbefoto adressiert wird. Da aber auch Angehörige anderer Generationen – hier von den unter Zwanzigjährigen bis hin zu den 60 bis 65-Jährigen – dem hier propagierten Stil oder Lifestyle nicht befremdlich, sondern anerkennend und sehr freundlich gegenüberstehen und selbst einige Elemente des Burberry Style aufweisen, versichern sie den Bildbetrachter(inne)n und potentiellen Kund(inn)en, dass diese – sollten sie sich für diesen (generationsspezifischen) Stil entscheiden – auch die Zustimmung und das Wohlwollen anderer Generationen auf ihrer Seite haben. Damit erhält der Burberry Style das Attribut, zugleich modern und zeitlos zu sein.

#### *Die Alltagsvermittlung oder Alltagskontextuierung*

Zur Funktion der Vermittlung zwischen den Generationen kommt aber noch eine weitere Vermittlungsfunktion hinzu, diejenige der Vermittlung zwischen Pose bzw. Hyperritualisierung und Alltag, also die Funktion der Kontextuierung im Alltag, der Veralltäglichen. Denn die linke Gruppe wirkt (obschon auch sie die Kleidung von Burberry trägt) von ihrer Bekleidung bis hin zur gestisch-mimischen und körperlichen Präsentation weniger gestylt und weniger ent-persönlicht und somit weniger posenhaft. Sie macht (mit Ausnahme des Mannes im Hintergrund) einen eher alltäglichen Eindruck, vermittelt somit zwischen der rechten Gruppe und dem Betrachter.

Dies lässt sich noch einmal genereller formulieren: Die stereotypisierende und ent-individualisierende Stilisierung und somit das Posieren wirken – obschon für die Werbung unumgänglich – befremdlich, insbesondere dann, wenn Individualität zur Pose wird. Um diese Befremdung, diese Ent-Alltäglichen zu vermitteln, kann die Werbung mit der ironischen Brechung arbeiten (vgl. dazu meine Interpretation in Bohnsack 2007b). Dies geschieht hier nicht. Stattdessen wird in diesem Fall auf dem Bild selbst eine vermittelnde Instanz zwischen dem zentralen Werbeträger teil des Fotos (der rechten Gruppe) und dem Betrachter positioniert.

Trotz der Ausdifferenzierung in zwei Gruppen zerfällt das Bild planimetrisch nicht in zwei Teile, weil die Frau in der Mitte rechts ihrerseits zwischen der vermit-

telnden und der vermittelten Gruppe vermittelt: Dies ist zunächst in ihrem Kleidungsstil (weißer Rock, geblünte Jacke) fundiert und in ihrer Ähnlichkeit zur stehenden jungen Frau hinsichtlich der Mimik und Frisur wie auch hinsichtlich des Blicks auf den Betrachter.

Diese Mittlerposition der knienden Frau ist auch in der Planimetrie fundiert: Sie gehört planimetrisch zu beiden Gruppen, da sich die Kreise (bzw. Kreis und Ellipse) bei ihr überschneiden. Ihre Zugehörigkeit zur größeren Gruppe wird insofern bestätigt, als sie mit ihrem Oberkörper dort hingewandt ist, und insofern, als sie kniet und somit eine Verbindung schafft zu den Kindern und den Angehörigen der größeren Gruppe, die teilweise auf dem Boden sitzen oder liegen. Auch durch ihr Alter stellt sie eine Verbindung zu den älteren Angehörigen der linken Gruppe her. Mit ihrer knienden Haltung schafft sie zugleich auch eine Verbindung zu den am Boden sitzenden Kindern. (Mit diesen verbindet sie auch noch einmal ein gesonderte planimetrische Komposition: eine weitere Ellipse; ohne Abb.).

Die multiple Vermittlungsfunktion der knienden Frau ist zusätzlich noch darin fundiert, dass sie direkt, d.h. auf herausfordernde Weise, Blickkontakt zum Bildbetrachter aufnimmt. Sie ist dadurch, dass sie vielfältig verortbar ist bzw. in vielfältige Relationen eingebunden ist, eine Figur, die sich für mannigfaltige Identifikationen und Projektionen eignet. – Diese Offenheit der Identifikation betrifft der Tendenz nach insgesamt die abgebildeten Personen und zwar aufgrund der für das Bild konstitutiven Irritationen hinsichtlich der Einschätzung der Verwandtschaftsverhältnisse wie hinsichtlich der Altersschätzungen der einzelnen Personen.

#### 4.2.3 Bild-Text und Bild-Logo

Nach dem Prinzip der Suspendierung des textlichen Vor-Wissen interpretieren wir auch den zum Bild gehörenden Text und ggf. das Logo erst *nach* der eigentlichen Bildinterpretation. Bild-Text ist hier zum einen der Schriftzug „Burberry“, darunter „London“. Darunter wiederum finden wir in der amerikanischen Version: „9 East 57th Street New York“ und in der russischen Version: „Москва, Столешников пер., 10; ГУМ, Красная площадь, 3“ (Moskau Stoleschnikov-Gasse 10 GUM Roter Platz 3).

Parallel zur Heftfalz bzw. Bildmittelsenkrechten findet sich – von unten nach oben geschrieben – in beiden Versionen der Schriftzug: „Burberry Limited 2005 800 284 8480 Burberry Com.“ In der russischen Version fehlt die Nummer nach der Jahreszahl.

Der Schriftzug „Burberry“ bzw. „Burberry London“ ist zugleich das Logo der Firma. Er ist in die rechte Ellipse integriert bzw. ruht diese auf ihm. Dies bestätigt, dass diese Gruppe der eigentliche Werbeträger ist.

#### 4.2.4 Zusammenfassung

Zusammenfassend lassen sich also folgende Botschaften des Bildes unterscheiden. Zunächst diejenigen, welche weniger den propagierten Lifestyle selbst, als vielmehr dessen Vermittlung und Verortung betreffen.

#### Vermittlung und Verortung des propagierten Lifestyle

Diese Funktionen der Vermittlung und Verortung werden wesentlich durch das Arrangement bzw. die Relation der beiden unterschiedlichen Gruppen – der Gruppe der Betrachtenden und der Betrachteten – zueinander geleistet: Der hier zu vermittelnde Lifestyle wird – in der Relationierung mit anderen abgebildeten Generationen – *generationsspezifisch* verortet: Primärer Adressat ist eine Gruppe bzw. Alterskohorte, welche sich zwar zum einen noch in der Jugend- oder vor-familialen Phase befindet, sich aber allmählich (noch unentschieden) an der Eltern- und Ernährerrolle orientiert.

Zugleich wird diese Gruppe *in andere Generationen* integriert. Der betrachteten Gruppe wird von Seiten der betrachtenden Gruppe, der Angehörigen anderer Generationen, eine freundliche Aufmerksamkeit zuteil. Diese – hier sowohl die 60- bis 65-Jährigen bis hin zu den unter Zwanzigjährigen – stehen dem propagierten Stil und Lifestyle nicht befremdlich, sondern wohlwollend gegenüber. Indem die Bindung an die ältere Generation im Bild verankert ist, kann dies auch als ein Element einer *Traditionsbindung* und einer Sicherung von Kontinuität im Bereich der Tradition des Burberry Style und als Element seiner Zeitlosigkeit gelesen werden.<sup>43</sup> Darüber hinaus repräsentiert die betrachtende Gruppe aber nicht nur die generationsspezifische Kontextuierung und die Bindung an Tradition, sondern auch eine Veralltäglichsung, eine *Alltags-Kontextuierung*.

Während die linke Gruppe die Funktion der Vermittlung zwischen der rechten Gruppe und den Bildbetrachter(inne)n hat, vermittelt die Frau im Zentrum (Kate Moss) noch einmal zwischen der linken und der rechten Gruppe.

#### Komponenten des propagierten Lifestyle

Die *zentrale* Komponente des hier propagierten Lifestyle lässt sich – anknüpfend an die oben bereits dargelegte Interpretation der rechten Gruppe – folgendermaßen formulieren: Der Burberry Style schafft *Zugehörigkeit und Zusammengehörigkeit bei gleichzeitiger Wahrung der Individualität*. Die zentrale Botschaft hat also – und dies ist im Sinne von Imdahl (1996a: 107) typisch für eine im Medium des Bildes, der Ikonik vermittelte tiefer liegende Semantik (bzw. für deren sprachliche Fassung) – den Charakter einer „Sinnkomplexität des Übergegensätzlichen“.

Eine *weitere* Komponente, die zugleich aber auch die Vermittlung dieses Lifestyle betrifft, lässt sich so formulieren: Der Burberry Style *vermittelt zwischen den Generationen und somit zwischen Mode und Tradition*. Die Vermittlung zwischen den Generationen dokumentiert sich in der Akzeptanz des Stils der 20 bis 35-Jährigen durch die anderen – insbesondere die älteren – Generationen. Die Vermittlung zwischen den Generationen dokumentiert sich darin, dass der Burberry Style, wenn

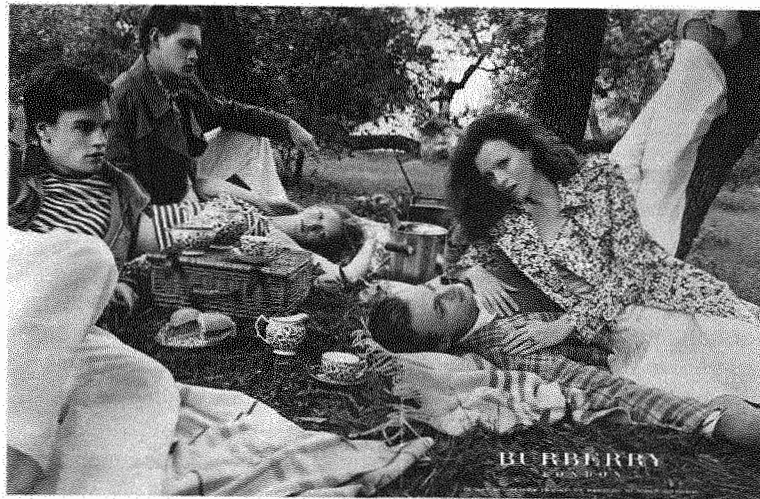
43 Als weitere Komponente kommt hier ein Element des Konservatismus hinzu durch den nahezu in der Bildmitte positionierten 60 bis 65-jährigen Mann im dunklen Anzug mit Weste, Krawatte und weißem Stecktuch. Er nimmt nicht nur durch seine Kleidung sondern, wie dargelegt, auch durch seine choreografische Position eine Sonderstellung ein. Beides – die Förmlichkeit der Kleidung wie auch die ‚dezenste‘ Positionierung im Hintergrund – vermitteln den Eindruck, dass es sich hier um einen Butler handelt. Dies unterstreicht die Atmosphäre von Luxus und stellt zugleich eine gewisse Verbindung zu konservativen Stilelementen her, öffnet sich diesen gegenüber.

auch in einer anderen Variante, sich ebenso in der Bekleidung anderer Generationen wieder findet.

#### 4.2.5 Komparative Analyse: der propagierte Lifestyle im Werbefoto eines anderen nationalen Marktes

Wenn wir dieses Foto aus der russischen Ausgabe der Zeitschrift Vogue 2005 mit einem Foto aus der deutschen Ausgabe dieser Zeitschrift vergleichen (Abb. 7), so zeigen sich zunächst deutliche Unterschiede:<sup>44</sup> Das Foto für den russischen Markt ist daran orientiert, auf dem Wege von Kontextuierungen auf die Akzeptanz seitens der Rezipient(inn)en der Werbung hinzuwirken: durch eine Alltagskontextuierung des propagierten Lifestyle sowie durch eine Kontextuierung im Rahmen der abgebildeten Betrachter/innen anderer Generationen, welche ihr Wohlwollen demonstrieren.

Abbildung 7: Burberry London. Entnommen aus: Vogue 2005 (deutsche Ausgabe)



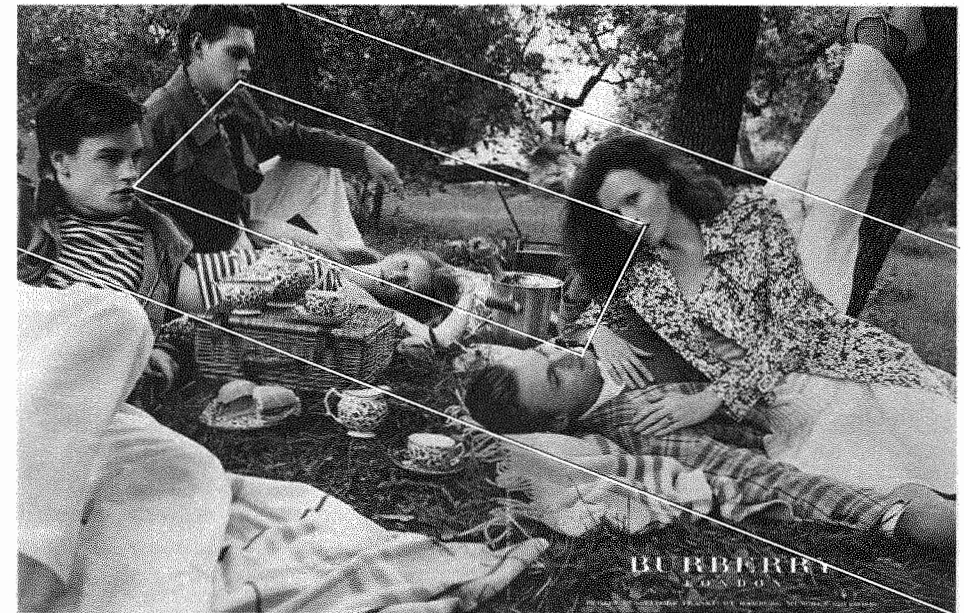
Auf dem Foto für den deutschen Markt finden wir diese Kontextuierungen nicht. Es finden sich keine abgebildeten Betrachter/innen, die zwischen den Posen der Betrachteten und den Bildbetrachter/innen vermitteln würden. Es ist hier – im deutschen Kontext – offensichtlich zum einen nicht notwendig, letztere behutsam zur Präsentation des Lifestyle hinzuführen. Zum anderen fehlt hier die durch die Relationierung der beiden Gruppen demonstrierte Einbindung in den *Generationen- und*

<sup>44</sup> Die Schritte der vor-ikonografischen und ikonografischen Interpretation können hier nicht abgedruckt werden. Sie sind selbstverständlich auch bei den Vergleichsfällen notwendiger Bestandteil der Fallanalyse.

*Familienzusammenhang*, welchem im deutschen Kontext offensichtlich nicht so viel Wert beigemessen wird.

Jenseits dieser – möglicherweise kulturell bedingten – *Unterschiede* zwischen den nationalen Angeboten zeigen sich jedoch *Übereinstimmungen* hinsichtlich der zentralen Komponente des propagierten Lifestyle: Die Sozialität des Bildes ist dadurch charakterisiert, dass die Personen auf dem Bild zwar körperlich relativ nahe beieinander positioniert sind (sich in einem Fall auch berühren) und vor allem durch die Planimetrie in eine deutliche Zugehörigkeit gebracht werden (Abb. 8), gleichwohl aber keineswegs kommunikativ einander zugewandt sind:

Abbildung 8: Burberry London. Entnommen aus: Vogue 2005 (deutsche Ausgabe): Planimetrie (Linien von R.B.)



Obschon die Frau rechts auf den Oberkörper des liegenden Mannes gelehnt ist und ihn mit beiden Händen berührt, schaut der liegende Mann an ihr vorbei und die Frau sieht nicht zu ihm, sondern nimmt – aufmerksam und beobachtend, aber auch eine gewisse Aufmerksamkeit erheischend – Blickkontakt mit dem Betrachter auf. Sie sieht in die Kamera. Zudem hat es den Eindruck, dass sich die Frau mehr auf dem Mann abstützt, als dass sie seine Nähe sucht. Auch die junge Frau links oben, die auf dem Rücken zwischen den beiden jungen Männern liegt, schaut nicht auf diese beiden Männer bzw. auf einen von ihnen. Sie blickt eher in Richtung des liegenden Mannes zur rechten, der seinerseits in die entgegen gesetzte Richtung schaut. Der eine der beiden jungen Männer links schaut in die Ferne, der andere teilnahmslos zu Boden. Auch hier fehlt (wie in der rechten Gruppe des Fotos für den russischen Markt) in evidenter Weise nicht nur der Blickkontakt, sondern die Personen nehmen auch hier einander nicht einmal optisch wahr.

Es ergibt sich somit hier ebenfalls ein Kontrast zwischen dieser Nicht-Wahrnehmung einerseits sowie der körperlichen Nähe der beteiligten Personen und ihrer planimetrischen Zugehörigkeit andererseits. Wir haben es also auch hier mit jener Botschaft zu tun, welche sich durch die bereits identifizierte Übergegensätzlichkeit einer *Zugehörigkeit* und *Zusammengehörigkeit bei gleichzeitiger Wahrung der Individualität* auszeichnet.<sup>45</sup>

Die Botschaft, mit der hier für Kleidung geworben wird, bezieht sich damit auf eines der zentralen Probleme der Identitätskonstitution, der Selbstpräsentation im Alltag, welches Habermas (1973: 230) mit Bezug auf Goffman bezeichnet hat „als das paradoxe Verhältnis, dem Anderen gleich und doch von ihm absolut verschieden zu sein“, indem der einzelne „gleichzeitig seine soziale Identität und seine personale Identität wahr“. – Fast interessanter noch als die Botschaft selbst allerdings ist das, was sich im Foto darüber offenbart, wie diese Botschaft durch die Werbung vermittelt bzw. nicht vermittelt werden kann. In dem der Werbung eigenen Medium der Pose kann Individualität nur ex negativo, d.h. auf dem (Um-) Weg fehlender sozialer Bezogenheit dargestellt werden.

Die zentrale Botschaft dieser Werbung, nämlich dass der Burberry Style einen Beitrag zur Bewältigung der Paradoxie von Zugehörigkeit und Individualität zu leisten vermag, wird hier im Medium des Bildes transportiert. Hierin bestätigt sich, dass das Bild zur Vermittlung von Ambiguitäten prädestiniert ist (siehe auch Kap. 3.5 sowie 4.3). Ein derartiges Verständnis der tiefer liegenden Semantik des Bildes findet sich bei Roland Barthes (1990: 54) dort, wo er den „stumpfen Sinn“ charakterisiert, und bei Umberto Eco dort, wo dieser sich mit der „ästhetischen Botschaft“ des Bildes befasst, deren tiefer gehende Semantik durch eine „produktive Ambiguität“ (Eco 1994: 146) gekennzeichnet sei.

In einer empirisch evidenten und für die qualitative Methodik anschlussfähigen Weise ist diese Leistung, dieses Potential des Bildes, welches – im Kontext der Werbung – zugleich sein Beeinflussungs-Potential ausmacht, bei Imdahl (1996a: 107) aufgewiesen worden. Er sieht, wie in Kap. 3.4 bereits dargelegt, das Spezifische der tiefer gehenden Semantik des Bildes, des „ikonischen“ Sinnes, in einer „Sinnkomplexität des Übergegensätzlichen“.

45 Es ist wohl offensichtlich, dass eine derartige Botschaft den Produzenten dieses Fotos nicht als schlicht bewusste Intention, also als „intendierter Ausdruckssinn“ (Mannheim 1964), zugerechnet werden kann. Vielmehr ist diese Botschaft als das Produkt eines Habitus zu verstehen, welcher im Sinne von Bourdieu (1976: 207) weder als ein vollkommen bewusster noch als ein vollkommen unbewusster modus operandi einzustufen ist. Loïc D. Wacquant (1996: 41) spricht hier auch von der „Intentionalität ohne Intention“. Diese „objektive Intention, die sich niemals auf die bewußte Absicht des Künstlers beschränkt, ist eine Funktion der Denk-, Wahrnehmungs- und Handlungsmuster, die der Künstler seiner Zugehörigkeit zu einer bestimmten Gesellschaft oder Klasse verdankt“ bemerkt Bourdieu (1970: 154) mit Bezug auf Panofsky. Die Frage, welcher Art diese Zugehörigkeit im Falle unseres Fotos ist, welche corporate identity oder welcher ‚corporate habitus‘ hier ihren Ausdruck finden (etwa derjenige der Firma Burberry?), lässt sich auf dem Wege der Interpretation des Werbefotos allein nicht beantworten.

### 4.3 Exemplarische Interpretation von Familienfotos und Methodentriangulation

Während in Kap. 4.2 professionelle Fotos aus dem öffentlichen Bereich Gegenstand der Analyse waren, befassen wir uns in diesem Kapitel mit Amateurfotos aus dem privaten Bereich. Es handelt sich um Interpretationen von Fotos zweier Familien. Dabei steht die eine, die wir Familie *Schiller* genannt haben, im Zentrum, und die andere, die Familie *Telchow*, hat eher den Charakter eines Vergleichshorizonts. Die Fotointerpretation wird mit der Interpretation von Tischgesprächen und Gruppendiskussionen kombiniert. Wir haben es also mit einer sogenannten Methodentriangulation (siehe u.a. Flick 2008; Marotzki 1999) zu tun. Diese kann nur dann erfolgreich sein, wenn die unterschiedlichen methodischen Zugänge innerhalb eines übergreifenden methodologischen Rahmens integriert werden können (dazu. Bohnsack et al. 1995: Kap. 7), wie dies auf der Grundlage der dokumentarischen Methode möglich ist.

Die hier interpretierten Familienfotos sind im Jahre 2002 im Rahmen einer Untersuchung mit dem Titel „Erziehung und Tradition. Tradierungsprozesse in Familien“ erhoben, d.h. von den Familien aus deren Fotosammlung für uns ausgewählt worden.<sup>46</sup> Eine der zentralen Fragestellungen des Projekts, der ich auch im Folgenden nachgehen werde, zielt auf den familienspezifischen Habitus, wie er u.a. in Kommunikations-, Sozialisations- und Erziehungsstilen seinen Ausdruck findet. Dabei geht es zunächst darum, die Stile oder Komponenten des Habitus in der Interaktion zwischen den Eltern und ihren Kindern zu rekonstruieren, um dann auch der Frage nachzugehen, ob und inwieweit Elemente dieser Stile bereits in der Herkunftsfamilie der Eltern, also in den Familien der Großelterngeneration, zu beobachten sind und somit möglicherweise tradiert wurden.

Die Analyse stützt sich neben der Interpretation von Familienfotos auf die Auswertung von Tischgesprächen (von Gesprächen während des Mittag- und/oder Abendessens und teilweise des Frühstücks, die von den Familien selbst aufgezeichnet wurden) und Gruppendiskussionen mit den Eltern unter Anwesenheit eines der beiden Großelternpaare (mütter- oder väterlicherseits). Auch die Tischgespräche und Gruppendiskussionen sind im Jahr 2002 erhoben worden.

#### *Atheoretisches, konjunktives und kommunikativ-generalisiertes Wissen*

Die kollektiven familienspezifischen Stile sind uns entweder auf dem Wege über die Rekonstruktion der in den Gesprächen, Diskussionen und Fotos implizierten Wissensbestände, über die Rekonstruktion des *impliziten* Wissens also, zugänglich oder über die Beobachtung des *inkorporierten* Wissens, zu dem wir lediglich über die Interpretation der Fotos einen validen Zugang finden. Implizites wie inkorporiertes Wissen fasse ich unter Mannheims (1964a) Begriff des *atheoretischen* Wissens zusammen, welches als handlungsleitendes Wissen den Habitus oder auch Orientierungsrahmen eines Individuums oder Kollektivs konstituiert (vgl. dazu Kap. 2).

46 Das ursprüngliche Vorhaben (Bohnsack/Gebhardt/Kraul/Wulf 2001), in dem Familien aus Brandenburg und Winzerfamilien aus dem Raum Koblenz in eine komparative Analyse einbezogen wurden, ist als Pilotprojekt von der DFG für ein Jahr gefördert worden.

Mit Bezug auf das kollektive handlungsleitende Erfahrungswissen einer Gruppe, eines Milieus oder Generation spricht Karl Mannheim (1980) vom „konjunktiven Erfahrungsraum“. Ein konjunktiver Erfahrungsraum par excellence ist derjenige der Familie (mehr noch derjenige der Mutter/Kind-Beziehung). Die Familie ist aber nicht nur konjunktiver Erfahrungsraum, nicht nur Milieu, sondern auch Institution, ein Gefüge von gesellschaftlich generalisierten Rollenerwartungen.

Während wir alle über ein derartiges *kommunikativ-generalisiertes Wissen* (Mannheim 1980; Bohnsack 2001: 329ff.) verfügen (bspw. ein Wissen um generelle stereotype Rollenerwartungen an Mütter und Väter), teilen nur diejenigen, welche in die Gemeinsamkeiten und Traditionen einer konkreten familialen Praxis miteinander eingebunden sind, darüber hinaus auch ein *konjunktives Erfahrungswissen*. Wir haben es also mit einer Doppelstruktur, einer „Doppeltheit“ (Mannheim 1980: 296) alltäglicher Erfahrungs- und Begriffsbildung, beispielsweise im Hinblick auf den Begriff der Familie, zu tun. Denn Bezeichnungen und Äußerungen haben – ganz generell – einerseits eine öffentliche oder gesellschaftliche und andererseits eine nicht-öffentliche oder milieuspezifische Bedeutung.

Bereits im Bereich der Textinterpretation reduzieren wir unser in die Interpretation eingebrachtes *Vor-Wissen* auf das kommunikativ-generalisierte Wissen, auf das wir bei der formulierenden Interpretation (notwendigerweise) zurückgreifen. Das konjunktive Wissen, bei dem wir uns in den Erfahrungsraum der Erforschten hinein begeben müssen, ist – im Zuge der reflektierenden – Interpretation am Text selbst zu erarbeiten und zu belegen. Vor-Annahmen hinsichtlich des konjunktiven Wissens gilt es zunächst zu suspendieren.

Im Bereich der Bildinterpretation gehen wir nun noch einen Schritt weiter. Vorab der Interpretation gilt es hier das konjunktive Wissen selbst dann zu suspendieren, wenn es in empirisch valider Form auf der Grundlage von Textinterpretationen gewonnen worden ist, da wir, wie u.a. bei Foucault, aber auch bei Imdahl ausgeführt worden ist (siehe: Kap. 3.3), dazu neigen, unser textliches Vor-Wissen in das Bild hineinzuprojizieren. Das hat die Konsequenz, dass im Bereich der Methodentriangulation die Verfahren der Textinterpretation bzw. deren Ergebnisse, hier: die Ergebnisse der Auswertung von Tischgesprächen und Gruppendiskussionen, erst nach vollzogener Bildinterpretation einbezogen werden. Das an die Bildinterpretation herangetragene Vor-Wissen reduziert sich somit, wie auch bereits in 4.1 dargelegt, auf die kommunikativ-generalisierten Wissensbestände.

#### *Kommunikativ-generalisiertes (ikonografisches) Vor-Wissen*

Die beiden hier analysierten Familien stammen aus dem Land Brandenburg und weisen, von ihrem Bildungs- und Ausbildungshintergrund her gesehen, viele weitere Gemeinsamkeiten auf: Der Sozialisationshintergrund aller Elternteile reicht in die DDR zurück, und sie sind in der Zeit von 1964 bis 1970 im Land Brandenburg geboren, die Kinder zwischen 1989 u. 1996. Zur Familie Schiller gehören neben dem Vater (geb. 1969) und der Mutter (geb. 1970) die Kinder Rebecka (geb. 1989), Tobias (geb. 1990) und David (geb. 1996). Die beiden älteren Kinder besuchen (zum Zeitpunkt der Erhebung der Tischgespräche und Gruppendiskussionen) die Grundschule, der jüngere Sohn den Kindergarten. Zur Familie Telchow gehören neben der Mutter (geb. 1968)

und dem Vater (geb. 1964) die beiden Söhne Daniel (geb. 1993) und Olaf (geb. 1996). Der jüngere besuchte den Kindergarten und der ältere die dritte Klasse der Grundschule.

Alle Eltern haben den Abschluss der zehnten Klasse erworben, eine Ausbildung im Bereich erzieherischer Berufe absolviert und sich zum Teil noch über eine Hochschulausbildung weiterqualifiziert. Alle sind zum Zeitpunkt der Erhebung beruflich in pädagogischen Einrichtungen tätig.

Was die *Herkunftsfamilien der Schillers* anbetrifft, so hat Frau Schiller keine Geschwister, ihr Ehemann einen Bruder und eine Schwester. Die Großelterngeneration besteht zum einen aus der Mutter von Frau Schiller (geb. 1935) und ihrem zweiten Ehemann (geb. 1935). Ihren ersten Ehemann hat die Mutter von Frau Schiller 1970 geheiratet und wurde 1975, als Frau Schiller 5 Jahre alt war, von ihm geschieden. Zu ihm besteht nur noch wenig Kontakt, sowohl von Frau Schillers Seite als auch ihrer Mutter. Zur Großelterngeneration gehört weiterhin die Mutter von Herrn Schiller (der Vater ist 1997 verstorben). Die Großelterngeneration verfügt insgesamt über eine nicht-akademische Berufsausbildung und war in entsprechenden Berufen tätig (Buchhaltung, Tontechnik, Haushaltsführung). Bis auf den Ehemann von Frau Schiller hat sich die Familie bereits zu DDR-Zeiten und bis heute der katholischen Glaubensgemeinschaft zugeordnet.

Was die *Herkunftsfamilien der Telchows* anbetrifft, so hat Herr Telchow keine und Frau Telchow vier Geschwister, zu denen ein enger Kontakt besteht. Die Großelterngeneration bestand zum Zeitpunkt der Erhebung aus der Mutter von Frau Telchow und Herrn Telchows Vater. Frau Telchows Mutter (geb. 1933) hat den Beruf der Kindergärtnerin erlernt und später als Horterzieherin gearbeitet. 1954 hat sie ihren Mann geheiratet, der 1978 verstorben ist. Herrn Telchows Vater ist 1928 geboren und war als Lehrer tätig. Herrn Telchows Mutter ist 1992 verstorben. Wie die Eltern sind auch die Großeltern der Familie Telchow konfessionslos.

Während wir – als Suchstrategie – bei der Auswahl der beiden Familien (im Hinblick auf die Sozialisation in der DDR-Kultur) kulturspezifische wie auch bildungsmilieutypische Gemeinsamkeiten vermuten können, lassen sich im Hinblick auf die lediglich bei der Familie Schiller zu beobachtende Bindung an die katholische Religionsgemeinschaft insbesondere vor dem Hintergrund der Familiengeschichte in der DDR milieuspezifische Differenzen vermuten.

#### *4.3.1 Zur Auswahl der Fotos*

Die Familien wurden in mündlicher und schriftlicher Form während der Kontaktaufnahme um Fotos gebeten. Schriftlich wurde ihnen mitgeteilt, dass für uns alle Bilder interessant seien, welche von ihnen „selbst als wichtig für Ihre Familiengeschichte“ angesehen würden. Darüber hinaus wurde gebeten „um Fotos von Hochzeiten und Kommunionen, Konfirmationen, bzw. Jugendweihen sowie von Einschulungen (...), insbesondere von der Kommunion (bzw. Jugendweihe, Konfirmation) und Hochzeit der Großelterngeneration. Uns interessieren sowohl ‚offizielle‘ Fotos von professionellen Fotografen als auch selbst gemachte Schnappschüsse.“ Dabei haben wir eine Anzahl von ca. 30 Bildern genannt. Familie Telchow hat schließlich insgesamt 52 Fotos für uns ausgewählt und Familie Schiller 29.

Wenn uns die Familien ihre Fotos zur Verfügung stellen, so haben diese mehrere Etappen der Selektivität im Rahmen des familienspezifischen Erfahrungsraums durchlaufen: die Selektivität durch die Wahl des Momentes der Ablichtung, ggf. die technische Bearbeitung durch die familialen Amateurfotograf(inn)en nach der Ablichtung, dann die Auswahl der Fotos durch Angehörige der Familie, die sich darin vollzieht, dass nur bestimmte Fotos für die weitere Aufbewahrung (in elektronischer Form oder als Papierbilder) ausgewählt werden. Schließlich stellt ein weiterer wesentlicher Schritt in diesem mehrstufigen Selektionsprozess jene Auswahl dar, welche die Familienmitglieder (hier in erster Linie die Eltern) aus diesem Bestand (dem ‚Fotoalbum‘ im materiellen oder metaphorischen Sinne) für die Forscherinnen und Forscher vornehmen.

Die im Moment der Ablichtung in ihrer je spezifischen Simultaneität hergestellte selektive Ausrichtung des Sujets, also der Szenerie und die selektive Ausrichtung der Komposition (siehe auch: Kap. 3.9) ist nur der erste Schritt, in dem sich die stilistischen Präferenzen der abbildenden und der abgebildeten Bildproduzent(inn)en dokumentieren, die in der Regel beide zum Erfahrungsraum der Familie gehören und somit deren Orientierungsrahmen oder Habitus repräsentieren. Die stilistischen Präferenzen verdichten und steigern sich in ihrer Selektivität zunehmend, indem aus den ‚Schnappschüssen‘ dann – wiederum durch Mitglieder der Familie und somit innerhalb ihres Erfahrungsraums – einige für ein ‚Fotoalbum‘ ausgewählt und somit *authentisiert* werden. In einer weiteren Steigerung der Selektivität und Authentisierung und der damit verbundenen Verdichtung der stilistischen Präferenzen wird aus dem Bestand des ‚Fotoalbums‘ schließlich noch einmal für die Selbstpräsentation gegenüber den Forscher/innen ausgewählt. In den für die Forscher/innen ausgewählten Bildern findet also eine zunehmende Verdichtung der stilistischen Präferenzen des Erfahrungsraums der Familie und somit des familialen Orientierungsrahmens oder Habitus seinen Ausdruck (siehe dazu auch: Nentwig-Gesemann 2007b: 225).

Nach Durchsicht haben wir die Fotos, die, wie von uns gewünscht, alle im weitesten Sinne die Familie als Sujet haben, in folgende Rubriken bzw. Untersujets eingeteilt:

*Rituale und Übergänge:*

Familie Telchow: eigene Trauung (2); Einschulung des älteren Sohnes Tobias (3); Sylvester-Serie 1996; 1997; 1998; 1999; 2000; 2002 (6)

Familie Schiller: Trauung (1); Polterabend (2); Kinder als Neugeborene (2); Kommunion Rebecka u. Tobias (22)

*Urlaub, Ausflug, kleine Feiern:*

Familie Telchow: Gartenfest (1); Sommer in den Bergen (6); Winter in den Bergen (6); an der See (2)

Familie Schiller: im Kindermuseum (1); Grillabend (1); Ponyreiten (1); Radtour (2); in den Bergen (2); im Freizeitpark (1)

*Alltag mit Kindern:*

Familie Telchow: (23); Familie Schiller: (13)

*Alltag ohne Kinder (nur Eltern):*

Familie Telchow: keine

Familie Schiller: (2)

*Familiengeschichte (Herkunftsfamilie/Großelterngeneration):*

Familie Telchow: Trauung der Eltern von Frau Telchow (1); Kommunion der Schwester von Frau Telchow (1); Trauung der Schwester von Frau Telchow (1); nicht identifizierbar (1)

Familie Schiller: Kommunion Frau Schiller (1)

Die Auswahl aus diesen Fotos für eine intensive Interpretation orientiert sich an den Erfordernissen einerseits einer *horizontalen* komparativen Analyse von Stilelementen der Familie Schiller mit denjenigen der Familie Telchow. Dabei gilt es insbesondere das (Generationen-) Verhältnis von Eltern und Kindern bzw. das Verhältnis von Eltern- und Kinderwelt innerhalb der Familie in den Blick zu nehmen. Andererseits wird im Sinne einer *vertikalen* komparativen Analyse innerhalb der jeweiligen Familientraditionen das Verhältnis der beiden Familien, genauer der Mütter, zu deren Herkunftsfamilie, also der Großelternfamilie, beleuchtet. In beiden Dimensionen der komparativen Analyse – der horizontalen wie der vertikalen – wird aber auch den Geschlechterverhältnissen und dem Verhältnis von Kinder- und Erwachsenenwelt Aufmerksamkeit gewidmet.

Ausgewählt für eine erste vergleichende Interpretation der Familien Telchow und Schiller wird zunächst je ein Foto nach dem Prinzip der *Fokussierung*. Im Sinne der dokumentarischen Methode gehen wir davon, dass die Grundstruktur des Falles, d.h. der modus operandi oder Habitus der Bildproduzent(inn)en, sich grundsätzlich in all deren Produkten zu dokumentieren vermag – allerdings mehr oder weniger deutlich bzw. mehr oder weniger leicht zugänglich. Somit geht es wesentlich darum, einen möglichst leichten oder direkten und in diesem Sinne forschungsökonomischen Zugang zu finden. Deshalb folgen wir – wie auch im Falle von Passagen im Bereich der Textinterpretation, also u.a. der Gesprächs- und Interviewanalyse – sozusagen als Suchstrategie dem Auswahlkriterium der *Fokussierung* (siehe auch: Bohnsack 2006c), indem wir solche Fotos auswählen, die sich durch besonders markante und somit auffällige Elemente der Formalstruktur oder auch der Szenerie oder auch durch Brüche und Diskontinuitäten in diesen Bereichen auszeichnen und somit (zumindest auf den ersten Blick) besonders aussagekräftig erscheinen (vgl. auch 6.3.2). Dies ist im Bereich der Bildinterpretation auch deshalb wichtig, da wir hier noch über vergleichsweise wenig Erfahrungen verfügen.

Im Bereich der Bildinterpretation betrifft die Fokussierungsqualität sowohl Auffälligkeiten im Bereich der formalen Strukturen der Performanz der *abbildenden* Bildproduzent(inn)en (Wahl der Perspektivität; die mit der Wahl des Ausschnitts, der Kadrierung, verbundene planimetrische Komposition) wie auch der Performanz und Szenerie der *abgebildeten* Bildproduzent(inn)en (szenische Choreografie, d.h. der interaktive Bezug aufeinander und die Positionierung im Raum, die Struktur der Gebärden, Auffälligkeiten der Bekleidung und der Requisiten).

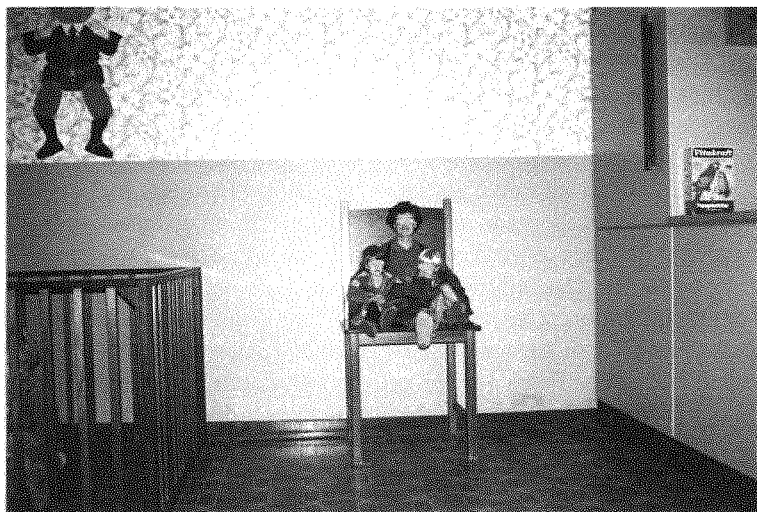
Die beiden ausgewählten Fotos „Museum“ aus der Familie Telchow und „Gartenfest“ aus der Familie Schiller fallen vor allem durch die Be- oder Verschiebung in den Proportionen (Museum) auf. Sie sind beide der Rubrik „Urlaub, Ausflug, kleine Feste“ zu zuordnen, in der jene Fotos versammelt sind, die hinsichtlich ihres Sujets aus dem Alltag herausfallen, ohne aber der Kategorie „Rituale und Übergänge“, also den ‚großen Festen‘, zuzugehören, welche für die

Geschichte und Tradierung der Familie und ihrer Beziehung zur Herkunftsfamilie aufschlussreich sind.

Den ‚großen Festen‘ werden wir uns mit den Fotos der Kommunion und Jugendweihe unter 4.3.5 u. 4.3.6 zuwenden. Die Fotos zu den Ritualen der Familientradition sind auch deshalb von besonderer Bedeutung, weil sie nicht schlicht eine Ablichtung von Ritualen darstellen, sondern, wie Erving Goffman (1979:10) bemerkt hat, als eine Verdichtung ritueller oder zeremonieller Akte gelten können, also selbst Bestandteil dieser Rituale sind und somit durch ihre ikonische Kraft ihrerseits Einfluss auf die Familientradition gewinnen.

#### 4.3.2 Familie Schiller: „Museum“

Abbildung 9: Foto der Familie Schiller „Museum“



Ausgewählt wurde dieses Foto, wie bereits angesprochen, aufgrund seiner Fokussierungsqualität, d.h. insbesondere aufgrund von Auffälligkeiten im Bereich der Formalstruktur oder der Szenerie (in Relation zu den anderen Fotos der Familie). Diese Auffälligkeiten betreffen hier vor allem die Proportionen in der Relation der abgebildeten Gegenständlichkeiten zu den abgebildeten Personen sowie Auffälligkeiten der Planimetrie, die den Eindruck vermitteln, das Bild würde planimetrisch zerfallen.

#### Formulierende Interpretation

##### Vor-ikonografische Interpretation

Aus Gründen des Umfangs ist die vor-ikonografische Interpretation hier nicht abgedruckt (siehe aber als Beispiel weiter unten die Interpretationen 4.3.4 und 4.3.6 sowie die Beispiele in: Bohnsack 2006a u. 2007b).

#### Ikonografische Interpretation

Aufgrund der Kontextinformationen, die uns die Familie zur Verfügung gestellt hat, wissen wir, dass dieses Foto während des Familienurlaubs 1995 aufgenommen wurde. Es handelt sich bei den abgebildeten Personen um Herrn Schiller, damals 26, und seine Kinder Rebecka, damals sechs, sowie Tobias, damals fünf Jahre alt.

Da die auf dem Bild abgebildeten Accessoires, also insbesondere die Einrichtungsgegenstände, in Relation zu den Personen einen vergrößerten Maßstab aufweisen, lässt sich vermuten, dass das Foto in einem Museum oder einer Freizeiteinrichtung aufgenommen worden ist, in dem die Größenrelation von Erwachsenen zu den Einrichtungsgegenständen jener Relation entspricht, wie sie für die Größenrelation von Kleinkindern zu diesen Accessoires typisch ist. Fotografarin, also abbildende Bildproduzentin, ist Frau Schiller.

#### Reflektierende Interpretation

##### Formale Komposition

##### Planimetrie

Der Bildmittelpunkt befindet sich nahezu im Zentrum der abgebildeten Personen-Gruppe, nahe des Kopfes der Tochter Rebecka, sodass die (Teil-) Familie hierdurch planimetrisch fokussiert ist.

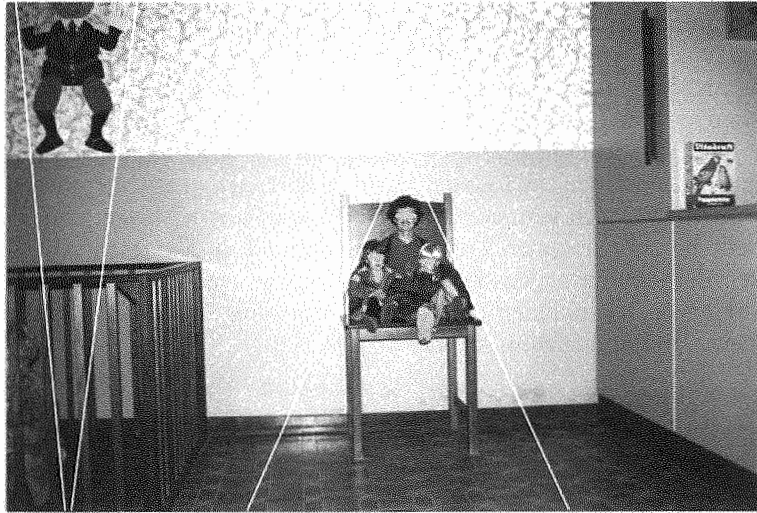
Das Bild wird insgesamt dominiert durch waagerechte und senkrechte Linien, die wie eine Art Gitterstruktur wirken: Zusätzlich zu den Karos des Fußbodens, den waage- und senkrechten Linien der Umrisse des (Einbau-) Schrankes auf der rechten Seite sowie des Stuhles in der Mitte sind es vor allem die waagerechte Linie der Rückwand (die Fußleiste und die Unterschiede der Wandgestaltung im deren oberem und unterem Teil), welche das Bild in gewisser Weise zerschneiden und die abgebildeten Personen eingrenzen. Verstärkt wird dies noch durch die Gitterstäbe des links ins Bild ragenden ‚Laufgitters‘ (da diese Linien auf den ersten Blick evident sind, erscheint es in diesem Fall überflüssig, sie in das Foto einzuzeichnen).

##### Szenische Choreografie

In gewisser Weise wird diese Gitterstruktur durch zwei (imaginative) gleichschenklige und spitzwinklige Dreiecke durchbrochen. Das eine erhalten wir, wenn wir über die Umrisse von Herrn Schiller mit den beiden Kindern auf dem Stuhl ein nach unten geöffnetes gleichschenkliges Dreieck zeichnen, das andere, wenn wir die Umrisse der Hände und Füße der ‚Quasi-Person‘ des Hampelmanns im Gewande eines Polizisten links oben durch Linien miteinander verbinden. Die abgebildeten Personen Hampelmann und Familie bzw. Hampelmann und Vater geraten durch die gegenläufige Bewegung (Öffnung) der beiden gleichschenkligen spitzwinkligen Dreiecke in gewisser Weise in ein antipodisches Verhältnis zueinander.

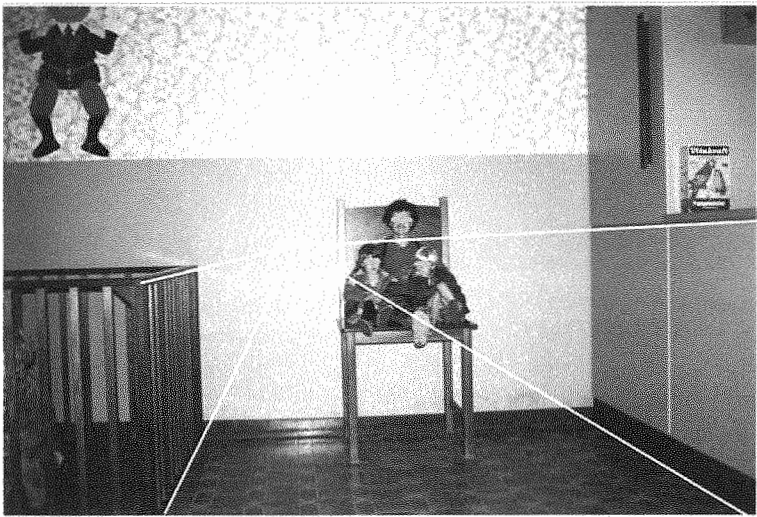
Die Familiengruppe ist so nahe zueinander gerückt, dass die Kinder nahezu auf dem Vater sitzen. Dadurch, dass die Kinder eng ihn gelehnt sind bzw. er sie an die Brust gedrückt und die Arme um sie gelegt hat, wird uns der Eindruck großer Nähe vermittelt.

Abbildung 10: Foto der Familie Schiller „Museum“: szenische Choreografie



### Perspektivität

Abbildung 11: Foto der Familie Schiller: „Museum“: Perspektivität



In Relation zum Bildmittelpunkt ist das perspektivische Zentrum, der Fluchtpunkt, weit nach links verschoben. Er befindet sich zwischen der Familie und dem ‚Gitterkasten‘, dem Laufgitter. Erst indem die Personen nicht perspektivisch fokussiert sind, wird es der abbildenden Bildproduzentin, der Fotografin, möglich, das Laufgitter wie auch den ‚Hampelmann-Polizisten‘ mit ins Bild zu nehmen. Die abbildende Bildproduzentin, Frau Schiller, hat hierauf – und damit auf die Abbildung

der Relation von Familie und Hampelmann-Polizist – offensichtlich Wert gelegt. Eine gewisse perspektivische Fokussierung der abgebildeten Familie ergibt sich dadurch, dass sich die Horizontlinie in Augenhöhe der Familienmitglieder befindet, genauer: zwischen der Augenhöhe der Kinder und derjenigen des Vaters.

### Ikonologisch-ikonische Interpretation

Aufgrund der szenischen Choreografie erscheint Herr Schiller als beschützender Vater. Zugleich steht das gesamte Arrangement dieser (Freizeit-) Einrichtung, in der Herr Schiller und die beiden Kinder (sowie auch Frau Schiller als Fotografin) sich befinden, für die Haltung, sich in die Welt der Kinder zu begeben und die Objekte dieser Welt aus ihrer Perspektive wahrzunehmen. Damit wird Herr Schiller als jemand abgebildet, der sich in die Welt der Kinder begibt, gleichzeitig aber beschützender Vater bleibt.

Ihren besonderen Charakter erhält diese beschützende Haltung dann noch einmal in Relation zu dem in einigen Hinsichten bedrohlichen Charakter der Umwelt. Da die Kinder auf dem Stuhl, dessen Sitzfläche in einer Höhe angebracht ist, die in etwa ihrer gesamten Körpergröße entspricht, gerade noch Platz finden, gewinnt man den Eindruck, dass die Umarmung des Vaters, diese vor dem Absturz zu bewahren sucht, indem er sie sozusagen unter seine ‚Fittiche‘ nimmt.

Auch erscheint die (Teil-) Familie in der Größe des Raumes ein wenig verloren oder isoliert. Und dennoch, obschon viel Raum vorhanden ist, wirkt die Gruppe, durch die planimetrische Gitterstruktur des Fotos eingegrenzt, beengt oder sogar eingesperrt, sodass ihr Beisammensitzen den Charakter eines Zusammenrückens in dieser – in doppelter Hinsicht – bedrohlichen Umgebung erhält.

Diese Bedrohlichkeit bzw. der Eindruck des Eingesperrt-Seins oder des potentiellen Eingesperrt-Werdens verstärkt bzw. bestätigt sich durch das Laufgitter als einer Art Käfig, welches links massiv ins Bild ragt bzw. dadurch ins Bild gerückt wird, dass das perspektivische Zentrum (gegenüber dem Bildmittelpunkt) weit nach links verschoben ist und sich zwischen der Familie und diesem Gitterkasten befindet. Die Umarmung des Vaters bewahrt die Kinder somit nicht nur vor dem Absturz und wirkt der Isolation im großen Raum entgegen, sondern scheint sie auch vor dem Eingesperrt-Werden in dem großen Käfig zu beschützen (in dem auch bereits – allerdings nur undeutlich zu erkennen – ein Kind eingesperrt ist). Dieser Gitterkasten kann als Metapher für eine kontrollierende und einengende Pädagogik gelten.

Bestätigt wird diese Metaphorik durch den links oben ins Bild ragenden Hampelmann im Gewande des Polizisten (oder des Polizisten in der Funktion des Hampelmanns). Diese Figur, zu dem Herr Schiller und seine Kinder den Gegenpol bilden, repräsentiert also jenen Typus des Erwachsenen, der den Kindern in einer Mischung aus Clownerie und autoritärer Kontrolle begegnet bzw. hinter der Fassade der Clownerie die disziplinierende Kontrolle verbirgt.

Die wesentlich durch die Planimetrie bedingte Eingrenzung und Einengung wird durch die szenische Choreografie ein wenig aufgebrochen. Herr Schiller formiert mit den beiden Kindern auf dem Stuhl ein nach unten geöffnetes gleichschenkliges spitzwinkliges Dreieck, welches den Gegenpol bildet zu dem nach oben geöffneten Dreieck des Polizisten in der Funktion des Hampelmanns oder des



Hampelmanns im Gewande des Polizisten, welches in das Laufgitter weist. Die beschützende Haltung des Vaters vermittelt somit zugleich Sicherheit gegenüber jenem Typus oder jenen Komponenten von Erziehung, wie sie durch den linken Teil des Bildes, den Gitterkasten und den Hampelmann-Polizisten repräsentiert werden.

Jenseits dieser beschützenden und Intimität vermittelnden Haltung wirkt Herr Schiller in dem vorgegeben Arrangement ein wenig lächerlich – und dies notwendigerweise. Denn dieses Arrangement bringt die erwachsenen Besucher dazu, sich in der Beziehung zu den Objekten, den „Aktanten“ im Sinne von Latour (1998), wie ein Kind, somit also *kindisch* zu verhalten bzw. zu positionieren (indem beispielsweise die Füße des Vaters nach vorne über die Sitzfläche hinausragen, da die Beine, wie bei einem Kleinkind, in der Sitzposition zu kurz sind, um sie in den Knien zu beugen). Damit wird durch das Arrangement der Einrichtung die Grenze zwischen einer Kindzentrierung im Sinne einer *Übernahme der kindlichen Perspektive* und einer Imitation des kindlichen Verhaltens, also einem *kindischen* Verhalten, systematisch verunklart bzw. wird beides in eins gesetzt.

Dies hängt auch damit zusammen, dass durch das Arrangement der Unterschied zwischen Erwachsenen und Kindern auf die Differenz in ihren Größenrelationen *reduziert* wird. Die Kinderwelt erscheint als Erwachsenenwelt im verkleinerten Maßstab, und die Kinder werden letztlich als kleine bzw. zu kleine und somit in gewisser Weise unzulängliche oder defizitäre Erwachsene konstruiert. Und Erwachsene, die sich in diesem Kontext wie Kinder verhalten, machen sich diese Defizite zu eigen und erscheinen somit ein wenig lächerlich.

Somit ist die gesamte Einrichtung und damit auch deren Besuch dazu angetan, den Erwachsenen Erfahrungen vor allem im Sinne kindlicher Unzulänglichkeiten oder *Defizite* zu vermitteln (indem die Erwachsenen bspw. die Erfahrung machen, nicht an den Griff der Kühlschranktür heran zu reichen und nur schwer auf einen Stuhl zu gelangen). Für die Kinder selbst haben diese Erfahrungen allerdings gar keinen Neuigkeitswert. Potentielle (neue) Erfahrungen für die Kinder finden sich allenfalls auf einer höheren Stufe der Reflexivität, indem diese etwas darüber lernen können, wie Erwachsene (insbesondere die für die Einrichtung Verantwortlichen) die Welt der Kinder konstruieren.

Dass diese Einrichtung für Kinder selbst wenig anregend ist, findet im Übrigen auch in Mimik und teilweise Gestik von Rebecka und Tobias seinen Ausdruck, indem beide (nach der Maskierung allerdings kaum mehr zu erkennen) gelangweilt bis abwartend schauen und Tobias dabei seinen Kopf (in einer demonstrativen Geste des Gelangweilt-Seins) mit der rechten Hand abstützt.

Insoweit als das Arrangement dieser Einrichtung einen spaßigen oder belustigenden Charakter aufweist, resultiert dieser wesentlich aus der Belustigung über Unzulänglichkeiten der Kinder, wenn sie sich in der Welt der Erwachsenen bewegen, also über den defizitären Charakter der kindlichen Welt sowie über das Kindische jener Erwachsenen, wenn sie sich in diesem Kontext ähnlich unbeholfen oder defizitär wie Kinder verhalten. Den übergreifenden oder primären Rahmen bildet hier also die Perspektive der Erwachsenen und deren Maßstäbe.

Frau Schiller hat ihren Mann und ihre Kinder somit auf einem Foto abgelichtet, in dem die komplexe Verhandlung über das Verhältnis von Kinder- und Erwachsenenwelt, bei der tendenziell der defizitäre Charakter der kindlichen Welt im Vor-

dergrund steht, durch die Intimität und den Schutz seitens des Vaters bewältigt wird. Es ist also der Vater, welchem hier die Funktion zukommt, bei Komplikationen hinsichtlich des Verhältnisses von Kinder- und Erwachsenenwelt durch Intimität und Nähe zu vermitteln.

### *Homologien zum Orientierungsrahmen der Tischgespräche*

Die Analyse der Tischgespräche und teilweise der Gruppendiskussionen betätigt dieses komplexe Verhältnis von Kinder- und Erwachsenenwelt einerseits sowie die (vermittelnde) Funktion des Vaters in diesem Kontext andererseits. Wie diese Diskurse zeigen, ergeben sich dort einige Spannungen im Verhältnis der Perspektive der Eltern zu denen der Kinder, die mit Defizitzuschreibungen an diese verbunden sind. Es werden dort hohe, d.h. der Welt der Erwachsenen entsprechende, Anforderungen und Maßstäbe hinsichtlich (zweckrationaler) biografischer Planung und Entscheidung an die Kinder gestellt, welchen diese, vor allem der ältere Sohn, nicht gerecht werden (können oder wollen). Dies bildet die Grundlage für Konfrontationen, bei denen der Vater eine vermittelnde Funktion einnimmt. Wobei gleichwohl insbesondere von Seiten der Mutter die Tendenz besteht, in entscheidenden Punkten den Kindern den Bezugsrahmen der Eltern in gewisser Weise ‚über zu stülpen‘.

Wir können uns bei der Erarbeitung von Homologien zum Familiengespräch in der Familie Schiller vor allem auf ein Tischgespräch stützen, dessen Transkript (*Schiller TG: Differenzierungsstunde*) und Auswertung an anderer Stelle (Bohnsack 2008a: Kap. 12.2) bereits umfassend abgedruckt worden sind.

Es geht um eine Entscheidung bzw. Weichenstellung in der weiteren schulischen Ausbildung des Sohnes hinsichtlich der Fächerwahl, für den diese eine Frage der „Lust“ darstellt. Ihm geht es um die Wahl des „Lustfaches“ (*Transkript Schiller TG: Differenzierungsstunde: 118*). Diese resultiert nicht primär aus der Neigung zu einem Fach, sondern ist im Wesentlichen in der persönlichen Beziehung zur Lehrperson wie auch in der Einbindung in das Peer-Netzwerk begründet.

Demgegenüber wird von Seiten der Eltern, insbesondere der Mutter, dieser Orientierungsrahmen vollständig und mit einer gewissen Rigidität abgelehnt. So sei es „total unentscheidend ob du Frau Ullrich (die Lehrerin; R.B.) gut findest oder nicht“ (*Transkript Schiller TG: Differenzierungsstunde 187 u. 190*). Der Mutter erscheint die Fächerwahl als eine Entscheidung von biografischer Tragweite. Dabei wird die schulisch relevante Entscheidungssituation von den Eltern im Rahmen biografischer Planung unter dem Aspekt betrachtet, welche der (zu wählenden) Schulfächer eine Zukunftssicherheit für die weitere Ausbildung, aber auch darüber hinaus, zu vermitteln vermag. Die schulische Ausbildung wird dezidiert in einen zweckrationalen Rahmen gestellt: „man geht in die Schule (...) um später n Beruf zu haben, in dem man gut arbeiten kann“ (*Transkript Schiller TG: Differenzierungsstunde 173-174*). Diese rigide Orientierung an biografischer Planungssicherheit, mit der sich die Mutter recht hart gegen eine durch „Lust“ motivierte Wahl des Sohnes wendet, findet auch in anderen Gesprächen und in der Gruppendiskussion mit Eltern und Großeltern ihren Ausdruck.

Im Einzelnen lässt sich an dem Tischgespräch zeigen, dass in der Familie Schiller zwei Orientierungsrahmen aufeinander treffen, die nicht miteinander vereinbart werden können. Auf der einen Seite haben wir den *Orientierungsrahmen der Eltern*, der im Wesentlichen von der Mutter getragen wird. Auf der anderen Seite haben wir den *Orientierungsrahmen des älteren Sohnes*, der zum Teil von der Tochter unterstützt wird. Im Zuge des konfrontativen Diskurses kommt es zu „Falschrahmungen“, indem die Mutter an entscheidenden Punkten insbesondere dem älteren Sohn Orientierungen unterstellt, die eher ihrem eigenen Rahmen entsprechen und somit Differenzen zwischen den Orientierungsrahmen verdeckt oder verwischt.<sup>47</sup>

Der Vater übernimmt weitgehend den Orientierungsrahmen der Mutter, leistet dabei aber eine Vermittlung zwischen diesem und demjenigen der Kinder (vgl. auch 4.3.6). Die Intimität in der Beziehung des Vaters zu den Kindern, wie sie im Foto ihren Ausdruck findet, zeigt sich auch in den Tischgesprächen. So gelingt es dem Vater sogar, auch den jüngsten Sohn, der dem Gespräch propositional (inhaltlich) kaum folgen kann, durch den Bezug auf dessen Prosodie (Intonation und Satzmelodie) performativ in den Diskurs zu integrieren (genauer dazu: Bohnsack 2008a: Kap 12.2)

In den Tischgesprächen wird somit deutlich, dass sich die Familie zwar auf der einen Seite in hohem Maße durch Offenheit auszeichnet, was die behandelten Themen und Probleme anbetrifft. Andererseits aber werden Differenzen in den Orientierungsrahmen in einer Weise bearbeitet, welche diese Unterschiede verdeckt und/oder der Perspektive der Kinder nur mit Einschränkungen Rechnung trägt. Dies resultiert daraus, dass in der Perspektive der Eltern die Kinder bei ihren Entscheidungen deren biografische Relevanz und Tragweite nicht zu überblicken vermögen. Sie erachten den eigenen Bezugsrahmen für überlegen und so selbstverständlich, dass mehr oder weniger unbemerkt die Orientierungen der Kinder letztlich nur noch in diesem eigenen Bezugssystem diskutiert, abgehandelt und gerahmt werden (Falschrahmung).

Diese Einschränkungen resultieren also aus der Sorge und Unsicherheit der Eltern hinsichtlich der Zukunft der Kinder, sodass es zu tiefer gehenden Auseinandersetzungen und Konfrontationen mit den Orientierungen und Argumenten der Kinder kommt. Durch diese tiefer gehende Auseinandersetzung werden insbesondere die Kinder dazu gebracht, ihre Orientierungen zur Explikation zu bringen. Dadurch kommt es zu einer für diese Familie typischen Nähe, einer „Nähe durch Konfrontation und Reflexion“, wie man dies im Sinne einer „Übergegensätzlichkeit“ in der Sprache von Imdahl (vgl. Kap. 3.5) formulieren könnte.

Es zeichnen sich somit deutliche Homologien ab zwischen dem Foto „Museum“ und den Tischgesprächen hinsichtlich der Komplexität im Bereich des Verhältnisses von Kinder- und Erwachsenenwelt und hinsichtlich der Konstruktion der kindlichen Welt als in gewisser Weise unzulänglich und defizitär bei gleichzeitig – insbesondere vom Vater getragener – Intimität und Nähe.

47 Dies lässt sich als eine Form eines ‚machstrukturierten‘ Diskurses verstehen, dessen Erfolg davon abhängig ist, dass die Rahmeninkongruenz und die Falschrahmungen verdeckt oder latent bleiben (siehe auch Bohnsack 1983: Kap. 2 sowie Luhmann 1975).

### 4.3.3 Familie Telchow: „Gartenfest“

Abbildung 12: Foto der Familie Telchow: „Gartenfest“



Ausgewählt wurde dieses Foto, wie auch das Foto „Museum“ der Familie Schiller, aufgrund seiner Fokussierungsqualität. Diese resultiert hier vor allem aus den Auffälligkeiten der ungewöhnlichen Be- bzw. Verkleidung der abgebildeten Personen.

#### *Formulierende Interpretation*

#### *Vor-ikonografische Interpretation*

Aus Gründen des Umfangs ist diese hier nicht abgedruckt.

#### *Ikonografische Interpretation*

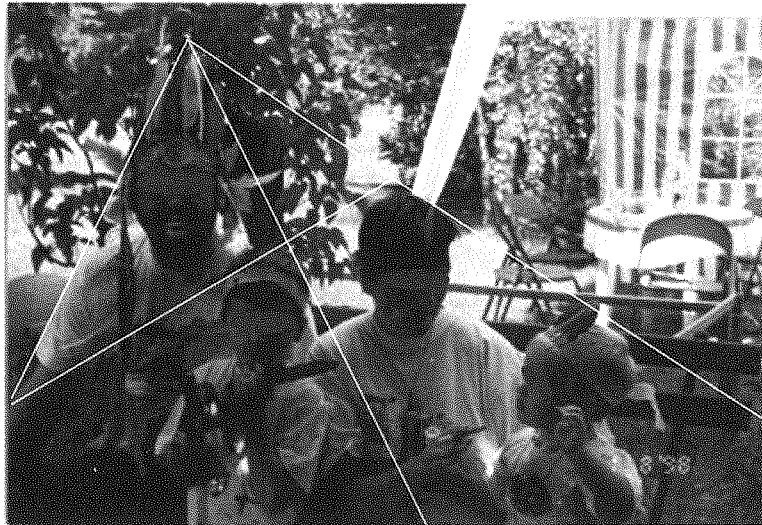
Aufgrund von Kontextinformationen seitens der Familie Telchow wissen wir, dass das Foto auf einem von der Schwester von Frau Telchow veranstalteten Gartenfest im Sommer 1998 aufgenommen wurde, und entweder die Schwester oder ihr Mann, also Familienangehörige im weiteren Sinne, fotografiert haben. Frau Telchow war damals 30, Herr Telchow 34 und die Söhne Daniel 5 und Olaf 2 Jahre alt.

#### *Reflektierende Interpretation*

#### *Formale Komposition*

#### *Planimetrie*

Abbildung 13: Foto der Familie Telchow: „Gartenfest“: Planimetrie u. szenische Choreografie



Wie häufig bei Gruppenfotos wird die Planimetrie durch das Arrangement der Personen zueinander, also die szenische Choreografie, entscheidend mitbestimmt, sodass uns die Rekonstruktion der Planimetrie zugleich wesentliche Aufschlüsse über die sozialen Beziehungen der abgebildeten Bildproduzent(inn)en zu vermitteln vermag.

Die Planimetrie wird durch drei ineinander integrierte (imaginative) Dreiecke bestimmt, welche wir erhalten, wenn wir die Spitzen der Köpfe, genauer: des Kopfschmucks der beteiligten Personen, miteinander verbinden. Die unteren Seiten der Dreiecke sind nicht mehr im Bild. Die obere Spitze des umfassenden (spitzwinkligen) Dreiecks wird durch die Spitze des Kopfschmucks von Herrn Telchow gebildet. Die linke Seite des Dreiecks führt zum Schnittpunkt des Armes von Herrn Telchow mit der linken Bildkante. Die rechte Seite verbindet die (Spitze der) Köpfe von Herrn Telchow, Frau Telchow und Olaf. Die linke Seite dieses Dreiecks gehört zugleich zu einem gleichschenkligen (und spitzwinkligen) Dreieck, dessen rechter Schenkel die Spitze des Kopfschmucks von Herrn Telchow mit den Umrissen des Kopfes von Daniel (dem älteren Sohn) verbindet. Letzterer rechter Schenkel wiederum gehört zu einem gleichschenkligen (und stumpfwinkligen) Dreieck, dessen linker Schenkel die Spitze des Kopfes von Frau Telchow mit den Umrissen des Kopfes von Daniel (dem jüngeren Sohn) verbindet.

Durch die drei Dreiecke werden drei verschiedene soziale Szenarien miteinander verbunden und integriert: die Szenerie der Gesamtfamilie und innerhalb dieser die Mutter mit beiden Söhnen und der Vater mit dem älteren Sohn. Das gleichschenklige Dreieck, welches Frau Telchow mit beiden Söhnen verbindet, liegt mit seiner Symmetrieachse in etwa auf der Bildmittelsenkrechten und steht dadurch im Zentrum, sodass die Beziehung von Frau Telchow mit ihren beiden Söhnen planimetrisch fokussiert ist. Mehr noch aber ist Frau Telchow selbst fokussiert, deren

Kopf sich darüber hinaus im Bildmittelpunkt und deren Körperachse sich nahe der Bildmittelsenkrechten befindet.

#### Perspektivität

Abbildung 14: Foto der Familie Telchow: „Gartenfest“: Perspektivität



Wir haben es mit der Übereckperspektive (Schrägperspektive) zu tun (siehe Kap. 7.4: Anhang), welche durch zwei Fluchtpunkte gekennzeichnet ist, die hier beide rechts und links außerhalb des Bildes liegen. Die Horizontlinie befindet sich in Augenhöhe von Herrn Telchow, welcher seinerseits (ebenso wie der ältere Sohn) den Blick auf den Fotografen bzw. die Fotografin gerichtet hat.

Die Kinder sind im Vordergrund, die Eltern im Hintergrund, positioniert, wobei Herr Telchow aufgrund der nach links hinten weisenden Perspektivität am weitesten in den Hintergrund, genauer: in den hinteren Vordergrund, gerückt ist.

#### Szenische Choreografie

Da die Planimetrie auf diesem Bild durch die Positionierung der Körper zueinander wesentlich konstituiert wird, wird mit ihr auch zugleich das soziale Arrangement der Personen, also deren szenische Choreografie, deren soziales Beziehungsmuster, zum Ausdruck gebracht (siehe Abbildung 13).

Die Mutter steht im planimetrischen Zentrum des Bildes. Auch wird durch die ins Bild ragende und direkt auf ihren Kopf weisende Seitenwand (bzw. verkleidete Stützsäule) des Partyzeltes die Aufmerksamkeit auf ihren Kopf gelenkt. Die Kernszenerie innerhalb der szenischen Choreografie, markiert durch das gleichschenklige (und stumpfwinklige) Dreieck, wird durch die Mutter mit ihren beiden Kindern gebildet.

Die planimetrische Komposition, welche durch die mehrfach miteinander integrierten Dreiecke entscheidend bestimmt wird, verleiht der szenischen Choreografie, also dem sozialen Beziehungsmuster Familie, eine ausgeprägte soziale Harmonie und Ausgewogenheit. Dies wird dadurch verstärkt, dass sich diese soziale Szenerie im Vordergrund vom kompositorischen Durcheinander und von der Buntheit des Hintergrundes rechts abhebt.

Hinzu kommt der Eindruck von Ruhe und Entspannung. Denn die Familie stellt sich mit diesem Foto beim Spiel, aber dennoch in Ruheposition – auf der Bank sitzend – dar. Dieser Eindruck von Ruhe und Entspannung wird einerseits verstärkt dadurch, dass die Familie im Schatten sitzt (und nicht bspw. an dem in der Sonne stehenden Tisch im Hintergrund), und andererseits durch das Rauchen der Friedenspfeife.

Das Fehlen von Spannung findet aber auch noch in anderer Weise seinen Ausdruck. Die Eltern lassen sich hier wie selbstverständlich, d.h. ohne Distanz, auf ein Spiel bzw. eine Welt ein, die sie mit den Kindern teilen. Wir haben es hier also mit einer spielerischen und spannungsfreien Verbindung von Kinder- und Erwachsenenwelt zu tun. Diese resultiert hier daraus, dass alle Beteiligten, wie es einem Spiel entspricht, in eine dritte Welt eintauchen bzw. sich in diese zurückziehen: in die Welt der Indianer.

Alle Beteiligten entledigen sich auf diese Weise spielerisch ihrer angestammten sozialen und persönlichen Identität, was in der Verkleidung und Bemalung seinen unmittelbaren Ausdruck findet. Sie sind nicht mehr Erwachsene, Eltern, Mutter, Vater und Kinder (soziale Identität) und nicht mehr Olga und Ferdinand, Daniel und Olaf (persönliche Identität), sondern Indianer. Sie suspendieren somit – zumindest auf den ersten Blick – im Medium des Spiels, welches hier als kindliches Spiel, als Spiel innerhalb der Welt der Kinder, Bedeutung gewinnt, die institutionalisierten Rollenanforderungen. Wobei das Indianerspiel mit seiner konventionellen Kostümierung und seinen konventionellen Rollen und Geschichten allerdings bereits selbst eine Art Institution darstellt.

Einer genaueren Betrachtung hält dieser Eindruck einer Suspendierung der familialen Rollenmuster allerdings nicht stand. Denn der Vater – von der konventionellen Rollenverteilung her das Oberhaupt der Familie – ist auch hier der Häuptling. Mehr aber noch als das Outfit gibt uns die Planimetrie Auskunft darüber, dass diese Szenerie im Sinne einer Verteilung traditionaler Familienrollen strukturiert ist.

Indem, wie dargelegt, die Mutter im planimetrischen Zentrum des Bildes steht und die Kernszenerie innerhalb der szenischen Choreografie (markiert durch das gleichschenklige stumpfwinklige Dreieck) durch die Mutter mit ihren beiden Kindern gebildet wird und die Mutter auch in der größten körperlichen Nähe zu den Kindern positioniert ist, wird ihr die Funktion im *Innenraum* der Familie und im Bereich der Herstellung von Intimität zugewiesen, wie es – in Relation zur Funktion des Vaters – der konventionellen geschlechtsspezifischen Rollenverteilung entspricht.

Denn der Vater hat hier die Funktion der *Außenvermittlung* inne. Zunächst wird dies in seiner ‚flankierenden‘ Stellung deutlich. Er ist *neben* dem durch Intimität charakterisierten ‚Familienkern‘ von Mutter und Kinder (und zugleich auch *über*

diesem) positioniert. Sodann befindet sich das Kameraauge und somit das Auge des Fotografen in Augenhöhe von Herrn Telchow. Indem dieser seinerseits (ebenso wie der ältere Sohn) den Blick auf den Fotografen bzw. die Fotografin und damit auch auf potentielle (spätere) Bildbetrachter/innen gerichtet hat, ist er es, der den Außenkontakt herstellt.

Im Unterschied zum Foto der Familie Telchow ist es im Foto „Museum“ der Familie Schiller der Vater, dem die Funktion der Herstellung von Intimität zukommt. Er befindet sich – von der szenischen Choreografie her betrachtet – in einer ähnlichen Position seinen Kindern gegenüber wie Frau Telchow den ihrigen. Deutlicher aber noch als Frau Telchow nimmt er die Kinder unter seine ‚Fittiche‘.

Die Familie Telchow erscheint somit also auf den ersten Blick als Raum kindlichen Spiels, in dem die Erwachsenen sich auf die Welt der Kinder einlassen. Während also – etwas grobkörnig formuliert – im Falle der Familie Schiller die Kinder als verkleinerte Erwachsene erscheinen und somit die Welt der Erwachsenen den primären Rahmen bildet, erscheinen hier die Erwachsenen als vergrößerte Kinder, indem der Bezugsrahmen der Kinder als primärer zugrunde gelegt wird. Gleichwohl wird aber deutlich, dass die Rollenerwartungen der Institution Familie in Kraft bleiben, das Spiel strukturieren und der Kinderwelt somit klare Grenzen setzen.

### *Homologien zum Orientierungsrahmen der Tischgespräche*

Die Tischgespräche in der Familie Telchow sind im Jahre 2002 aufgenommen worden, also 4 Jahre später als das Foto „Gartenfest“. Die Söhne sind somit zu diesem Zeitpunkt bereits 6 und 9 Jahre alt.<sup>48</sup>

Während im Tischgespräch in der Familie Schiller, bei dem es um die Weichenstellung hinsichtlich der Fächerwahl in der weiteren schulischen Ausbildung des Sohnes geht, vom älteren Sohn gefordert wird, über diese Fächerwahl Rechenschaft abzulegen, werden in der Familie Telchow (in einer thematisch vergleichbaren Passage) die vom jüngsten Sohn eingebrachten und vom älteren Sohnes unterstützten Fragen nach den Bedingungen und Voraussetzungen weiterführender schulischer Bildung von der Mutter (M) zunächst nur zögerlich beantwortet.

Schließlich wird betont, dass es bei Kindern, denen es „schwer fällt“, weiterführende Schulen zu besuchen, unter bestimmten Bedingungen „in Ordnung“ ist, wenn sie „nach der zehnten Klasse Schluss machen können“ (*Transkript Telchow TG: Schullaufbahn 51-53*):

- 51 M: (3) °na ja° (2) es gibt aber Kinder, denen es schwer fällt die Schule  
52 und die freuen sich dann wenn sie nach der zehnten Klasse Schluss machen  
53 können ne; (2) und wenn se dann ein gutes Zeugnis haben, ist dit schon in Ordnung.

Entscheidend sind also Neigung und Wohlbefinden der Kinder selbst. Dies steht in deutlichem Kontrast zur Familie Schiller, in der vor allen durch die Mutter Proble-

<sup>48</sup> Da Frau und Herr Telchow im Schichtdienst arbeiteten, war bei der überwiegenden Zahl der Tischgespräche lediglich ein Elternteil anwesend, sodass uns die Gespräche über das Verhältnis der Elternteile untereinander kaum Auskunft geben können.

me der (berufs-) biografischen Sicherheit der Kinder ins Zentrum gestellt werden. In der Diskussion um die Weichenstellung bei der Fächerwahl während der weiteren schulischen Ausbildung des älteren Sohnes erscheint dessen „Lust“ für ein Fach vollkommen irrelevant. Die Eltern Schiller verstehen sich als diejenigen, welche die (von ihnen so gesehene) Rationalität des Bildungssystems als Teil der Logik der Erwachsenenwelt gegenüber den Kindern zu vertreten haben, damit diese erfolgreich sind. Wobei dieser Erfolg den Eltern prekär erscheint und somit rigide Grenzbeziehungen notwendig macht.

Demgegenüber verstehen Frau Telchow sowie (dies wird in anderen Tischgesprächen deutlich) auch ihr Mann die Familie als eine Instanz jenseits der Leistungsanforderungen und Selektivitätskriterien der Institution Schule. Im Vergleich zur Familie Schiller sind die Verhältnisse in der Familie Telchow geradezu umgekehrt: Während in der Familie Schiller sich die Mutter unter dem Druck sieht, den Kindern den Planungshorizont der weiteren schulische Laufbahn und deren zweckrationale Anforderungen vor Augen zu halten, ist es, wie bereits angesprochen, in der Familie Telchow der jüngere Sohn (S2), dem es nur mit Nachdruck gelingt, das Thema des Übergangs in eine weiterführende Schule in den Diskurs einzubringen. (*Transkript Telchow TG: Schullaufbahn 85-106*):

- 85 S2: Mama, und wenn die Kinder nicht wollen, (1) hm:: in die  
 86 zeh- wenn da in der zehnten (Klasse) sind (.) zu gehen? Was  
 87 ist dann;  
 88 M: Wenn se dann aufhören wollen mit der Schule?  
 89 S2: Nein wenn se weiter machen wollen.  
 90 M: Na dann müssen sie an eine (.) andere Schule, wie eine kleine  
 91 Bewerbung schreiben dass sie noch weiter lernen möchten  
 92 und dann entscheidet der Direktor von der anderen Schule, ob  
 93 die Kinder weiter lernen dürfen. guck mal bei Anton zum  
 94 Beispiel, war dis so, bei Sven, bei Jan,  
 95 S2: [ Und was haben se da  
 96 gesagt?  
 97 S1: [ Wie jetzt?  
 98 M: [ Die durften alle. (.) die durften alle weiter machen.  
 99 S1: [ Bis zur. bis zur drei- dreizehnten.  
 100 M: Hm (3) die warn erst auf einer Schule, auf einer Realschule  
 101 wo es bis zur zehnten Klasse geht, (2) und dann durften se  
 102 ähm weiter machen. auf einem Gymnasium heißt dis denn.  
 103 (3) ne?  
 104 S2: Hhm. aber wenn nicht dann iss es mir au- (.) is es mir egal.  
 105 (1)  
 106 M: Na klar. hast ja auch noch Zeit.

S2, der jüngere Sohn, stellt die Frage danach, was mit denjenigen „Kindern“ geschieht, die nach der 10. Klasse nicht die Schule verlassen, sondern „weitermachen“ (89) wollen. Indem die Mutter die Entscheidung über den weiteren Schulbesuch ausschließlich der Entscheidungsmacht des Direktors der weiterführenden Schule überträgt, sie sozusagen vertrauensvoll in dessen Hände legt, bleibt der Beitrag der Schüler/innen selbst und auch jener der Familie vollkommen ausgeklammert.

Die innerfamiliäre Sphäre wird also gegenüber anderen Institutionen und deren Diskursen durch die Eltern abgeschirmt. Dies geschieht im Rahmen der konventio-

nellen Rollenverteilung, wie sie auch bereits im Foto erkennbar wurden. Die Eltern verstehen sich als diejenigen, welche der familialen Umwelt und ihren Anforderungen und Konflikten nur gefiltert durch ihre Vorstellungen Zugang zur Welt der Familie gewähren.

Dies steht im deutlichen Kontrast zur Familie Schiller, in der die Kinder im innerfamiliären Diskurs mit den Anforderungen außerfamiliärer Institutionen sowie mit Problemen biografischer Planung und Sicherheit regelrecht konfrontiert werden. Während also in der Familie Telchow die Umweltkomplexität in stark reduzierter Form in die familieninterne Kommunikation und Wirklichkeit Eingang findet, zeichnet sich der familieninterne Diskurs in der Familie Schiller durch eine hohe Komplexität aus, deren Bewältigung eine familieninterne konfrontative Auseinandersetzung mit sich bringt, bei der den Kindern die Kompetenz zur Bewältigung dieses Problems nicht vollständig zuerkannt wird.

Schlagwortartig lassen sich die familialen Orientierungsrahmen oder Habitus derart zusammenfassen, dass wir es in der Familie Schiller mit demjenigen der *familialen Intimität im Kontext von Konfrontation und Reflexivität* zu tun haben und in der Familie Telchow mit demjenigen einer *familialen Intimität im Kontext von institutioneller Absicherung und Abgrenzung*.

Die in der Familie Telchow zu beobachtende Begrenzung dessen, was innerfamiliär verhandelt wird, führt dazu, dass der familiäre Diskurs sehr stark in sich selbst kreist. Im Unterschied zum provokativen und konfrontativen Diskurs der Familie Schiller haben wir es hier mit einer Art selbst-referentiellen Bestätigungs- und Absicherungsdiskurs zu tun, der nicht jene Komplexität und jenen Tiefgang wie in der Familie Schiller erreicht und bisweilen Züge einer „inhaltslosen“ Wiederholung annimmt. (*Transkript Telchow TG: Schwimmen und Matheunterricht 01-07*)

- 01 V: Und was gab's bei dir Danny? (1)  
 02 S1: Hm, war heute wieder sehr schön,  
 03 V: Oh:: das war die schöne Woche wa,  
 04 S1: [ @ (.) @  
 05 V: Gab's eigentlich einen Tag der nicht schön war bei dir?  
 06 S1: Nöö.  
 07 V: Nöö, aha, na is ja is ja schön wenn's schön war.

Nachdem einer der Söhne aufgefordert worden ist, seine Erfahrungen des Schulalltags zu resümieren, gipfelt der Diskurs sehr bald in der Konklusion von Herrn Telchow: „na ja, is ja, is ja schön wenn's schön war“, womit die Reduktion des Gesprächs auf seine reine Bestätigungsfunktion von ihm selbst demonstriert und nahezu karikiert wird.

Diese Orientierung an der Familie als selbst-referentiellen Freiraum gegenüber exterioren, insbesondere institutionellen Zwängen resultiert nicht aus einer kritischen Distanz gegenüber der Institution Schule, sondern geht mit einem Vertrauen in die Gerechtigkeit, die Selektionskriterien sowie insgesamt den Sinn der Ausbildungsinstitutionen einher, was erst die Voraussetzung dafür ist, das Thema Schule im familialen Diskurs weitgehend suspendieren zu können.<sup>49</sup> Die Bedeutung, wel-

<sup>49</sup> Auch die Hausaufgabenkontrolle wird, wie in der Gruppendiskussion deutlich wird (*Transkript Telchow GD: Schule 210-212*), grundsätzlich an die Horterzieher delegiert.

che die Telchows den schulischen Ritualen beimessen, kommt auch darin zum Ausdruck, dass sie uns eine Serie von drei Fotos zur Einschulung des älteren Sohnes zur Verfügung gestellt haben. In den Familienfotos der Familie Schiller gewinnen derartige schulische Rituale bzw. Übergangsrituale keine Relevanz. An die Stelle derartiger ritueller Vermittlungen tritt hier eher die (meta-) kommunikative Auseinandersetzung mit den Anforderungen der Institutionen.

Während die Eltern in der Familie Schiller diese grundlegend als Ort der Vermittlung von Leistungsorientierung verstehen, halten die Eltern in der Familie Telchow deren Erfahrungsraum frei von derartigen Erwartungen. *Aufstiegs- und Leistungsorientierung* sowie *konfrontative Auseinandersetzungen* stellen insgesamt zentrale Komponenten des Orientierungsrahmens in der Familie Schiller dar (vgl. auch: 4.3.6), während man bei der Familie Telchow von einem *In-Sich-Ruhen im Kontext institutioneller Sicherheit* sprechen kann.

Es lässt sich hier allerdings letztlich nicht valide klären, inwieweit es sich nicht allein um milieuspezifischen Differenzen zwischen den Telchows und den Schillers, sondern auch um Differenzen in der (Kompetenz-) Entwicklung ihrer Kinder aufgrund derer Altersunterschieden und damit zusammenhängenden unterschiedlichen Phasen der Familienentwicklung handelt, inwieweit wir es also (in der Sprache der dokumentarischen Methode) mit einer Überlagerung von Entwicklungs- und Milieutypik zu tun haben (siehe: Bohnsack 2009a). Während das Alter der Kinder zum Zeitpunkt der fotografischen Aufnahmen in beiden Familien noch im vergleichbaren Bereich liegt (in der Familie Schiller sind sie 5 u. 6 Jahre und in der Familie Telchow 5 u. 2 Jahre alt), ist dies bei der Aufzeichnung der Tischgespräche in dem Maße nicht mehr der Fall (in der Familie Schiller haben wir die Altersspanne von 12, 11 u. 5 und in der Familie Telchow von 9 u. 6 Jahren).

Somit könnten die am Vergleich der Tischgespräche erarbeiteten Unterschiede der Orientierungsrahmen beider Familien auch auf Differenzen im Bereich der Kompetenz der Kinder bzw. der Unterstellung bzw. Antizipation von Kompetenzdifferenzen (etwa im Bereich biografischer Planung) seitens der Eltern zurückzuführen sein (in dem Sinne, dass in der Familie Telchow Gespräche auf der Basis berufsbiografischer Perspektiven verfrüht erscheinen). Um valide kontrollieren zu können, inwieweit wir es auch mit einer Überlagerung von Milieu- und Entwicklungstypik zu tun haben, wäre eine weiterreichende komparative Analyse von Familien mit Kindern unterschiedlicher Alterstufen erforderlich.

Allerdings wird durch diese fehlende Kontrolle die Gültigkeit von Milieudifferenzen nicht grundlegend in Frage gestellt. Denn zum einen findet die Differenz der Orientierungsrahmen auch in der komparativen Analyse der Familienfotos ihren Ausdruck, bei denen (wegen des Zeitpunktes der Aufnahme) Altersdifferenzen zwischen Kindern beider Familien kaum von Bedeutung sind. Und zum anderen betreffen die unterschiedlichen Anforderungen an die Planungs- und Argumentationsfähigkeit der Kinder lediglich eine Komponente des Orientierungsrahmens der Familie, deren andere das Verhältnis zur Leistung und das Vertrauen in die gesellschaftlichen Institutionen darstellt, hinsichtlich derer sich ebenfalls erhebliche Differenzen zwischen den Orientierungsrahmen der Familien Telchow und Schiller zeigen.

Diese Kontraste zwischen einem *In-Sich-Ruhen* in der Familie Telchow und einer *leistungsbezogenen Bewegung* in der Familie Schiller lassen sich sehr deutlich

am Thema Urlaub, also am Sujet der Urlaubsfotos, vergleichen. Dort sind die Fotos der Familie Schiller zwar in Ruhepositionen aufgenommen (bzw. nicht in Phasen aktiver Bewegung), allerdings handelt es sich um Phasen der Ruhe während aktiver Bewegungen. So rastet die Familie beim Bergwandern im Gebirge im Schnee hoch am Berg oder im Sommer im tiefen Wald, wobei erkennbar ist, dass alle über ein solides Outdoor-Outfit verfügen und auch die Kinder feste Bergstiefel tragen. Sehr eindrucksvoll bringt die Familie ihre Haltung in zwei Fotos zum Ausdruck, die während der Radtour aufgenommen worden sind.

Abbildung 15: Familie Schiller: Urlaubsfoto „Fahrradrast an der Straße“



Auf einem der Fotos, auf das ich hier nur kurz eingehe, steht die gesamte Familie am Rande einer Straße neben bzw. über ihren gut ausgerüsteten Trekkingrädern, die alle mit speziell für Radtouren konzipierten Gepäcktaschen schwer beladen bzw. behängt sind. Es handelt sich somit also um eine mehrtägige Tour und eine nur kurze Rast. Der jüngste Sohn sitzt im Kindersitz hinten auf den Fahrrad von Frau Schiller, die beiden älteren Kinder tragen Fahrradhelme. Die feste Regenbekleidung und die wasserfeste Ausrüstung zeigen, dass die Familie relativ unabhängig vom Wetter in Bewegung ist.

Die Familie Telchow präsentiert sich auf den 14 Urlaubsfotos, die sie uns zur Verfügung gestellt hat, im deutlichen Kontrast zu den Urlaubsfotos der Familie Schiller. So sind die Familie oder einige ihrer Mitglieder auf drei dieser Fotos im Zusammenhang mit unterschiedlichen Seilbahnen dargestellt, die ihnen den Aufstieg erleichtert haben: entweder befinden sie sich im Zugang zur Seilbahn, vor dieser oder in deren Gondel. Auf den übrigen (mit einer Ausnahme in den Bergen aufgenommenen) Urlaubsfotos der Familie Telchow ist an der Kleidung, insbesondere am Schuhwerk (Sandalen) erkennbar, dass sich die Familie nicht weit von der Seilbahn oder dem Auto entfernt hat.

Abbildung 16: Familie Telchow: Urlaubsfoto „In der Seilbahngondel“



Diese Kontraste zwischen den beiden Familien Telchow und Schiller, welche neben der Leistungsorientierung im weiteren Sinne insbesondere die Orientierung an biografischer Sicherheit und das Verhältnis zu den staatlichen Institutionen betreffen, verweisen also auf milieuspezifische Differenzen. Dass dies etwas mit den *Herkunftsmilieus* der Eltern resp. der Mütter zu tun hat, ist uns im Forschungsprozess zunächst auf der Grundlage zweier Fotos aus den Herkunftsfamilien von Frau Schiller und Frau Telchow evident geworden: dem Foto der Kommunion von Frau Schiller (4.3.4) und demjenigen der Jugendweihe in der Familie Telchow (4.3.5). Für die männliche Linie der beiden Familien stehen uns entsprechende Fotos nicht zur Verfügung.

#### 4.3.4 Kommunion von Frau Schiller

Es handelt sich um ein Foto zur Kommunionsfeier von Frau *Schiller* im Jahre 1981, welches ich dann (in 4.3.5) mit dem Foto der Jugendweihe der Schwester von Frau Telchow in etwa aus dem Jahre 1978 vergleichen werde. Wir haben auf dieses Foto der Telchows zurückgegriffen, da uns Frau Telchow ein Foto ihrer eigenen Jugendweihe nicht zur Verfügung stellen konnte. Da es in unserer Analyse um das Familienmilieu, genauer um das Milieu der Herkunftsfamilie, von Frau Telchow und nicht um deren individuelle Biografie geht, ist diese Differenz für unsere Auswertung nicht problematisch.

Abbildung 17: Foto zur Kommunion von Frau Schiller



#### Formulierende Interpretation

##### Vor-ikonografische Interpretation

##### Zum Bildvordergrund:

Im Vordergrund des Bildes sind 7 Personen zu sehen: links drei Frauen und ein Mädchen und rechts zwei Männer und ein männlicher Jugendliche. Die Gruppe steht auf einem mit Platten gepflasterten Weg. Dieser wird von einer nur spärlich mit Gras bewachsenen Fläche durch einen auf Pfählen gespannten Draht abgetrennt, wie er zur Absperrung von Baustellen verwendet wird.

*Alter der abgebildeten Personen:* Das Mädchen ist 10–12 Jahre alt, die (vom Bildbetrachter aus gesehen) links hinter ihr stehende Frau ca. 60 Jahre, die direkt hinter ihr stehende Frau ca. 35–40, und die rechts neben ihr stehende Frau ca. 65 Jahre alt. Der männliche Jugendliche rechts daneben ist ca. 16–17, der Mann wiederum rechts von ihm ca. 45 und der Mann ganz rechts ca. 65–70 Jahre alt.

*Bekleidung:* Die Frauen sind mit hellen Sommermänteln und weißen Schuhen, also sonn- oder festtäglich, bekleidet. Die beiden Frauen vorne (vermutlich auch die Frau hinten) haben die Hände in den Manteltaschen. Das Mädchen trägt ein weißes Kleid, eine helle Strickjacke und weiße Strumpfhosen mit schwarzen Schuhen und hat einen Blumenkranz im Haar. Die Männer und der männliche Jugendliche sind mit Sakkos und schwarzen Schuhen bekleidet. Die Männer tragen weiße Hemden und Krawatten, der junge Mann ein weißes Hemd mit Pulli.

*Gebärden und Mimik:* Die Mimik (die aufgrund der Maskierung für den Leser allerdings kaum erkennbar ist) zeigt insbesondere bei den älteren Frauen und beim 65–70-jährigen Mann eine spezifische Strenge oder auch Härte, ohne aber direkt ab-

weisend zu sein. Die Körperhaltung, insbesondere wiederum der beiden älteren Frauen und des 65–70-jährigen Mannes, zeichnet sich durch eine strikt senkrechte Körpermittelachse aus, die in Verbindung mit der Mimik als „starr“ bezeichnet werden kann.<sup>50</sup>

#### *Zum Bildhintergrund:*

Im *vorderen* Bildhintergrund ist – parallel zur Bildebene wie auch parallel zur Plattenreihe – eine Baumreihe zu sehen, die entlang eines Grabens oder Baches durch das gesamte Foto verläuft. Im *entfernteren* Bildhintergrund zieht sich wiederum parallel hierzu und ebenfalls durch das gesamte Foto ein fünfgeschossiger Neubau, der als Plattenbaukonstruktion identifizierbar ist.

#### *Ikonografische Interpretation*

Frau Schiller, also das Kommuniionskind, ist auf dem Foto 11 Jahre alt. Die ca. 35-jährige Frau hinter ihr ist ihre Mutter. Bei der ca. 65-jährigen Frau in der Mitte des Bildes handelt es sich um die Großmutter von Frau Schiller, also um die Ur-Großmutter (von der heutigen Familiengeneration betrachtet), bei dem 65–70-jährigen Mann um den Großonkel mütterlicherseits von Frau Schiller. Der ca. 45-jährige Mann links von ihm ist der Bruder der Mutter von Frau Schiller, und der junge Mann links neben ihm dessen Sohn. Bei der ca. 65-jährigen Frau ganz links im Bild handelt es sich um die Großtante von Frau Schiller. Der Vater des Kommuniionskindes fehlt auf dem Bild. Wir wissen, dass die Eltern geschieden sind und somit fehlen auch Angehörige des zum Vater gehörigen Teils der Familie. – Fotografin ist die Frau des Bruders der Mutter des Kommuniionskindes, der heutigen Frau Schiller, also die Tante von Frau Schiller.

#### *Reflektierende Interpretation*

##### *Formale Komposition*

##### *Planimetrie*

Die Planimetrie wird insgesamt sehr dominant durch die senkrechten und waagerechten Linien bestimmt, welche mit dem Plattenbau im Hintergrund vorgegeben sind und die dann durch die ganz harten *waagerechten* Linien der Bäume konturiert hervorgehoben werden sowie durch die senkrechten Linien des Bachverlaufs und der Bodenplatten im Vordergrund. Die abbildende Bildproduzentin und die abgebildeten Bildproduzent(inn)en haben nicht nur den Plattenbau mit seiner Geometrie als Hintergrund, sondern auch die Standfläche der Platten im Vordergrund gewählt.

<sup>50</sup> Die vor-ikonografische Beschreibung und dann auch die ikonologische Interpretation der *Mimik* stellen uns allerdings – hierauf werde ich in 5.4.2 noch einmal genauer eingehen – vor große Probleme. Es fehlt, ebenso wie bei den Gebärden, eine Beschreibungssprache in den Sozialwissenschaften. Darüber hinaus ist eine ikonologische Interpretation des mimischen Ausdrucks in valider Weise nur möglich unter Berücksichtigung des gesamten Gesichtsfeldes und mehr noch auch der Rekonstruktion des Zusammenspiels, des Kontexts simultaner Ausdrucksbewegungen, welche den gesamten Körper umfassen. Die Interpretation von Gebärden und insbesondere der *Mimik* hat sich an der „whole body“ conception“ zu orientieren, wie Ray Birdwhistell (1952: 8) dies genannt hat.

Dadurch dominiert die *waagrecht-senkrechte Strukturierung* das Gesamtfoto und gibt ihm eine *Strenge* und eine *rigide Ordnung*.

Abbildung 18: Foto zur Kommunion von Frau Schiller: Planimetrie



Die abgebildeten Bildproduzent(inn)en fügen sich in diese planimetrische Struktur bzw. betonen sie deren Senkrecht-Strukturierung noch einmal durch ihre strikt senkrechten Körpermittelachsen, die den Eindruck einer starren Körperhaltung vermitteln. Dies wird in konturierter Weise sichtbar und gewinnt seine interpretative Evidenz vor dem Vergleichshorizont der Fotos der Jugendweihe in der Familie Telchow (4.3.5) und der Kommunion der Kinder von Frau Schiller (4.3.6).

Planimetrisch ins Zentrum gerückt ist die Großmutter, weil sie zum einen vor dem (den Hintergrund strukturierenden) stark hervorgehobenen schwarzen Mittelteil der Plattenbaureihe positioniert und zum anderen vom rechten und linken Rand der Personengruppe ungefähr gleich weit entfernt ist.

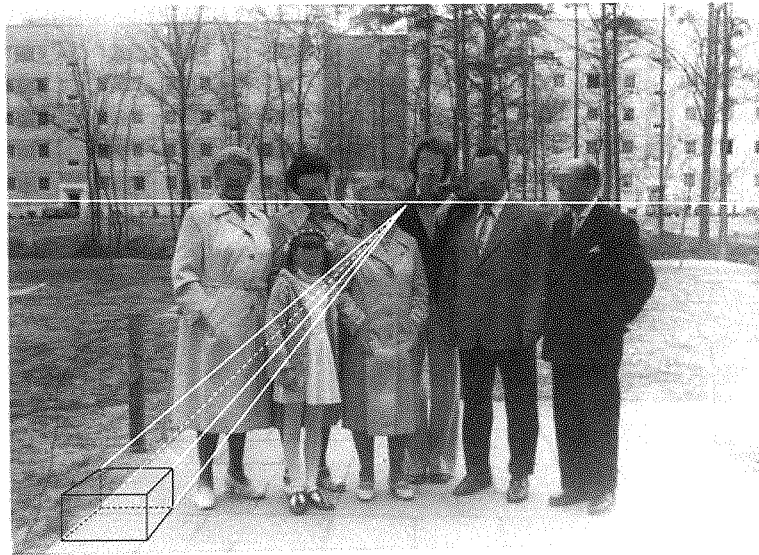
##### *Perspektivität*

Es ist auch die Großmutter, die perspektivisch fokussiert ist, und nicht etwa das Kommuniionskind. Die Horizontlinie (also die Linie, auf welcher der Fluchtpunkt liegt) befindet sich in Höhe der Augen der Großmutter, der Fluchtpunkt neben ihrem Kopf rechts, sodass sich der Fotograf, der abbildende Bildproduzent und somit auch der Betrachter, gleichsam „Auge in Auge“ mit der (Ur-) Großmutter befinden.

Die Horizontlinie ist zugleich planimetrisch hervorgehoben, d.h. sie bildet im Bereich der planimetrischen Komposition eine dominante Linie, da sie (zumindest nahezu) mit der Grundlinie des Plattenbaus im Hintergrund übereinstimmt. Die (Ur-) Großmutter ist somit auch deshalb sowohl planimetrisch wie auch perspektivisch ins Zentrum gerückt, weil ihre Augen auf jener Linie platziert sind, durch die Planimetrie und Perspektivität miteinander vermittelt sind.



Abbildung 19: Foto zur Kommunion von Frau Schiller: Perspektivität



### Szenische Choreografie

Abbildung 20: Foto zur Kommunion von Frau Schiller: szenische Choreografie



Das Kommunionkind ist zwar weder perspektivisch noch planimetrisch fokussiert, wohl aber durch die szenische Choreografie. Denn die drei Frauen haben das Mädchen in die Mitte genommen (wobei die Mutter direkt hinter ihr steht, so dass auch diese, obschon in den Hintergrund gerückt, durch die beiden älteren Frauen flan-

kiert wird). Damit bildet der weibliche Teil der Gesamtfamilie ein in sich geschlossenes Ensemble, welches nahe zusammengedrückt ist: Die jüngere und jüngste Generation steht innen, die ältere außen.

Die Umrisse dieses Ensembles bilden eine symmetrische Haus- oder Zeltform. An dieses das Bild strukturierende Ensemble sind die Männer – eher neben- als miteinander stehend – rechts angereiht. Während (bedingt durch die Halbkreisbildung) die Frauen und das Mädchen in der Weise leicht schräg zur Bildebene stehen, dass die (von den abgebildeten Bildproduzent(inne)n aus gesehen) linke Schulter nach hinten genommen wird, ist es bei den Männern umgekehrt. Der Jugendliche, der Sohn der Fotografin, steht demgegenüber nahezu frontal zur Bildebene und nimmt somit eine ‚vermittelnde‘ Position ein zwischen der linken und der rechten Gruppe. Dies wird unterstrichen dadurch, dass er sich mit dem linken Ellenbogen einerseits auf die Schulter des ca. 45-jährigen Mannes links neben ihm und andererseits mit der rechten Hand auf die Schulter der ca. 65-jährigen Frau rechts neben ihm stützt.

Der Körper des ca. 65–70-jährigen Mannes ist noch stärker als bei den anderen Männern von einer frontalen Position zur Bildebene nach rechts weggedreht. Zudem schaut er als einziger nicht zum Fotografen. Auch überlappt sich seine Körperfläche nicht mit der der anderen. Insgesamt vermittelt er den Eindruck eines gewissen Unbeteiligtseins. Er steht gewissermaßen ‚daneben‘ in einer flankierend-beobachtenden Position, welche (durch die auf dem Rücken verschränkten Arme) an diejenige eines militärischen Vorgesetzten neben seiner Truppe erinnert.

### Ikonologisch-Ikonische Interpretation

Die Strenge der planimetrischen Komposition mit ihrer Rigidität der waagrecht-senkrechten Strukturierung sowohl des Hintergrundes als auch des Vordergrundes korrespondiert mit der Strenge der Mimik der abgebildeten Bildproduzent(inne)n und ihrer relativen starren Haltung. Diese wird noch dadurch unterstrichen, dass die Frauen ihre Hände in den Taschen ‚vergraben‘ haben. Diese strenge und starre Haltung wird vor allem durch die (Ur-) Großmutter präsentiert, welche in mehrfacher Hinsicht – planimetrisch wie perspektivisch – ins Zentrum gerückt wird. Lediglich der ca. 16-jährige Jugendliche bricht aus dieser Haltung nicht nur durch eine freundliche Mimik, sondern auch eine lässige Haltung (die Körpermittelachse ist nicht strikt senkrecht) ein wenig aus.

Innerhalb dieser strengen Komposition bilden die drei Frauen mit dem Mädchen in ihrem haus- oder zeltförmigen Ensemble noch einmal einen in sich geschlossenen Block, der vor allem das Mädchen, in gewisser Weise aber auch ihre Mutter, zugleich abschirmt wie auch vereinnahmt und einkesselt. Der Ensemble-Charakter der Frauen bzw. der Kontrast zwischen den Geschlechtern wird auch durch ihre helle Kleidung der Frauen im Kontrast zur dunklen Kleidung der Männer hervorgehoben ebenso wie auch dadurch, dass die Frauen mit Mänteln bekleidet sind, die Männer lediglich mit Anzug bzw. Sakko.

Indem die Frauen die Hände in den Manteltaschen halten, dokumentiert sich – neben dem Eindruck von Starrheit und Strenge – auch, dass sie sich nicht allzu behaglich fühlen, dass sie dieser unbehaglichen Situation aber zugleich mit einer gewissen Entschlossenheit begegnen.

Die Männer wirken hier wie hinzu- oder angefügt. Die flankierende Position der Männer, indem diese weniger *in als neben* der Familie positioniert sind, welche in den uns vorliegenden Familienfotos allgemein zu beobachten ist (davon hier abgedruckt: 4.3.3, 4.3.5 u. 4.3.6) ist, wird im Habitus des Ur-Großvaters ganz rechts (durch seine ‚Vorgesetzten-Haltung‘) sozusagen auf die Spitze getrieben. Dadurch wird die Atmosphäre der Rigidität und Strenge noch einmal unterstrichen.

Ein weiteres – wenn auch weitaus weniger dominantes – Ensemble bildet die Großmutter mit ihren Enkeln – aufgrund der dominanten Schräge, die deren Köpfe miteinander verbindet. Das Kommunionkind ist halb zur Großmutter gewandt und gewissermaßen an sie gelehnt; der ältere Enkelsohn hat die Hand um ihre Schulter gelegt. Dadurch wird der fokussierte Charakter der Großmutter noch einmal bestärkt: Ihr Einfluss auf die jüngste Generation – und damit der Tradierungszusammenhang – erscheinen gefestigt.

Vor dem Hintergrund unseres ikonografischen Vor-Wissens um die Scheidung der Mutter erhält diese Solidarität der Frauen, durch welche Mutter und Tochter (auch gegenüber den Männern) abgeschirmt oder auch eingeschlossen werden, ihre besondere Bedeutung. Und dies gilt auch für die Atmosphäre der Rigidität und Strenge.

Im gewissen Kontrast zu dieser Komponente der Atmosphäre des Bildes steht jene andere Komponente, wie sie durch den *provisorischen* Charakter des sonstigen Vordergrundes hervorgerufen wird. Zum einen ist der Plattenweg noch unvollendet. Er scheint nirgendwo hinzuführen bzw. ist unklar, wohin. Er hat den Charakter einer Baustelle, deren ehemalige Absperrung noch vorhanden ist. Das gibt dem Bild den Charakter des *Unfertigen* und *Ungesicherten*, eben des Provisorischen. Verstärkt wird dies durch den weit nach hinten gerückten Hintergrund bzw. den fehlenden Mittelgrund. Die vergleichsweise kleine Schar wirkt hierdurch in besonderer Weise *isoliert*, d.h. aus Bezügen, die sie einbetten und halten könnten, herausgenommen und somit ein wenig ‚verloren‘. Auch hier deuten sich Bezüge zur biografischen Situation der Familie an wie auch zur besonderen Situation der Kommunionfeier im Kontext der DDR-Gesellschaft, denen es auf der Grundlage einer komparativen Analyse mit dem Foto der Jugendweihe (4.3.5) in der Familie Telchow wie auch in Triangulation mit den Gesprächsanalysen nachzugehen gilt.

Auch ist zu bemerken, dass ein Bezug zu einer Institution – vermittelt über die Ablichtung eines Gebäudes wie etwa einer Kirche – hier nicht gegeben ist. Diese Beobachtung erhält ihre Evidenz allerdings erst durch eine komparative Analyse mit Fotos anderer Familien aus derselben Generation, wenn diese etwa vor einem Gebäude abgelichtet sind, welches eine staatliche oder kirchliche Institution repräsentiert (siehe dazu auch die komparative Analyse mit dem Foto der Jugendweihe: 4.3.5).

Diese Relation bzw. dieses Spannungsverhältnis zwischen dem Charakter des *Provisorischen* und *Ungesicherten* und der (*sozialen*) *Isolation* der kleinen Schar einerseits und der *Rigidität* und *Strenge* (in Planimetrie, Mimik und Gestik) andererseits macht die zentrale Stimmung des Bildes aus. In diesem Sinne lässt sich diese Stimmung bzw. der familiäre Habitus in sprachlich-textlicher Form als eine *Übergegensätzlichkeit* (vgl. auch Kap. 3.5) formulieren – als diejenige von *Rigidität* und *Strenge* im Kontext des *Provisorisch-Ungesicherten*.

#### 4.3.5 Jugendweihe in der Familie Telchow

Zur Konturierung des Orientierungsrahmens der Herkunftsmilieus von Frau Schiller wird dieser – zunächst in horizontaler komparativer Analyse – mit demjenigen des Herkunftsmilieus von Frau Telchow kontrastiert, ehe wir (in Kap. 4.3.6) zur vertikalen komparativen Analyse mit der nächsten Generation *innerhalb* der Familie Schiller kommen und dort ein Foto der Kommunion der Kinder von Frau Schiller interpretieren.

Um die Darstellung zu straffen, wird hier neben der knappen ikonografischen Interpretation nur die ikonologisch-ikonische wiedergegeben.

Abbildung 21: Foto der Feier zur Jugendweihe der Schwester von Frau Telchow



#### *Ikonografische Interpretation*

Es handelt sich um die Jugendweihe der ca. 5 Jahre älteren Schwester von Frau Telchow, ungefähr im Jahre 1978. Die Schwester ist auf dem Bild 14–15 Jahre alt. Obwohl die junge Frau ca. 6 Jahre älter ist als Frau Schiller, liegt (da es sich ja um die Jugendweihe, nicht um die Kommunion handelt) die Entstehungszeit dieses Fotos der Familie Telchow (1978) nicht weit von derjenigen des Kommunionfotos von Frau Schiller (1981) entfernt.

Wir wissen ferner, dass die Schwester von Frau Telchow (weiß bekleidet) in der Mitte der vordersten Reihe steht und links neben ihr die Großmutter. Die Identität der Person rechts neben der Schwester kennen wir nicht. Wiederum rechts neben dieser Unbekannten stehen die Eltern von Frau Telchow. Diese selbst ist vermutlich die zweite von rechts in der Reihe der Kinder ganz vorne.

Bei dem Haus im Hintergrund handelt es sich offensichtlich um ein öffentliches Gebäude. Da der Vater von Frau Telchow als Lehrer arbeitet und die Mutter als Erzieherin, könnte die Gruppe vor einer Schule oder einem Kindergarten positioniert sein.

### *Ikonologisch-ikonische Interpretation*

Wie auf dem Kommunionfoto von Frau Schiller sind auch hier die abgebildeten Bildproduzent(inn)en, also Familie und Verwandtschaft, nicht – wie etwa eine Familiengeneration später auf dem Foto der Kommunion der Kinder der Familie Schiller (4.3.6) – ‚im Grünen‘, sondern auf einem Plattenweg und vor einem Gebäude positioniert, worin sich (wie der Vergleich mit anderen Fotos von Familienfeiern aus dieser Zeit bestätigt) auch zeitgeschichtliche Elemente dokumentieren. Entsprechend sind auch hier einige Elemente einer Waagrecht-senkrecht-Strukturierung zu erkennen.

Von der szenischen Choreografie und der leiblich-gestischen Positionierung der abgebildeten Bildproduzent(inn)en her vermittelt dieses Bild allerdings einen vollständig anderen Eindruck. Vor dem Vergleichshorizont des Fotos der Kommunion von Frau Schiller (4.3.5) erscheint hier die körperliche Positionierung nahezu ungeordnet. Die gesamte Körper- und Kopfhaltung ist hier – mit Ausnahme vielleicht der Hauptperson und ihrer Eltern – keineswegs durch eine strikt senkrechte und in diesem Sinne starre Körpermittelachse charakterisiert. Entsprechend variiert auch die Positionierung der Rumpfe und Köpfe der abgebildeten Bildproduzent(inn)en zueinander und deren Entfernung voneinander.

Die Variationsbreite in der Körperhaltung wird durch die Kinder in der vordersten Reihe noch gesteigert, die in der Hocke und überwiegend auf einem Knie positioniert, eher den Eindruck einer Startposition in einem Wettlauf als den einer feierlichen Haltung vermitteln.

Ist bereits die szenische Choreografie und die leiblich-gestische Positionierung in ihrer Ungezwungenheit weit entfernt von Rigidität und Strenge, so gilt dies auch für die Mimik, die sich – vielfach mit einem Lächeln verbunden – insgesamt durch eine gewisse Heiterkeit auszeichnet. Diese Ungezwungenheit und Gelöstheit in szenischer Choreografie, Gebärden und Mimik wird durch die Variationsbreite in der Kleidung bekräftigt, die bis dahin geht, dass ein Vertreter der älteren Generation (links außen in der ersten Reihe hinter den Kindern) ein kariertes Hemd trägt. Er bricht aus einigen wenigen Elementen von Feierlichkeit, die durch einige schwarze Anzüge und das weiße Kleid der Hauptperson ins Bild gebracht werden, auch dadurch aus, dass er während der fotografischen Aufzeichnung, welche ja ihrerseits Element des Rituals ist, eine Zigarette raucht.

Die Gruppe steht hier nicht auf freiem Feld bzw. einem baustellenartigen Gelände, wie die Angehörigen von Frau Schiller auf deren Kommunionfoto, sondern im Schutz eines Gebäudes. Der Fotograf hat Wert darauf gelegt, die beiden Fenster des Gebäudes und die gesamte Tür (also auch deren oberen Bereich) mit auf dem Bild zu haben, obschon er die Gruppe aus größerer Nähe und somit in besserer Qualität hätte ablichten können, wenn er auf diese weitere Perspektive verzichtet hätte. Hierin dokumentiert sich, welche Bedeutung dem Gebäude beigemessen wird. Diese findet ihren Ausdruck auch in der Planimetrie, indem die Mitte der Tür auf der Bildmittelsenkrechten liegt, und die Fenster auf der Bildmittelwaagerechten aufrufen. Tür, Eingangsbeleuchtung und Fenster des Gebäudes bilden den Rahmen für die Gruppe bzw. bilden Gebäude und Gruppe auf diese Weise gewissermaßen eine Einheit.

Abbildung 22: Foto zur Jugendweihe der Schwester von Frau Telchow: Planimetrie und szenische Choreografie



Da es sich bei diesem Gebäude offensichtlich um ein öffentliches oder staatliches, um ein Amt oder (wie das ikonografische Vor-Wissen vermuten lässt) um eine Schule handelt und indem die Gruppe sich im Schutz dieses Gebäudes und im Eingangsbereich positioniert hat (die Tür ist leicht geöffnet), dokumentiert sich in diesem Foto eine besondere Nähe und auch ein – im wahrsten Sinne des Wortes – besonderer Zugang zu den staatlichen Institutionen. Vor dem Vergleichshorizont der Gruppe auf dem Foto der Kommunion von Frau Schiller, die in ihrer Positionierung auf dem freien Feld in gewisser Weise *isoliert* wirken, d.h. herausgelöst aus Bezügen, die sie einbetten und halten könnten, vermittelt uns diese Foto eine Sicherheit der Anbindung oder Einbindung. Dieser Eindruck wird gestützt durch die (im Vergleich zur Kommunion von Frau Schiller) wesentlich größere Gruppe.

Deutlich wird hier also eine Sicherheit und Eingebundenheit (wenn nicht Geborgenheit), die etwas mit der Nähe zu staatlichen Institutionen zu tun hat, den Institutionen der DDR-Gesellschaft. Diese Bindung an die DDR-Gesellschaft findet ja auch in dem Anlass der Feier und des vorliegenden Fotos ihren Ausdruck: die Jugendweihe, mit der als solche bereits die biografische Bindung an die DDR-Institutionen bekräftigt wird – ganz im Gegensatz zu Frau Schillers Kommunionfeier, deren Zelebrierung als Bekenntnis der Zugehörigkeit zur katholischen Glaubensgemeinschaft der kleinen Schar bereits als solche eine Randstellung und Isolation in der DDR-Gesellschaft zuweist. Dieser Frage gilt es auf der Grundlage der Gesprächsanalysen weiter nachzugehen.

Diese Bindung an die DDR-Institutionen erhält eine weitere Plausibilität auf der Grundlage unseres ikonografischen Vor-Wissens, dass Herr Telchow als Lehrer und Frau Telchow als Erzieherin in die DDR-Institutionen eingebunden waren, wohingegen die Mutter von Frau Schiller als Buchhalterin tätig war und – ebenso

wie die Eltern von Herrn Schiller – auch nicht über eine Hochschulausbildung verfügt.

Im Unterschied zum Foto der Kommunion von Frau Schiller steht hier die junge Frau, welche den Anlass für das Ritual bietet, auch im perspektivischen Zentrum (aus Raumgründen habe ich auf die grafische Darstellung der Rekonstruktion der Perspektivität verzichtet). Es dokumentiert sich somit in der planimetrischen und perspektivischen Komposition eine mit dem abgebildeten Ritual selbst verbundene Fokussierung und nicht – wie im Falle der Großmutter auf dem Foto der Schillers – eine spezifische Dominanz in der Familienhierarchie.

Was die szenische Choreografie anbetrifft, so stehen auch hier die Frauen im Zentrum, die Männer am seitlichen oder hinteren Rande. Diese Positionierung ist, wie gesagt, ein verallgemeinerbares Phänomen, welches sich der Tendenz nach überwiegend auf den uns zur Verfügung stehenden Familienfotos zeigt. Es deutet sich an, dass dies als ikonischer Ausdruck konventioneller Geschlechterverhältnisse gelten kann, denzufolge die familialen Binnenverhältnisse eine Angelegenheit der Frauen und die Außenvermittlung eine Angelegenheit der Männer ist.

Eine andere Dimension dieses Fotos betrifft die Jugendweihe als Übergangsritual, also als Ritualisierung des Übergangs von der Kindheit- zur Jugendphase. Die fünf Kinder auf dem Foto haben sich in einer insgesamt nicht sehr bequemen Haltung (die an die Startposition in einem Wettlauf erinnert) in die Hocke begeben. Dadurch wird der schon bereits vorhandene Größenunterschied zu den Erwachsenen noch einmal besonders markiert. Markiert wird insbesondere der Unterschied zur jungen Frau im Mittelpunkt, deren Statuspassage vom Kind zur Jugendlichen auf diese Weise optisch-ikonisch veranschaulicht wird. In diesem Sinne begeben sich die Kinder durch die Haltung die sie dabei einnehmen, schon einmal in die Startposition zum Übergang in die Jugendphase.

Hierin dokumentiert sich wiederum sehr deutlich, dass das Foto der Jugendweihe nicht allein als bildliche Repräsentation der Ritualisierung des Übergangs von der Kindheit zur Jugend zu verstehen ist, sondern selbst Bestandteil des Rituals ist, zu dessen Verdichtung beiträgt, wie Erving Goffman (1979:10) bemerkt hat.

#### *Komparative Analyse der Familientraditionen von Schiller und Telchow unter Einbeziehung der Gruppendiskussionen und Tischgespräche*

Wie in der Gruppendiskussion mit den Eltern Schiller sowie der Mutter von Frau Schiller und ihrem Lebensgefährten zum Ausdruck gebracht wurde, war ihrer Erfahrung nach die Zugehörigkeit zur katholischen Religionsgemeinschaft, wie sie in der Kommunion ihren Ausdruck findet, mit hohen Risiken der Ausgrenzung verbunden. Zugleich hatte sich die Mutter von Frau Schiller durch ihre Scheidung, die zum Zeitpunkt des Fotos wenige Jahre zurücklag, aber auch innerhalb der katholischen Religionsgemeinschaft marginalisiert. Diese Marginalisierung oder auch Isolation sowie die damit verbundene Unsicherheit vermögen sich hier ebenso zu dokumentieren wie der damit verbundene Druck und in gewisser Weise auch die Reaktion auf diesen Druck durch Strenge und Rigidität. Dies fügt sich zu dem in dem Foto der Kommunionfeier von Frau Schiller zu beobachtenden Habitus der *Rigidität und Strenge im Kontext des Provisorisch-Ungesicherten*.

In der Gruppendiskussion mit den Schillers und den Eltern von Frau Schiller wird betont, dass sowohl Herrn Schiller als auch Frau Schiller aufgrund ihres mit der Kommunion verbundenen Bekenntnisses der Zugang zum EOS-Abschluss verwehrt blieb. Hierin dokumentiert sich eine grundlegende Unsicherheit gegenüber den Institutionen der Ausbildung, die mit der doppelten Marginalisierung der Familie zu Zeiten der DDR aufgrund ihres christlichen Bekenntnisses allgemein und ihrer Zugehörigkeit zur katholischen Kirche im besonderen im Zusammenhang zu sehen ist.

Herr und Frau Schiller hatten erst auf dem zweiten Bildungsweg nach der Wende die Chance zu einer akademischen Ausbildung, die sie dann auch nutzten. Die (bereits erwähnte) ausgeprägte nahezu ‚verkrampfte‘ oder rigide Orientierung an biografischer Sicherheit ganz besonders auch im Bereich der Ausbildungsinstitutionen gewinnt vor diesem Hintergrund an Plausibilität: Wie das Tischgespräch zeigt, wird bereits die Wahl von Fächern in der Grundschule im Rahmen des Erwartungshorizonts der gesamten weiteren schulischen Ausbildung und einer unterstellten beruflichen Sicherheit abgehandelt.

Hiermit einher geht die Orientierung an Leistung zum Zwecke der biografischen Absicherung. Diese Leistungs- und Aufstiegsorientierung findet, wie dargelegt, auch in den Fotos ihren Ausdruck, die den Freizeit- und Urlaubsbereich betreffen, und strukturiert somit umfassend die Weltanschauung der Schillers.

Im Kontrast zur Orientierungsfigur des Erwerbs biografischer Sicherheit durch Leistung in der Familie Schiller bleiben Verhandlungen über Leistungsorientierungen aus dem Diskurs der Familie Telchow regelrecht ausgeklammert. Über die Möglichkeit, eine weiterführende Schule zu besuchen, entscheidet nicht die individuelle Leistung, sondern der Direktor.<sup>51</sup> Dieses Vertrauen in die Institutionen gewinnt seine Plausibilität vor dem Hintergrund der Einbindung der Großelterngeneration in die DDR-Institutionen, welche auch in dem Foto der Kommunion der Familie Telchow sich dokumentiert.

51 Dies verweist darüber hinaus auch auf den durchaus prekären Charakter des Themas individuelle Leistung in der Familie Telchow. Hinweise darauf finden sich auch in der Gruppendiskussion mit den Telchows und der Mutter von Frau Telchow. Damit zeigen sich aber – durch die Kontraste zwischen den Familien Telchow und Schiller hindurch – auch wiederum Gemeinsamkeiten. Denn auch in dem etwas rigiden und ‚verkrampften‘ Umgang mit den Leistungserwartungen in der Familie zeigt sich ein prekäres Verhältnis diesen gegenüber. Inwieweit dessen Wurzeln in der DDR-Sozialisation zu suchen sind, hätte in valider Weise erst in der komparativen Analyse mit weiteren (Familien-)Milieus in den neuen Bundesländern geklärt werden können (die, wie erwähnt, aufgrund fehlender Finanzierung allerdings nicht bis zur Auswertung durchgeführt werden konnte).

#### 4.3.6 Kommunion der Kinder der Familie Schiller

Abbildung 23: Foto zur Kommunion der Kinder der Familie Schiller



#### Formulierende Interpretation

##### Vor-ikonografische Interpretation

##### Zum Bildvordergrund:

Auf dem Foto sind 21 Personen abgebildet. Im Vorderfeld der Gruppe sind 7 Kinder, davon 4 Jungen und drei Mädchen im Alter von 4 bis 12 Jahren abgebildet. Im hinteren Feld sehen wir 13 Erwachsene, davon 7 Frauen und 5 Männer. Von einer weiteren, offensichtlich erwachsenen, Person sind lediglich die Kopfhare zu sehen.

Die Erwachsenen sind im Alter zwischen 30 und 45 Jahren. Der Mann links am Rand ist eher um die 50 Jahre alt. Zwei weibliche und zwei männliche Jugendliche im Alter zwischen 16–18 Jahren sind zu identifizieren. Eine der Frauen hält einen Säugling auf dem Arm.

Bis auf zwei Kinder, die eine Kerze in der Hand halten und von denen das Mädchen mit einem weißen langen Kleid und der Junge mit weißem Hemd und schwarzer Hose bekleidet ist, sind die sonstigen abgebildeten Personen eher alltäglich, also nicht festtäglich, gekleidet.

##### Zum Bildhintergrund:

Dieser wird durch einen sehr dichten Busch- und Baumbewuchs gebildet, der direkt hinter der Gruppe beginnt, sodass lediglich ein vorderer, aber nicht ein entfernterer Bildhintergrund gegeben ist. Durch diesen dunkelgrünen Hintergrund schimmert an wenigen Stellen der blaue Himmel durch.

#### Ikonografische Interpretation

Das Foto ist aus Anlass der Kommunion beider Kinder, Rebecka und Tobias Schiller, also der Doppelkommunion im Mai des Jahres 2000 aufgenommen worden. Rebecka ist 11 und Tobias 10 Jahre alt.

Bei den Personen auf dem Foto handelt es sich (nach Angaben von Frau Schiller) – von links nach rechts gehend – um folgende: Der Onkel von Frau Schiller (der auf dem Foto ihrer eigenen Kommuniionsfeier als zweiter von rechts dargestellt ist); der Sohn des Cousins von Frau Schiller (auf dem Foto der Kommuniionsfeier von Frau Schiller ist der Cousin der Dritte von rechts); weitläufige Verwandtschaft von Frau Schiller (die Frau mit Baby auf dem Arm und der Ehemann links neben ihr), mit der Frau Schiller auch befreundet ist (im Vordergrund: Tochter dieser Eheleute); Frau Schiller und davor ihre Tochter Rebecka; Freundin von Rebecka; Mutter von Herrn Schiller (der Vater von Herrn Schiller ist verstorben). Onkel von Herrn Schiller; Patentante von Tobias Schiller, dem männlichen Kommuniionskind (Bekannte von Frau Schiller aus ihrer Ausbildungszeit); Tante von Herrn Schiller; davor: Tobias Schiller; (im Hintergrund) Bruder von Herrn Schiller; links neben ihm sind lediglich Kopfhare einer offensichtlich erwachsenen Person zu sehen; vor dem Bruder von Herrn Schiller (halbverdeckt): Schwester von Herrn Schiller; davor: dessen Cousin; davor: Sohn des befreundeten Ehepaars von Frau Schiller; ganz rechts: Herr Schiller; links hinter ihm: dessen Cousine.

Die Mutter von Frau Schiller (auf dem Foto ihrer eigenen Kommuniionsfeier hinter ihr positioniert) und deren zweiter Ehemann waren (laut Informationen aus der Gruppendiskussion) zwar auf der Feier dabei, sind aber nicht auf dem Foto abgebildet, da sie gerade mit dem jüngsten Sohn David spazieren gehen. – Hinsichtlich der Person der Fotografin/des Fotografen besteht keine Sicherheit. Möglicherweise fotografiert auch hier, wie bereits auf dem Foto der Kommunion von Frau Schiller, wiederum ihre Tante.

#### Reflektierende Interpretation

##### Formale Komposition

##### Planimetrie

Die planimetrische Komposition könnte – im Vergleich zum Foto der Kommunion von Frau Schiller – kaum unterschiedlicher sein. Während das Foto der Kommunion von Frau Schiller durch waagerechte und senkrechte Linien gleichsam gitterartig strukturiert, ist es hier unmöglich, gerade Linien oder gar 90°-Winkel zu identifizieren. Auch der durch und durch unebene – nach links und nach hinten ansteigende – Boden bietet, wie gesagt, keine waagerechten Linien. Demgegenüber dominieren runde Formen, kreis-, ellipsen- oder eiförmige Gebilde. Diese werden einerseits durch die Laubbäume und Büsche konstituiert sowie den durch das Laub geformten Lichteinfall. Hieraus resultiert die Platzierung der Gruppe wie in einer Höhle. Andererseits – und vor allem – werden diese Formen durch das räumlich-körperliche Arrangement der Personen zueinander gebildet.

Abbildung 24: Foto zur Kommunion der Kinder der Familie Schiller: Planimetrie und szenische Choreografie



Dieses Arrangement bietet vergleichsweise klare Linien für die Rekonstruktion der planimetrischen Komposition. So lässt sich die gesamte dargestellte Gruppe in ihrer formalen Komposition durch zwei gleich große und einander teilweise überschneidende Kreise rekonstruieren, in deren Überlappungsbereich das weibliche Kommunionkind steht. Dieses ist auch vom linken und rechten Rand der Gruppe gleich weit entfernt.<sup>52</sup>

Wie häufig bei Gruppenfotos (vgl. auch 4.3.3) wird auch hier die Planimetrie durch das körperlich-räumliche Arrangement der Personen zueinander, also die szenische Choreografie, entscheidend bestimmt oder mit bestimmt, sodass uns die Rekonstruktion der Planimetrie zugleich wesentliche Aufschlüsse über die sozialen Beziehungen der abgebildeten Bildproduzent(inn)en zu vermitteln vermag. Die Rekonstruktion der planimetrischen Komposition führt uns zu den Besonderheiten der szenischen Choreografie.

#### Szenische Choreografie

Wenn wir unser ikonografisches – genauer: unser kommunikativ-generalisiertes – Vorwissen mit einbeziehen, kommen wir zu folgender Interpretation: Rechts vom

52 Allerdings ergibt sich in planimetrischer Hinsicht insofern eine Verschiebung des Bildausschnitts, der Kadrierung, gegenüber der abgebildeten Szenerie, der Gruppe, als zwar der linke Bildrand mit der Gruppe abschließt, nicht aber der rechte, und somit rechts ein Stück Busch- und Baumbewuchs sowie die von der Sonne beschienene Wiese mit abgelichtet ist (dadurch wird deutlich, dass der linke Teil der Gruppe nicht zwangsläufig im Schatten hätte stehen müssen). Der kompositorische Beitrag der *abbildenden* Bildproduzentin resp. des *abbildenden* Bildproduzenten ist somit nicht vollständig kongruent mit dem kompositorischen Aufbau, wie er durch die Gruppierung, die szenische Choreografie, hergestellt wird, welche wesentlich eine Leistung der abgebildeten Bildproduzent(inn)en darstellt (die sich z.T. im Schatten positioniert haben). Es gibt also eine gewisse Inkongruenz zwischen abgebildeten und abbildenden Bildproduzent(inn)en, die wir ggf. genauer interpretieren könnten, wenn wir sicher wüssten, wer hier fotografiert hat.

weiblichen Kommunionkind und auch von seiner Mutter, von Frau Schiller, steht (mit Ausnahme der besten Freundin des weiblichen Kommunionkindes und der Patentante des männliche Kommunionkindes) die Verwandtschaft von Herrn Schiller. Links von Frau Schiller finden wir die Angehörigen ihrer Verwandtschaftslinie, die eindeutig in der Minderheit sind, vertreten lediglich durch ihren Onkel und den Sohn ihres Cousins.

Abbildung 25: Foto zur Kommunion der Kinder der Familie Schiller: szenische (Teil-)Choreografie



Als ein weiteres Element der szenischen Choreografie findet sich jenseits der Verwandtschaftslinien bzw. der verwandtschaftlichen Kreise und sozusagen quer zu ihnen ein Ensemble, welches aus drei Frauen (Frau Schiller, einer Freundin von ihr und der Mutter von Herrn Schiller) sowie dem Kommunionkind besteht und sich als ein haus- oder zeltförmiges Gebilde charakterisieren lässt. Hier zeigen sich Übereinstimmungen zum Foto der Kommunion von Frau Schiller (4.3.5). Auch hier stehen drei Frauen hinter und neben dem weiblichen Kommunionkind.

In dieses Ensemble sind hier aber – im Unterschied zum Foto der Kommunion von Frau Schiller – abgesehen von dem Säugling noch zwei weitere weibliche Kinder aufgenommen: das kleine Kind der Verwandtschaft und die Freundin von Rebecca. Dennoch bleibt das Ensemble in seiner sozialen Grund-Konstellation und auch in seinen zelt- oder hausförmigen Umrissen (vgl. 4.3.5) erhalten.

#### Perspektivität

Da sich keine waage- oder senkrechten Linien und infolge auch keine rechten Winkel identifizieren lassen und der Boden durch und durch uneben ist, ist die Identifikation des perspektivischen Zentrums, des Fluchtpunkts nicht möglich.

Die Atmosphäre dieses Fotos ist gegenüber derjenigen der Kommunion von Frau Schiller (4.3.5) vollständig verändert. Im Kontrast zu dem älteren Foto, in dem die Gruppe gleichsam isoliert auf freier Fläche ohne Mittelgrund steht, ist hier die Gruppe durch den üppig grünen Mittelgrund aus Bäumen und Hecken umfassend eingerahmt. Dies wird durch die Schattenwirkung noch verstärkt. Nicht nur der Hintergrund ist schattig, sondern auch nach vorne hin ist die Gruppe durch Schatten abgeschirmt. Sie steht wie in einer Höhle, in deren Eingang von außen Licht hereinfällt.

Dieses Foto vermittelt somit den Eindruck einer Geschlossenheit und auch Geborgenheit, die auf dem älteren Foto der Kommunion von Frau Schiller vollständig fehlt und mehr noch in Kontrast zum dortigen Eindruck der sozialen Isolation und Ungesicherheit steht.

Die Familie Schiller ist – im Unterschied zur Aussage des nicht lange nach der Scheidung der Eltern des Kommunionkindes aufgenommenen Foto der Herkunftsfamilie von Frau Schiller – eine vollständigere, geborgene und ‚intaktere‘ Familie. Dies findet seinen Ausdruck auch in der größeren Verwandtschafts- und Freundesgruppe (wobei allerdings die Doppelkommunion zu berücksichtigen ist).

Es zeigt sich aber zugleich auch, dass es die Linie von Herrn Schiller ist, die den entscheidenden Beitrag zu diesem intakten und geborgenen Familien- und Verwandtschaftsgefüge leistet. Den beiden Verwandtschaftslinien ist – obschon sie zahlenmäßig völlig ungleichgewichtig vertreten sind – von der planimetrischen Aufteilung her gleich viel Raum gelassen. Die Angehörigen der Linie von Frau Schiller sind nur sehr unvollständig präsent. Die Mutter von Frau Schiller und ihr Stiefvater haben (wie wir wissen) zwar an der kirchlichen Kommunionfeier teilgenommen, sich aber der gemeinsamen Präsentation im Verwandtenkreis entzogen. Sie gehen mit dem jüngsten Sohn spazieren.

Diese ‚Lückenhaftigkeit‘ in der Linie von Frau Schiller findet auch in der Art und Weise ihren Ausdruck, wie sich die dennoch Anwesenden aus dieser Linie positioniert haben – in relativ großer Distanz und in diesem Sinne vergleichsweise ohne Verbindung zueinander. Demgegenüber stehen die Angehörigen der Verwandtschaftslinie von Herrn Schiller nahe beieinander, mit engem Körperkontakt zueinander.

Die Verwandtschaftslinien sind zudem separiert, wobei das Foto und seine Atmosphäre durch die Linie von Herrn Schiller dominiert werden. Diese steht auch im Licht, während die hier kaum repräsentierte Linie von Frau Schiller noch dazu im Schatten positioniert ist. Wobei, wie bereits erwähnt, dies zu vermeiden gewesen wäre, da rechts neben der Gruppe die Wiese von der Sonne beschienen ist.

Trotz des Eindrucks einer relativen Geschlossenheit und Geborgenheit (zumindest des rechten Teils der Gruppe) ist die Mimik insbesondere der Frauen eher angestrengt als heiter und gelöst. Dies gilt insbesondere für die einzig abgebildete Großmutter: die Mutter von Herrn Schiller. Diese Begegnung zweier Familienlinien und die rituelle ikonische Präsentation und Verdichtung dieser Begegnung scheinen anstrengend zu sein, wie ja auch die Separierung der Gruppen und das Fehlen der anderen Großmutter und ihres zweiten Ehemanns vermuten lässt.

Was die *Geschlechterverhältnisse* anbetrifft, so stehen auch hier – wie im Foto zur Kommunion von Frau Schiller (4.3.5) und zur Kommunion der Schwester von

Frau Telchow (4.3.6) – die Männer außen und am Rande. Entweder bleiben sie im Hintergrund oder sie flankieren das Geschehen wie ganz rechts Herr Schiller oder ganz links der Onkel von Frau Schiller (der nahesten der anwesenden Verwandten von Frau Schiller). Diese Außenpositionierung der Männer, die Flankierung der Familie nach außen, ist, wie bereits angesprochen, ein verallgemeinerbares Phänomen, welches sich der Tendenz nach auf allen uns zur Verfügung stehenden Familienfotos zeigt (hier abgedruckt in 4.3.4 u. 4.3.5) und somit – in Relation zur Innenpositionierung und Zentrierung der Frauen – als ikonischer Ausdruck konventioneller Geschlechterverhältnisse gelten kann.

Hinzu kommt im Bereich der Geschlechterverhältnisse hier in der Familie Schiller jedoch eine Fokussierung des weiblichen gegenüber dem männlichen Kommunionkind durch die Strukturierung im Bereich der planimetrischen Komposition und der szenischen Choreografie. Damit einher geht eine Fokussierung der anwesenden weiblichen Kinder.

Wie dargelegt, lässt sich im Bereich der szenischen Choreografie ein Ensemble von drei Frauen um das weibliche Kommunionkind herum als ein haus- oder zeltförmiges Gebilde ausmachen. In diese ‚Schutzbegleitung‘ sind – im Unterschied zum Foto der Kommunion von Frau Schiller – abgesehen von dem Säugling noch weitere weibliche Kinder aufgenommen: eine kleine Tochter der Verwandtschaft und eine Freundin von Rebecka, sodass die Peer-Bindung in dieser Präsentation Bedeutung gewinnt. In beiden Fotos kommt eine Dominanz der Frauen bei gleichzeitiger Solidarität untereinander zum Ausdruck, wobei dies im jüngeren Foto durch eine Dominanz der Mädchen noch ergänzt wird.

Denn obschon hier auch die Kommunion des Sohnes Tobias gefeiert wird (wie auch an der Kommunionkerze in seiner Hand zu erkennen ist), erscheint das männliche Kommunionkind doppelt marginalisiert: Er bleibt einerseits von der gesamten szenischen Choreografie und der durch sie bestimmten planimetrischen Komposition her eher unbeachtet. Andererseits bildet sich um Tobias herum kein Ensemble. Auch sein Vater steht vergleichsweise weit entfernt von ihm.

Im Unterschied zur Kontinuität im Bereich der *Geschlechterverhältnisse*, in dem die Dominanz der Frauen sich weiterhin abzeichnet, zeichnet sich im Bereich der *innerfamiliären Generationsverhältnisse*, also der Eltern-Kindbeziehungen, im Vergleich mit dem Foto der Kommunion von Frau Schiller (4.3.5) eine stärkere Kindzentrierung ab (die allerdings eben durch eine Fokussierung der weiblichen Kinder überlagert wird). Durch die Anwesenheit der vergleichsweise großen Zahl von Freunden und Freundinnen der Kommunionkinder wird deutlich, dass die Kommunion in dieser Generation nun nicht mehr primär ein Fest der Erwachsenen, sondern zunehmend eines der Kinder ist.

### *Homologien zum Orientierungsrahmen der Gruppendiskussion und Tischgespräche*

In der Gruppendiskussion betont Frau Schiller in Bezug auf die Doppelkommunion ihrer beiden Kinder, dass der Sohn noch „nicht reif genug“ gewesen sei (GD Kommunion 2; 86). Da er sich so „unmöglich“ verhalten habe, würde sie ihn nachträg-

lich exkommunizieren: „würd ich exkommunizieren@wenn ich dazu die Macht hätte@der is unmöglich ja? also et=is nich dass ich so überzogen äh (2) religiös bin oder g-so irgendwie äh“ (*Transkript Schiller GD Kommunion 2: 87-89*).

Zugleich erfahren wir in der Gruppendiskussion aber auch, dass der Sohn sich (mehr oder weniger selbstverständlich) dem Kommunionstermin seiner Schwester anzupassen hatte (und nicht umgekehrt) und somit zugunsten seiner Schwester ein Jahr ‚zu früh‘ die Kommunion empfangen hat. Abgesehen davon, dass die „Unreife“ des Sohnes ja möglicherweise eben gerade darauf zurückzuführen ist, wird hier noch einmal das bestätigt, was auf dem Foto zur Kommunion in den unterschiedlichen Positionierungen der beiden (Kommunions-) Kinder seinen Ausdruck gefunden hat: die Vorrangigkeit der Tochter vor dem Sohn (im Bereich familialer Rituale).

Während „Rebecka n exklusivt Kleid ge-genäht bekommen“ hat (114) (ein Kleid im mittelalterlichen Stil), hatte der Sohn Tobias „nur son Anzug an den wir zum Glück nur geliehen hatten weil er fand den total schrecklich“. Der Sohn musste sich mit einem geliehenen Anzug begnügen, der „eigentlich n bisschen zu kurz“ war. Demgegenüber wurde das Kommunionkleid für die Tochter zum „wirklichen Erlebnis“ (*Transkript Schiller GD Kommunion 2: 201-217*):

- 201 M: Ja und dann hat=er halt und dann is dat ja auch noch was anderet Tobias der hatte  
 202 nur son Anzug an, den wir zum Glück nur geliehen hatten weil er fand den total  
 203 schrecklich und (.) wenn er die Arme n bisschen hob, @( ) verrutscht@ also  
 204 war der eigen- eigentl n bisschen zu kurz aber er war so so schmal dass er  
 205 eigentlich da (.) sehr gut drin aussah solange die Arme halt unten waren und er h-  
 206 sagte immer (.) oh dieser Anzug also der fühlte sich wahrscheinlich schon in  
 207 Y2: | Mhm  
 208 M: seiner Haut nicht wohl. mit diesen Sachen. und sie hatte nun sich dieses Kleid da  
 209 V: | Mhm  
 210 Y2: | Mhm:  
 211 Y1: | Hm hm  
 212 M: war ja da waren wir ja zichmal nach X-stadt gefahren zum (.) äh weil dat genäht  
 213 V: | Mhm.  
 214 M: wurde und dann zum Abstecken und dat war ja wirklich nu n Erlebnis und ich  
 215 Y2: | Mhm  
 216 M: meine wenn=de son Kleid anhat dat erste Mal (.) und dann hat se son ganz tollen  
 217 Myrtenkranz dazu gehabt

Zum einen bestätigt sich hier das, was in den Fotos der Familienrituale generell seinen Ausdruck findet: dass die familialen Feste eine Angelegenheit der Frauen sind. Nicht nur, weil sie von diesen organisiert werden, sondern auch, weil es die Frauen sind, die sich hier präsentieren, hier im Fokus stehen, während die Männer diese ‚flankieren‘. Dass dies nicht allein auf familiale Rituale begrenzt ist, zeigen andere Fotos (vgl. „Gartenfest“ der Familie Telchow und „Fahrradrast“ der Familie Schiller; 4.3.3). Diese innerfamiliale Dominanz der Frauen strukturiert – wie die Tischgespräche in der Familie Schiller und die anderer Familien zeigen<sup>53</sup> – den Alltag insgesamt.

In der Familie Schiller zeigt sich allerdings in dieser evidenten Ungleichheit der Positionierung der beiden von ihrer rituellen Funktion her gleichgestellten Kin-

53 Siehe dazu die im Rahmen der Lehrforschung entstandenen Abschlussarbeiten von Klar (2007) und Foltys (2006). Die Tischgespräche der Familie Telchow eignen sich für derartige Analysen nicht, da die Eheleute – bedingt durch den Schichtdienst – selten gemeinsame Mahlzeiten einnehmen.

der eine darüber hinaus gehende Tendenz zu einer geschlechtsspezifischen Polarisierung, die einerseits in der Gruppendiskussion ihre unmittelbare Bestätigung findet. Andererseits dokumentiert sich dies auch in den Tischgesprächen, vor allem dort, wo in der Auseinandersetzung der Mutter mit dem älteren Sohn Mutter und Tochter an einem Strang ziehen, der jüngere Sohn (5 Jahre) sich demgegenüber (mit dem ihm zu Verfügung stehenden Mitteln) mit dem älteren Sohn (11 Jahre) solidarisiert und dabei auch explizit geschlechtsspezifische Gemeinsamkeiten anspricht (Sport ist etwas für die Jungs bzw. die Männer, denn dadurch „wird man schön stark“; *Transkript Schiller TG Differenzierungsstunde: 296*). Der Vater nimmt wiederum eine vermittelnde Funktion in dieser Auseinandersetzung ein.

Wie bereits angesprochen (4.2.3), ist es die Mutter in der Familie Schiller, welche die biografisch relevanten Situationen strukturiert und die Entscheidungen trifft. In jener bereits erwähnten Situation während eines abendlichen Tischgesprächs, in der es um die Wahl des Sohnes hinsichtlich einer spezifischen Art des Schulfachs geht, bereitet der Vater die Situation für die Entscheidung in der Weise vor, dass er entscheidungsnotwendige Themen einbringt und die Problem Darstellungen initiiert, wobei er aber in seinen Redebeiträgen die zu treffenden Entscheidungen bzw. die den Kindern zu vermittelnden Orientierungen systematisch ausklammert (d.h. er initiiert neue Themen, ohne Propositionen zu erbringen).

Der Vater überlässt der Mutter die Stellungnahmen und die pädagogischen Interventionen, welche die biografischen Orientierungen im Bereich der Schule und Ausbildung betreffen – und dies, obschon er, wie der Diskurs zeigt, in mancher Hinsicht besser über Fragen der Unterrichtsorganisation informiert ist als seine Frau.

Homolog zu unserer Beobachtung der familialen Interaktion zeigt unsere Rekonstruktion der (retrospektiven) Darstellungen der biografisch relevanten Entscheidungssituationen in der Gruppendiskussion mit den Eltern und Großeltern (mit der Mutter und dem Stiefvater von Frau Schiller), dass diese Darstellungen fast ausschließlich von Frau Schiller und ihrer Mutter getragen worden sind. Dies gilt insbesondere für die Rekonstruktion der Heirat der Eheleute Schiller. Frau Schiller betont zu Beginn der gesamten Diskussion, dass sie ja eigentlich den „Schein wahren“ und ihrem Mann den Vortritt bei der Darstellung der Situation der Hochzeit lassen müsse (*Transkript Schiller GD Hochzeit: 09-12*):

- 09 M: | Naja ick (.) muss hier aber den  
 10 Schein wahren und da @ (heute vielleicht mal)@ den Vortritt lassen  
 11 V: | Nee, nee, (1) nee nee wieso? (1)  
 12 fühl ich mich doch gar nicht so berufen also

Herr Schiller fühlt sich nicht „berufen“, d.h. biografisch nicht disponiert, den „Schein zu wahren“ und den konventionellen Rollenvorstellungen zu entsprechen (diese Interpretation gewinnt allerdings erst unter Berücksichtigung des Kontextes Validität), indem er als Vater die Strukturierung der Darstellung übernimmt. Bemerkenswerterweise übernimmt Frau Schiller die Strukturierung des Diskurses erst, nachdem ihre Mutter sie in ironisierender Weise (die Tochter hat im Zuge der Verhandlung über die Strukturierung des Diskurses einen ebenso ironisierenden Unterton wie ihre Mutter) mehr oder weniger dazu aufgefordert hat.



Hierin dokumentiert sich, dass die Großmutter diejenige ist, die in letzter Instanz für die Strukturierung des Diskurses zuständig ist. Es sind die Frauen und vorzugsweise die älteren Frauen, die – wie ihnen von Seiten der Männer ohne weiteres zugestanden wird – zumindest in Bezug auf familienspezifische Angelegenheiten den Diskurs strukturieren. Dies erweist sich, wie der Vergleich mit dem Kommunionfoto 4.3.4 zeigt, als ein über die Generationen hinweg gültiges (tradirtes) Muster.

In Bezug auf den Entschluss zur Heirat wird deutlich, dass dieser durch die Schwangerschaft von Frau Schiller zwar beschleunigt, vorgezogen oder „vorge-schoben“ wurde (30), an der wechselseitigen Zuneigung und der grundlegenden Heiratsabsicht aber wohl keinerlei Zweifel bestand. Im Unterschied zu dieser emotionalen Sicherheit in Bezug auf die Zuneigung zu ihrem Mann artikuliert Frau Schiller durchaus ihre „Angst vor dem was kommt oder so“, d.h. ihre Unsicherheit im Hinblick auf das biografische Risiko einer Heirat überhaupt.

Auch hier wird der Erwartungshorizont einer *starken Orientierung an biografischer Sicherheit, an Planungssicherheit* deutlich, wie dieser sich (im Rahmen der Tischgespräche) vor allem auch in Bezug auf die schulische Ausbildung und dort in einer in gewisser Weise rigiden Orientierung an biografischer Absicherung dokumentiert und dessen Soziogenese in der doppelten Marginalisierung (in der DDR durch die Zugehörigkeit zur katholischen Glaubensgemeinschaft und durch die Scheidung der Mutter) zu suchen ist.

Indem die Großmutter mehrfach betont, wie „stolz“ sie ist, dass die Ehe erfolgreich verlaufen sei, wird zweierlei deutlich: *Zum einen* wird die Bewältigung der Ehe und damit des Lebens als *Leistung* zugerechnet. Diese Betonung der eigenen praktischen Lebensführung als Leistung, welche, wie dargelegt (4.3.2) den Orientierungsrahmen der Familie Schiller prägt, gewinnt *zum anderen* vor dem Hintergrund der eigenen Scheidung an Bedeutung. Diese muss somit als ein Versagen erscheinen, insbesondere der Ehefrau, der ja offensichtlich (zumindest in dieser Familie) die entscheidende Bedeutung bei der biografischen Strukturierung zukommt. Die Angst, dass auch die Tochter in dieser Hinsicht versagen könnte, sitzt tief.

Damit wird auch der ausgeprägte „Stolz“ plausibel, den die Großmutter nachdrücklich zum Ausdruck bringt. Die Großeltern können „stolz“ sein, weil die Kinder das Problem, bei dem sie selbst (auch der Stief-Großvater ist geschieden) versagt haben, meistern konnten und damit quasi einen Generationenauftrag erfüllt oder eingelöst haben, wie sich vor allem auch in dem Foto zur Kommunion der Kinder hinsichtlich Geborgenheit und sozialer Eingebundenheit dokumentiert.

Explizit thematisiert wird die Angst vor dem ungewissen Zukunftshorizont in der Familie Schiller also lediglich von Frau Schiller, deren Versuch der Elaboration dieses Themas von den Großeltern dann recht massiv unterbunden wird. Auch in dieser Hinsicht will Frau Schiller nicht „den Schein wahren“, den aufrecht zu erhalten, die Großeltern sich bemühen. Dies zeigt sich in der ausgesprochen offenen Haltung von Frau Schiller während der gesamten Gruppendiskussion, in ihrer Bereitschaft, auch brisante Themen diskursiv zu verhandeln. Dieser Bruch mit der Wahrung des Scheins findet dann auch in der Diskursorganisation der Tischgespräche seinen Ausdruck sowie in der Offenheit hinsichtlich der zu verhandelnden *Themen*, die allerdings mit dem konfrontativen und provokativen Stil von Frau Schiller

einhergeht und einer geringen Offenheit dem *Orientierungsrahmen* der Kinder gegenüber (genauer dazu: Bohnsack 2008a: 12.2).

Der Weg von der Zuneigung zwischen Frau und Herrn Schiller hin zur Heirat, auf dessen Darstellung sich die Erzählungen konzentrieren, wird von beiden Frauen (Mutter und Großmutter) als ihre Sache dargestellt. Insbesondere die Großmutter bemerkt eher am Rande, dass ihr Mann ja „überhaupt nichts gewusst“ habe (*Transkript Schiller GD Hochzeit: 94-111*):

94 MM: [Naja er hats ja erst mal überhaupt nicht gewusst, dat war ja erst mal erst mal  
95 haben wir dat ja untereinander ausgemacht, da hatte ich ja die Schwierigkeit,  
96 weil ja Juliane im Grunde jenommen meine Tochter is ja und ((EA)) (da hat  
97 ?: [ ((flüstert)): jetzt kommt's raus |  
98 M: [ Also dis weeß  
99 ich schon gar nicht mehr.  
100 MM: man erst), darüber gar nicht gesprochen hab, für mich war dat natürlich  
101 besonders schwierig, die ganze Situation, ja, und (1) seine Eltern, die waren  
102 V: [ Hm,  
103 Y1: [ Kuchen?  
104 M: (Nee noch nich mehr) @(. )@  
105 MM: ja och nicht so begeistert davon, da hatte ich ja och noch n' ganz schönen  
106 M: [ ((atmet tief ein)) Hoach: @(. )@  
107 MM: Kampf auszustehen (.) bei seinen Eltern zu Hause da bin ich dann alleine  
108 hingefahren ja, und habe da mit ihnen und sie warn ja eigentlich (1) strikt  
109 dagegen.  
110 M: [ Ach so, (.) @(. )@  
111 det weiß ick ja überhaupt nich mehr,

Ohne Wissen ihres Mannes hatte die Großmutter den „Kampf auszustehen“ mit den Eltern von Herrn Schiller. Somit ist sie es und nicht Herr Schiller, die dessen Eltern zu überzeugen und möglicherweise auch zu informieren hatte. Herrn Schiller, der auch ansonsten keine Rolle spielt im Zuge der Anbahnung der Heirat, wird somit nicht einmal die Kompetenz zuerkannt, seine eigenen Eltern zu überzeugen. Die Großmutter hat alle Fäden in der Hand.

Und in dieser Hinsicht, d.h. im Hinblick auf eine derartige Komponente einer weiblichen Dominanz in der Familienstruktur, zeigen sich, wie gesagt, Homologien zwischen Mutter und Großmutter, die – wie das Foto der Kommunion von Frau Schiller (3.4.3) zeigt – bis zur Ur-Großmutter reichen und die insgesamt in den Fotos der Familie Schiller ihren Ausdruck finden.

#### 4.3.7 Schluss: Triangulation, Vertiefung und Validierung

Es lässt sich somit zeigen, dass es bereits auf der Basis der Interpretation weniger Fotos möglich wird, grundlegende Komponenten des familialen Orientierungsrahmens zu identifizieren, die in der Triangulation mit den Gesprächsanalysen dann validiert, vertieft und differenziert werden können. Die Bildinterpretation ermöglicht es, einen Einblick in hochkomplexe Zusammenhänge des Familienmilieus zu gewinnen, sodass von hier gezielte Fragen an die Gesprächsanalysen bzw. Textinterpretationen gestellt, zugleich aber auch die Interpretationen der Gespräche in ei-

nem präziseren, vor allem aber komplexeren, Rahmen gefasst werden können. Voraussetzung hierfür ist eine gemeinsame metatheoretische und erkenntnistheoretisch-methodologische Basis, wie wir sie in der dokumentarischen Methode und der praxeologischen Wissenssoziologie finden. Die hierin fundierte Analyseeinstellung und die methodischen Standards können gleichermaßen der Bild- wie der Textinterpretation zugrunde gelegt werden, ohne aber die je unterschiedlichen Eigenlogiken der Bild- und Textdimension zu verwischen.

In der durch die Bildinterpretation zunächst angeleiteten komparativen Analyse ließen sich deutliche milieuspezifische Differenzen zwischen zwei Familien rekonstruieren, welche auf den ersten Blick, d.h. hinsichtlich formaler Indikatoren im Bereich der Ausbildung und beruflichen Tätigkeit der Eltern, weitgehende Übereinstimmungen aufweisen. Diese milieuspezifischen Differenzen betreffen das Verhältnis von Kinder- und Erwachsenenwelt, also der innerfamilialen Generationen zueinander, ebenso wie die Geschlechterverhältnisse und vor allem auch die Sicherheit resp. Unsicherheit gegenüber den gesellschaftlichen Institutionen (insbesondere der Ausbildung) und das Verhältnis zur Leistungsorientierung. In letzterer Hinsicht bietet bereits die Interpretation weniger Fotos in der vertikalen und horizontalen komparativen Analyse auch erste Hinweise auf die Soziogenese dieser milieuspezifischen Differenzen in der DDR-Sozialisation. Es zeigen sich Differenzen zwischen der Herkunft aus einem in der DDR integrierten und einem in der DDR marginalisierten Milieu.

Die Methodentriangulation, hier also die Triangulation der Interpretation von Fotos mit textbasierten Verfahren, ermöglicht eine Validierung und Vertiefung der an den Bildern erarbeiteten Komponenten der familialen Orientierungsrahmen und ihrer milieuspezifischen Differenzen und umgekehrt. Möglichkeiten der Generalisierung einer derart sich abzeichnenden Milieutypik (vgl. Bohnsack 2009a) wie auch die Rekonstruktion des Verhältnisses der Milieutypik zu einer (wesentlich vom Alter der Kinder abhängigen) Entwicklungstypik von Familien und somit erste Ansätze zur Bildung einer Typologie, würden allerdings eine weitergreifende komparative Analyse voraussetzen.

## 5. Dokumentarische Video- und Filminterpretation

In diesem Kapitel wird es darum gehen, vor dem Hintergrund der Auseinandersetzung mit vorliegenden Theorien und Methoden zur Filminterpretation aus den Bereichen der qualitativen Sozialforschung, also der Sozialwissenschaften, aber auch der Filmwissenschaften und der Filmtheorie, ein eigenes Auswertungsverfahren auf der Grundlage der dokumentarischen Methode zu entwickeln und zu begründen, welches dann im nächsten Kapitel (6) anhand einer Fernsehshow forschungspraktisch demonstriert wird.

### 5.1 Video und Film als Erhebungsinstrument oder Alltagsdokument

Wie bei Einzelbildern so lassen sich auch im Bereich von Filmen und Videos grundsätzlich zwei Wege unterscheiden, auf denen diese in der empirischen Forschung Verwendung finden können: Auf der *einen Seite* haben wir es mit solchen Filmdokumenten zu tun, die in den jeweiligen kulturellen Lebenszusammenhängen oder Milieus, welche Gegenstand der sozialwissenschaftlichen Analyse sind, *selbst* produziert worden sind. Dies gilt einmal für den *privaten* Bereich – beispielsweise für Familienvideos, die von Angehörigen der Familie aufgenommen worden sind. Diese geben uns dann Aufschluss über die Handlungspraxis und den Erfahrungsraum, das Milieu bzw. den Habitus einer Familie.

Ähnliches gilt aber auch für Filme und Bilder aus dem *öffentlichen* Bereich, also für den Bereich der *Massenmedien*, beispielsweise für Fernsehsendungen oder Werbefilme. Hier erhalten wir u.a. Einblicke in die durch die Medien transportierten oder propagierten Lifestyles bzw. in den Stil oder Habitus, welcher durch die Akteurinnen und Akteure, die Moderator(inn)en und Bildproduzent(inn)en präsentiert und transportiert wird. Diese Akteure der öffentlichen Bildmedien erscheinen, wie beispielsweise in der Analyse der Fernsehshow von Stefan Raab in Kapitel 6 deutlich wird, als Repräsentanten und Identifikationsfiguren von (konjunktiven) Erfahrungsräumen und rudimentären Milieus, die sie mit den Zuschauern teilen, und bestätigen somit einen spezifischen Habitus (beispielsweise hinsichtlich des Umgang mit der Konstruktion des kulturell Fremdem).

Von jenen visuellen Grunddaten, die insgesamt Eigenprodukte der Erforschten sind, lassen sich – auf der *anderen Seite* – solche unterscheiden, die zu Forschungs-

zwecken produziert werden. Letzteres ist dort der Fall, wo Foto, Film und Video als *Erhebungsinstrument* der (sozial-) wissenschaftlichen Forschung Verwendung finden. Für diesen Bereich der Videoanalyse liegen auf der Grundlage der dokumentarischen Methode bereits exemplarische Analysen vor. Beispielsweise finden sich Videografien von Schülerinteraktionen im Klassenraum (siehe Wagner-Willi 2004, 2005, 2006 u. 2007) oder von Patientinnen mit „psychogenen Bewegungsstörungen“ während der Chefarztvisite (siehe Klambeck 2007) oder von Vorschulkindern, die sich untereinander und mit Erwachsenen über ihre alltäglichen Praktiken des Spiels verständigen (Nentwig-Gesemann 2006a u. 2007a).

Dort, wo Videos und Filme zu Forschungszwecken erstellt werden, interessieren in der Regel lediglich die Gestaltungsleistungen der *abgebildeten* Bildproduzenten. Videografien und Filmdokumente, die von den Erforschten selbst produziert worden sind, interessieren demgegenüber sowohl hinsichtlich der Gestaltungsleistungen der *abgebildeten* wie auch der *abbildenden* Bildproduzent(inn)en, also der Personen vor und hinter der Kamera in ihrem Zusammenspiel. Für diesen Bereich der Videoanalyse liegen auf der Grundlage der dokumentarischen Methode erst seit kurzem exemplarische Analysen vor. So hat Astrid Baltruschat (2008 u. 2009) Videos ausgewertet, in denen Lehrer/innen und Schüler/innen (im Rahmen eines Wettbewerbs) Problemen ihres eigenen Alltags in verdichteter Weise Ausdruck verliehen haben (vgl. auch Kap. 5.6.4). Im vorliegenden Band gehe ich selbst in Kapitel 6 am Beispiel der Interpretation einer Fernsehshow exemplarisch auf diesen Bereich der dokumentarischen Videoanalyse ein.

Die grundlegenden Unterschiede des methodischen Zugangs sind also weniger davon abhängig, ob es sich um Produkte der Massenmedien oder um außermediale Produkte handelt, sondern davon, ob die Produkte eigens zu *Forschungszwecken* erstellt worden sind oder ob sie als *Alltagsprodukte* zum Gegenstand der Forschung werden.<sup>54</sup> Wenn auch bei den zu Forschungszwecken erstellten Produkten die Gestaltungsleistungen der *abbildenden* Bildproduzent(inn)en, also der Forscher/innen, nicht zum eigentlichen Forschungsgegenstand gehören, müssen diese gleichwohl in Rechnung gestellt, also methodisch kontrolliert werden. Dies betrifft u.a. die Kameraführung, die Wahl der Perspektivität und der Einstellungsgröße wie auch eventuelle Montageleistungen.

54 Der Begriff des Films als „Alltagsdokument“ umfasst hier also alles, was nicht zu Forschungszwecken erstellt worden ist, somit auch den ‚artifizell gestalteten Film‘, den Faulstich (2002: 16) vom Film als Instrument der Aufzeichnung und Konservierung von Alltagssituationen unterscheidet. Für den hier gewählten methodischen Zugang ist diese Unterscheidung nicht von Bedeutung. Auch Marotzki/Schäfer (2006: 62) betonen: „Generell gelten für die Auseinandersetzung mit Film und Video in der qualitativen Sozialforschung die gleichen filmsprachlichen Codes und Filmstandards wie beim fiktionalen Film.“ Die in Kap. 3.9 mit Bezug auf die Fotografie aufgeworfene Frage, ob die dokumentarische Methode, die sich ja teilweise auf die Kunstgeschichte stützt, auf profane ebenso wie künstlerische Produkte Fotografie Anwendung finden kann, stellt sich hier in ähnlicher Weise. Wie dargelegt, sind die wesentlichen Fortschritte der qualitativen Methoden im Bereich der Textinterpretation darin zu sehen, dass sie diese Texte mit ähnlicher Sorgfalt und Systematik rekonstruiert hat, wie sie die Literaturwissenschaft künstlerisch anerkannten Produkten angedeihen lässt. Einen derartigen Respekt vor der profanen Alltagskultur, der „profane culture“, den die dokumentarische Methode im übrigen mit den Cultural Studies teilt, ist nicht nur im Bereich der Bild-, sondern auch der Filminterpretation geltend zu machen.

Wie Ray L. Birdwhistell (1970: 151), der Klassiker der Gebärdenanalyse, betont, ist die filmische Aufzeichnung zu Forschungszwecken – im Vergleich zu anderen Aufzeichnungsmedien – in ihrer spezifischen Selektivität außergewöhnlich schwer zu kontrollieren, da sie uns in ihrer Ausrichtung in vielen Bereichen vollkommen natürlich erscheint: „Es ist außerordentlich schwierig, ständig auf der Hut zu sein hinsichtlich des Ausmaßes an Kontrolle, welche durch die Fokussierung und die Selektion des Kameramannes und seines Aufnahmeteams ausgeübt wird. Nahaufnahmen erscheinen den erfahrenen Betrachtern passend, der Schwenk der Kamera von den Sprechern zum Hörer oder von Sprecher zu Sprecher erscheint ebenso natürlich. Wir alle, die wir in westliche und vor allem amerikanische dramaturgische Konventionen eingeübt sind, werden dadurch beeinflusst.“

Die methodische Kontrolle der Einflüsse der Forscher/innen als abbildende Bildproduzent(inn)en gelingt also nur dann, wenn auch im Falle von Fotos oder Videos, die zu Forschungszwecken erstellt worden sind, Erkenntnisse hinsichtlich Kameraführung, Perspektivität, Einstellungsgröße etc., wie sie vor allem im Bereich der Filmwissenschaft ausgearbeitet worden sind, zumindest in Ansätzen in die methodische Reflexion einbezogen werden. So betont Erving Goffman (1990: 31): „jede einzelne Handlung einer fotografischen Aufzeichnung oder einer Filmaufzeichnung ist eine interpretative Handlung, weil sie notwendigerweise selektiv ist.“

Die Auseinandersetzung darüber, inwieweit Forscherinnen oder Forscher auch durch eine auf einem Stativ vollkommen unbeweglich positionierte Filmkamera bereits einen Einfluss ausüben, geht bis in die Ursprünge der sozialwissenschaftlichen bzw. kulturanthropologischen Nutzung des Mediums zurück. Margaret Mead und Gregory Bateson (2002) haben darüber bereits in den 40er Jahren des 20. Jh. Kontroversen ausgetragen. Howard S. Becker (1986) bemerkt dazu: „Mead schien nicht gesehen zu haben, dass auch die Positionierung der Kamera an einem festen Platz und ihr Verbleiben dort eine Voreingenommenheit des Forschers repräsentiert.“

Allerdings ist hier geltend zu machen, dass wir zwar – analog zu Interventionen im Bereich sprachlicher Forschungsmedien – davon ausgehen müssen, dass jede Intervention einen Einfluss ausübt und dass Interventionen auch nicht zu vermeiden sind, dass aber eine Kontrolle des Einflusses dieser Interventionen eher möglich ist, wenn sie in Bezug auf alle Fälle *gleichartig* und somit *vergleichbar* sind. In ähnlicher Argumentation heißt es bei Collier u. Collier (1986: 163) in Bezug auf die Fotografie: „Diese Fotografien sollten in einer konsistenten Weise erstellt werden, um die Möglichkeit für eine komparative Analyse zu schaffen.“

Bisher ging es – wie überwiegend insgesamt in diesem Kapitel – um die Produktanalyse, um die Analyse des Produkts der abgebildeten und abbildenden Bildproduzent(inn)en, die uns Aufschluss über deren Habitus oder den von ihnen propagierten Lifestyle vermitteln soll.

Im Bereich der Filmanalyse als Medienforschung gilt es dann die *Produktanalyse* von der *Rezeptions- oder Mediennutzungsforschung* noch einmal zu unterscheiden. Diese stellt ein relativ umfangreiches Forschungsfeld der sozialwissenschaftlichen Analyse dar und hat mit der (in den Sozialwissenschaften weniger verbreiteten) Interpretation visueller Grunddaten nur indirekt etwas zu tun. Gegenstand der Rezeptions- oder Mediennutzungsforschung sind nicht – oder nur sekundär – die Fotos und Filme selbst, sondern eben deren Rezeption, Wahrnehmung, Nutzung und Aneignung.

## 5.2 Zum Verhältnis von Produkt- und Rezeptionsanalyse

In der Forschungspraxis finden sich viele Berührungspunkte zwischen Produkt- und Rezeptionsanalysen von Filmen. Methodologisch bleibt das Verhältnis – wie auch im Folgenden gezeigt werden wird – vielfach ungeklärt. Somit kommen auch Abhandlungen zur Produktanalyse, welche das eigentliche Thema dieses Kapitels darstellt, nicht umhin, das Verhältnis der beiden Bereiche der Filmanalyse im Ansatz zu klären. Dabei wird es auch darum gehen, die Unabhängigkeit von Rezeptions- und Produktanalyse in methodologischer Hinsicht zu begründen. Vor diesem Hintergrund kann dann auch erläutert werden, unter welchen Bedingungen der Zusammenhang von Produkt- und Rezeptionsanalyse herausgearbeitet werden kann.

Was die qualitative Rezeptionsforschung anbetrifft, so ist allen wesentlichen Ansätzen in diesem Bereich die Prämisse gemeinsam, dass die Rezipientinnen und Rezipienten als aktive, in soziale und interaktive Beziehungen eingebundene, Akteure betrachtet werden (vgl. u.a. Aufenanger 1995 u. Winter 1995) und in diesem Sinne als „konstruktiv-realitätsverarbeitend“ (Neumann-Braun/Schneider 1993: 197). Dies gilt auch für die Rezeptionsanalyse im Rahmen der dokumentarischen Methode, in deren Zentrum der Habitus der Rezipierenden steht. Allerdings ist hier, wie Michel (2006a: 122) formuliert, folgendes geltend zu machen: „Als Produkte des Habitus sind Rezeptionsprozesse zwar ebenfalls als *aktiv* zu konzeptualisieren, nicht jedoch als autonom, zielgerichtet, absichtsvoll, individuell und bewusst.“

In den Anfangsphasen der qualitativen Rezeptionsforschung in der Bundesrepublik hat die „strukturanalytische Rezeptionsforschung“ (Neumann/Charlton 1989) größeren Einfluss gewonnen. Dort wurde die Rezeptionsforschung grundsätzlich mit einer Produktanalyse verbunden. Deren Resultat erscheint dort als „objektive Sinnstruktur des Medienprodukts“ (Neumann-Schneider 1993: 197) und wird den ‚subjektiven‘ Interpretationen der Rezipient(inn)en gegenübergestellt. Letztere werden somit in selektiver Weise vor dem Vergleichshorizont einer Sinnstruktur rekonstruiert, wie sie von den wissenschaftlichen Interpret(inn)en entworfen worden ist – mit den darin implizierten Normalitätsmustern. Diese Normalitätsvorstellungen betreffen insbesondere die Kompetenzen der Rezipierenden. Die Vorgehensweise der strukturanalytischen Rezeptionsforschung ist an diejenige der objektiven Hermeneutik angelehnt, welche von mir an anderer Stelle (vgl. Bohnsack 2003b) umfassend diskutiert und kritisiert worden ist.

### 5.2.1 Exkurs: Produkt-, Rezeptions- und Diskursanalyse in den Cultural Studies

Auch die Vertreter der Cultural Studies, die wohl als derzeit populärste Richtung in der Film- und Filmrezeptionsanalyse gelten können und die stark durch die Semiotik beeinflusst sind, wenden sich gegen derartige objektivistische Zugänge. So ordnet Rainer Winter (1997b: 81) den Cultural Studies eine „anti-objektivistische Sicht des Wissens“ zu. An anderer Stelle formuliert Winter (1997a: 54), indem er – wie die Vertreter der Cultural Studies ganz allgemein – Bilder wie Filme unter dem Oberbegriff des Textes subsumiert: „Die Rezeption und Aneignung von Texten

wird zu einer kontextuell verankerten gesellschaftlichen Praxis, in der Texte als Objekte nicht vorgegeben sind, sondern erst auf der Basis sozialer Erfahrung produziert werden.“ Es ist dies, was von Ang (1997) als „radikaler Kontextualismus“ bezeichnet wird. Und bei Hepp (2004: 133) heißt es: „Am Ende dieser stark semiotisch orientierten Phase steht eine umfassende Kontextualisierung des Medientext- bzw. Produktbegriffs.“

Die Produktanalysen im Bereich der Cultural Studies sehen sich somit schon früh vor ein methodologisches Probleme gestellt, welches von ihnen (zuerst: Hall 1980: 134; siehe auch: Winter 2003b) mit dem – vor allem von Roland Barthes (1990: 34) ausgearbeiteten – Begriff der *Polysemie*, also der Vieldeutigkeit des Textes, Filmes oder Bildes bezeichnet wird (dazu auch: Kap. 3.8).

In diesem Zusammenhang wird dann auch deutlich, dass das Problem der Polysemie – als ein Problem mit der unbewältigten Mehrdeutigkeit auch der wissenschaftlichen Analyse selbst – es ist<sup>55</sup>, welches dazu führt, dass die Vertreter und Vertreterinnen der Cultural Studies sich bei der Analyse des Produkts *entweder* zunehmend auf die Interpretationen der Rezipienten und Rezipientinnen stützen *oder* die Produktanalyse in eine *theoretische* resp. auch historische Analyse überführen, in der eine übergreifende gesellschaftliche Betrachtung von Prozessen der medialen Produktion von Wissen und deren institutioneller Einbindung angestrebt wird.

Der erstere Weg stützt sich vor allem auf die „kritische Ethnografie“, die ursprünglich im Bereich der Jugendstudien der Cultural Studies entwickelt worden ist (Willis 1979 u. 1981, Clarke et al. 1979)<sup>56</sup> und dann als methodologischer Rahmen der Rezeptions- resp. Aneignungsforschung übernommen wurde (siehe u.a. Morley 1999). Der zweite Weg ist derjenige der „Diskursanalyse“. Diese Differenzierung im Bereich der Entwicklung der Cultural Studies findet sich bei Andreas Hepp (2004: 256; siehe auch Geimer/Ehrenspeck 2009), bei dem es auch heißt (a.a.O.: 153): „Die zentrale Stellung, die in früheren Publikationen Fiskes und anderer Vertreter der Cultural Studies die Kategorien der Polysemie und der Offenheit haben, wird nun von der des Diskurses eingenommen.“ Ich möchte mich zunächst kurz der Diskursanalyse zuwenden, um dann auf die „kritische Ethnografie“ einzugehen und schließlich auf das Problem der Polysemie der Produktanalyse und auf die Gründe für dessen unzureichende Bewältigung in den Cultural Studies.

Die Diskursanalyse der Cultural Studies hat sich zunehmend an Foucault orientiert, ist aber durchaus in gewisser Weise auch als Weiterentwicklung des sogenannten Ecoding/Decoding-Modells von Stuart Hall zu verstehen, welches „als der zentrale Ausgangspunkt der Medienstudien der Cultural Studies gelten“ (Hepp 2004: 117) kann und demzufolge „Medienkommunikation stets als ein Prozess gedacht wird, in

55 Wobei die Probleme der Vertreter der Cultural Studies mit der Polysemie nicht oder weniger dahin gehen, dass der Unterschied zwischen den Rezeptionsweisen der Zuschauer/innen und den ‚Rezeptionen‘, d.h. den Interpretationen, der Forscher/innen nicht bestimmt und somit die Differenz von Common Sense und wissenschaftlicher Interpretation nicht begründet werden kann, wie dies von Seiten der dokumentarischen Methode geltend gemacht werden muss, sondern der damit verbundene Verlust des kritischen Potentials der Cultural Studies, indem „durch das Feiern einer Offenheit medialer Texte der Raum für jede Form von Kritik genommen würde“, wie Hepp (2004: 140) zusammenfasst.

56 Für eine Auseinandersetzung mit diesen Studien vor dem Hintergrund der dokumentarischen Methode siehe: Bohnsack 2000, 2004a u. b sowie 2008a: 153.

dem der Medientext auf nicht hintergehbare Weise zwischen „encoding“ (Produktion) und „decoding“ (Rezeption) lokalisiert ist“ (a.a.O.: 111f.).

Dieses Encoding/Decoding-Modell und dessen Fortentwicklung im Sinne einer Diskursanalyse sucht also den gesamten Zusammenhang von Produkt/Produktion/Rezeption in den Blick zu nehmen und darüber hinaus das Produkt in Relation zu anderen Produkten im kulturellen und gesellschaftlichen Kontext zu verorten. In diesem Sinne sind, wie Winter (2004: 207) ausführt, „z.B. eine semiotische Analyse eines Hollywoodfilms oder die ethnographische Erforschung kultureller Welten ohne Bezug zum Verhältnis von Kultur und Macht keine Cultural Studies.“ Dieses Kernanliegen der Cultural Studies, die Produkt- und Rezeptionsanalysen in einen übergreifenden Zusammenhang zu verorten, gehört zu deren wichtigen Beiträgen und Verdiensten und ist auch für die dokumentarische Methode von zentraler Bedeutung.<sup>57</sup>

Allerdings stellt sich gerade deshalb umso mehr die Frage nach der *empirischen Basis* einer derartigen übergreifenden Analyse. Und hier ist geltend zu machen, dass eine empirisch fundierte Diskursanalyse ohne eine empirische Produktanalyse (hier also bspw. die Analyse des Hollywood-Films selbst) nicht denkbar ist, durch diese gleichsam hindurch muss, indem sie sich als Voraussetzung für die Diskursanalyse des Produkts selbst empirisch zu vergewissern hat. In diesem Sinne können sich auch die Vertreter der Cultural Studies dem Problem der Produktanalyse im engeren Sinne letztlich nicht entziehen.

Dies zeigt sich bereits daran, dass auch von Forschern, die sich in der Tradition der Cultural Studies verorten, immer wieder Ansätze zu einer empirischen Produktanalyse im engeren Sinne vorgelegt werden. So ist die Produktanalyse von Mikos (2003b: 146), die er auch als „struktur-funktionale Analyse“ bezeichnet, forschungspraktisch zunächst weitgehend auf die „Analyse ästhetischer Strukturen“ (a.a.O.: 140) reduziert, genauer eigentlich auf die Rekonstruktion der formal-technischen „Konventionen der Darstellung“ wie vor allem die Kameraarbeit (mit der Wahl von Perspektivität und Einstellungsgröße) und die Selektivität von Montage und Schnitt (vgl. dazu Kap. 5.5.3 u. 5.6).

Die Analyse von Mikos hat gegenüber vielen sozialwissenschaftlichen Ansätzen, wie etwa der Konversationsanalyse (und auch gegenüber der *dominanten* Vorgehensweise in den Cultural Studies), zwar den Vorzug, dass der Bildhaftigkeit, der Eigenlogik des Bildlichen, Rechnung getragen wird<sup>58</sup>, jedoch verbleibt diese struktur-funktionale Analyse im Bereich der formalen Strukturen, sozusagen der *Syntax* des Filmes. Mit der eigentlichen *semantischen* Analyse wird hier nicht angeschlossen, vielmehr wird der Bereich der Produktanalyse verlassen und auf die Rezeptionsanalyse verwiesen: „Ziel der struktur-funktionalen Analyse ist es daher nicht, den Film an sich zu analysieren, sondern ihn als Teil eines Kommunikationsprozesses zu sehen. Ihr geht es um die Bedingungen und Strukturen (gemeint sind die formal-ästhetischen; R.B.), die die Geschichte im Kopf der Zuschauerinnen und

Zuschauer entstehen lassen. Sie setzt damit am Filmverstehen und Filmleben der Rezipienten an, wenn sie Filme analysiert“ (Mikos 2003b: 147).

Damit sind wir bei dem anderen Versuch der Bewältigung des Problems der Polysemie, indem hier an die Stelle der Produktanalyse bzw. an die Stelle entscheidender Phasen der Produktanalyse die Analyse der Rezeption, die „kritische Ethnografie“ tritt, die sich auf die Prozesse des Verstehens oder Interpretierens seitens der Zuschauer/innen richtet.<sup>59</sup> Dabei ist die Ethnografie der Cultural Studies mit einer konsequenten Hinwendung zur Alltagskultur und deren Rehabilitierung im Sinne einer „Popular Culture“ oder „Profane Culture“ (Willis 1981) verbunden, in der sich große Übereinstimmungen zeigen mit der dokumentarischen Methode, der praxeologischen Wissenssoziologie sowie auch mit der Ethnomethodologie und deren formaler Hinwendung zu den Strukturen, zur Eigenlogik, zum kunstvollen Charakter alltäglicher Praxis, zu den „ordinary, artful ways“ of „daily life“, wie Harold Garfinkel (1967: vii) dies genannt hat.

Wenn wir auf das bereits angeführte Zitat von Winter (1997a: 54) zurückkommen, so wird Rezeption dort folgendermaßen definiert: „Die Rezeption und Aneignung von Texten wird zu einer kontextuell verankerten gesellschaftlichen Praxis, in der Texte als Objekte nicht vorgegeben sind, sondern erst auf der Basis sozialer Erfahrung produziert werden.“ Wie in dem Zitat bereits deutlich wird, zeichnen sich nun innerhalb der Rezeptionsanalysen der Cultural Studies Unklarheiten dahingehend ab, inwieweit die „Texte“ jeweils kontextgebunden im eigentlichen Sinne *rezipiert* oder in der Situation der Rezeption nicht vielmehr überhaupt erst *produziert* werden. Dabei geht die Vorstellung, dass die Texte in der Situation der Rezeption kontextuell erst „produziert“ bzw. „konstituiert“<sup>60</sup> werden, über einen „Kontextualismus“ hinaus bzw. hinter diesen zurück und wird – ähnlich wie in der Ethnomethodologie – zu einem „methodologischen Situationismus“ (Ang 1997: 90) hingeführt bzw. mit diesem in eins gesetzt (vgl. zur Kritik am Situationismus der Ethnomethodologie auch: Bohnsack/Nohl 2001b).

Wenn ich aber diese Kontexte als „kulturelle und gesellschaftliche“ verstehe, wie dies von den Vertretern der Cultural Studies selbst ja ganz deutlich beansprucht wird (vgl. u.a. Mikos 2003a: 249f.), so haben wir es nicht mit situationistischen Strukturen zu tun, sondern mit sozialen Formationen von einiger Dauerhaftigkeit, denen dann auch dauerhafte soziale Erfahrungen und Bedeutungszusammenhänge auf Seiten der Mitglieder oder Angehörigen dieser kulturellen oder gesellschaftlichen Kontexte entsprechen. Bedeutungen werden somit nur oberflächlich betrachtet je situativ, d.h. in der Situation der Rezeption, konstituiert oder produziert. Vielmehr lässt sich zeigen, dass im Zuge der Bedeutungszuschreibung dauerhafte, milieuspezifisch verankerte Erfahrungskontexte *aktualisiert* werden.

57 Zum Zusammenhang von Diskursanalyse und dokumentarischer Methode siehe auch v. Rosenberg 2008.

58 Wie Mikos (2003b: 146) betont, versucht die „struktur-funktionale Filmanalyse“, „vor allem die visuelle, bildliche Seite zu analysieren“.

59 Ein anderer Weg, wie er von Rainer Winter (2003a) beschritten wird, weist letztlich in dieselbe Richtung, also in diejenige der Rezeptionsanalyse bzw. bewegt er sich zwischen Rezeptions- und Diskursanalyse, indem dort an die Stelle einer Produktanalyse im eigentlichen Sinne die Analyse der *Diskurse über* den Film tritt, wie sie uns bspw. als öffentliche Diskurse in Form von Filmrezensionen in deren unterschiedlichen Lesarten in den Medien begegnen.

60 An anderer Stelle (1993: 156) heißt es bei Winter: „Die Bedeutung, die ein Text gewinnt, lässt sich durch die Analyse des Textes allein nicht bestimmen, sondern nur durch die Berücksichtigung der sozialen Kontexte, in denen er rezipiert und interpretiert wird. So wird ein Film erst im Moment der Interaktion mit den Zuschauer konstituiert.“

Die Schwierigkeiten des Zugangs zu dauerhafteren sozialen Strukturen und Erfahrungskontexten führen dazu, deren Existenz überhaupt in Frage zu stellen, wie sich dies etwa bei Fiske (1989: 23) zeigt, wenn er resümiert: „Die Leute bilden eine sich verändernde Menge von Bindungen (allegiances), die quer zu allen sozialen Kategorien liegen; verschiedene Individuen sind zu unterschiedlichen Zeiten unterschiedlichen Formationen zugehörig und wechseln zwischen diesen häufig in einer ständigen Bewegung.“ Der Zugang zu jenen Mustern der Sinn- und Erfahrungskonstitution, die den oberflächlichen, den ‚wörtlichen‘ Sinngehalt von Äußerungen zu transzendieren und zu tiefer liegenden und dauerhafteren habituellen und kollektiven Orientierungsstrukturen vorzudringen vermögen, ist allerdings vom Einsatz komplexer Verfahren der Textinterpretation abhängig, wie sie in den Cultural Studies bisher kaum entwickelt bzw. genutzt worden sind.

Eine derartige Kritik am fehlenden empirischen Zugang zu dauerhaften Strukturen ist im Bereich der Gesprächsanalyse und des Gruppendiskussionsverfahrens (vgl. Bohnsack 2008b) in Auseinandersetzung mit dem interpretativen Paradigma und auch mit Bezug auf Vertreter/innen der Cultural Studies, wie etwa David Morley (Bohnsack 2004a u. b, 2007g), umfassender ausgearbeitet worden, sodass ich hierauf an dieser Stelle nicht genauer einzugehen brauche (siehe dazu auch Geimer 2008).

Es ist somit also zwischen zwei Arten von Kontexten zu unterscheiden: den *interaktiven* Kontexten, in der die Rezeption je situativ stattfindet, und den Kontexten im Sinne *sozialer Lagerung*. Innerhalb derart – wie auch immer gearteter und wie auch immer heterogener – dauerhafter (kultureller und gesellschaftlicher) Erfahrungskontexte im Sinne sozialer Lagerung, die wir im Rahmen der dokumentarischen Methode als *Erfahrungsräume* oder Milieus bezeichnen, findet sowohl die Herstellung der Texte, Bilder und Filme wie auch deren Rezeption statt. Inwieweit es im Zuge dieser Rezeption zu einem „Verstehen“ kommt (siehe dazu weiter unten: 5.2.3), ist von Kongruenzen dieser Erfahrungsräume oder Erfahrungskontexte auf Seiten der Rezipierenden mit denjenigen auf Seiten der Herstellenden abhängig. Zuzustimmen ist den Vertretern der Cultural Studies wiederum dahingehend, dass wir aufgrund der Pluralität derartiger Erfahrungskontexte immer eine „multikontextuell bestimmte Situation“ in Rechnung zu stellen haben, wie Mikos (2003a: 250) mit Bezug auf Ang (1997: 89) formuliert.

### 5.2.2 Die Polysemie in der Produkt- und Rezeptionsanalyse

Wenn wir nun zu der Frage nach dem Verhältnis von Produkt- und Rezeptionsanalyse zurückkehren, so habe ich deutlich zu machen versucht, dass das Problem der Polysemie – als ein Problem mit der unbewältigten Mehrdeutigkeit auch der wissenschaftlichen Analyse – es ist, welches die Konsequenz hat, dass im Kontext der Cultural Studies die Produktanalyse entweder sehr schnell in eine *theoretische* resp. auch historische Analyse im Sinne der Diskursanalyse überführt wird, mit der eine übergreifende gesellschaftliche Betrachtung von Prozessen der medialen Produktion von Wissen und deren institutioneller Einbindung angestrebt wird, oder die Analyse des Produkts sich auf die Interpretationen der Rezipienten und Rezipien-

tinnen stützt und die Produktanalyse sich somit hin zu einer Rezeptionsanalyse verlagert.

Eine derartige Substitution der Produkt- durch die Rezeptionsanalyse bzw. eine Verlagerung von der Produkt- zur Rezeptionsanalyse erscheint vor allem aus zwei Gründen problematisch: *Zum einen* erfahren wir dann etwas über die Rezipient(inn)en und nicht über die Produzent(inn)en des Films. Im Sinne der dokumentarischen Methode ist uns der Habitus der Produzent(inn)en am Produkt selbst – und letztlich auch nur dort – unmittelbar und in valider Weise zugänglich. Das bedeutet aber auch, dass – im Sinne des Encoding/Decoding-Modells – wir uns die Sinnstruktur, den *modus operandi* des Encoding auf der Grundlage einer empirischen Analyse empirisch zu erarbeiten vermögen, deren Gegenstand das Produkt selbst ist. Hier zeigen sich dann auch durchaus Übereinstimmungen des Encoding/Decoding-Modells mit der dokumentarischen Methode (siehe: Michel 2006: 182ff.).

*Zum anderen* erscheint die Verlagerung von der Produkt- zur Rezeptionsanalyse deshalb problematisch, weil damit das Problem der Polysemie und der Objektivität oder Gültigkeit der Interpretation oder Lesart der Forschenden nicht bewältigt, sondern lediglich verlagert wird: Denn dieses Problem stellt sich dort erneut, wo wir als Forschende dann wiederum die Interpretationen oder Darstellungen der Rezipient(inn)en, also deren verbale Äußerungen, d.h. deren Texte (in ihrer Polysemie) zu interpretieren haben.

Der Unterschied zwischen den Rezeptionsweisen der Zuschauer/innen und den ‚Rezeptionen‘, d.h. den Interpretationen, der Forscher/innen vermögen die Cultural Studies nicht genauer zu bestimmen und somit auch die Differenz von Common Sense und wissenschaftlicher Interpretation nicht zu begründen. Dies wird dann auch von Winter (1997b: 83) explizit gemacht, wenn er formuliert: „Der Forscher nimmt nicht die Rolle des unabhängigen Beobachters ein. Er ist eher unterstützender Mitspieler.“ Diese Diffusion von Beobachter- und Teilnehmer-Haltung führt dann auch zu einer Diffusion von wissenschaftlichem und moralischen Diskurs: „Die ‚neue‘ Ethnografie im Rahmen der Cultural Studies begreift sich nicht nur als wissenschaftlicher, sondern auch als moralischer Diskurs“ (a.a.O.: 90).

Erst die genauere Bestimmung der Relation der Alltagsinterpretationen zu den Interpretationsleistungen der Forscher/innen vermag den Anspruch einer Wissenschaftlichkeit zu begründen und damit schließlich auch den Aufwand an Ressourcen zu legitimieren, der darin investiert wird. Die Wissenschaftlichkeit von Interpretationen lässt sich letztlich nur von ihrem methodischen Potential her begründen, über die Interpretationen des Alltags, des Common Sense (in welcher Weise auch immer) hinaus zu weisen, diese systematisch und methodisch kontrolliert zu *transzendieren*.

Damit ist – im Sinne der dokumentarischen Methode wie aber u.a. auch der Luhmannschen Systemtheorie – nicht der Anspruch auf eine ‚höhere Rationalität‘ wissenschaftlicher Interpretationen gemeint, sondern der aus einem *Wechsel der AnalyseEinstellung* resultierende Gewinn an Erkenntnispotential. Der Wechsel der AnalyseEinstellung lässt sich (wie in 3.2. und 4.3.1 genauer dargelegt wird) formulieren als derjenige von der Frage nach dem, *was* gesellschaftliche Tatsachen, Äußerungen und Handlungen *sind*, zur Frage danach, *wie* diese *hergestellt* werden, also zur Frage nach dem *modus operandi*, der performativen Struktur, dem genetischen Prinzip, dem *Habitus*.

Dieser Wechsel der AnalyseEinstellung ist, wie bereits angesprochen (Kap. 3.2), der Wechsel von der Ikonografie zur Ikonologie. Wie in Kapitel 3.8 genauer dargelegt, ist auf der *ikonografischen* Ebene (sofern ich mich auf die konjunktiven, die milieu- und individuenspezifischen Wissensbestände einlasse) eine nahezu unendliche Polysemie gegeben, da ich das Bild (oder auch den Text) in immer neue Kontexte zu stellen vermag. Demgegenüber kann ich auf der *ikonologischen* Ebene, auf der ich nach dem *Wie*, nach dem modus operandi der Herstellung der Darstellung, frage, auf ein Kontextwissen vollständig verzichten. (Allerdings ist – siehe dazu auch: Kap. 3.8 und die Ausführungen weiter unten – bei der Beobachtung des modus operandi die Dimension, welche mir dabei in den Blick gerät, abhängig von meinen *Vergleichshorizonten*, die ich jedoch methodisch zu kontrollieren vermag).

Die im Kontext der Cultural Studies durchgeführten Studien verbleiben in systematischer Weise auf der ikonografischen Ebene. Dies wird auch explizit dort zum Ausdruck gebracht, wo betont wird, dass diese Studien – indem sie sich methodologisch und methodisch an der Semiotik vor allem von Roland Barthes und Umberto Eco orientieren – ihre Analyse auf der *konnotativen* Ebene ansiedeln (zusammenfassend: Hepp 2004: Kap. 2.1 u. 4.1). Die konnotative Ebene lässt sich aber – und dies wird bei Eco (1994: 243) explizit hervorgehoben – analog zur ikonografischen Ebene bei Panofsky verstehen, also der Ebene von Common Sense-Interpretationen, auf der wir mit einer nahezu unendlichen Polysemie konfrontiert sind. Roland Barthes (1990: 34) hat, wie bereits erwähnt, den Begriff der Polysemie wohl am umfassendsten ausgearbeitet und korrespondierend spricht der andere Klassiker der Semiotik, Umberto Eco (1977), von dem „offenen Kunstwerk“ (siehe dazu auch: Kap. 3.8 sowie Michel 2006: 76ff. und zur Kritik an Eco auch: Kap. 5.4.2).

Die Probleme im Umgang mit der Polysemie, wie wir sie im Bereich der Cultural Studies beobachten können, stehen also im Zusammenhang mit jenem Weg der semiotischen Analyse, die systematisch auf der konnotativen Ebene verbleibt und somit die Differenz zum Common Sense nicht zu bestimmen vermag. Wenn führende Vertreter der Cultural Studies Abstand nehmen von der Differenz zwischen Beobachter- und Teilnehmer- (resp. „Mitspieler“-) Haltung und zwischen wissenschaftlichem und moralischem Diskurs, so stellt sich die Frage, ob dies nicht vielleicht – zumindest zu einem Teil – auch darin begründet ist, dass hier aus einer (methodologischen) ‚Not‘ eine Tugend gemacht wird.<sup>61</sup>

Den Wechsel der AnalyseEinstellung von der Ikonografie, von der Ebene der Konnotationen, hin zur Ikonologie, wie er für die dokumentarische Methode konstitutiv ist, hat gleichermaßen für die Produkt- wie die Rezeptionsanalyse Gültigkeit. Im Rahmen der dokumentarischen Methode weisen beide Bereiche methodologisch insofern Gemeinsamkeiten auf, als sie einen derartigen Wechsel der AnalyseEinstellung vornehmen und somit die Frage nach dem *Habitus* stellen.

61 Ähnliche Probleme zeigen sich auch im Bereich der Evaluationsforschung bei jenen Forschern und Forscherinnen, die sich im Rahmen des interpretativen Paradigmas verorten (siehe Bohnsack 2006c). Eine derartige Verbindung von interpretativem Paradigma und Cultural Studies findet ganz zentral in der Person von Norman K. Denzin ihren Ausdruck, der als Mitbegründer des interpretativen Paradigmas sich zugleich im Rahmen der Cultural Studies verortet (siehe u.a.: Denzin 1991)

Diesen Gemeinsamkeiten auf der methodologischen und methodischen Ebene steht aber die klare Differenz des Untersuchungsgegenstandes gegenüber: im Falle der Produktanalyse zielt das Erkenntnisinteresse auf den Habitus und den Modus der Erfahrungskonstitution bzw. den Erfahrungsraum der *Produzent(inn)en* und im Falle der Rezeptionsanalyse auf den Habitus, Modus der Erfahrungskonstitution und Erfahrungsraum der *Rezipierenden* sowie auf mögliche Veränderungen innerhalb desselben (für die Rezeptionsanalyse auf der Grundlage der dokumentarischen Methode siehe: Michel 2003, 2004, 2005, 2006a, 2006b u. 2007; Selke 2005b u. 2007; Schäffer 2009). Damit ist bereits ausgeschlossen, dass, wie wir dies etwa im Falle der oben erwähnten „strukturanalytischen Rezeptionsforschung“ beobachten, die von den wissenschaftlichen Interpret(inn)en entworfene Sinnstruktur des Produkts als Normalitätsfolie herangezogen wird, um vor diesem Vergleichshorizont dann die Sinnstruktur der Rezipient(inn)en und insbesondere deren Kompetenzen (in selektiver Weise) zu beleuchten und zu beurteilen.

Die Produktanalyse vermag insofern einen ergänzenden Beitrag zu leisten zur Rezeptionsanalyse, als sie eine – über die Rezeptionsanalyse hinausgehende, diese unterstützende und präzisierende – Klärung der Beziehung, der Relation zwischen dem Erfahrungsraum bzw. den Erfahrungsräumen der Bildproduzent(inn)en einerseits und der Rezipient(inn)en andererseits ermöglicht. Jürgen Wittpoth (2003: 83) hat dazu formuliert: „Wenn es bei der ikonologischen *Interpretation* darum geht, habitusspezifische Aspekte zu identifizieren, die im Bild zum Ausdruck kommen, ginge es bei der *Rezeptionsforschung* um die Frage, ob und wie Rezipienten nach Maßgabe ihres Habitus auf solche habitusspezifischen Aussagen reagieren.“ So etwa, wenn wir der Frage nachgehen, ob der in einem Werbespot propagierte Lifestyle mit dem Habitus potentieller Rezipierender in Übereinstimmung zu bringen ist oder Anschlussmöglichkeiten an diesen eröffnet. Oder wenn wir beispielsweise die Frage stellen, ob der Habitus der Protagonist(inn)en in einem Kinofilm zum Habitus bestimmter Milieus, Generationen oder Altersphasen (etwa Jugendlichen in der Adoleszenzkrise) Homologien, Übereinstimmungen oder Anschlussmöglichkeiten aufweist.

Je nach spezifischem Erfahrungshintergrund und Sozialisationsgeschichte der Rezipient(inn)en wird diesen ein verstehender Zugang zu je unterschiedlichen – einander aber nicht ausschließenden – Erfahrungsräumen des Bildes oder Filmes möglich.<sup>62</sup> Beispielsweise eröffnet sich zu einem Bild oder Film einer Veranstaltung der Jugendweihe in der DDR ein jeweils anderer Zugang je nachdem, ob wir es mit Rezipient(inn)en mit einer Sozialisationsgeschichte in der DDR oder in der BRD zu tun haben. Im Bild dokumentieren sich spezifische Orientierungsgehalte, zu denen nur diejenigen einen unmittelbar *verstehenden* Zugang haben, die Gemeinsamkeiten eines spezifischen Erfahrungsraums teilen.

In unserem Beispiel ist dies der Erfahrungsraum der DDR-Kultur (kulturspezifischer Erfahrungsraum), der allerdings wiederum – je nach Generations- und Milieuzugehörigkeit der Rezipierenden – durch einen generations- und milieuspezifischen

62 Im Zusammenhang des (alltags-)praktischen Umgangs mit den Medien – etwa des Computers – spricht Burkhard Schäffer (2003) von diesen Erfahrungsräumen auch als „Medienpraxiskulturen“.

schen Erfahrungsraum überlagert wird.<sup>63</sup> Die entsprechend der Sozialisationsgeschichte und der sozialen Herkunft der Rezipierenden je unterschiedliche Überlagerung dieser Erfahrungsräume führt zu je differenten Rezeptionsweisen. Die Akteure im Alltag, im Common Sense, rezipieren, d.h. verstehen und interpretieren, vor dem Vergleichshorizont ihrer jeweils eigenen (milieu-, generations-, geschlechtsspezifischen) Erfahrungsräume. Dabei gibt es immer auch „Bereiche des Bildes, die verstanden werden können und andere, die interpretiert werden müssen“, wie Burkhard Schäffer (2009) formuliert, da in manchen Bereichen eine Übereinstimmung mit eigenen Erfahrungsräumen gegeben ist und sofern ein (unmittelbares) Verstehen möglich ist und in anderen nicht, sodass ein Interpretieren notwendig wird.<sup>64</sup>

Jene soziogenetisch fundierten Vergleichshorizonte, welche ich in der alltäglichen Interpretation intuitiv an das Bild herantrage, lassen sich allerdings empirisch-methodisch kontrollieren und variieren. Dies gelingt insoweit, als an die Stelle intuitiver Vergleichshorizonte solche treten, die empirisch fundiert sind, die uns also ebenfalls in Form von Bildern vorliegen, welche in die empirische Analyse einbezogen, also selbst zum Gegenstand empirischer Analyse werden. Wir sprechen in diesem Sinne (auch im Bereich der Textinterpretation) von *komparativer Analyse* (siehe u.a. Bohnsack 1989, 2006a: Kap. 11, 2009a; Nohl 2007).

Die Produktanalyse der dokumentarischen Methode unterscheidet sich von den Common Sense-Interpretationen also zum einen durch den Wechsel der Analyseeinstellung vom *Was* zum *Wie* und zum anderen auch durch die *systematische* Variation explizit eingeführter Vergleichshorizonte, also durch die komparative Analyse, welche einen methodisch kontrollierten Umgang mit der Polysemie und zugleich eine mehrdimensionale Typenbildung ermöglicht. Auf dem Wege der komparativen Analyse wird die Polysemie, die Mehrdeutigkeit, nicht eliminiert, sondern methodisch kontrolliert und somit der Blick auf die Mehrdimensionalität des Bildes eröffnet.

Je nach *Vergleichshorizont* geraten unterschiedliche – einander aber nicht ausschließende – Dimensionen oder Erfahrungsräume des Bildes in den Blick. In der dokumentarischen Methode wird dem (auch im Bereich des Textes) mit der *Mehrdimensionalität* der Sinn- und Typenbildung Rechnung getragen (vgl. Bohnsack 2007c). Die Polysemie des Bildes und Filmes erscheint somit nicht – wie bei Roland Barthes und auch in den Cultural Studies – primär als methodologisches und methodisches Hindernis, sondern als entscheidendes Potential für die methodische Weiterentwicklung. Zugleich ermöglicht die mehrdimensionale komparative Analyse auch die Möglichkeit der Reflexion auf die ‚Subjektivität‘ der Forscher/innen oder genauer: auf deren Standortgebundenheit. Eine derartige Reflexion oder Reflexivität wird auch von Seiten der Cultural Studies gefordert (siehe u.a. Morley 1999; Winter 1997b), aber forschungspraktisch-methodisch dort letztlich nicht überzeugend realisiert. Nach Einschätzung von Bettina Fritzsche (2007: 37) „gibt es jedoch innerhalb der Cultural Studies keine weitergehenden Überlegungen dazu, wie eine derartige Selbstreflexivität methodisch umsetzbar wäre.“

63 Zum generationsspezifischen Erfahrungsraum siehe auch Bohnsack/Schäffer 2002.

64 Für die Rezeptionsforschung auf der Basis von Interviews und Gruppendiskussionen bedeutet dies, dass in den derart erhobenen Texten ein „Oszillieren zwischen verstehenden Sinnaktualisierungsprozessen und interpretierenden Sinnbildungsprozessen“ zu erwarten ist (Schäffer 2009).

Was die Mehrdimensionalität der Typenbildung auf der Grundlage einer systematischen komparativen Analyse, einer systematischen Variation empirischer Vergleichshorizonte, anbelangt, so ist dies im Bereich der Textinterpretation forschungspraktisch umfassend erprobt und elaboriert worden (siehe dazu: Kap. 2.4 und die dort angeführte Literatur). Im Bereich der Bild- und Filminterpretation stehen wir hier noch ganz am Anfang.

### 5.2.3 Zur Differenzierung des Rezeptionsbegriffs: Interpretation, Verstehen und Aneignung

Nach der Differenzierung von Produkt- und Rezeptionsanalyse geht es weitergehend nun um eine Differenzierung innerhalb der Rezeptionsanalyse bzw. des Rezeptionsbegriffs, nämlich diejenige zwischen einem *interpretierenden*, d.h. einem ausschließlich kognitiven, Zugang einerseits und der *Aneignung* andererseits, welche einen *verstehenden* Zugang voraussetzt. Auf die Unterscheidung von Verstehen und Interpretieren bin ich in Kapitel 2.3 bereits eingegangen. Entscheidend für eine Aneignung im Sinne der dokumentarischen Methode ist es, ob die im Bild oder im Film dargestellte Praxis bzw. der ihr zugrunde liegende modus operandi, also der im Bild oder Film vermittelte Habitus (der Protagonisten und Protagonistinnen), anschlussfähig ist an denjenigen der Rezipient(inn)en, an eine Praxis, welche deren Sozialisationsgeschichte betrifft, ob die im Bild oder Film dargestellte Praxis also anschlussfähig ist an die Erfahrungsräume der Rezipient(inn)en.

Mit anderen Worten geht es darum, ob der sich in dem Bild oder Film dokumentierende modus operandi integrierbar ist in die handlungsleitenden Wissensbestände der Rezipierenden, also in jene ihre Handlungspraxis orientierenden Wissensbestände, welche diese in der selbst *ge-* und *er-*lebten Alltagspraxis erworben haben. Karl Mannheim (1964a: 534) spricht hier auch von einem „selbsterworbenen“ Wissen im Unterschied zu einem lediglich durch Dritte vermittelten Wissen: „Nur wirklich selbst erworbene Erinnerung, in aktuellen Situationen wirklich erworbenes ‚Wissen‘ besitze ich wahrhaft. Nur dieses Wissen ‚sitzt fest‘, aber auch nur dieses bindet wirklich.“

Die „bindende“ Eigenart dieses Wissens ergibt sich aus dessen *handlungsleitendem*, die *Praxis* orientierendem, Charakter. Dieses Erfahrungs-Wissen hat zugleich einen *ver-*bindenden Charakter im Sinne eines *konjunktiven* Wissens insofern, als mich etwas Besonderes – nämlich ein konjunktiver Erfahrungsraum – mit jenen verbindet, die dieses (in der selbst erlebten Alltagspraxis erworbene) Erfahrungs-Wissen mit mir teilen. Unter denjenigen ist ein unmittelbares *Verstehen* möglich – im Unterschied zu einem *Interpretieren* (vgl. Kap. 2.3).

Vom konjunktiven Erfahrungswissen zu unterscheiden ist ein Wissen, welches mir nur über die Darstellung seitens Dritter, also über deren kommunikativ vermittelte Darstellung einer Praxis, zugänglich ist. Es handelt sich um ein Wissen, welches wir mit Karl Mannheim (1980) als *kommunikatives* – im Unterschied zu einem *konjunktiven* – Wissen bezeichnet haben.

Mannheim (1964b) hat dies am Beispiel des Generationenzusammenhangs erläutert (siehe auch Bohnsack/Schäffer 2002). So verfügen bspw. diejenigen, die ge-



gen Ende des Zweiten Weltkriegs geboren worden sind und deren Kindheit somit in die Nachkriegsphase fällt, über ein *gemeinsames*, ein *konjunktives* Erfahrungswissen, welches daraus resultiert, dass sie die Wirren dieser zeitgeschichtlichen Phase gemeinsam *ge-* und *er-*lebt haben. Dadurch gehören sie zu einem konjunktiven Erfahrungsraum und bilden einen Generationszusammenhang, hier: denjenigen der „Kriegskinder“ (der dann – in milieuspezifischer Ausprägung – seinen Ausdruck in der Generationsgestalt der „68er“ gefunden hat; siehe Preuss-Lausitz et al. 1983). Insofern wie die medialen Angebote nicht nur kommunikativ-generalisiertes Wissen vermitteln, sondern auch den Transfer konjunktiver Wissensbestände leisten (und damit an die konjunktiven Wissensbestände und Erfahrungsräume der Rezipient(inn)en anzuschließen vermögen), spricht Burkhard Schäffer (1998: 36ff.) von den „transkonjunktiven Wissensbeständen“ der Medien.

Diese sind, wie gesagt, Voraussetzung für Prozesse der Aneignung, also für ein (unmittelbares) „*Verstehen*“ seitens der Rezipient(inn)en – im Unterschied zu einem „*Interpretieren*“, welches sich auf einer rein kognitiven Ebene bewegt. Verstehen bzw. Aneignung auf der einen und Interpretieren auf der anderen Seite stellen also die beiden grundlegenden Modi der Rezeption dar (für eine Differenzierung unterschiedlicher Modi des Interpretierens in einer empirischen Rezeptionsanalyse siehe Geimer 2008 u. 2009).

*Aneignung* basiert also auf der Chance von Kongruenzen zwischen den konjunktiven Erfahrungsräumen der Medienproduzent(in)en einerseits und der Rezipient(inn)en andererseits. Diese Kongruenzen bewegen sich auf der Ebene eines handlungsleitenden, also die Praxis orientierenden (Erfahrungs-) Wissens, welches weitgehend implizit oder atheoretisch bleibt. Eine *Aneignung* von Medieninhalten setzt ein *Verstehen* voraus, d.h. Aneignung ist dann möglich, wenn es den Rezipient(inn)en gelingt, mit ihrem eigenen konjunktiven Erfahrungswissen an das medial vermittelte Wissen anzuschließen – wenn also beispielsweise in einem Film Situationen einer Alltagspraxis inszeniert werden, in denen sich Handlungsprobleme und Handlungsorientierungen einer Praxis dokumentieren, die Homologien, also strukturelle Ähnlichkeiten, zu selbst erfahrenen Situationen der eigenen Praxis der Rezipient(inn)en aufweisen.

Derartige Situationen werden dann nicht lediglich kognitiv rezipiert und in dem Sinne auf einer theoretisierenden Ebene erörtert, sondern sie können auf dem Wege der in dieser Praxis implizierten Orientierungen und Wissensbestände in die eigene Handlungspraxis – genauer: in die Wissensbestände, welche diese Praxis orientieren – hinein genommen und in diesem Sinne angeeignet werden. Wenn etwa ein Jugendlicher mit einem Film zu Konflikten in der Familie konfrontiert wird, in dem er seine eigenen Konflikte in ihrer Struktur wiederzuerkennen vermag, dann ist es wahrscheinlich, dass die Reaktionen des jugendlichen Protagonisten im Film das Handeln des jungen Rezipienten beeinflussen können.

Allerdings werden hier einige Differenzierungen notwendig. *Zum einen* führt nicht jedes *Verstehen* auch zur *Aneignung*. Denn ich kann durchaus einen verstehenden Zugang zu den in einer (filmischen) Praxis implizierten Orientierungen und Wissensbeständen finden, ohne diese in die eigene Praxis zu integrieren. *Zum anderen* lassen sich mindestens zwei unterschiedliche Wege der Aneignung differenzieren (siehe dazu die empirische Analyse von Geimer 2008 u. 2009): Die im Film

transportierten impliziten (handlungsleitenden) Wissensbestände können dem eigenen Erfahrungsraum der Rezipient(inn)en entweder mehr oder weniger umstandslos integriert werden und somit eine *affirmative* Funktion gewinnen (reproduktive Aneignung), oder der Erfahrungsraum des Rezipienten wird durch sie verändert, so dass (im Sinne einer produktiven Aneignung) Entwicklungs- oder auch Bildungsprozesse in Gang gesetzt werden.<sup>65</sup>

Wesentlicher noch als die Differenzierung unterschiedlicher Modi der Aneignung aber ist die Unterscheidung von Interpretation und Aneignung (Verstehen). Dies deshalb, weil hierin grundlagentheoretische Voraussetzungen für eine Wirkungsforschung angelegt sind. Die Aneignung einer Praxis gelingt nicht, wie dargestellt, auf dem Weg der Interpretation und (theoretischen) Reflexion, sondern auf dem Wege des Verstehens einer durch Bild und Ton vermittelten Handlungspraxis und deren Interpretation in die eigene Praxis der Rezipient(inn)en, also u.a. auf dem Wege der „*Mimesis*“ (vgl. Bourdieu 1992: 105 sowie Wulf 1998 u. 2006). Burkard Michel (2006a: 394) spricht hier von einer „*praxeologischen Rezeptionsforschung*“, für welche die dokumentarische Methode (hier: bei der Auswertung von Gruppendiskussionen mit Rezipienten und Rezipientinnen) besonders geeignet erscheint: „Rezeption wird hier nicht als Akt ‚reiner‘ Erkenntnis, als philologisches ‚Entziffern‘ oder geistig-reflexives ‚Dechiffrieren‘ begriffen, sondern als prä-reflexive, atheoretische und vorsprachliche *Praxis*“ (siehe auch: Michel 2005 u. 2006b).

Dabei gilt es aber den Praxisbegriff genauer zu klären: Eine in den Medien, in Bild und Film, *vermittelte* Praxis, also eine Praxis, die *Gegenstand* medialer Darstellung ist (und als solche ggf. Homologien zu anderen Situationen der Alltagspraxis der Rezipient(inn)en aufzuweisen vermag) ist nicht zu verwechseln mit der Praxis des *Umgangs mit den Medien* in der Situation der Rezeption. Bei Angela Keppler (2006: 35) sehe ich Probleme einer derartigen Verwechslung: „Wer sich eine Sportübertragung ansieht, nimmt an einem Ereignis teil, das weit mehr als nur die *Wiedergabe* eines wirklichen Ereignisses ist: Die Dramaturgie der Übertragung im Fernsehen bewirkt vielmehr eine eigene Art der *Präsenz* dieses Ereignisses“. Hier wird die Präsenz in der Situation der medialen Übertragung einer Praxis, also die Präsenz vor dem Fernseher, gleichgesetzt mit der Präsenz in der Situation der Praxis selbst, hier also: der aktiven Teilnahme am Sportereignis. Der Begriff der Praxis wird hier nicht mit ausreichender Genauigkeit gefasst bzw. nicht hinreichend differenziert, obschon im Bereich der Rezeptionsforschung bereits sehr früh von Bachmair (1984) auf ein ähnliches Problem hingewiesen und gefordert wurde, dass die Rezeptionsforschung nicht allein die Situation der Rezeption selbst, sondern die

65 Es erscheint hier – nicht nur in erziehungswissenschaftlicher, sondern in grundlagentheoretischer Perspektive – sinnvoll, Anchlüsse an die Diskussion um den Bildungsbegriff in der Weise zu suchen, wie dieser insbesondere von Winfried Marotzki (1990) und Hans-Christoph Koller (1999) ausgearbeitet worden ist. Arnd-Michael Nohl (2006a) hat in diesem Zusammenhang – theoretisch und empirisch untermauert – argumentiert, dass Bildungsprozesse gerade auch durch vorreflexive Erfahrungen und Handlungspraktiken in Gang gesetzt werden können. In derartigen spontanen Handlungspraktiken tritt das Neue in den Erfahrungsraum und die Biographie von Menschen völlig unerwartet ein und führt zu Wandlungsprozessen, die als spontane Bildung verstanden werden können (siehe auch v. Rosenberg 2008). Nohl zeigt aber auch, dass nur bestimmte neue Erfahrungen – solche, für die man eine implizite Sensibilität besitzt – bildsam sein können.

Alltagssituationen der Rezipient(inn)en in den Blick zu nehmen habe (siehe dazu auch Fritzsche 2003 u. 2007).

Ein ähnliches Problem finden wir bei Filmwissenschaftlern, die in der Tradition der Cultural Studies stehen. So u.a. bei Lothar Mikos, der sich dem Problem der „Aneignung“ mit Bezug auf den Praxisbegriff bei Bourdieu und dessen Konzept des „praktischen Sinns“ zu nähern versucht<sup>66</sup>: „Der praktische Sinn eines Objekts ist durch den Zusammenhang, den es in einer Handlung hat, bestimmt“, heißt es bei Mikos (2003a: 31). Und soweit kann ich ihm folgen. Die direkt anschließende Exemplifizierung von Mikos, also das von ihm angeführte Beispiel, zeigt allerdings, dass sein Praxisbegriff nicht weit genug greift und als solcher auch demjenigen von Bourdieu nicht entspricht: „Auf diese Weise kann z.B. das Kino im Rahmen des praktischen Sinns dadurch bestimmt sein, dass man immer nur am Wochenende mit Freunden ins Multiplex geht.“

Auch hier wird die (performative) Praxis der Rezeptionssituation (Besuch des Multiplex-Kinos) nicht unterschieden von jener (propositionalen) Praxis, wie sie Gegenstand der filmischen Darstellung ist bzw. dieser als modus operandi zugrunde liegt und als solche möglicherweise Anschlüsse an die Praxis der Rezipient(inn)en bzw. an deren Habitus eröffnet.<sup>67</sup> Der Praxisbegriff greift hier also deshalb nicht weit genug, weil lediglich die Alltagspraxis der *Rezeptionssituation* in den Blick gerät, nicht aber die Lebens- und Alltagspraxis der Rezipientinnen im *umfassenden* Sinn bzw. genauer: jener modus operandi, welcher als Habitus die Alltagspraxis der Rezipient(inn)en übergreifend strukturiert<sup>68</sup> und als solcher ggf. Gegenstand filmischer Darstellungen ist, d.h. der Praxis der abgebildeten Bildproduzent(inn)en zugrunde liegt.

Ein ähnliches Problem findet sich bei David Morley, einem prominenten Vertreter der Cultural Studies, der im Zusammenhang der Rezeptionsforschung die „Ethnografie“ nicht nur fordert, sondern auch forschungspraktisch realisiert. Wenn dieser (1999: 313) „von Kommunikationspraktiken im Alltag“ als dem „zentralen Thema“ einer „Ethnographie des Fernsehpublikums“ spricht, wird bei genauerer Lektüre deutlich, dass auch hier lediglich die Praxis des Umgangs mit den Medien gemeint ist, der Begriff der (Alltags-) Praxis somit auf die Situation der Medienrezeption verkürzt wird. Voraussetzung dafür, den Bezug zur Lebens- und Alltagspraxis der Rezipientinnen in einem *umfassenden* Verständnis herzustellen, ist ein Zugang zu jenen Wissensbeständen der Rezipientinnen, welche deren Handlungs- und Alltagspraxis orientieren, also ein Zugang zu den handlungsleitenden Wissensbeständen, auf deren Grundlage eine (mimetische) *Aneignung* medialer Darstellungen dann möglich ist, wenn die medial dargestellte Alltagspraxis in ihrem modus

66 Mikos (2003b: 142) differenziert zwischen „Rezeption“ und „Aneignung“, wobei die Rezeption hier auf die schlechte (physische) Wahrnehmung reduziert wird.

67 In diesem Sinne lassen sich jene Medien, die lediglich eine Funktion im Kontext der performativen Praxis haben (u.a. der Computer, das Handy), von jenen anderen unterscheiden, die darüber hinaus eine dargestellte, propositionale Praxis vermitteln (u.a. Film, Fernsehen). Wenn Burkhard Schäffer (2003) von „Medienpraxiskulturen“ spricht, so ist lediglich ersterer Bereich gemeint (hier: die Praxis des Umgangs mit dem Computer).

68 In den klassischen Studien der Cultural Studies, die allerdings eher der Jugend- als der Medienforschung zuzurechnen sind (siehe v.a. Willis 1979 u. 1981 sowie Clarke 1979), findet sich ein derartiger umfassender Zugang zur Handlungspraxis.

operandi Bezüge und Kongruenzen aufweist zu den die Alltagspraxis orientierenden Wissensbeständen der Rezipient(inn)en.

Alexander Geimer (2008 u. 2009) hat in diesem Zusammenhang in einer empirischen Analyse von Filmnacherzählungen auf der Grundlage der dokumentarischen Methode zwischen diesen beiden – einander nicht ausschließenden – Modi des Umgangs mit Filmen unterschieden, „in denen Filme zum einen als Ressource für soziale Interaktion und zum anderen als Ressource zur Welterfahrung herangezogen werden.“ Entsprechend müssen dann auch zwei Arten des *Medien-Wissens* unterschieden werden. Ein durch die Medien angeeignetes Wissen als ein Wissen *über* etwas ist zu unterscheiden von einem in die Praxis des Umgangs mit den Medien investiertes und/oder in dieser Praxis des Umgangs erworbenes Wissen, ein Wissen *um* etwas (siehe auch Schäffer 2003 u. Bohnsack/Schäffer 2002).

### 5.3 Ansätze der Film- und Videointerpretation in der aktuellen qualitativen Forschung

Die Unklarheiten bei der Bestimmung der Relation von Produkt- und Rezeptionsanalyse bzw. deren fehlende Differenzierung, wie sie im vorherigen Kapitel kritisiert wurde, sind in vieler Hinsicht die Konsequenz einer fehlenden Präzisierung der Differenz zwischen den Prozessen des Verstehens, Interpretierens und Theoretisierens im Alltag, im Common Sense (auf denen die Rezeptionsprozesse basieren), einerseits und den sozialwissenschaftlichen Interpretationen und Theoriebildungen andererseits.

#### 5.3.1 Alltagsinterpretation und wissenschaftliche Interpretation

Die Ansätze der Bild- und Filminterpretation im Rahmen des interpretativen Paradigmas wie auch der Cultural Studies haben mit der dokumentarischen Methode den *rekonstruktiven* Charakter ihrer empirischen Forschung gemeinsam. Das bedeutet, dass sie die Produkte und Konstruktionen des Common Sense als „Konstruktionen ersten Grades“, wie Alfred Schütz (1971) dies genannt hat, auf dem Wege ihrer Re-Konstruktion einer sorgfältigen Analyse unterziehen und somit diese ‚profanen‘ Produkte der „Popular Culture“ oder „Profane Culture“ (Willis 1981) mit einem Respekt behandeln, wie dies bis dahin nur den Werken der Literatur und Kunst vorbehalten war.

So formuliert Jürgen Raab (2008: 166) als Vertreter des interpretativen Paradigmas, der sich mit der von ihm umfassend ausgearbeiteten „wissensoziologischen Bildhermeneutik“ innerhalb der phänomenologischen Soziologie in der Tradition von Alfred Schütz, Berger/Luckmann und Hans-Georg Soeffner verortet, mit Bezug auf seine eigene empirische Forschung: „Anders jedoch als die Kunstgeschichte, die sich vornehmlich für diejenigen Bilder interessiert, die sie als künstlerisch wertvoll einstuft und die ‚profanen‘ visuellen Alltags- und Subkulturen ausklammert, richtet sich die vorliegende Untersuchung gezielt auf solche lokalen und

vermeintlich marginalen skopischen Ordnungen innerhalb der visuellen Kultur der Moderne.“

In dieser Hinsicht bestehen weitgehende Übereinstimmungen mit der dokumentarischen Methode und der ihr zugrunde liegenden praxeologischen Wissenssoziologie. Allerdings differenzieren wir im Bereich der dokumentarischen Methode zwischen den *atheoretischen* als den habituellen und handlungsleitenden Wissensbeständen des jeweiligen kulturellen Kontextes und dessen Common Sense-*Theorien*, und der rekonstruktive Zugang zum handlungsleitenden Wissen ist zugleich mit einer Distanz den *Theorien* des Common Sense gegenüber verbunden. Die Differenz von Common Sense-Theorien und wissenschaftlichen Theorien und somit der „Bruch mit den Vorannahmen des *common sense*“ (Bourdieu 1996: 278) ist in der dokumentarischen Methode von zentraler Bedeutung für ihre Begründung als ein *wissenschaftliches* Verfahren (siehe. Kap. 2). Die Differenz zwischen Common Sense und Wissenschaft kann allerdings, wie bereits angesprochen, nicht (mehr) mit dem Anspruch einer ‚höheren‘ Rationalität der sozialwissenschaftlichen Analyse begründet werden, wohl aber mit dem Anspruch auf einen Zugang im Modus einer *anderen* Art der Rationalität, also im Sinne eines *Wechsels der Analyseeinstellung*.

Dieser Wechsel der Analyseeinstellung, wie er mit der dokumentarischen Methode vollzogen wird, entspricht in vieler Hinsicht dem Wechsel von der Beobachtung erster zur Beobachtung zweiter Ordnung im Sinne von Luhmann (1990: 86f.). Auch hier verfügt der Beobachter zweiter Ordnung nicht über eine ‚höhere‘, gleichwohl aber eine *andere* Rationalität resp. eine andere Analyseeinstellung als der Beobachter erster Ordnung. Der Wechsel der Analyseeinstellung lässt sich als derjenige von der Frage nach dem *Was* zur Frage nach dem *Wie* bezeichnen, wie ich dies im Anschluss an Heidegger (1986) und später auch Luhmann (1990), vor allem aber eben an Mannheim (1964a: 134; ursprünglich: 1921/22) bezeichnen möchte.

Was den Bereich der Bildinterpretation anbetrifft, so ist Erwin Panofsky, wie in Kapitel 3 bereits dargelegt, in seiner für die Kunstgeschichte Bahn brechenden Differenzierung zwischen der *ikonografischen* und der *ikonologischen* Interpretation durch die dokumentarische Methode seines Zeitgenossen Mannheim und durch eben jenen damit verbundenen Wechsel der Analyseeinstellung beeinflusst worden.

Mit dem Common Sense zu brechen, bedeutet im Sinne der dokumentarischen Methode, sich von den Common Sense-*Theorien* zu lösen. Dies setzt zum einen voraus, dass wir die Inhalte dieser Alltagstheorien, also die begrifflich-theoretischen Explikationen hinsichtlich ihres Geltungscharakters einklammern und sie in Relation zum Handeln der Akteurinnen und Akteure setzen, d.h. in Relation zu deren handlungsleitenden impliziten Wissensbeständen. Zum anderen bedeutet der Bruch mit dem Common Sense, dass die in den Common Sense-Theorien implizierten Methoden und Konstruktionsprinzipien zwar *Gegenstand* der sozialwissenschaftlichen Analyse sind, nicht aber deren *Methode* (siehe auch: Bohnsack 2006). In der Methodologie der Cultural Studies wie auch in deren Forschungspraxis fehlt eine begründete Differenz zum Common Sense und damit fehlt auch ein entsprechender Wechsel der Analyseeinstellung. Dies gilt analog für andere Bereiche der sog. interpretativen Sozialforschung, wie u.a. der sozialwissenschaftlichen Her-

meneutik und der hermeneutischen Wissenssoziologie, und für deren Film- und Videointerpretation – jenseits der sonstigen Unterschiede zwischen Cultural Studies und diesen anderen Bereichen interpretativer Sozialforschung.

Im Bereich des interpretativen Paradigmas wird das Verhältnis von Sozialwissenschaft und Common Sense u.a. bei Jürgen Raab und Dirk Tänzler (2002: 222f.) explizit formuliert.: „Wie bei der Analyse jeglicher (...) Texte bedient sich auch die Sinnrekonstruktion audiovisueller Produktionen systematisch und methodisch kontrolliert jener Mittel, über die auch jeder ‚gewöhnliche‘ Fernsehzuschauer für die Entschlüsselung medialer Inhalte verfügt.“ Die Autoren formulieren dies im Zusammenhang einer Analyse von Werbespots.

Indem im Rahmen des interpretativen Paradigmas wie auch der Cultural Studies der in den Common Sense-Theorien implizierte Zugang zu den audiovisuellen Produktionen nicht transzendiert wird, verbleibt die Interpretation des Bild- und Filmmaterials auf der ikonografischen Ebene bzw. vermag sie diese nicht systematisch zu transzendieren. Sie findet somit keinen systematischen Zugang zu den impliziten Wissensbeständen, deren Rekonstruktion als wesentliche Voraussetzung angesehen werden kann für den Zugang zu visuellem Material in seiner Eigenlogik.

### 5.3.2 Grenzen des interpretativen Paradigmas

Wenn Uwe Flick im Jahre 2002 für die Analyse von Film- und Videomaterial zusammenfassend feststellte: „Bislang gibt es keine unmittelbar auf die visuelle Ebene abzielenden Auswertungsverfahren für solches Material“ (Flick 2002: 231f.), so sind wir auch heute in diesem Bereich qualitativer Methoden, was deren Breite anbetrifft, kaum weiter fortgeschritten. Dies hängt, wie bereits mehrfach angesprochen, ganz wesentlich damit zusammen, dass die fortgeschrittenen qualitativen Methodologien, für welche elektromagnetische oder elektronische Aufzeichnungen die Grundlage darstellen, zunächst vor allem im Bereich der Gesprächsanalyse, also auf der Grundlage der Audiografie, entwickelt worden.

Diese sind, wie überwiegend die qualitative Forschung ganz allgemein, dem interpretativen Paradigma zuzuordnen, stehen insbesondere in der Tradition der Ethnomethodologie und Konversationsanalyse. Die methodologisch fundierten Ansätze auf der Grundlage der Videografie sind hieran anschließend entwickelt worden. So heißt es bei Christian Heath (1997: 184), einem der Wegbereiter dieser Richtung der Videoanalyse: „Ethnomethodologie und Konversationsanalyse haben deshalb die Mittel zur Verfügung gestellt, mit Hilfe derer es möglich wird, Videos für soziologische Absichten zu nutzen“.

Charles Goodwin (2001: 157) formuliert die Prämissen und Grenzen einer derartigen Videoanalyse, die sich selbst in der Tradition der Ethnomethodologie und Konversationsanalyse verortet, sehr deutlich: „in den Arbeiten, die es hier zu beschreiben gilt, werden weder das Visuelle („vision“) noch die Bilder oder andere Phänomene, die Gegenstand der Betrachtung seitens der Teilnehmer sind, als kohärente, in sich geschlossene Bereiche („self-contained domains“) behandelt, die in ihrer eigenen Begrifflichkeit zum Gegenstand der Analyse werden können. (...) Der Fokus der Analyse ist somit nicht die Repräsentation des Visuellen *per se*, sondern

vielmehr die Rolle, welche visuelle Phänomene in der Produktion sinnvollen Handelns spielen.“

Goodwin vermag hier klar die Grenzen der eigenen Analyse zu markieren: Bilder werden hier nicht – wie die Konversationsanalyse dies in Bezug auf Texte zu leisten imstande ist – als selbst-referentielle Systeme behandelt (vgl. dazu Kapitel 3). Goodwin (ebenda) führt dazu weiter aus, „dass visuelle Phänomene nur erforscht werden können, indem ein vielfältiges Set von semiotischen Ressourcen berücksichtigt wird (...). Viele von diesen, wie z.B. die Struktur, die durch laufende Gespräche beigesteuert wird, sind in keinerlei Hinsicht visuell, aber die sichtbaren Phänomene (...) können ohne sie nicht zureichend analysiert werden.“

Während Goodwin es hier also nicht für möglich hält, sichtbare Phänomene ohne Berücksichtigung von Gesprächen, von verbalen Äußerungen zu analysieren, findet sich andererseits in der Konversationsanalyse eine lange Tradition, Gespräche, also verbale Phänomene, ohne Berücksichtigung anderer semiotischer Ressourcen, insbesondere sichtbarer Phänomene, zu analysieren. Die Ausklammerung dieser anderen Ressourcen wird dort sogar zum methodischen Prinzip erhoben.

Die Konversationsanalyse war bahnbrechend für die paradigmatische Orientierung qualitativer resp. rekonstruktiver Textinterpretation, indem sie zuallererst einen derartigen empirischen Zugang zum Text postuliert hat, also einen Zugang zum Text als einem selbst-referentiellen System, wie ich dies bereits zu Beginn des Kapitels 3 mit dem Zitat von Harvey Sacks (1995: 536) zu verdeutlichen suchte: Es ist aber eben dies, was von Seiten der konversationsanalytisch beeinflussten Videointerpretation im empirischen Zugang zum *Bild* nicht nur nicht realisiert, sondern auch methodologisch ausgeschlossen wird. Das zentrale Problem dabei ist, dass ich weder hier noch in anderen Veröffentlichungen der in der Tradition der Konversationsanalyse sich verortenden Videoanalyse eine explizite Begründung für diese Differenz im Bereich des methodologischen und (erkenntnis-) theoretischen Status von Bildern im Unterschied zu Texten finden konnte.

Diese methodologische und forschungspraktische Orientierung dieser Videoanalyse, die sich in der Tradition der Ethnomethodologie und Konversationsanalyse verortet, hat zum einen die Konsequenz gehabt, dass die hier entwickelten Verfahren primär in ihrer *ergänzenden Funktion zur Gesprächsanalyse* entwickelt worden sind. Darüber hinaus bindet die Orientierung an der Konversationsanalyse und der Ethnomethodologie diese Videointerpretation aber auch an weitere methodologisch-theoretische Prämissen des *interpretativen Paradigmas*.

Aus der Perspektive der dokumentarischen Methode ist eine der wesentlichen Begrenzungen des interpretativen Paradigmas darin zu sehen, dass eine Differenzierung von (unmittelbarem) *Verstehen* einerseits und *Interpretieren* andererseits nicht möglich ist. Verstehen vollzieht sich auf der Basis gemeinsam geteilter *impliziter* Wissensbestände („konjunktiver Erfahrungen“ im Sinne von Mannheim). Im Unterschied dazu setzt das *Interpretieren* eine *Explikation* dieser Wissensbestände voraus und ist ausschließlich im Medium der Sprache möglich (was selbstverständlich nicht bedeutet, dass nicht auch im Medium der Sprache implizite Wissensbestände vorausgesetzt und vermittelt werden).

Die Analyse im Bereich des interpretativen Paradigmas setzt, wie der Name schon sagt, erst auf jener Ebene ein, auf der Verständigung sich nach dem Modell

des wechselseitigen *Interpretierens* vollzieht. Dies wird vor allem auch darin deutlich, dass die Interpretationen ganz wesentlich auf der Ebene des *wörtlichen* Sinngehalts der Erforschten verbleiben. In den Blick geraten damit die Alltagstheorien, die Common Sense-Theorien der Erforschten, nicht aber deren handlungsleitendes Wissen, welches als implizites Wissen die Praxis strukturiert.

Eine Verständigung *im Medium des Bildes* selbst vollzieht sich aber eben – in den Bereichen, in denen sie als solche gelingt und sich nicht (ergänzend) des Mediums Sprache bedienen muss – auf der Basis impliziter (gemeinsam geteilter) Wissensbestände. In diesem Sinne, indem nämlich handlungsleitendes Wissen in ihnen ‚gespeichert‘ ist, bilden sie eine *Handlungspraxis* ab bzw. repräsentieren sie handlungsleitendes Orientierungswissen. Das handlungsleitende Potential der Bilder erschließt sich mir, indem ich den in ihnen sich dokumentierenden *modus operandi* bzw. Habitus rekonstruiere, wie er in den abgebildeten Gebärden und der räumlichen Positionierung der *abgebildeten* Bildproduzent(inn)en zueinander, also der szenischen Choreografie, seinen Ausdruck findet, ebenso aber auch in den Gestaltungs- und Selektionsleistungen der *abbildenden* Bildproduzent(inn)en.

Diese Interpretation des Habitus ist, wie bereits in Kapitel 3 angesprochen (siehe auch Kap. 5.4), kategorial verschieden von einer Interpretation nach Art der Unterstellung von Motiven, die sich derart vollzieht, dass ich Geschichten konstruiere, in welche die Akteure auf dem Bild eingebunden sind. Ich verbleibe dann, wenn wir der Unterscheidung von Panofsky folgen, auf der ikonografischen Ebene. Es ist dies die Ebene, auf der wir uns im Alltag, also in unseren Common Sense-Interpretationen, den Bildern nähern. Und es ist eben diese Ebene, welche das interpretative Paradigma nicht grundsätzlich zu transzendieren vermag.

### 5.3.3 Videoanalyse in ergänzender Funktion zur Gesprächsanalyse

Die aktuellen Video-Analysen im Bereich der interpretativen Soziologie und der Ethnografie sind vor allem im Bereich der sogenannten „work-place-studies“ entstanden und stehen wesentlich in der Tradition der Konversationsanalyse, worin, wie bereits skizziert, angelegt ist, dass diese Verfahren der Videoanalyse primär in ihrer *ergänzenden Funktion zur Gesprächsanalyse* entwickelt worden sind. Somit stellt hier das Bild lediglich eine Ergänzung des Textes dar, also eine Ergänzung des Transkripts des gesprochenen Wortes.

Das gilt auch für die „Video-Interaktions-Sequenzanalyse“ oder „Video-Interaktions-Analyse“, einem Verfahren, welches von Hubert Knoblauch (2004 u. 2005) ausformuliert wurde. Die Analyse des Visuellen stellt hier ein weitgehend ungelöstes Problem dar: „Für jede Sequenzanalyse stellt dieser Horizont synchroner Bildverweise ein eingestandenes Problem dar, da er die Sequentialität der Daten und der Vorgehensweise unterbricht.“ (Knoblauch 2005: 272). Da im Rahmen der Video-Interaktions-Analyse das Video nicht den Stellenwert eines Alltagsprodukts hat, sondern dies lediglich als Erhebungsinstrument genutzt wird (vgl. Kap. 5.1), sind die Beeinträchtigungen aufgrund des beschränkten Zugangs zur Simultanstruktur des Bildes im Bereich der Analyse der Leistungen der abbildenden Bildproduzent(inn)en weniger problematisch als diejenigen im Bereich der Leistungen der

abgebildeten Bildproduzent(inn)en, also der Gebärden in ihrer Simultanstruktur (vgl. Kap. 5.5). Eine besondere Sensibilität für letzteres Problem findet sich – abgesehen von der bereits erwähnten Studie auf der Basis der dokumentarischen Methode von Wagner-Willi (u.a. 2005 u. 2007) auch in den auf die Forschung im Bereich der Erwachsenenbildung bezogenen Ausführungen von Nolda (2007: 483f.) unter dem Begriff der „Konstellationsanalyse“. Diese bearbeitet das „Phänomen der Simultaneität und ist anders als die Sequenzanalyse in besonderer Weise mit der visuellen Ebene verbunden.“

In Bereichen, in denen die Videografie als Erhebungsinstrument genutzt wird und allzumal dort, wo das Erkenntnisinteresse auf die Gesprächsanalyse gerichtet ist, ist es selbstverständlich legitim, der Bildanalyse einen lediglich ergänzenden Stellenwert zu zuweisen, wie dies im Bereich der dokumentarischen Methode etwa in der bereits erwähnten Studie von Iris Nentwig-Gesemann (2006a u. 2007a) der Fall ist, die der Frage nachgeht, wie sich Vorschulkinder untereinander und mit Erwachsenen über ihre alltäglichen Praktiken des Spiels verständigen. Analog gilt dies auch für Analysen im Bereich der Unterrichtsforschung (vgl. Krummheuer/Naujok 1999; Brandt/Krummheuer/Naujok 2001 sowie andere Arbeiten in: Aufschneider/Wenzel 2001).

Ebenso wie die Video-Interaktions-Sequenzanalyse von Hubert Knoblauch steht auch die Video- resp. Filmanalyse von Angela Keppler, in welcher das Video als (mediales) Alltagsprodukt Gegenstand der Analyse ist, in der Tradition der Konversationsanalyse. Sie bemerkt mit Bezug auf diese Tradition: „Allerdings bleibt auch hier bislang die Rolle der visuellen Präsentation – und mit ihr die non-verbale Seite der Kommunikation – ausgeblendet, obwohl sich die Konversationsanalyse grundsätzlich als ein Programm der Untersuchung verbaler wie non-verbaler Kommunikation versteht“ (Keppler 2006: 104).

Knoblauch (2005: 273) verweist als eine „Möglichkeit, das Problem der Unbestimmbarkeit synchroner visueller Bezüge zu lösen“, auf Arbeiten im Bereich der „sozialwissenschaftlichen Hermeneutik“, da diese „eine umfassende Bilddeutung visueller Aspekte vornimmt“. Er nennt vor allem die Arbeit von Jürgen Raab (2002), eine Analyse eines professionellen Hochzeitsvideofilms. Diese setzt sich in ausgesprochen gründlicher Weise mit den technischen Voraussetzungen dieses Bild-Mediums auseinander, sowohl mit der „Kamerahandlung (Wahl des Standorts, der Perspektiven, der Kadrierung usw.)“ wie auch mit den „Nachbearbeitungstechniken (Schnitt, Montage, Verfremdungen, Vertonung)“. Die Analyse von Raab bleibt auch nicht auf die Produkte der abbildenden Bildproduzenten beschränkt, sondern nimmt auch „Wechselwirkungen zwischen den Darstellungen vor und den Handlungen hinter der Kamera“ in den Blick (Raab 2002: 472).

Auch diese Studie stützt sich auf die „Verfahrensweisen der sozialwissenschaftlichen Hermeneutik sowie der im Kern die gleiche Forschungsperspektive bezeichnenden hermeneutischen Wissenssoziologie“ (ebenda), welche, wie bereits diskutiert, wenig Möglichkeiten bietet, den bereits im Common Sense erreichten Zugang zu Bild und Film auf der Grundlage eines Wechsels der AnalyseEinstellung in der Weise zu transzendieren, dass es in systematischer Weise gelingt, zur Rekonstruktion der impliziten Wissensbestände vorzudringen. Eine umfassende und fundierte theoretische sowie methodologische Begründung der in der hermeneutischen

Wissenssoziologie fundierten „wissenssoziologischen Bildhermeneutik“ findet sich in Raab 2008 (Kap. 5).

In Kontext der „hermeneutischen Wissenssoziologie“ ist auch die einschlägige Studie von Reichertz (2000) mit dem Titel „Die frohe Botschaft des Fernsehens“ zu verorten. Jo Reichertz hat solche Fernsehshows untersucht, in denen Kandidat(inn)en Einblicke in intime und krisenhafte Lebenszusammenhänge geben und die er „performative Fernsehauftritte“ nennt (Reichertz 2000: 41).<sup>69</sup>

Was die methodische Verfahrensweise im Hinblick auf die Bildinterpretation anbetrifft, so erfahren wir lediglich, dass Reichertz „vorschlägt“, „den vom Betrachter erstellten Bildtext als Feldprotokoll zu betrachten“ und dass dieses Feldprotokoll dann im Sinne der Sequenzanalyse der objektiven Hermeneutik interpretiert wird. Allerdings wird der Schritt vom Bild zum „Bildtext“ nicht beschrieben oder methodisch und methodologisch durchleuchtet bzw. begründet. Indem Reichertz den Bildtext der sequenzanalytischen Verfahrensweise der Textinterpretation unterzieht, wird der Besonderheit und Eigensinnigkeit des Bildes gegenüber dem Text kaum eine Chance gelassen. Somit „erfolgt die Interpretation und Analyse auch hier vor allem anhand des transkribierten *sprachlichen* Texts, visuelle Kompetenzen werden allenfalls am Rande berücksichtigt“, wie Angela Keppler (2006: 103) die Vorgehensweise von Reichertz kritisch resümiert.

Keppler selbst (2006) hat eine weit fortgeschrittene Methode der Film- bzw. Fernsehanalyse von sozialwissenschaftlicher Relevanz vorgelegt. Auch bei ihr bleibt jedoch *zum einen* die Sprache das zentrale Medium ihrer Analyse. Es zeigt sich auch hier die Tendenz, der Bildanalyse eine „ergänzende“ Funktion zuzuschreiben. So heißt es in kritischer Auseinandersetzung mit der Konversationsanalyse, der sie sich in vieler Hinsicht selbst verpflichtet weiß: „An dieser Stelle ist der ergänzende Einsatz *bildanalytischer* Methoden dringend notwendig. Eine sozialwissenschaftliche Filmanalyse sollte stärker als bisher eine Detailanalyse der verbalen und non-verbalen Elemente sowie der bildsprachlichen Mittel mit einschließen“ (Keppler 2006: 104).

Den bildanalytischen Verfahren lediglich eine „ergänzende“ Funktion zukommen zu lassen, wie Keppler hier ihre eigene Analyse charakterisiert, vermag von unserem Standpunkt der dokumentarischen Methode der Filminterpretation den Besonderheiten der Bildmedien (und auch der Rekonstruktion non-verbaler Handelns) nicht gerecht zu werden. Angela Kepplers Analyse (2006) ist aber eben gerade auf die Eigenarten der Bildmedien gerichtet, wie bereits der Untertitel ihres Buches („Eine Theorie des Fernsehens...“) verdeutlicht. Hier kann die Filmanalyse die „bildsprachlichen Mittel“ nicht lediglich „mit einschließen“. Vielmehr sollten diese einen gleichberechtigten Stellenwert erhalten. Dies aber setzt wiederum voraus, dass sie nicht vor dem Hintergrund des sprachlich-textlichen Vor-Wissens interpretiert werden, sondern dass – wie in Kapitel 2.3 dargelegt – eine Verfahrensweise

69 Dies sind Fernsehshows wie z.B. „Nur die Liebe zählt“, „Surprise, Surprise“ und „Traumhochzeit“ (vgl. Reichertz 2000: 10). Sein Erkenntnisinteresse definiert er folgendermaßen: „Von Interesse sind hier (in einer ersten Annäherung) die sozialen Gründe, welche *Menschen dazu bewegen, sich als Kandidaten für performative Fernsehauftritte zu bewerben*“ (a.a.O.: 41), allerdings nicht „subjektive“ Motive, sondern „sozial typisierte, also objektive Beweggründe“ (a.a.O.: 43).

methodisiert wird, bei der es gelingt, bestimmte Bereiche des sprachlich-textlichen Vor-Wissens zu suspendieren.<sup>70</sup>

Die Grenzen der Methodik von Keppler, deren Erkenntnisinteresse auf die Bildmedien gerichtet ist, sehe ich zum einen, wie bereits dargelegt, darin, dass bei ihr die Sprache das zentrale Medium der Analyse bleibt, und zum anderen – und dies ist auch ganz allgemein bei den Filmwissenschaften im engeren Sinne zu beobachten – darin, dass dort, wo die Bildanalyse Bedeutung gewinnt, diese weitgehend auf den *modus operandi* der *abbildenden* Bildproduzent(inn)en fokussiert. Dieser Bereich wird in umfassender und übersichtlicher Weise dargestellt (vgl. Keppler 2006, Kap. 3.4.: Visuelle Dimensionen). Demgegenüber bleiben die Bewegungen der *abgebildeten* Bildproduzent(inn)en, also ihre Gebärden und somit der Bereich der nonverbalen Kommunikation (im Unterschied zur Analyse der von ihnen produzierten Texte) aus der Betrachtung und der genaueren empirischen Analyse weitgehend ausgeschlossen.<sup>71</sup>

Bedeutung gewinnen die Bewegungen der *abgebildeten* Bildproduzent(inn)en eher unter dem Gesichtspunkt des Einflusses der *abbildenden* Bildproduzent(inn)en auf diese – also im Sinne einer „Bewegungsregie“ (Keppler 2006: 119). Der „Mimik, Gestik und weiteren Bewegungen“ kommt darüber hinaus lediglich der Stellenwert zu, dass sie „dem Gesagten eine zusätzliche Kontur verleihen“ (a.a.O.: 121). Entsprechend schenkt Keppler auch der genaueren Analyse von Fotogrammen (vgl. Kap. 3.4) und ihrer formalen Komposition kaum Beachtung. In letzterer Hinsicht befindet sie sich, wie gesagt, in Übereinstimmung mit den aktuellen filmwissenschaftlichen Betrachtungen, aber in einem Kontrast zu den klassischen Filmtheorien von Balázs, Kracauer und Panofsky, in denen diese Ebene der Analyse, die in der Sprache der Kunstgeschichte als vor-ikonografische bezeichnet werden kann, im Zentrum stand.

#### 5.4 Die Bewegungen der abgebildeten Bildproduzent(inn)en

Erwin Panofsky hat sich nicht allein mit der Bild-, sondern auch mit der Filminterpretation befasst und betont, „dass ein Film (*moving picture*), auch wenn er zu sprechen gelernt hat, ein Bild bleibt, das sich bewegt (*a picture that moves*), und nicht zu einem Werk der Literatur wird, das man aufführt.“ (1999: 27; ursprünglich: 1949). Und der Klassiker der Filmtheorie Belá Balázs (2001a: 26) hatte schon 1924 betont: „Feinheit und Kraft der Bildwirkung und der Gebärde machen die Kunst des Films aus. *Darum hat er nichts mit der Literatur zu schaffen.*“

70 Gewisse – allerdings methodisch noch ganz am Anfang stehende – Ansätze zu einer „Emanzipation“ des Bildmaterials vom Sprachmaterial“ zeigen sich bei Kade/Nolda (2007: 170) und Nolda (2007) in ihrer Analyse von Kursen der Erwachsenenbildung und in der im selben Arbeitskontext entstandenen – an die objektive Hermeneutik anknüpfenden – Arbeit von Herrle (2007).

71 im Kap. 3.4.5: „Bildkomposition“ (Keppler 2006:117ff.) widmet sie diesem Bereich im theoretischen Teil ihres Buches knapp drei Seiten von insgesamt 340.

Werke der Literatur aufzuführen, wie dies in der Perspektive von Balázs und Panofsky im Theater geschieht<sup>72</sup>, stellt für diese eine ganz andere Form des Kunstschaffens dar als der Film: „Diese Künste (des Theaters; R.B.) agieren sozusagen von oben nach unten. Sie beginnen mit einer Idee, die in die gestaltlose Materie projiziert werden soll, nicht mit den Objekten, aus denen die äußere Welt besteht“ (Panofsky 1999: 53). Diese „Idee“ ist nach Panofsky im Bereich des Theaters das wesentlich im Medium der „Literatur“, also der Sprache, entworfene, in seiner Dramaturgie textlich und narrativ vorformulierte und auf der Ebene der Programmatik und der Skripte angesiedelte ‚Stück‘<sup>73</sup>.

##### 5.4.1 Der Zugang des Films zu elementaren Ausdrucksformen und Ebenen sozialer Realität

Wenn wir diese Sinndimension des sprachlich und narrativ Formulierten, die nach Ansicht von Panofsky im Theater in den Vordergrund tritt, in dem Kategoriensystem zu fassen suchen, welches von Panofsky in seinen methodologischen Schriften (1932 u. 1975) ausgearbeitet worden ist, so ist dies die *ikonografische* Sinnenebene, diejenige der ikonografischen Narrationen, aus der die Akteure auf der Bühne ihr Treiben deduktiv („von oben nach unten“) ableiten sollen. In Abgrenzung von dem derart charakterisierten deduktiven Logik des Theaterstücks operiert der Film also von ‚unten nach oben‘.<sup>74</sup>

Panofsky (1999: 54) betont: „Der Stoff des Films ist die äußere Realität als solche“, Während Panofsky hier von „äußerer Realität“ spricht, verwendet Kracauer, der, wie seine Zitate (1964: 400) und seine Korrespondenz mit Panofsky<sup>75</sup> zeigen, durch Panofsky beeinflusst ist, den Begriff der „physischen Realität“, deren „Errettung“ durch den Film gefördert werden soll, wie Kracauer (1964: 389) pathetisch formuliert.<sup>76</sup> Wenn Filme der „filmischen Einstellung genügen“ (Kracauer 1964: 391), also „dem Medium gemäß sind, werden sie nicht von einer vorgefaßten Idee zur materiellen Welt herabsteigen, um diese Idee zu erhärten; umgekehrt beginnen sie damit, physische Gegebenheiten auszukundschaften, und arbeiten sich dann in

72 In der hier angestrebten qualitativen Methodologie des Films geht es nicht darum, die Ausführungen von Panofsky, Balázs und Kracauer auf ihre theaterwissenschaftliche Plausibilität hin zu diskutieren, sondern ihre *sozialwissenschaftliche*, insbesondere handlungstheoretische, Relevanz auszuloten.

73 In diesem Sinne warnt auch der Klassiker der Filmtheorie Pudowkin (2003b: 75) vor der „Theatralisierung der Filmkunst“.

74 Einen derartigen Herstellungsprozess des Filmes beschreibt David Lynch in einem Interview mit der Süddeutschen Zeitung (Süddeutsche Zeitung 2007) in markanter Weise: „Ich habe keine globale Vision, bevor ich zu drehen beginne. Alles hängt von der Bewegung ab, vom Fluss der Ideen. In gewisser Weise existieren meine Filme schon vor dem Dreh ...ich sammle einfach die Fragmente, Gesichter, Wörter, Töne und Räume, um sie wie in einem Puzzle zusammenzufügen. Das Verständnis ist eine Abstraktion, die aus der Intuition hervorgeht“.

75 Zwischen Panofsky und Kracauer, der – ebenso wie Mannheim – hinsichtlich seiner Existenz als (Zwangs-) Emigrant ein Schicksalsgenosse von Panofsky ist, hat eine umfangreichere Korrespondenz stattgefunden; siehe als Überblick: Panofsky 2003: 1234.

76 Mit Bezug auf die Fotografie spricht Kracauer (1964: 45) sogar von ihrem Wesen, „Natur im Rohzustand wiederzugeben, so wie sie unabhängig von uns existiert.“

der von ihnen ausgewiesenen Richtung nach oben, zu irgendeinem Problem oder Glauben hin“ (Kracauer 1964: 253).

Wenn wir Kracauer und ebenso Panofsky nicht objektivistisch oder essentialistisch interpretieren wollen<sup>77</sup>, so lässt sich geltend machen, dass es ihnen hier nicht, wie die Begriffe der „äußeren Realität“ und der „physischen Realität“ und andere Formulierungen missverstanden werden könnten, um einen Realismus auf der *ikonografischen Ebene* geht, auf der sich die Frage stellt, ob die vom Film *erzählte Geschichte* eine realistische ist.<sup>78</sup> Vielmehr geht es Kracauer wie auch Balázs und Panofsky darum, dass das Medium Film einen bisher nicht gekannten Zugang zu einer elementaren Ebene sozialer Verständigung und somit sozialer Realität eröffnet.

Es geht Kracauer, Balázs und Panofsky, anders formuliert, nicht darum, den *Herstellungscharakter* von Realität zu negieren oder zu unterlaufen, sondern darum, den Zugang zu gewinnen zu einer *elementaren* und grundlegenden Dimension der Herstellung von Realität, nämlich derjenigen auf der Ebene der Körperlichkeit, der inkorporierten Gesten, der Gebärden und der Mimik, wie sie uns in den Gestaltungsleistungen der abgebildeten Bildproduzent(inn)en begegnen.<sup>79</sup> Insofern ist auch die Betonung der „physischen“ oder „materiellen“ Elemente dieser Ebene in dem Sinne zu verstehen, dass es sich hier ganz wesentlich um körpergebundene, um *inkorporierte* Praktiken und um die Positionierung von Körpern und physischen Objekten im Raum handelt.

In seinen methodologischen Schriften zur Kunstinterpretation spricht Panofsky (1975) von der Ebene „primärer oder natürlicher Bedeutungen“ und zieht damit schon in der Formulierung eine stärkere Grenze gegenüber objektivistischen Sichtweisen. Diese Ebene primärer Bedeutungen, für die der Film den Blick zu schärfen vermag, hat Panofsky auch als *vor-ikonografische Ebene* bezeichnet (siehe dazu auch Kap. 3), die er bekanntlich von der *ikonografischen*, an Text und Narration gebundenen, Sinnenebene unterscheidet. Und es ist die *vor-ikonografische Ebene*, deren genaue Beobachtung und Beschreibung die wesentliche Grundlage der ikonologischen Interpretation darstellt.

Dass im Verständnis von Panofsky die den Film in seiner Eigenart konstituierende Sinnenebene nicht die ikonografische ist, wird auch darin deutlich, dass Panofsky (1999: 27) der Sprache, die für diese Ebene zentral ist, einen lediglich mar-

77 Dass wir dies nicht tun sollten, ergibt sich bei Panofsky aus seinen anderen methodologischen Schriften (1932 u. 975), bei Kracauer aus seiner Kritik an den „Realisten des 19. Jahrhunderts“ im ersten Kapitel desselben Buches (1964: 40f.). Allerdings ist Kracauer in vieler Hinsicht wieder-sprüchlich. So verwendet er auf S. 395 die Formulierung: „Unter allen existierenden Medien ist es allein das Kino, das in gewissem Sinne der Natur den Spiegel vorhält“. Auf S. 40 hatte er dies als „Glaubenssatz der extremen Realisten des 19. Jahrhunderts“ kritisiert: „In Wirklichkeit gibt es einen solchen Spiegel gar nicht.“ – Eine objektivistische oder naturalistische Interpretation Panofskys und auch gleich Kracauers findet sich bei Prange (1992: 186): „Mit dem Verzicht auf eine Berücksichtigung des technischen Mediums wird die Realität selbst zum Regisseur erkoren“.

78 Als Realismus auf der ikonografischen Ebene versteht Keppler (2006: 160) die Ausführungen von Kracauer: „Abweichend von Siegfried Kracauer und anderen nehme ich an, dass der Film in Unterschied zur klassischen Fotografie kein von Haus aus realistisches Medium ist“. Keppler (2006: 64) interpretiert Kracauer dahingehend, er habe Panofsky missverstanden.

79 „Wir können nur dann darauf hoffen, der Realität nahezukommen, wenn wir ihre untersten Schichten durchdringen“, heißt es bei Kracauer (1964: 387).

ginalen Stellenwert beimit: „Jeder Versuch, im Film Vorstellungen und Empfindungen ausschließlich oder hauptsächlich durch Sprache mitzuteilen, hinterlässt ein Gefühl der Verlegenheit oder Langeweile oder beides.“ Somit ist auch, wie wiederum Kracauer (1964: 393) deutlich macht, nicht die „Story“, also das Narrative des Films, die zentrale Sinnenebene, sondern es sind die „kleinen Einheiten“ als „Momente des täglichen Lebens“: „Man betrachte irgendein Element eines solchen Story-Films. Zweifellos hat es die Aufgabe, der Story zu dienen, zu der es gehört; aber gleichzeitig affiziert es uns auch stark, vielleicht sogar in erster Linie, als ein fragmentarisches Moment der sichtbaren Realität, umgeben von einem Hof unbestimmbarer Bedeutungen“ (ebenda).

„Unbestimmbar“ und „fragmentarisch“ erscheint diese Sinnenebene, wenn wir sie mit der spezifischen Bestimmtheit des *ikonografischen* Sinns vergleichen, welcher die Elemente dadurch bestimmbar, d.h. explizierbar, macht, dass er sie in einen Ordnungszusammenhang stellt, in dem diese als Teile von (motivierten) *Handlungen*, also von Um-zu-Motiven (vgl. auch die Ausführungen in Kap. 3) und schließlich von Handlungsverläufen, also Geschichten, Narrationen konstruiert werden. Wenn wir im Alltag über einen Film zu berichten bzw. Rechenschaft über ihn abzulegen versuchen, bewegen wir uns zumeist auf dieser Ebene. Die ikonografischen Typisierungen und Konstruktionen entsprechen den Konstruktionsprinzipien der Theorien des Common Sense, der *theoretischen* Alltagskonstruktionen (s. dazu auch: Kap. 2 sowie genauer: Bohnsack 2009a).

Nun lässt sich allerdings die Narration auf der *ikonografischen Ebene* unterscheiden von derjenigen auf der *vor-ikonografischen* oder *denotativen Ebene*. Auch Metz (1972: 138) betont die Bedeutung der denotativen oder vor-ikonografischen Ebene für das, was die Besonderheit des „Kinos“ ausmacht: „Das Kino ist *in erster Linie* wegen seiner Verfahren der Denotation eine spezifische Sprache.“<sup>80</sup> Die denotative Ebene umfasst aber nicht nur den Bereich der Bewegungs- und Handlungsverläufe der abgebildeten Bildproduzent(inn)en, sondern auch und insbesondere die Montage und die Einstellungsveränderungen, also den Bereich der Leistungen der abbildenden Bildproduzent(inn)en, wie sie zur Konstitution dieser Ebene der Erzählung beitragen.

Jene Elemente, welche im Kontext des ikonografischen Sinngehalts, durch den unsere Alltagstheorien strukturiert sind, als Fragmente oder Bruchstücke erscheinen, bilden im Kontext des (auf der denotativen oder vor-ikonografischen Ebene aufbauenden) ikonologischen oder dokumentarischen Sinngehalts „neuartige Totalitäten“ (Mannheim 1964: 123) oder Ganzheiten. Nach Mannheim (1964: 120) „muß beim Verständnis des Dokumentarischen nicht das ganze Werk in seinen objektiven Sinnbezügen erfasst sein, das Dokumentarische kann vielmehr auch an einem Bruchstück in Erscheinung treten, z.B. an einer charakteristischen Linienführung, an einer eigentümlichen Gliederung des Raumes oder Farbenbehandlung, mit einem Wort: auch an unselbständigen Teilen des Werkes.“

80 Metz (1972: 138) bezeichnet die Erzählung auf der denotativen Ebene mit dem griechischen Begriff der „Diegese“ und führt aus, dass dieser Begriff in seinen griechischen Ursprüngen „speziell einen der obligaten Teile der judizialen Rede bezeichnete, nämlich die Darlegung der Fakten. Im Zusammenhang mit dem Kino wurde der Begriff von Etienne Souriau aufgegriffen; er bezeichnet die dargestellte Instanz des Films, d.h. im Grunde die Gesamtheit der filmischen Denotation.“

Die dokumentarische oder ikonologische Interpretation ist nicht darauf angewiesen, die Einheit des zu interpretierenden Falles in der Weise zu rekonstruieren, wie diese auf der ikonografischen Ebene verstanden wird: also den gesamten Film als Einheit zu nehmen oder sich an den Themen und Unterthemen zu orientieren, die durch die Einheiten der Film-Narration segmentiert sind. In den Common Sense-Theorien gehen wir von solchen an die sprachlich-narrative Ebene gebundenen Vorstellungen der Einheit des Falles aus. Demgegenüber ist im Sinne der dokumentarischen Methode die Einheit des Werkes dann und insoweit gegeben, wenn an dessen unterschiedlichen Teilen – und auch an „Bruchstücken“ – immer wieder homologe Muster im Sinne eines *modus operandi*, einer performativen Struktur identifiziert werden können.

Im Falle der Filminterpretation auf der Basis der dokumentarischen Methode wird – im chronologischen Ablauf der Analyseschritte – ein derartiger *modus operandi* und eine in diesem Sinne verstandene Einheit oder Ganzheit primär im Bereich des Bildes und erst sekundär im Bereich des Textes und der ikonografischen Narrationen gesucht.<sup>81</sup> Es geht darum, dass „die Filmwissenschaft nach der Erzählung die Bildhaftigkeit der Filme in den Vordergrund stellt“ (Bredenkamp 2004: 17). Dies gilt auch für die Aufteilung und Auswahl der zu interpretierenden Sequenzen (dazu: 5.8.2). Die Aufteilung orientiert sich an Einheiten auf der Ebene der vor-ikonografischen Beschreibungen der abgebildeten (Szenen) und der abbildenden Bildproduzent(inn)en (Einstellungen). Die sich aus der Relationierung von Einstellungen und Szenen ergebende Einheit ist die „Sequenz“ (genauer dazu: 5.6.1).

#### 5.4.2 Gebärden, Operationen und institutionalisierte Handlungen

Das Dokumentarische oder Ikonologische findet sich nicht nur an Bruchstücken eines Werkes, sondern auch an Bruchstücken einer Handlung. Denn im Unterschied zur Interpretation des objektiven oder immanenten bzw. ikonografischen Sinngehalts, welche danach fragt, *was* das für eine Handlung ist (bspw. ein „Gruß“) und somit Motive unterstellen muss (bspw.: „ich hebe die Hand, um zu grüßen“), stellt sich auf der dokumentarischen oder ikonologischen Sinnesebene die Frage nach dem *Wie* der Herstellung dieser Handlung (bspw. ist das Heben der Hand „präventiv“, „unsicher“, „starr“).

Hinsichtlich der Bedeutung dieser auf der vor-ikonografischen Ebene zu vertretenden Elemente von Bewegungsabläufen, auf die sich dokumentarische Interpretation ganz wesentlich bezieht, zeigen sich, wie erwähnt, in einigen Punkten Übereinstimmungen zu der von Erving Goffman (1979: 24) vorgelegten Fotointerpretation, deren zentraler Gegenstand Bewegungen sind, die er als „*small behaviors*“ bezeichnet und die unterhalb der Ebene einer Interpretation von *Handlungen* angesiedelt sind.

81 Selbstverständlich ist gleichwohl die Einheit des gesamten Produkts, des Filmes erst dann gegeben, wenn Homologien zwischen dem Dokumentsinn des Bildes und demjenigen des Textes identifizierbar sind. (Wobei Homologien auch bedeuten kann, dass in homologer Weise ‚Brüche‘ oder Diskontinuitäten in beiden Dimensionen – im Text wie im Bild – identifizierbar sind.)

Für den Klassiker der Bewegungsanalyse Ray L. Birdwhistell (1968: 380), durch den auch Goffman beeinflusst ist, steht außer Frage, dass im Bereich von „Körperbewegung und Mimik“ („*body motion and facial expression*“) eine „Sprache“ der Bewegung existiert, die derjenigen der gesprochenen Sprache vergleichbar ist, sowohl in ihrer Struktur als auch in ihrem Beitrag zu einem systematisch geordneten Kommunikationssystem.“ Birdwhistell (1970: 79f.) erläutert die Bedeutung dieser Sprache der Bewegung am Beispiel des militärischen Grußes. Die Beobachtung, dass diese Handlung – obschon hochstandardisiert – durch Körperbewegung und Mimik, durch das *Wie* ihrer Herstellung, eine enorme Variabilität ihrer Bedeutung erhält, habe für ihn einen der ersten „Durchbrüche“ in der von ihm entwickelten Gebärdenanalyse, der „*kinesics*“ stimuliert: „Durch den Wechsel in Haltung, Gesichtsausdruck, der Geschwindigkeit oder Dauer der Bewegung des Grüßens und sogar in der Wahl ungeeigneter Kontexte für die Handlung kann der Soldat den Empfänger des Grußes ehren, herabwürdigen, zu gewinnen versuchen, beleidigen oder befördern.“

Ich möchte die Bewegungsabläufe auf der vor-ikonografischen Ebene noch einmal differenzieren in *Gebärden* einerseits und *operative Handlungen* andererseits. Mit letzterem Begriff lehne ich mich an den Begriff der „Operation“ bei Habermas (1981) an. So unterscheidet er zwischen „Handlungen“ einerseits und „Körperbewegungen“ oder „Operationen“ andererseits: Eine Körperbewegung oder Operation „ist Element einer Handlung, aber keine Handlung“ (Habermas 1981: 146). Allerdings wird im Rahmen der Handlungstheorie von Habermas, die an das zweckrationale Modell und somit an die ikonografische Ebene gebunden bleibt, den Operationen eine eigenständige Semantik nicht zuerkannt: Operationen „gewinnen nur als Infrastruktur anderer Handlungen einen Bezug zur Welt. Operationen berühren die Welt nicht“ (a.a.O.:147).

Träger der *Gebärden* können die Extremitäten (bspw. ‚Ausstrecken des Armes‘) sein, der Rumpf (bspw. ‚Drehen oder Beugen des Rumpfes‘), der Kopf (bspw. ‚Senken des Kopfes‘), aber auch die Mimik (bspw. ‚Lächeln‘). Die Gebärden lassen sich noch einmal in ihre *Elemente* differenzieren, die ich mit Birdwhistell (1952) als „*Kineme*“ bezeichne. Die vorliegenden Videoanalysen auf der Grundlage der dokumentarischen Methode setzen auf der vor-ikonografischen Ebene mit der genauen Rekonstruktion von Gebärden und operativen Handlungen und teilweise der *Kineme* an. Dies gilt für die bereits erwähnten Analysen von Monika Wagner-Willi (2004, 2005, 2006 u. 2007) über Schülerinteraktionen im Klassenraum oder von Amelie Klambeck (2007) zu Patient(inn)en mit „psychogenen Bewegungsstörungen“ und deren Bewegungen während der Chefarztvisite. Letztere Analyse wendet sich insbesondere auch den Elementen von Gebärden, den *Kinemen* zu.<sup>82</sup>

Die Beschreibung der *Mimik* stellt uns vor die größten Probleme im Bereich der vor-ikonografischen Interpretation. Sie erscheint in besonderer Weise polysem oder „polyphon“ und dies insbesondere aufgrund ihres simultanen Charakters, wie der Klassiker der Filmtheorie Bela Balázs (2001a:45) betont: „Der Gesichtsausdruck ist überhaupt polyphoner als die Sprache. Das Nacheinander der Worte ist wie das Nacheinander der Töne einer Melodie. Doch in einem Gesicht können die verschiedensten

82 Amelie Klambeck (2007) verwendet hier die Begriffe „Bewegungen“ anstelle von „Gebärden“ und „physische Teilhandlungen“ anstelle von „operativen Handlungen“.



Dinge *gleichzeitig* erscheinen wie in einem Akkord, und das Verhältnis dieser verschiedenen Züge zueinander ergibt die reichsten Harmonien und Modulationen.“

Der Versuch einer eindeutigen Charakterisierung des mimischen Ausdrucks (bspw. als ein ‚Lächeln‘, ein ‚Stirnrunzeln‘ oder als ‚starr‘) ist in valider Weise erst unter Berücksichtigung des Kontexts gleichzeitiger Ausdrucksbewegungen des gesamten Gesichtsfeldes möglich, also unter Berücksichtigung der Ganzheit: „Wir reagieren auf ein Gesicht als eine Ganzheit: Wir sehen ein freundliches oder mürrisches Gesicht, ein eifriges oder ein gleichgültiges Gesicht, Aufrichtigkeit oder Verschlagenheit, lange bevor wir feststellen können, welche Züge oder Zusammenstellungen von Zügen diesen intuitiven Eindruck hervorrufen“, betont der Kunsthistoriker Gombrich (1978: 367).

Notwendig ist also die Rekonstruktion des Zusammenspiels elementarer Bewegungseinheiten in ihrer Simultaneität. Hiermit ist zunächst das simultane Zusammenspiel innerhalb abgegrenzter Bereiche des Körpers gemeint (vgl. Birdwhistell 1952: 17). Idealerweise umfasst die Rekonstruktion des simultanen Zusammenspiels aber den gesamten Körper, orientiert sich an der „whole body‘ conception“ wie Ray Birdwhistell (1952: 8) betont (dazu genauer: Kap. 4.5) und bezieht schließlich auch die räumliche *Positionierung der Körper zueinander* mit ein („A dreht B den Rücken zu“). Diese nimmt eine Sonderstellung im Bereich der Gebärden ein. Auf der ikonologischen Ebene lässt sie sich in Anknüpfung an Imdahl (1996a) als  *szenische Choreografie* bezeichnen (vgl. Kap. 3). Hubert Knoblauch (1998: 312) hat auf die Bedeutung dieser „Choreographie der Körper“ am Forschungsbeispiel Bezug genommen und bezeichnet sie in Anknüpfung an Adam Kendon (1990) als „Körperformation“.

*Operative Handlungen* (bspw. ‚Sich-Hinsetzen‘, ‚Gehen‘, ‚Hose hochziehen‘) umfassen in der Regel mehrere Gebärden in ihrer Sequenzialität. Wesentliches Unterscheidungsmerkmal gegenüber den Gebärden im elementaren Sinne ist aber, dass sie – und aus diesem Grund möchte ich sie bereits als Handlungen bezeichnen – mit (wenn auch nur rudimentären) zweckrationalen Motivkonstruktionen, d.h. mit Konstruktionen von Um-zu-Motiven, versehen werden können. Dies geschieht bspw. derart, dass die einzelne Gebärde (‚Beugen des Rumpfes‘) „selbst nur Mittel im Sinnzusammenhang eines Entwurfes“ ist (Schütz 1974: 119), nämlich des Entwurfes ‚Sich-Hinsetzen‘.

Durch diesen Entwurf, durch diese zweckrationale Konstruktion eines Um-zu-Motivs (‚sie beugt den Rumpf, um sich hinzusetzen‘) wird das Sich-Hinsetzen zu einer *Handlung*, wenn wir Handlung im Sinne von Alfred Schütz (1974: 115ff.) verstehen. Hier lassen sich dann jeweils weitere Hierarchien von Um-zu-Motiven konstruieren, also bspw.: ‚Die Lehrerin setzt sich hin, um den Unterrichtsbeginn zu signalisieren‘. Der Unterschied dieser Motivkonstruktion zu den bisher genannten besteht aber darin, dass die bisher genannten am physischen Bewegungsverlauf beobachtbar sind (und auf dieser Grundlage unmittelbar falsifiziert oder verifiziert werden können)<sup>83</sup>, wohingegen der Handlungsentwurf beim letzteren Beispiel nicht

83 Obschon auch hier eine vollständige Sicherheit der Motivkonstruktion auf der Grundlage der Beobachtung des Handlungsvollzugs selbst nicht gegeben ist: Wenn wir bspw. beobachten, dass A im Gehen gegen B stößt, so ist nicht so ohne weiteres zu klären, ob A den B (mit Absicht) ‚anrempelt‘ oder (unabsichtlich) gegen ihn stolpert.

direkt am Handlungsverlauf *beobachtet* werden kann, sondern lediglich als Entwurf, als Um-zu-Motiv *unterstellt* werden muss. Die Bedeutung dieser operativen Handlung der Lehrerin erschließt sich in ihrer Signifikanz (im Sinne von G. H. Mead 1968) erst in der *Reaktion der Schüler/innen* auf diese Handlung.

Sobald also die Realisierung oder Enaktierung des Entwurfs am Handlungsverlauf *selbst* nicht mehr beobachtbar ist, sondern unterstellt oder attribuiert werden muss, bewegen wir uns bereits auf der *ikonografischen Ebene*. Die Unterstellung oder Attribuierung macht es erforderlich, dass wir auf Erwartungserwartungen (im Sinne der Idealisierung der Reziprozität der Perspektiven nach Alfred Schütz (1971) und auf ein narrativ vermitteltes Wissen zurückgreifen. Vorausgesetzt ist ein Wissen auf der Ebene von *Institutionen*: bspw. der Institution des ‚Unterrichts‘ oder auch elementarer: der Institution des ‚Großes‘ oder der ‚Begrüßung‘. Ich möchte deshalb von *institutionalisiertem Handeln* (auch: Rollenhandeln) sprechen. Dieses ist bereits auf der *ikonografischen Ebene* angesiedelt. Ich erinnere an das (in Kap. 3) bereits angeführte Beispiel von Panofsky (1975), der den Unterschied am Fall des ‚Hut-Ziehens‘ (genauer eigentlich: ‚Abnehmen des Hutes‘) erläutert, welches erst in ikonografischer Interpretation dann zum ‚Grüßen‘ wird.

Die *Validität* der Zuschreibung von Motiven, bei denen – auf der vor-ikonografischen Ebene – die Enaktierung des Entwurfes am Handlungsvollzug selbst beobachtbar wird (operative Handlung), ist unvergleichlich viel größer als diejenige der Konstruktionen von Erwartungserwartungen auf der ikonografischen Ebene. Und das Bild ist das einzige Medium, welches uns einen derart methodisch kontrollierbaren, also validen Zugang zu dieser elementaren Ebene der Handlungsinterpretation zu vermitteln vermag.

### Bewegungen der abgebildeten Bildproduzent(inn)en

vor-ikonografische Ebene	Beispiele	Motivkonstruktion
<i>Kineme</i> ( <i>Elemente von Gebärden</i> )	Kopf und Schultern gehen nach vorne, Becken nach hinten etc.	
<i>Gebärden:</i> <i>Gestik und Mimik</i> ( <i>Kinemorpheme</i> )	Beugen des Rumpfes	
<i>Operative Handlungen</i>	Sich-Setzen	A beugt den Rumpf, um sich zu setzen (Um-zu-Motiv ist am Bewegungsverlauf beobachtbar)
ikonografische Ebene	Beispiele	Motivkonstruktion
<i>institutionalisierte Handlungen</i> ( <i>Rollen</i> )	Lehrerin setzt sich ans Pult	A setzt sich, um den Unterricht zu beginnen (Um-zu-Motiv ist am Bewegungsverlauf nicht beobachtbar)

#### 5.4.3 Motivkonstruktion versus Rekonstruktion des Habitus

Eine Eigenart der ikonologischen oder dokumentarischen Interpretation kann nun deutlich werden, welche von elementarer Bedeutung ist: Dieselbe *Gebärde* (bspw. ‚Beugen des Rumpfes‘) kann immer auf zwei Sinnebenen zugleich interpretiert werden: zum einen zweckrational im Rahmen der Konstruktion eines Um-zu-Motivs (‚Sich-Setzen‘), mit der wir uns auf die Suche nach dem *subjektiv gemeinten Sinn* begeben. Zum anderen und zugleich kann die Gebärde auch (unter dem Aspekt des *Wie* der Herstellung dieser Gebärde) als *Dokument* für das *Wesen* oder den *Habitus* des Akteurs (‚Unsicherheit‘, ‚Gebrechlichkeit‘) interpretiert werden. Dies gilt auch für die Interpretation *operativer Handlungen* (‚Sich-Setzen‘). Auch diese können zweckrational, also im Kontext von Um-zu-Motiven (‚um Unterrichtsbereitschaft zu signalisieren‘) oder auch dokumentarisch interpretiert werden. Entscheidend ist dann, *wie* sich jemand hinsetzt, oder *dass* er oder sie sich (gerade jetzt) setzt: „Nicht das ‚Was‘ eines objektiven Sinns, sondern das ‚Daß‘ und das ‚Wie‘ wird von dominierender Wichtigkeit.“ (Mannheim 1964a: 134).

Die Strömungen qualitativer Sozialforschung, die sich auf die Suche nach dem subjektiv gemeinten Sinn in der Tradition einer verstehenden Sozialwissenschaft von Max Weber begeben, nehmen lediglich die erstgenannte Dimension in den Blick. Dies ist vor allem die in der Tradition der phänomenologischen Soziologie von Alfred Schütz stehende Strömung, die durch die sogen. Wissenssoziologie von Berger/Luckmann beeinflusst ist und als „hermeneutische Wissenssoziologie“ oder „wissenssoziologische Hermeneutik“ bezeichnet wird.

Im Unterschied zur Soziologie von Schütz und auch zur hermeneutischen Wissenssoziologie kann in der Wissenssoziologie von Harold Garfinkel (1961) und auch in der Wissenssoziologie in der Tradition von Mannheim, die ich als praxeologische Wissenssoziologie (Bohnsack 2006b) bezeichne, die Konstruktion von Motiven im Sinne der Unterstellung eines subjektiv gemeinten Sinns zwar der zentrale *Gegenstand* wissenschaftlichen Interpretierens sein, nicht aber deren *Methode* und auch nicht die Grundlage für die Entwicklung einer sozialwissenschaftlichen Methode.

Die phänomenologische Soziologie samt der Ansätze, die in dieser Tradition verwurzelt sind, wie auch die ethnografischen Ansätze verbleiben im Wesentlichen auf der *ikonografischen Ebene* und somit auf der Ebene (der Rekonstruktion) von Common Sense-Konstruktionen. Sie vermögen zwar die Common Sense-Theorien, die Konstruktionen ersten Grades im Sinne von Alfred Schütz (1971), und auch die ‚Methode‘ dieser Konstruktionen, wie sie auf den subjektiv gemeinten Sinn gerichtet ist, zu *beschreiben*, *nachzuzeichnen*. Sie stellen diese Theorien und Methoden aber nicht in Frage bzw. zeigen sie deren Grenzen nicht auf und bleiben in diesem Sinne deskriptivistisch und in gewisser Weise unkritisch gegenüber dem Common Sense.

Demgegenüber konvergieren so unterschiedliche methodologische Positionen wie diejenige der Wissens- und Kultursoziologie von Bourdieu, des Konstruktivismus im Sinne von Luhmann und eben auch der Wissenssoziologie Mannheims in dem Punkt, dass eine Beobachterhaltung, welche ihre Wissenschaftlichkeit unter Beweis stellen will, nicht lediglich die Common Sense-Theorien nachzeichnen

kann. Vielmehr muss sie die Differenz ihrer eigenen AnalyseEinstellung zum Common Sense und somit die methodologische Differenz der eigenen Theorien ihm gegenüber, also den „Bruch mit den Vorannahmen des *common sense*“ (Bourdieu 1996: 278) bestimmen und methodisch-forschungspraktisch realisieren können.<sup>84</sup>

#### 5.4.4 Die AnalyseEinstellung auf das Performative

Methodologisches und methodisches Prinzip der dokumentarischen Interpretation ist es, auf unterschiedlichen Dimensionen desselben Falles zugleich anzusetzen, sodass diese unterschiedlichen Interpretationsdimensionen einander wechselseitig zu validieren vermögen. In Falle der Interpretation der Bewegungen der abgebildeten Bildproduzent(inn)en sind hier vor allem die unterschiedlichen Ebenen der Gebärden, der operativen Handlungen und der institutionalisierten Handlungen gemeint, die in der empirischen Analyse in ihrem Bezug aufeinander interpretiert werden.

Auf allen diesen Ebenen lässt sich die Frage nach dem *Wie* stellen, nach dem *modus operandi* der Herstellung. Diese AnalyseEinstellung der dokumentarischen Methode lässt sich auch als die AnalyseEinstellung auf das *Performative* bezeichnen. Im Sinne dieser Methode, wie sie in der praxeologischen Wissenssoziologie fundiert ist, wie auch im Sinne von Panofsky und Bourdieu, ist Performativität nicht primär das Produkt einer Inszenierung oder Selbststilisierung, also nicht primär ein *Darstellungsmodus*, eine Metapher des Theatralischen. Vielmehr bezeichnet sie primär eine *Existenzweise*, zu der die Selbststilisierung, der „intendierte Ausdruckssinn“, wie Mannheim (1964a) dies nennt, ggf. noch hinzutritt.

Möglicherweise muss die Aneignung, die Inkorporierung des habituellen Handelns, auch durch eine solche Selbststilisierung hindurch: Sie ist dann zunächst ein intentionales Produkt und wird erst allmählich routinisiert und habitualisiert. Der weitaus überwiegende Teil unserer Alltagspraxis wird jedoch bereits als habitualisierter *modus operandi* angeeignet. Wie Bourdieu (1982: 740) mit Bezug auf Leibniz behauptet, sind wir Menschen „in Dreiviertel unserer Handlungen Automaten“. Eine Selbststilisierung oder Selbstinszenierung markiert dann immer schon einen gewissen Verlust bzw. Mangel an Authentizität des Habitus, des habituellen Stils. Denn – wie es bei Erving Goffman (1975: 290) heißt – „Stil kommt uns unecht vor, wenn er absichtsvoll ist.“

Die Differenzierung von *Existenzweise* und *Darstellungsmodus* und die Rekonstruktion ihrer *Doppelstruktur* ist charakteristisch für die *praxeologische* AnalyseEinstellung, wie sie der dokumentarischen Methode entspricht, und unterscheidet

84 In eine ähnliche Richtung zielen wohl die Überlegungen von William Mitchell (2002: 171) zu den „Visual Studies“, wenn er deren Programm der Erforschung der visuellen Kultur formuliert als „the overcoming of what has been called the natural attitude“. Es geht darum, „to overcome the veil of familiarity and self-evidence that surrounds the experience of seeing, and to turn it into a problem for analysis, a mystery to be unraveled“ (a.a.O.: 166). Auch bei Mitchell geht es somit um einen Bruch mit Selbstverständlichkeiten und Vorannahmen der Konstruktionen auf der ikonografischen Ebene – wenn auch wohl nicht in jene *praxeologische* Richtung, wie sie für Bourdieu und die dokumentarische Methode charakteristisch ist.

diese von der AnalyseEinstellung auf das Performative, die in der Theaterwissenschaft und in der historischen Anthropologie eingenommen wird (vgl. Wulf/Zirfas 2001). Eine derartige analytische Differenzierung und eine entsprechende empirische Rekonstruktion dieser Art der Doppelstruktur alltäglichen Handelns finden wir dort nicht.<sup>85</sup>

#### 5.4.5 Die unterschiedlichen Ebenen dokumentarischer Interpretation und die Primordialität der vor-ikonografischen Ebene

Zwar ist es, wie dargelegt, methodisches Prinzip einer Analyse des *modus operandi*, der performativen Struktur nach Art der dokumentarischen Interpretation, auf den *unterschiedlichen* Ebenen *desselben* Falles zugleich anzusetzen (also u.a. der Ebene von Gebärden, operativen Handlungen und institutionalisierten Handlungen) und den Versuch zu unternehmen, Homologien zwischen Ebenen herauszuarbeiten, die dann einander wechselseitig zu validieren vermögen.

Allerdings stellt im Bereich der Bild- und Filminterpretation die *vor-ikonografische* Ebene die *primordiale* Ebene dar: zum einen, weil sie durch das textlich-narrative Vorwissen weniger vorstrukturiert ist und somit den elementarsten Zugang zu jener dokumentarischen Sinnebene, zur performativen Struktur, erschließt, die dem Bild eigentümlich ist (die Art und Weise, wie jemand den ‚Rumpf beugt‘, also der *modus operandi*, erscheint als *Dokument* für das *Wesen* oder den *Habitus* des Akteurs, bspw. für seine ‚Unsicherheit‘ oder ‚Gebrechlichkeit‘). Zum anderen vermittelt uns diese Ebene aber auch im Hinblick auf die *Konstruktion von Motiven*, die zweckrationale Konstruktionen der Unterstellung von Um-zu-Motiven (Beispiel: Er beugt den Rumpf, um sich zu setzen) einen validen Zugang. Denn die *Validität* der Zuschreibung von Motiven, bei denen die Enaktierung des Entwurfes am Handlungsvollzug selbst beobachtbar ist (operative Handlung), ist, wie dargelegt, unvergleichlich viel größer als die Validität von Zuschreibungen auf der Ebene institutionalisierter Erwartungserwartungen.

Es hat somit einen doppelten Grund, dass die „Leibsprache“ – wenn wir Mannheim (1964a: 136) folgen – auch in besonderer Weise dazu geeignet ist, uns einen Zugang zur „Struktur der atheoretischen Sinngebilde“ zu eröffnen: „Um dies klarer zu sehen, verweisen wir auf die Leibsprache, insbesondere auf die *Physiognomik*, die viel geeigneter ist, die Struktur der Kulturgebilde zu verraten als der vornehmliche Träger des theoretischen Sinnes, die Wortsprache.“

Die Bild- und Filminterpretation und somit auch die sozialwissenschaftlichen, aber auch die nicht-sozialwissenschaftlichen Disziplinen, die auf Bildinterpretation

<sup>85</sup> Neben der analytischen Differenzierung von Existenzweise und Darstellungsmodus finden wir im Kontext der dokumentarischen Methode noch weitere wichtige Unterscheidungen: Wir differenzieren zwischen der *Performanz*, der im *performatorischen* Vollzug ihrer aktuellen Durchführung oder Herstellung aufgezeichneten verbalen und nonverbalen Äußerungen, einerseits und der *Performativität*, der *performativen Struktur*, welche sich in einer derartigen Herstellung oder Darstellung für den Beobachter dokumentiert, andererseits. Die performative Struktur oder Performativität ist die in dokumentarischer Interpretation rekonstruierbare *Prozessstruktur*, der *modus operandi* oder *Habitus* im Sinne von Bourdieu (1976) und Panofsky (1989).

angewiesen sind, stehen allerdings vor dem Problem, dass eine Beschreibungssprache auf der *vor-ikonografischen* Ebene weitgehend fehlt. Der Filmtheoretiker Belá Balázs, hatte bereits 1924 betont: „Noch einige Jahre guter Filmkunst und die Gelehrten werden vielleicht daraufkommen, dass man mit Hilfe des Kinematographen das Lexikon der Gebärden und der Mienen zusammenstellen müsste wie das Lexikon der Worte.“ (Balázs 2001a: 19). Für die sozialwissenschaftliche Bild- und Filminterpretation wäre nicht ein Lexikon, aber eine Theorie der (vor-ikonografischen) Gebärden und Mimik und ein Categoriesystem zu ihrer Beschreibung wesentliche Voraussetzung für die Methodik der Bild- und Filminterpretation.

## 5.5 Fotogramm und Simultaneität

Nicht zuletzt die Bedeutung der Gebärden und der Mimik haben Roland Barthes veranlasst, dem Fotogramm, also dem Standbild, dem „still“, eine zentrale Bedeutung für die Semiotik des Filmes einzuräumen. Roland Barthes (1990: 64) erläutert dies insbesondere am Beispiel von Eisensteins Panzerkreuzer Potemkin (vgl. auch Kap. 3.4). Nach Barthes (1990: 64) „lässt sich in gewissem Maß (...) das Filmische paradoxerweise nicht im Film ‚am rechten Ort‘, ‚in der Bewegung‘, ‚in natura‘ erfassen, sondern bisher nur in einem wichtigen Artefakt, im Fotogramm.“ Barthes unterscheidet also zwischen dem „Filmischen“ und dem „Film“ und schreibt: „Das Filmische ist dasjenige im Film, das sich nicht beschreiben lässt (...). Das Filmische beginnt erst dort, wo die Sprache und die gegliederte Metasprache aussetzen“ (Barthes 1990: 63).

### 5.5.1 Die Bedeutung des Fotogramms für die Analyse von Gebärden oder Kinemorphemen

Die beiden wohl prominentesten Semiotiker, Roland Barthes und Umberto Eco, nehmen in geradezu konträrer Weise Stellung zur Bedeutung des Fotogramms. Denn Ecos Einschätzung des Fotogramms ist – ohne dass er auf Barthes explizit Bezug nehmen würde – von ganz anderer Art. In seinen Ausführungen zum Film geht Eco (1994: 250ff.) der Frage nach, in welcher Weise Gebärden durch den „kinematographischen Code“ repräsentiert werden können. Er bezeichnet Gebärden mit Birdwhistell (u.a. 1952 u. 1970) als „Kinemorpheme“. Eco (1994: 259) nennt als Beispiel für derartige Kinemorpheme „typische Kopfbewegungen wie das Zeichen ‚nein‘ und das Zeichen ‚ja‘“.

Wenn der Film in Standbilder, also Fotogramme, zerlegt wird, gehen nach Eco diese Signifikanz (Bejahungen und Verneinungen) verloren: „Wenn ich zwei typische Kopfbewegungen wie das Zeichen ‚nein‘ und das Zeichen ‚ja‘ in soundsoviele Photogramme zerlege, dann finde ich soundsoviele verschiedene Positionen, die ich nicht als Positionen der Kinemorpheme ‚nein‘ und ‚ja‘ identifizieren kann.“ Eco (a.a.O.: 258) betont: Die elementaren ikonischen Zeichen, „die Icone erzeugen durch eine diachronische Bewegung Kinemorpheme.“ Die ikonischen Zeichen

selbst ergeben nach Eco (1994: 258) allerdings keinen Sinn, „so daß viele kleine Bewegungseinheiten ohne Sinn verschiedene gestische Einheiten mit Sinn bilden können.“

Nun soll keineswegs bestritten werden, dass Gebärden Bewegungsabläufe und somit diochronischer Art sind. Birdwhistell hat jedoch in seinen bahnbrechenden empirischen Analysen das besondere – insbesondere auch zeitliche – Verhältnis der Gebärden, der Kinemorpheme<sup>86</sup> zu den sie konstituierenden elementaren Einheiten, den „Kinemen“ oder „Kinen“<sup>87</sup> in einer Art und Weise herausgearbeitet, aus der sich andere Schlussfolgerungen ergeben: Im Sinne von Birdwhistell (1952: 15ff.) konstituieren sich Kinemorpheme bzw. Kinemorphe als Muster („patterns“) von Kinemen bzw. Kinen. Der Sinngehalt von Kinemen erschließt sich – wie ganz generell derjenige aller von Birdwhistell untersuchten Bewegungseinheiten – erst aus dem Kontext heraus: „Keine Bewegungseinheit trägt per se Sinngehalte in sich. Sinn erwächst aus dem Kontext.“ (Birdwhistell 1952: 10).<sup>88</sup>

So lässt sich – als ein Beispiel – das Kinemorphem „Zwinkern“ („wink“) erst aus dem Kontext unterschiedlicher Kineme erschließen, bspw. der folgenden:

„a) Das rechte Auge ist geschlossen, während das linke geöffnet bleibt. – b) Die Mundhaltung ist „normal“. – c) Die Nasenspitze ist eingedrückt (Kaninchennase). – d) Der linke Augenhöhlenrand ist schräg („squinted“).“ (Birdwhistell 1952: 19).<sup>89</sup>

Wie das Beispiel deutlich macht, entfalten die Kinemorpheme oder Gebärden, hier: das „Zwinkern“, ihre Signifikanz nur dann, wenn ihre Konstituenten, die elementaren Bewegungseinheiten der Kineme, *synchron* oder *simultan* zum Ausdruck kommen. Die Relation der Kineme zueinander, der Modus ihrer Kontextuierung, der sie zu einem Kinemorphem werden lässt, ist also derjenige der Simultaneität, nicht – wie in dem Beispiel, an dem Eco den Zusammenhang von Kinemorphemen

86 Birdwhistell (1952: 17) spricht bei den Kinemorphemen auch von Gebärden („gestures“). Eine Variante von Gebärden sind nach Birdwhistell diejenigen, die in der Kultur bereits explizit als solche bezeichnet werden: bspw.: Hände schütteln, Nicken, Augenbrauen hochziehen („frown“).

87 Bei Birdwhistell (1952: 76) wird noch einmal unterschieden zwischen „Kinemorphem“ und „Kinemorph“. Letzteres ist die in einer Beobachtungssituation in ihrer konkreten Ausprägung beobachtbare Gebärde (bspw. das „Lächeln“ oder „Zwinkern“). Das auf der Grundlage des Vergleichs unterschiedlicher Ausprägungen abstrahierbare und somit – in einem Akt der Typenbildung sozusagen – verallgemeinerbare (Grund-)Muster, also bspw. das generelle Muster „Zwinkern“, nennt er Kinemorphem. Die Elemente, die elementaren Beobachtungseinheiten, aus denen sich Kinemorphe bzw. Kinemorpheme konstituieren, sind die „Kine“ oder „Kineme“ (bspw. die Augen- und Mundhaltung etc. beim Lächeln oder Zwinkern). Während das „Kin“ die in einer Beobachtungssituation empirisch gegebene kleinste unterscheidbare Einheit von Körperbewegungen darstellt („the least unit of body motion with discriminational meaning“; Birdwhistell 1952: 76), ist das „Kinem“ (analog zum Kinemorphem) die auf der Grundlage des Vergleichs unterschiedlicher Situationen bereits abstrahierte, also verallgemeinerte und somit typenhaft erfasste (Bewegungs-) Einheit. – Birdwhistell operiert also in elementarer Weise auf der Grundlage der komparativen Analyse.

88 Erving Goffman (1990) hat jene amerikanische Tradition der Interaktionsanalyse, die sich schon früh auf Filmaufzeichnungen gestützt hat – von Bateson und Mead über Birdwhistell bis zur Konversationsanalyse – und die ihre Ursprünge wesentlich in der Palo Alto-Schule hat, als Context Analysis bezeichnet. Allerdings wird die Differenz zwischen simultaner und sequenzieller Kontextuierung bei ihm nicht herausgearbeitet.

89 Es handelt sich um eine Variante, eine Ausprägung eines Kinemorphems, also um ein Kinemorph. In einer Versuchsreihe hat Birdwhistell (1952) Testpersonen unterschiedliche solcher Varianten vorgelegt: bspw. kann auch das linke Auge geschlossen und das rechte geöffnet sein.

und elementaren ikonischen Zeichen zu erläutern versucht – derjenige der Diachronizität oder Sequenzialität. Das bedeutet, dass Kinemorpheme sich nicht durch die Sequenzierung aufeinander folgender Fotogramme in ihrem Sinngehalt erschließen lassen, sondern zunächst allein durch die einzelnen Fotogramme, auf denen die Konstellation oder Komposition unterschiedlicher Kineme in ihrer Simultaneität sichtbar wird. Entscheidend ist also der simultane Kontext – im Unterschied zum sequentiellen Kontext.

Somit gilt es also auch, den Begriff des *Kontexts* genauer zu fassen. Einem der Wegbereiter der Videoanalyse im Bereich der qualitativen Methoden, Christian Heath (1997: 187), ist ohne weiteres zuzustimmen, wenn er betont, dass die Interpretation von Gebärden und körperlichen Ausdrucksformen von ihrer Kontextuierung abhängig, also kontextgebunden, sei: „Unglücklicherweise findet sich die weit verbreitete Tendenz unter den Forschern non-verbaler Kommunikation, anzunehmen, dass die Bedeutung einer Geste unentwirrbar an ihre physische Form gebunden ist, statt an den Kontext, in dem sie auftritt“.

Aber bei Heath wird die Kontextuierung dann – in Orientierung an der Kontextuierung sprachlicher Äußerungen – einseitig als eine sequenzielle verstanden: „Die emergente und sequentielle Organisation der Interaktion ist auch von Relevanz für die Betrachtung der kontextuellen oder ‚situativen‘ („in situ“) Bedeutung visuellen Verhaltens und der physischen Eigenschaften menschlicher Umwelten“ (Heath 1997: 187). Die *simultane* Kontextuierung gerät somit bereits bei diesem Vorreiter der sozialwissenschaftlichen Videoanalyse gar nicht in den Blick. Und dies ist in der aktuellen sozialwissenschaftlichen Videoanalyse insgesamt weit verbreitet.

Die Interpretation von Gebärden macht die Rekonstruktion von elementaren Bewegungseinheiten, den Kinen oder Kinemen, in ihrem simultanen und ganzheitlichen Zusammenspiel erforderlich: „Kinemorphologie ist die systematische Analyse der Muster von Kinen. Kein Kin steht jemals allein“ (Birdwhistell 1952: 15). Wenn wir einen interpretativen Zugang zu komplexen simultanen Relationen in einer Weise gewinnen wollen, welcher intersubjektiv überprüfbar und das bedeutet auch: immer wieder reproduzierbar ist, so ist die Analyse des Stand-Bildes, des Fotogramms notwendige Voraussetzung. Denn das Bild ist nach Imdahl (1996a: 23) eine „in seiner Ganzheitlichkeit invariable und notwendige, das heißt alles auf alles und alles aufs Ganze beziehende Simultanstruktur (...). Das Ganze ist von vornherein in Totalpräsenz gegeben.“

Im Zuge der Interpretation von Gebärden vollzieht sich die Rekonstruktion des simultanen Zusammenspiels der elementaren Bewegungseinheiten zunächst innerhalb abgegrenzter Bereiche des Körpers. Birdwhistell (1952: 17) unterscheidet acht (Körper-) Areale (u.a. den gesamten Kopf, das Gesicht, den Schulter-Arm-Bereich, die Hand etc.). Idealerweise umfasst die Rekonstruktion des Zusammenspiels aber den gesamten Körper, orientiert sich an der „whole body“ conception“ (Birdwhistell 1952: 8) und vollzieht sich nach dem Prinzip der Gestalt-Wahrnehmung: „perception takes place in Gestalten.“ (a.a.O.: 18).

Eine derartige Gestaltwahrnehmung entspricht in ihrer Architektur der dokumentarischen oder ikonologischen Sinninterpretation. Indem Umberto Eco die Konstitution von Signifikanzen einseitig an das Modell der Sequenzierung bindet, orientiert er sich jedoch eher an der Architektur ikonografischer Sinnkonstruktionen

mit ihrer instrumentellen oder zweckrationalen Grundstruktur: Elementare Bewegungseinheiten werden enacted, um zu „bejahen“, um zu „zwickern“. Hier wird, wie auch an anderen Stellen, deutlich, dass Eco nicht – zumindest nicht systematisch – über einen Sinnbegriff jenseits bzw. unterhalb der ikonografischen oder konnotativen Ebene verfügt, also jenseits oder unterhalb des „Bedeutungssinns“ bei Panofsky (Zu diesen Parallelen zwischen der konnotativen Sinnenebene im Bereich der Semiotik und der ikonografischen Sinnenebene bei Panofsky sowie zu den Grenzen der semiotischen Interpretation siehe auch Kap. 3).

### 5.5.2 Grenzen der Interpretation von Fotogrammen

Es ist also der Vorteil und die Chance des Fotogramms, dass dieses uns nicht an die Rekonstruktion von ikonografischen Einheiten bindet, also nicht auf der Ebene von Handlungen verbleibt, sondern uns vielmehr den Zugang zu den „Kinomorphemen“ im Sinne von Birdwhistell eröffnet – und als dessen Grundlage – den Zugang zum simultanen Zusammenspiel von Kinemen und somit zum *modus operandi* der Herstellung der Kinomorpheme und Gebärden.<sup>90</sup>

Bei Birdwhistell (1970: 147) selbst wird allerdings auch deutlich, dass eine Gebärdenanalyse, die sich *ausschließlich* auf die Interpretation von Fotogrammen („still photos“) stützen will, Probleme mit sich bringt und somit die Rekonstruktion der Bilderfolge notwendig wird: „Die Bilderfolge ergänzt die jeweilige Stellung („position“) um die Bewegung und reduziert nach meiner Ansicht die Wahrscheinlichkeit, dass ein im Übergang fotografiertes Zeitpunkts fälschlicherweise für eine Stellung gehalten wird. Allerdings konnten Byers und Hymen mir demonstrieren, dass in den Händen eines trainierten und sensitiven Beobachters, der sorgfältig die Bedingungen der Fotografie festhält, ein einziges Standfoto wertvoller ist als tausend Meter Film, welche von einem Filmemacher aufgenommen werden, der einen Film dreht und der nicht darin geübt ist, menschliche Ereignisse und Aktivitäten aufzuzeichnen.“

Birdwhistell weist hier also darauf hin, dass eine isolierte Betrachtung des Fotogramms uns das, was lediglich einen *Übergang* zwischen unterschiedlichen Bewegungs-Positionen (bspw. Gebärden bzw. Kinemorphemen) darstellt, fälschlicherweise selbst als eine Bewegungs-Position erscheinen lassen könnte. Somit wird deutlich, dass – schon allein im Hinblick auf die Gebärdenanalyse, die Analyse der Bewegungen der abgebildeten Bildproduzent(inn)en – die Interpretation von Fotogrammen allein nicht die Grundlage der Film- und Videointerpretation sein kann, ganz zu schweigen von einer Analyse, die dem Film insgesamt, also

<sup>90</sup> Die Bedeutung des Einzelbildes wird auch in der Filmwissenschaft betont. Allerdings wird hier typischerweise auf die im Einzelbild repräsentierte Gestaltungsleistung der *abbildenden* Bildproduzent(inn)en abgehoben: „Die Feststellung, dass Filme und Fernsehsendungen aus bewegten Bildern bestehen, ist nicht ganz richtig, denn tatsächlich setzen sie sich aus unbeweglichen Einzelbildern zusammen. Jedes einzelne Film- oder Fernsehbild bildet nicht nur etwas ab und stellt etwas dar, sondern ist in seiner ganz spezifischen Art und Weise gestaltet.“ (Mikos 2003: 51). Wenn es darum geht, diese Gestaltungsleistung systematisch zu erfassen und sie in ihrer Bedeutung für eine tiefer gehende Semantik des Filmes zu erschließen, steht die Filminterpretation allerdings wohl noch am Anfang.

auch den Bewegungen der abbildenden Bildproduzent(inn)en gerecht zu werden versucht.

Bereits eine videobasierte Gebärdenanalyse setzt also die Interpretation sowohl von *Simultaneität* als auch die Interpretation des *sequentiellen* Bewegungsablaufs voraus. Im letzteren Fall ist, wie auch Birdwhistell (1970) betont, die Zeitlupenanalyse sehr wichtig.

Die technische Bearbeitung des Films am Computer ermöglicht mit geringem Aufwand die Interpretation der sequenziellen Struktur sowohl auf der Grundlage der Zeitlupenbetrachtung dieses Ausschnitts als auch einer Betrachtung mit Hilfe einer Programmfunktion, welche die Sequenz in kleine Schritte (in Sekundenbruchteilen) zerlegt (bspw. Programmfunktion „Schritt vorwärts“ in Power DVD) oder mit Hilfe einer „Diashow“ (siehe auch Kap. 6.3.5).

Die sequentielle Betrachtung auf der Grundlage der Zeitlupe und des Abspielen in kleinen Schritten ist auch notwendig, um jene Momente identifizieren zu können, die es aufgrund der besonderen Dichte der Kineme, der Elemente von Gebärden, einer genaueren Rekonstruktion der Simultanstruktur dieser Kineme zu unterziehen gilt (siehe auch Kap. 6.3.5). Darüber hinaus vermittelt uns die Rekonstruktion der Gebärde in ihrem sequentiellen Ablauf den Einblick in die Motivstruktur des Handelns (Um-zu-Motive): zunächst der operativen Handlung (A nimmt den Hut) und schließlich auch der institutionalisierten Handlung (A begrüßt B) auf der ikonografischen Ebene.

Auf dieser Basis erhalten wir auch einen Zugang zur Narration des Films ebenso wie zu Formen nicht-narrativer filmischer Organisation (siehe dazu Bostnar/Pabst/Wulff 2002: 37ff.). Die sequentielle Abfolge der Einzelbilder und deren spezifische sequentielle Organisation in der Montage führt selbstverständlich nicht nur zur Darstellung chronologisch sequenzierter Handlungs- und Ereignisverläufe und somit zur Narration, sondern auch zur Darstellung zirkulärer Handlungs- und Ereignisverläufe wie etwa Beschreibungen.<sup>91</sup>

Narrationen wie auch andere durch die Montage hergestellte oder mit-hergestellte filmische Darstellungsformen können ebenso dokumentarisch interpretiert werden wie die Fotogramme. Wie bereits angesprochen, ist es von besonderer Bedeutung für die dokumentarische Interpretation und für deren Gültigkeit, dass sie auf den *unterschiedlichen* Dimensionen desselben Falles (der vor-ikonografischen wie der ikonografischen) zugleich ansetzen, diese zunächst unabhängig voneinander interpretieren kann, sodass diese Interpretationen auf den unterschiedlichen Ebenen daraufhin überprüft werden können, ob sie einander *wechselseitig zu validieren* vermögen.

Indem die Interpretation der Fotogramme zunächst getrennt von derjenigen der durch die Montage hergestellten oder mit hergestellten filmischen Darstellungsformen vollzogen wird, eröffnet sich dann auch hier die Möglichkeit, Homologien herauszuarbeiten und die beiden Interpretationsebenen wechselseitig zu validieren. Allerdings erscheint es wesentlich, dass die Interpretation der Fotogramme (und

<sup>91</sup> Und die sequentielle Organisation in der Montage hat auch nicht allein die Funktion der Strukturierung von Zeitlichkeit, sondern auch von Räumlichkeit. Dies wird in Kap. 5.6.3 genauer dargelegt.

damit der vor-ikonografischen Ebene und der Bildlichkeit im eigentlichen Sinne) der Analyse der Film-Narration und der durch die Montage hergestellten Darstellungsformen nicht untergeordnet, sondern in gleichberechtigter Weise einbezogen werden soll.

Wie in Kapitel 3 bereits zitiert, lässt sich im Verständnis von Roland Barthes (1990: 64) das „Filmische“, also das, was den Film in seiner Eigenlogik und eigensinnigen Semantik betrifft, „bisher nur in einem wichtigen Artefakt, im Fotogramm“ erfassen. Er nimmt damit eine radikale Position ein. Mit der Reduktion auf die Komposition des Fotogramms weist er der *Relation* zwischen den Fotogrammen bzw. Einzelbildern eine reduzierte Bedeutung im Hinblick auf das zu, was die Eigenlogik und Eigensinnigkeit des Films ausmacht. Damit erhält auch der andere wesentliche Bereich der Komposition des Films, nämlich die Komposition der Bildsequenzen, die *Montage* (dazu weiter unten) eine eher marginale Bedeutung. Auch Kracauer als Klassiker der Filmwissenschaft tendiert zu einer Position, welche, von der Fotografie ausgehend, deren Unterschiede zum Film in den Hintergrund treten lässt.

Innerhalb der Klassiker der Filmwissenschaft wird die gegenteilige Position durch Wsewolod Pudowkin (2003a: 72) repräsentiert, bei dem es heißt: „Die Montage ist das eigentliche schöpferische Moment, kraft dessen aus den leblosen Fotografien (den einzelnen Filmbildchen) die lebendige filmische Einheit geschaffen wird.“ Zu diesem Pol tendieren auch einige zeitgenössische filmwissenschaftliche oder filmtheoretische Analysen, in denen eine radikale Trennung von Fotografie und Film vorgenommen bzw. das Fotografische am Film als weniger relevant erachtet wird. Demgegenüber sollen hier in theoretischer wie in methodisch-forschungspraktischer Hinsicht sowohl die Differenzen wie auch die Gemeinsamkeiten von Fotografie und Film Berücksichtigung finden.

### 5.5.3. Zur Rekonstruktion von Einstellung und Perspektivität im Fotogramm

Die Betrachtungen von Kracauer und Panofsky stellen, wie dargelegt, ganz wesentlich den Beitrag bzw. den Habitus der *abgebildeten* Bildproduzent(inn)en im Film bzw. im Fotogramm ins Zentrum, obwohl sich bei beiden auch wesentliche Hinweise auf die Leistung von Einstellungswechsel und Montage finden, bei Panofsky (1999: 25) insbesondere unter dem Aspekt der „Dynamisierung des Raumes“ (siehe dazu die Ausführungen in 4.6.2). Aber auch die Komposition des Fotogramms ist durch den Habitus der *abbildenden* Bildproduzent(inn)en geprägt, wie er mit dem Begriff der „*Einstellung*“ gefasst wird: „das Filmbild formuliert damit umgekehrt eine innere Haltung zum Abgebildeten“ (Hickethier 1996: 57). Allerdings erscheint es terminologisch unglücklich, aber auch typisch für die neuere filmwissenschaftliche Literatur, nur den Beitrag der *abbildenden* Bildproduzenten als „Filmbild“ zu bezeichnen. Vielmehr setzt sich das Filmbild eben aus den Beiträgen beider – der abbildenden wie der abgebildeten Bildproduzent(inn)en – zusammen. Auch die Herstellung von Simultaneität ist nicht nur eine Leistung der abgebildeten, sondern ebenso auch der abbildenden Bildproduzent(inn)en.

Die Rekonstruktion der *Einstellung* bzw. der Einstellungsgröße (siehe dazu die Übersicht im Anhang: Kap. 7.4) „definiert sich an der Größe des abgebildeten Menschen im Verhältnis zur Bildgrenze“ (Hickethier 1996: 58)<sup>92</sup> und umfasst die Detail-, Groß- und Naheinstellung, halbnah, amerikanische, halbtotale, totale und Weit- oder Panorama-Einstellung (Hickethier 1996: 58f.; Borstnar/Pabst/Wulff 2002: 91). Diese ist zunächst eine Angelegenheit der Interpretation des Fotogramms und seiner *formalen Komposition* im Sinne von Imdahl (siehe dazu auch Kap. 3).

Neben der Einstellungsgröße erschließt sich auch der *Kamerastandort* auf dem Wege über das Fotogramm. In der filmwissenschaftlichen Literatur wird hier im Wesentlichen zwischen der Normal-, Auf- und Untersicht unterschieden (vgl. Hickethier 1996: 61f.). Mit dem Begriff des „Bildraums“ (Hickethier 1996: 71) wird in der filmwissenschaftlichen Literatur z. T. auf das Bezug genommen, was als Perspektivität bereits bei der dokumentarischen Einzelbildinterpretation herausgearbeitet wurde: die Rekonstruktion des perspektivischen Zentrums, des Fluchtpunkts („zentralperspektivischer Punkt“; Hickethier 1996: 71) und des perspektivischen Modus: Frontal- und Übereckperspektive sowie Aufsicht (Luftperspektive) (vgl. auch Kap. 7.4: Anhang).

Während die bisherigen Ausführungen zur Einstellung auch für die Fotografie relevant sind, kommt beim Film noch die „Kamerabewegung“ hinzu (die allerdings auf der Grundlage der einzelnen Fotogramme nicht zu erschließen ist). Grundsätzlich ist hier die Kamerabewegung auf einem unbeweglichen Stativ von derjenigen im Raum zu unterscheiden (vgl. u.a. Borstnar/Pabst/Wulff 2002: 96).

Ein weiteres Element der Komposition des Fotogramms ist die „Bildkomposition“ im engeren Sinne oder „Mise en Scène“ (Borstnar/Pabst/Wulff 2002: 97ff; siehe auch Keppler 2006: 117 sowie Marotzki/Schäfer 2006: 64), die Art und Weise, wie Personen und Objekte im Bildraum angeordnet, beleuchtet sind etc. Es handelt sich „um Gestaltungsmittel, die nicht eigentlich das fertige Bild selbst, sondern den Raum vor der Kamera im Aufnahmeprozess betreffen“ (Borstnar/Pabst/Wulff 2002: 99). Obschon die *Mise en Scène* somit das Produkt der Regie und der Akteure vor der Kamera, der Schauspieler darstellt, ist nach Ansicht der Autoren diese „im Grunde genommen eine noch ganz und gar vorfilmische Konstruktion von Weltzuständen“ (ebenda). Dies erscheint charakteristisch für die Perspektive der Filmwissenschaftler, für die der Bildraum erst als Leistung der Montage wirklich Bedeutung gewinnt.

Den durch die Montage geschaffenen Raum bezeichnet Hickethier (1996: 82f.) als den eigentlichen „filmischen Raum“ (siehe dazu genauer im nächsten Kapitel). Er unterscheidet ihn von dem „fotografierten optischen Raum“, den wir mit der Interpretation des Fotogramms rekonstruieren und den wir seinerseits dann noch ein-

92 Es stellt sich allerdings die Frage, ob es nicht richtiger ist, dass die jeweilige Einstellungsgröße sich in Relation zum „Establishing Shot“ oder „Master Shot“ bestimmt (ansonsten wäre die Einstellungsgröße ja dann auch gar nicht zu klären, wenn keine Personen abgelichtet werden). Der „Establishing Shot“ wird von James Monaco (2003: 58) definiert als „Einstellung zu Beginn einer Sequenz, die einen allgemeinen Überblick über Lokalität, Personal und Situation gibt“. Den „Master Shot“ definiert Monaco (2003: 103) in ähnlicher Weise als „meist in der Totale oder Halbtotale durchgedrehte Einstellung, die beim Schnitt als Grundlage für eine Szene genommen werden kann.“

mal von der Raumkonstitution im Sinne der „Mise en Scène“ unterscheiden können.

Noch weniger Beachtung als die Konstitution des fotografierten optischen Raums auf der Grundlage der gewählten Perspektivität findet in der Filmwissenschaft die Rekonstruktion der planimetrischen Komposition des Bildes, also die Rekonstruktion der Komposition der *Bildfläche*. Hierauf wird lediglich mit der Kategorie der „Kadrierung“ des „Bildfeldes“, also mit der Wahl des Bildausschnitts Bezug genommen. Während Deleuze (1997a: 35) die Bedeutung der Kadrierung für das Bild als ein „relativ und künstlich geschlossenes System“ theoretisch hervorhebt, fehlt aber in der Forschungspraxis der Filmwissenschaften die genaue Rekonstruktion der sich aus der Selektivität des Ausschnitts ergebenden Konsequenzen für die (formale) Komposition dieses Ausschnitts und der semantischen Bedeutung dieser Selektivität.

Demgegenüber stellt für den Kunsthistoriker Max Imdahl (vgl. dazu Kap. 3) gerade die planimetrische Komposition, wie sie u. a. durch die Wahl des Ausschnitts hergestellt wird, die wesentliche formale Grundlage dar, um den Weg zur Rekonstruktion des Bildes in seiner ihm eigentümlichen Logik und Eigensinnigkeit zu finden. In diesem Sinne gehört die Rekonstruktion der Planimetrie im Sinne einer dokumentarischen Filminterpretation zu den wesentlichen Dimensionen der Gestaltungsleistung des Films.

## 5.6 Montage, Einstellung und Sequenzialität

Während die Klassiker der Filmwissenschaft Balázs, Panofsky und Kracauer die Leistungen der abbildenden Bildproduzent(inn)en betonen, sie sogar mehr oder weniger in den Vordergrund stellen<sup>93</sup>, finden wir bei den Klassikern auch Positionen wie diejenige von Wsewolod Pudowkin (2003a: 72), der in der Montage „das eigentliche schöpferische Moment“ sieht, „kraft dessen aus den leblosen Fotografien (den einzelnen Filmbildchen) die lebendige filmische Einheit geschaffen wird.“

93 Gerade dort, wo Balázs in meisterhafter Weise die „Mikrophysiognomie“ der filmischen Akteure beschreibt und interpretiert, neigt er allerdings dazu, die Beiträge der abbildenden Bildproduzent(inn)en für die hier erreichte Komplexität des Ausdrucks auszublenden. Ein derartiges meisterhaftes Beispiel seiner Beobachtungskompetenz findet sich in seiner Beschreibung eines „Meisterstücks“ von Asta Nielsen: „Sie soll in einem Film aus intriganten Gründen einen Mann verführen. Sie heuchelt Liebe und spielt Komödie mit sehr überzeugenden Mienen. Doch während dieser Szene verliebt sie sich wirklich in diesen Mann. Ihre Gebärden (dieselben), ihre Mienen (dieselben) werden allmählich aufrichtig. Sie macht genau dasselbe wie vorhin, und es ist nicht zu sehen, woran es doch zu sehen ist, dass sie es anders meint. Aber es wird noch komplizierter. Sie merkt, dass ihr Spießgeselle sie hinter dem Vorhang beobachtet. Nun muß sie diesem glaubhaft machen, dass sie nur Komödie spielt, wie sie vorhin dem anderen glaubhaft machen wollte, dass sie aufrichtig war. Die doppelte Bedeutung ihres Mienenspiels hat sich umgekehrt. Auch jetzt spielt sie Komödie, auch jetzt ist ihr Ausdruck unecht. Bloß ist diesmal das Unehnte unecht geworden. Jetzt lügt sie, dass sie lügt“ (Balázs 2001b: 19 f.). Jenseits dieser großartigen dokumentarischen Interpretation der Mikrophysiognomie stellt sich allerdings die Frage, ob nicht Kameraführung, Perspektivität, Einstellungsgröße und schließlich auch Montage einen wesentlichen Beitrag zu dieser Komplexität und Mehrdimensionalität des Ausdrucks geleistet haben.

Auch in der zeitgenössischen Filmwissenschaft tritt die Bedeutung der Montage in den Vordergrund. So ist für Dieter Wiedemann (2005: 371) „die Montage das bestimmende Element der Filmkunst“. Allerdings ist die Begründung eine etwas andere als die von Pudowkin. Wiedemann argumentiert dahingehend, dass der Film im Bereich der Montage „keine Anleihen bei anderen Künsten“ (ebenda) machen müsse und somit ganz bei sich sei.<sup>94</sup> Hierin sieht Wiedemann auch den Grund dafür, dass die Montagetheorie innerhalb der Filmwissenschaft am weitesten entwickelt ist, wohingegen ansonsten die „Theorie der Bewegtbildgestaltung und damit auch die Analyse bewegter Bilder noch in einem eher embryonalen Stadium“ stecke (ebenda).

Nach Mikos (2003: 207) meint Montage „mehr als den reinen technischen Vorgang, mit dem einzelne Einstellungen zu Szenen zusammengefügt werden. *Montage* meint die Herstellung narrativer und ästhetischer Strukturen durch diesen technischen Vorgang“. Es geht „um die Verkettung der Bilder mit Hilfe der Montage, durch die Bedeutungen entstehen, die in den Bildern selbst nicht enthalten sind“ (Mikos 2003: 101). Beim Fernsehen spricht man statt von Montage auch von „Bildmischung.“ Nach Hickethier (1996: 150) adaptiert das Fernsehen „das erzählende Prinzip des filmischen Einstellungswechsels durch die gleichzeitige Aufnahme eines Geschehens durch mehrere Kameras (...). Die auf diese Weise auf dem Zuschauerbildschirm zustande kommende Mischung ist im Grunde die Simulation einer Montage.“

Da ich nunmehr stärker noch als bisher Bezüge zur Filmwissenschaft im engeren Sinne herstellen möchte, um Anchlüsse für die Erarbeitung einer sozialwissenschaftlichen Methodik der Film- und Videointerpretation zu suchen, erscheint der Versuch einer Begriffsklärung notwendig – zumal bei einigen Begriffen deren Verwendung und Definition auch innerhalb der Filmwissenschaft bei weitem nicht einheitlich ist.

### 5.6.1 Einstellungen, Sequenzen und Szenen

Dies zeigt sich bereits bei der Verwendung des Begriffes der „Einstellung“ („shot“). Dieser meint *zum einen* die elementare Filmeinheit und zwar die nächst größere nach den Einzelbildern. Damit ist „die von der Filmkamera ohne Unterbrechung gefilmte Aufnahme“ gemeint (Beller 2005: 10) bzw. ein „kontinuierlich beleuchtetes, ungeschnittenes Stück Film“ (Monaco 2003: 54). Eine Einstellung ist in dieser Definition also solange gegeben wie kein Schnitt erfolgt.

*Zugleich* wird der Begriff der Einstellung nun aber auch im Sinne der Art der Gestaltung des Filmausschnitts verwendet, wie er durch die Kamera geleistet wird. Diese konstituiert sich zum einen im Sinne der *Einstellungsgröße*. Zum anderen gehört, wie gesagt, die Wahl des *Kamerastandpunkts* dazu. Die Einstellung im Sinne von Einstellungsgröße und Kamerastandpunkt kann, wie oben dargelegt (5.3),

94 Wiedemann (2005: 371) zitiert hier Gerhard Schumm (2004: 167), der seinerseits Stanley Kubrick zitiert (ohne allerdings genauere Quellenangaben zu machen): „Die Montage ist der einzige Vorgang, bei dem der Film keine Anleihen bei anderen Künsten macht. Nur beim Schnitt ist der Film ganz bei sich.“

zunächst auf der Basis der Interpretation des Fotogramms und seiner *formalen Komposition* (im Sinne von Imdahl) rekonstruiert werden.

Während die bisherigen Ausführungen zur Einstellung, also zur Einstellungsgröße und zur Wahl des Kamerastandpunkts, einschließlich des Zooms, auch für die Fotografie relevant und auf der Grundlage des einzelnen Fotogramms zu erschließen sind, gilt dies nicht für die „Kamerabewegung“. Hierzu bedarf es der Rekonstruktion der Relation, der Sequenzierung der Fotogramme. Dies gilt selbstverständlich auch für den *Einstellungswechsel*, also den Wechsel der Einstellungsgröße und des Kamerastandpunkts und schließlich für den *Schnitt*, für die *Montage*. Um allerdings die durch Einstellungswechsel und Montage hergestellte Relation oder Sequenzierung der Einzelbilder genauer bestimmen zu können, bedarf es des Vergleichs der Einzelbilder, also der komparativen Analyse der Fotogramme. Das bedeutet auch, dass ich den Einstellungswechsel und die Leistung der Montage umso genauer rekonstruieren und hinsichtlich ihrer Bedeutung interpretieren kann, je genauer ich die Struktur der Fotogramme vor und nach Einstellungswechsel und Montage erfasst habe.

Zum einen wird der Begriff der *Einstellung* also im Sinne der Art der Gestaltung des Bildausschnitts verwendet, wie er durch die Kamera geleistet wird. Wenn im Folgenden von „Einstellung“ die Rede ist zumeist diese Bedeutung gemeint. Zum anderen ist damit aber – wie bereits angesprochen – auch die von der Filmkamera ohne Unterbrechung gefilmte und nicht geschnittene Aufnahme gemeint.

Neben den Unklarheiten bei der Definition des Begriffes der „Einstellung“ finden sich dann Unklarheiten bzw. Diskrepanzen bei der Definition der Begriffe „Szene“ und „Sequenz“. Dies wird auch von James Monaco (2003: 159) so gesehen, wenn er „Szene“ definiert als „allgemeine – und recht unpräzise – Bezeichnung für eine Einheit der Filmerzählung. Eine Szene besteht aus einer oder mehreren Einstellungen, die durch den Ort, die Handlung oder – wie im herkömmlichen Bühnenstück – durch die anwesenden Personen verbunden sind.“ Monaco verweist dann auf den Begriff der „Sequenz“ (2003: 148), den er aber in ähnlicher Weise definiert, nämlich als „eine Folge von inhaltlich zusammenhängenden Einstellungen“. Eine Differenz zwischen den Begriffen Szene und Sequenz zeichnet sich allerdings bereits hier ab – und zwar insofern, als der Begriff der Szene stärker an die Beiträge der abgebildeten Bildproduzent(inn)en gebunden wird: die „Handlung“ und die „anwesenden Personen“.

Der Begriff der „Sequenz“, mit dem in der Filmwissenschaft die Einheiten benannt sind, welche für die Analyse ausgewählt werden, wird überwiegend definitiv grundlegend an den der Einstellung gebunden. Eine „Sequenz“ ist dann ähnlich wie bei Monaco auch bei Beller (2005: 11) definiert als „die Vielzahl von Einstellungen (mindestens zwei), die einen gedanklichen und/oder formalen Kontext bilden“.

Eine Definition, welche den Begriff der Sequenz stärker an die Leistungen der *abgebildeten* Bildproduzent(inn)en und die *abgebildeten* Objekte bindet, findet sich bei Hickethier (1996:38). Dort wird unter Sequenz „eine Handlungseinheit verstanden, die zumeist mehrere Einstellungen umfasst und sich durch ein Handlungskontinuum von anderen Handlungseinheiten unterscheidet“. Mit dem Begriff der „Handlungseinheit“ wird bei Hickethier auf die Handlungen der abgebildeten Bild-

produzenten bzw. auf eine Einheit der abgebildeten Objekte Bezug genommen: „In der Regel werden Handlungseinheiten durch einen Ortswechsel, eine Veränderung der Figurenkonstellation und durch einen Wechsel in der erzählten Zeit bzw. der Erzählzeit markiert“ (ebenda).

Dieser eher negativen Definition von Sequenz, indem sie durch das bestimmt wird, was sie von anderen Sequenzen abgrenzt, steht bei Borstnar/Pabst/Wulf (2002: 139) eine eher positive Definition gegenüber: „Unter einer Sequenz versteht man ein Stück Film bzw. eine Episode, die grafisch, räumlich, zeitlich, thematisch und/oder szenisch zusammenhängt.“

Hier gewinnt, wie auch bei anderen Filmwissenschaftlern, der Begriff der „Szene“ eine Bedeutung. Diesen Begriff möchte ich aufgreifen und auf all das beziehen, was die Leistung der *abgebildeten* Bildproduzent(inn)en betrifft bzw. als Einheit der abgebildeten Objekte wahrgenommen werden kann, wie dies sich auch in der oben wiedergegebenen Definition von Monaco (2003: 159) impliziert ist.

Die von mir vorgeschlagene Definition, die eine Differenzierung in *abbildende* und *abgebildete* Bildproduzent(inn)en voraussetzt, wie wir sie in den Filmwissenschaften nicht finden, schafft eine gewisse, für die sozialwissenschaftliche Analyse notwendige, Klarheit: *Szenen* oder *Szenerien* stellen also Einheiten im Bereich der Bewegungen oder Handlungen der abgebildeten Bildproduzent(inn)en bzw. Kontinuitäten im Bereich der abgebildeten Objekte dar und können somit überwiegend als das Produkt der *abgebildeten* Bildproduzent(inn)en angesehen werden. Demgegenüber sind *Einstellung* und *Montage* das Produkt der *abbildenden* Bildproduzent(inn)en. Beides zusammen – also der Zusammenhang von *Szene* einerseits und *Einstellung* und *Montage* andererseits – konstituiert eine *Sequenz*.<sup>95</sup>

Wenn wir *Szenen* oder *Szenerien* als Einheiten im Bereich der Bewegungen oder Handlungen der abgebildeten Bildproduzent(inn)en fassen, so ist selbstverständlich entscheidend, welche Bewegungs- oder Handlungseinheiten letztlich zugrunde gelegt werden – jene auf der ikonografischen oder jene auf der vor-ikonografischen Ebene im Sinne von Panofsky (vgl. Kap. 3.3). Da im Sinne der dokumentarischen Film- und Videoanalyse die *vor-ikonografische* die elementare und primordiale Ebene ist, soll letztlich die Identifikation von Szenen auf dieser Ebene ausschlaggebend sein.

Die vor-ikonografische oder denotative Ebene der Filmerzählung (siehe auch: 5.4.1) umfasst nicht nur den Bereich der Bewegungs- und Handlungsverläufe der abgebildeten Bildproduzent(inn)en, sondern und insbesondere die Montage und die Veränderung der Einstellungen, also den Bereich der Leistungen der abbildenden Bildproduzent(inn)en, wie sie zur Konstitution dieser Erzählung beitragen. Wie ge-

95 Auch die Definition von Beller (2005: 11) lässt sich in gewisser Weise damit in Übereinstimmung bringen. Nach ihm ist „die durch Montage zusammengesetzte Szene filmterminologisch eine Sequenz“. Wobei die hier implizierte Unklarheit m.E. eben darauf zurück zu führen ist, dass sich das, was hier als „filmterminologisch“ gefasst wird, im wesentlichen auf die Leistung der abbildenden Bildproduzent(inn)en bezieht. Bordwell (2006: 154) verwendet den Begriff der *Szene* in dem Sinne, wie ich den Begriff der *Sequenz* verwenden möchte (ähnlich auch: Mikos 2003: 207). Und in gewisser Weise nimmt Bordwell auf das, was ich mit dem Begriff der Szene fassen möchte, mit dem Begriff des „Handlungsteils“ Bezug.



sagt, möchte ich den Begriff der *Sequenz* dort verwenden, wo beide Bereiche gemeint sind.

Der erste Interpretationsschritt einer dokumentarischen Filminterpretation ist derjenigen der Identifikation von *Hauptsequenzen* (HS), *Untersequenzen* (US) und *eingelagerten Sequenzen* (ES). Untersequenzen stellen in der Regel Einstellungsvarianten innerhalb von Hauptsequenzen dar. Eingelagerte Sequenzen sind solche Sequenzen, durch die andere Sequenzen unterbrochen und nach der eingelagerten Sequenz kontinuierlich fortgesetzt werden.

### 5.6.2 Montage und Einstellung als Produkt der abbildenden Bildproduzent(inn)en

Die Begriffe der Montage und der Einstellung bzw. des Einstellungswechsels beziehen sich also auf die Bewegungen und Konstruktionsleistungen der abbildenden Bildproduzent(inn)en, welche nicht nur von den Bewegungen der abgebildeten Bildproduzent(inn)en, sondern auch noch „strikt von der Bewegung der einzelnen Filmbilder während der Vorführung zu trennen“ sind (Sachs-Hombach 2003: 131). Es lassen sich also drei Dimensionen der Bewegung im Film unterscheiden:

- die Bewegungen bzw. Gestaltungsleistungen der abbildenden Bildproduzenten während und nach der Aufnahme, also die Sequenzierung der Einstellungen und die Montage
- die Bewegungen (Gebärden) der abgebildeten Bildproduzent(inn)en und
- die Bewegung der einzelnen Filmbilder während der Vorführung, welche das technische Medium darstellt, um in den anderen beiden Dimensionen den Bewegungen zur Darstellung verhelfen zu können.

Sowohl die modi operandi der abbildenden Bildproduzent(inn)en wie auch diejenigen der abgebildeten Bildproduzent(inn)en sind einer umfassenden Analyse erst zugänglich, wenn wir ihre Doppelstellung einerseits im Kontext von *Simultaneität* und andererseits im Kontext von *Sequenzialität* rekonstruieren. Die formale Komposition (Planimetrie, Perspektivität) des Fotogramms in ihrer Simultanstruktur ist zugleich das Produkt der abbildenden wie abgebildeten Bildproduzent(inn)en, wobei die Wahl der Perspektivität überwiegend den abbildenden Bildproduzent(innen) zuzurechnen ist, ohne dass aber eine klare Trennung möglich wäre. Die Sequenzierung der Fotogramme unter dem Aspekt des Einstellungswechsels und der Montage stellt eindeutig das Produkt der abbildenden Bildproduzent(inn)en dar und eröffnet uns den Zugang zu deren modus operandi.

Wie bereits angesprochen, ist die Reflexion auf Einstellung und Montage auch im Falle von Videografien bedeutsam, die zu *Forschungszwecken* erstellt worden sind, wenn die Videografie also als *Erhebungsinstrument* eingesetzt wird. Und weitergehend lässt sich Marotzki/Schäfer (2006: 62) zustimmen, wenn sie betonen: „Generell gelten für die Auseinandersetzung mit Film und Video in der qualitativen Sozialforschung die gleichen filmsprachlichen Codes und Filmstandards wie beim fiktionalen Film.“ Die methodische Kontrolle der Einflüsse der Forscher als abbildende Bildproduzent(inn)en gelingt nur dann, wenn auch in

diesem Fall die vor allem im Bereich der Filmwissenschaft ausgearbeiteten Erkenntnisse hinsichtlich Kameraführung, Perspektivität, Einstellungsgröße und schließlich auch Montage hier zumindest in Ansätzen in die methodische Reflexion einbezogen werden.

Für eine sozialwissenschaftliche Analyse, die, wie die dokumentarische Interpretation, das uns vorliegenden Produkts als Alltagsdokument und somit insgesamt als Dokument für die Weltanschauung der Erforschten nimmt, ist es (im Unterschied zur filmwissenschaftlichen Analyse) nicht notwendig, über die *technische* Herstellung im Bereich der Kameraführung, Einstellung und der Montage genauer Bescheid zu wissen. Welcher technischer Instrumente sich der Gestaltungswille der abbildenden Bildproduzent(inn)en (während und nach der filmischen Aufzeichnung) bedient hat, um die Gestaltung so und nicht anders zu realisieren, ist für die Interpretation nicht ausschlaggebend (zumal dies oft gar nicht mit Sicherheit zu rekonstruieren ist). Entscheidend sind das gestaltete Produkt und der sich in ihm dokumentierende Habitus.

### 5.6.3 Montage und Räumlichkeit

Die Elemente des Einzelbildes, des Fotogramms, gewinnen ihre gesicherte Bedeutung – insbesondere was die Gebärden und die Mimik betrifft – erst aus dem Kontext der Gesamtkomposition des Bildes. „Keine Bewegungseinheit für sich genommen ist Träger von Bedeutungen. Bedeutung erwächst aus dem Kontext“, heißt es bei Birdwhistell (1952: 10). Und das einzelne Fotogramm gewinnt seine *weitergehende* Bedeutung im Kontext der anderen Fotogramme, aus der Relationierung mit diesen, aus deren Sequenzierung, aus Einstellungswechsel und Montage. Auf diese Weise werden sowohl Zeitlichkeit als auch Räumlichkeit hergestellt. Die durch Einstellungswechsel und Montage herbeigeführte Konstruktion von Räumlichkeit bezeichnet Panofsky (1999: 25) als eine „Dynamisierung des Raumes“, eine Beobachtung, die er allerdings als „evident bis zur Selbstverständlichkeit“ ansieht: „Es bewegen sich nicht nur Körper im Raum, der Raum selbst bewegt sich, nähert sich, weicht zurück, dreht sich, zerfließt und nimmt wieder Gestalt an – so erscheint es durch wohlüberlegte Bewegung und Schärfenänderung der Kamera, durch Schnitt und Montage der verschiedenen Einstellungen.“<sup>96</sup>

Auch in der Konstruktion von Räumlichkeit finden sich die beiden grundlegenden Dimensionen filmischer Komposition: diejenige des Fotogramms als Kompo-

96 Inwieweit ist die durch Einstellungswechsel, Schnitt und Montage hergestellte Zeitlichkeit und Räumlichkeit grundlegend betrachtet nicht erst eine Erfindung des Filmes, sondern basiert auf intuitiven Kompetenzen, auf einer alltäglichen Praxis des Sehens, an welche der Film anschließen muss, um verstanden zu werden, die er dann allerdings zu verfeinern, aber auch zu verfremden vermag? Von Hans-Georg Soeffner und Jürgen Raab (1998: 131) wird dies eindeutig beantwortet. Es werde häufig „übersehen, dass der Schnitt eine Erfindung des Auges und das Schneiden eine der konstitutiven Aktivitäten des Sehens ist.“ Für David Bordwell (2006: 208) ist diese Frage nicht eindeutig zu beantworten: „Die einfachste Antwort auf diese Frage ist: ‚Weil diese Technik die geistigen Prozesse nachempfendet‘. Aber das ist vielleicht nicht die beste Art, sie zu verstehen. Es ist vielleicht so, dass wir uns diese Dinge, diese verschiedenen Traditionen sehr schnell zu eigen machen.“

sition *simultaner* Relationen und die Komposition im Sinne von Einstellungswechsel und Montage als Komposition von *Bildsequenzen*. In diesem Sinne umfasst die *Raumkonstruktion* im Film sowohl den „fotografierten optischen Raum“, der durch das Fotogramm zugänglich ist, wie auch den (durch die Montage geschaffenen) eigentlichen „filmischen Raum“, wie Hickethier (1996: 82f.) mit Bezug auf Pudowkin differenziert. Der filmische Raum wird auch „als ein diegetischer, narrativer-Raum verstanden, ein Raum also, dessen einzelne Segmente sich durch das im Gedächtnis des Betrachters gespeicherte Wahrnehmen der im Film gezeigten und damit ‚akkumulierten‘ Raumsegmente zu einem Ganzen zusammenschließen“ (Hickethier 1996: 83f.). Dass die Konstruktion von Räumlichkeit ohne Zeitlichkeit nicht möglich ist, stellt allerdings nicht eine Eigenart des Filmes dar.

In ähnlicher Weise hat Bordwell (1985: 117) zwischen dem durch die Montage hergestellten, durch das Arrangement der Einstellungen komponierten, Raum, dem „editing space“, also dem *montierten Raum*, und dem „shot space“, dem *Einstellungsraum*, unterschieden, welcher in unserem Verständnis auf der Grundlage des Fotogramms analysierbar ist und neben der Perspektivität eine Reihe weiterer Merkmale umfasst, u.a. das Überlappen von Umrissen, die Lichtwirkung, Farbgebung und Schärfentiefe oder Tiefenschärfe. Als weitere Dimension der Herstellung von Räumlichkeit nennt Bordwell den „sonic space“, den *Ton-Raum*, welcher durch akustische Phänomene hervorgerufen wird. Auf der Grundlage des Arrangements dieser drei unterschiedlichen Komponenten konstruiert der Rezipient den filmischen Raum in seiner Gesamtheit als „szenographischen Raum“ (Bordwell 1985: 113).

#### 5.6.4 Relationierung als Leistung dokumentarischer Interpretation

Setzt sich somit bereits die Konstruktion der Raum-Dimension selbst aus unterschiedlichen Komponenten zusammen (Einstellungsraum, montierter Raum, Ton-Raum), so gilt es in der dokumentarischen Filminterpretation darüber hinaus auch die Relationen zwischen den unterschiedlichen Dimensionen von Bild und Text in den Blick zu nehmen, so bspw. die Relation zwischen dem Text (bspw. einem Dialog) und der gewählten Einstellung. So erläutert der bekannte Cutter Walter Murch im Interview einige Möglichkeiten der „Komposition“, die er als „tiefergehender und geheimnisvoller“ bezeichnet: „Man kann den Mann die Frau ansehen lassen und ihr ins Gesicht sagen lassen: ‚Ich liebe dich.‘ Wenn der Bildausschnitt nicht stimmt, kann das bedeuten, dass er sie nicht wirklich liebt. *Sie* glaubt ihm, weil er es gerade gesagt hat. Aber die Filmemacher können die Einstellung so wählen, dass das Gegenteil verdeutlicht wird“ (Ondaatje 2002: 41f.).

Die bereits bei den Klassikern der Filmwissenschaft angelegte Kontroverse zwischen der Fokussierung der Filmanalyse auf das Einzelbild versus einer Fokussierung auf die Montage findet sich exemplarisch bei Kracauer (1964: 106), welcher auf ein Experiment von Kuleschow und Pudowkin Bezug nimmt: So „schnitt Kuleschow ein und dieselbe Aufnahme von Moskujins bewußt neutralem Gesicht in verschiedene Handlungszusammenhänge ein; das Ergebnis war, dass das Gesicht des Schauspielers bei einem traurigen Anlaß Schmerz und bei einer heiteren Gele-

genheit lächelnde Zufriedenheit auszudrücken schien.“<sup>97</sup> Kracauer hält dem entgegen, dass das Gesicht „bewußt neutral“ arrangiert worden sei und erst von daher die durch die Montage hergestellte „Story“ ihre starke Bedeutung entfalten konnte. Auch Balázs (2001b: 18) betont die eigenständige Kraft der Mimik, der „Mikrophysiognomie“ gegenüber der Story des Films: „Die innere Handlung, die nur an den Gesichtern zu sehen ist, interessiert heute mehr als jene, die in äußerer Bewegung sich kundtut.“

Gegenüber diesen Kontroversen und jenseits von ihnen ist aus der Perspektive der dokumentarischen Methode geltend zu machen, dass wir das ‚Ganze‘ eines Filmes nur dann in den Blick bekommen, wenn wir *Homologien*, also Strukturidentitäten herausarbeiten können zwischen jener Struktur, die sich innerhalb der Einstellungen dokumentiert, also in den *internen* Relationen des *Fotogramms* (die in wesentlichen Bereichen das Produkt der abgebildeten Bildproduzent(inn)en sind), und jener Struktur, die sich in der Relation der Einstellungen *zueinander* dokumentiert, also in Einstellungswechsel und Montage (die ausschließlich das Produkt der abbildenden Bildproduzent(inn)en sind. Zu leisten ist also – im Sinne der dokumentarischen Methode – die Relationierung der beiden Relationen: der Relationierung der im Fotogramm (simultan) hergestellten Relationen mit den durch Einstellungswechsel und Montage (sequenziell) hergestellten.

#### Zentrale Dimensionen dokumentarischer Film- und Videointerpretation

	abbildende/r Bildproduzent/in	abgebildete/r Bildproduzent/in
<b>simultane Performanz</b>	Formale Komposition des Fotogramms: – planimetrische Komposition – Perspektivität – Einstellungsgröße	Die tiefere Semantik einer Gebärde (ihr Dokumentsinn) kann in valider Weise nur in ihrer Relation zu simultanen anderen Gesten, also in Relation zum gesamten Körper und zur gesamten Interaktionsszene (szenische Choreografie), interpretiert werden
<b>sequenzielle Performanz</b>	– Kamerabewegung – Wechsel der Einstellung und Einstellungsgröße – Schnitt/Montage/Mischung	Die Film-Narration auf der <i>vorikonografischen</i> Ebene ist von größerer Bedeutung für die dokumentarische Interpretation als die Film-Narration auf der <i>ikonografischen</i> Ebene (story/plot)

#### Exkurs: die Relationierung der Relationen am Beispiel

Als ein Beispiel für die Relationierung von Relationen möchte ich auf die in Kapitel 6 vorgestellte Videointerpretation der Fernsehshow „Istanbul Total“ verweisen. Wesentliches Prinzip des Wechsels der Sequenzen, des Wechsels von Szenerie und Einstellung, also der Bildmischung dieser Show ist, dass diese Bildmischung durch den Moderator Stefan Raab zwar nicht selbst hergestellt, aber in ihrer Herstellung

<sup>97</sup> Ausführlicher ist dieses Experiment in Beller 2005: 22 dargestellt.

„dirigiert“ wird (vgl. 6.3.5). Der Wechsel folgt der Blick- bzw. Zeigerichtung und/oder der Bewegung von Stefan Raab – entweder durch harten Schnitt oder durch Kameraschwenk oder Zoom. Ausnahmsweise gibt St. R. auch *explizite* Anweisungen zum Wechsel der Szenerie.

Indem wir auf diese Weise das zu sehen bekommen, was Stefan Raab sieht bzw. wohin er blickt oder worauf er zeigt, wird hier Räumlichkeit (im Sinne des montierten oder filmischen Raumes) strikt vom Standort von Stefan Raab konstruiert (der damit zum abbildenden Bildproduzenten wird). Zugleich steht Stefan Raab als abgebildeter Bildproduzent in planimetrischer Hinsicht (formale Gestaltung der Fläche) wie auch perspektivisch (formale Gestaltung der Perspektivität) im Zentrum des fotografierten optischen Raumes.

Die Relationierung dieser beiden Dimensionen (und ihrer internen Relationen), verdeutlicht, dass Stefan Raab zugleich als abgebildeter wie als abbildender Bildproduzent die Performanz des Videos bestimmt. Ich spreche deshalb hier von einer *Monostrukturierung durch den Showmaster* und einer *Hyperzentrierung* auf seine Person, sodass der Orientierungsrahmen der Sendung und der Habitus des Moderators nicht mehr zu trennen sind (siehe: 6.3.4 sowie 6.5.1).

Als ein anderes Beispiel für eine derartige Relationierung von Relationen möchte ich auf eine kurze Sequenz aus einem Video eingehen, welches Astrid Baltruschat in meiner Forschungswerkstatt vorgelegt hat<sup>98</sup> und in dem Lehrer/innen Probleme ihres eigenen Alltags verdichtet zum Ausdruck zu bringen suchen. Dieses Amateur-Video mit dem Titel „Kammer des Schreckens“ erscheint auch deshalb interessant, weil das, was hier zunächst schlicht als das Produkt von Dilettantismus anmuten mag (u.a. das Abschneiden von Köpfen bei den abbildenden Bildproduzent/innen), sich durchaus als stilistisches Gestaltungselement der abbildenden Bildproduzent(inn)en und somit als Dokument ihres Habitus erweist.

Das Video ist auf der Grundlage des Transkriptionssystems MoViQ erstellt worden (dazu: Kap. 5.7.2; 6.2 u. 7.1). Ich habe daraus eine Sequenz von 23 Sek. mit vier Untersequenzen (US) ausgewählt: 0:57- 1:20.

Zunächst wird in einer *ersten* Untersequenz mit einer Dauer von 8 Sekunden eine voranschreitende weibliche Person abgebildet (die wir aufgrund des ikonografischen Vor-Wissens als vom Unterricht kommende Lehrerin identifizieren können). Die Kamera übernimmt ihre Laufrichtung, Perspektive und ihre Blickrichtung, welche nach unten auf ihre im Gehen sich bewegenden Füße und eine schwer ins Bild baumelnde Aktentasche zielt. Die Lehrerin ist somit zugleich abgebildete wie auch abbildende Bildproduzentin. Die Blickrichtung (der Kamerablick) bleibt starr: kein Schwenk nach oben oder zur Seite.

98 Es handelt sich um einen Beitrag, der im Kontext eines in der Region Nürnberg alle 2-3 Jahre ausgeschrieben Wettbewerbs mit dem Rahmenthema „Schulen im Wandel“ entstanden ist. Das Video wurde von Lehrer(inne)n eines Gymnasiums eingereicht zu der im Jahre 2004 gültigen Fragestellung: „Schule überdenken! – Muss die Institution Schule grundlegend verändert werden?“ (s. dazu Baltruschat 2008). In ihrer Dissertation (Baltruschat 2008) hat Astrid Baltruschat sowohl Videos, die von Schüler(inne)n, wie auch solche, die von Lehrer(inne)n produziert worden sind, interpretiert.

Abbildung 26: Sequenz aus Videotranskript „Kammer des Schreckens“ (Baltruschat 2008)

TC:	0:57	E12 0:58	0:59	1:00	1:01	1:02	1:03
Musik:	♩ ♩ ♩ ♩ ♩						
Geräusch:		Schritte					
Off:		Rausde Stimmen	aus	der	ferne	im	Hinter-
Kamera:		↑	↑	↑	↑	↑	↑

TC:	1:05	E13 1:06	1:07	1:08	1:09	1:10	E14 1:11	1:12
Geräusch:		lautes Stimmen-	gewirr				laute Stimmen	im
Off:								
Kamera:	↑	↑	↑	↑	↑	↑		

TC:	1:13	1:14	E15 1:15	1:16	1:17	1:18	1:19	1:20
Geräusch:			lautes Stimmen-	gewirr				
Off:	Hinter-	geräusch						

Darin dokumentieren sich zwei Komponenten des Orientierungsrahmens: Zum einen ist der Blick der Lehrerin selbstbezogen auf sich selbst gerichtet. Denn sie ist diejenige, die schaut (bzw. wird aus ihrer Perspektive geschaut) und zugleich ist sie diejenige, auf die geschaut wird (auf ihre Füße). Zum anderen ist der Blick gesenkt und nach unten gerichtet und z.T. auf die Last, die sie trägt und in der eine ‚Belastung‘ ihren Ausdruck findet. Es dokumentiert sich hier insgesamt eine Selbstreflexivität, die selbstbezüglich und monadisch bleibt, sich selbst nur selektiv in den Blick nimmt, indem sie auf die eigene Last gerichtet ist.

Die Szenerie in der *folgenden*, der *zweiten* Untersequenz (von 5 Sekunden Dauer) ist vollständig verändert und der Wechsel somit nach dem ersten Eindruck äußerst kontrastreich: Wir haben es mit einem Blick auf junge Menschen zu tun, auf (wie wir aufgrund des ikonografischen Vor-Wissens einordnen können) Schüler/innen, die in dichter und lebhafter Kommunikation miteinander stehen.

Dieser Kontrast zwischen der Selbstbezogenheit der Lehrerin, wie wir sie auf der Grundlage der Fotogramme in der ersten Untersequenz herausarbeiten können, und der Struktur der Fotogramme der zweiten Untersequenz, der kommunikativen Dichte und interaktiven Bezogenheit der Schüler/innen, lässt die Selbstbezogenheit konturierter hervortreten und ergänzt sie um die Komponente der Isolation bzw. Einsamkeit. Somit dokumentiert sich in der Relation (im Kontrast) von erster und zweiter Untersequenz, also in der Montage, auch etwas über das Verhältnis der Lehrerin zu den Schülerinnen und Schülern.

Dieses Verhältnis kann noch tiefer gehender beleuchtet werden, wenn wir in der zweiten Sequenz nicht allein die Szenerie, also die *abgebildeten* Bildproduzent(inn)en, sondern auch die Leistung der *abbildenden* Bildproduzentin in den Blick nehmen, die ja mit der Lehrerin identisch ist. Dies betrifft zum einen die Wahl des Bildausschnitts, die Kadrierung: Die Köpfe der in Gruppen und in enger

Kommunikation stehenden Schüler sind überwiegend abgeschnitten. Und der Kamerastandort ist so gewählt, dass diese Torsi in amerikanischer oder halbnaher Einstellung von unten, also aus der Untersicht, aufgenommen werden.

Im Abschneiden der Köpfe und in dieser Torsohaftigkeit im Allgemeinen dokumentiert sich, dass die Schüler/innen aus der hier eingenommenen Perspektive der Lehrerin (als abbildender Bildproduzentin) nicht als kommunikationsrelevante Gegenüber wahrgenommen werden. Und da die Lehrerin sich ihren Weg durch die dicht gedrängt stehenden Schüler/innen bahnen muss, erhalten diese (im Zusammenhang mit dem fehlenden Blickkontakt und kommunikativen Bezug) deutlich den Charakter eines Hindernisses. Die Schüler/innen werden nicht nur nicht als Gegenüber, sondern nicht einmal als Personen wahrgenommen. Vielmehr wirken sie wie eine Mauer und dadurch, dass die Untersicht eingenommen wird, zugleich übermächtig und bedrohlich.

In der Relationierung von erster und zweiter Untersequenz wird mittels des Wechsels von Szenerie und Einstellung auf der *einen Seite* also ein *Kontrast* zwischen den beiden Sequenzen hergestellt: der Kontrast zwischen der Selbstbezogenheit der Lehrerin einerseits und der interaktiven Bezogenheit und kommunikativen Dichte der Schülerinnen andererseits, welche die Selbstbezogenheit konturierter hervortreten lässt und sie um die Komponente der Isolation bzw. Einsamkeit ergänzt.

Auf der *anderen Seite* werden durch die Relationierung von erster und zweiter Untersequenz mittels des Wechsels von Szenerie und Einstellung aber zugleich auch *Gemeinsamkeiten*, also Homologien zwischen den Sequenzen, hergestellt: Denn wenn wir die Leistung der abbildenden Bildproduzentin einbeziehen, so weist die Selbstbezogenheit der Lehrerin in der ersten Sequenz Homologien auf zu der ihr eigentümlichen fehlenden Kommunikationsbereitschaft in der zweiten Sequenz.

Die dritte Untersequenz mit einer Dauer von 4 Sekunden stimmt mit der ersten weitgehend überein – sowohl in Szenerie wie auch in der Einstellung. Und die vierte Untersequenz mit einer Dauer von 5 Sekunden stimmt in dieser Hinsicht mit der zweiten überein, sodass erste und dritte sowie zweite und vierte Untersequenz weitgehend identisch sind. Durch diese Wiederholungen werden die durch die Montage hergestellten Homologien wie auch die Kontraste zwischen den Untersequenzen verfestigt. Es kommt zu einer Verdichtung der zentralen Strukturmerkmale, die auch in den Fotogrammen, d.h. in deren internen Relationen ihren Ausdruck finden: der übergreifenden Sinnstruktur einer Selbstbezogenheit und Isolation der Lehrerin, die auf ihre Last fixiert ist.

### 5.6.5 Sequenzielle und simultane Relationen und das Kontext-Wissen

Es kann hier wiederum deutlich werden, dass die dokumentarische Interpretation sich grundsätzlich nicht auf Einzelelemente, sondern auf die Relationen der Einzelelemente zueinander richtet. Genauer betrachtet ist somit das, was mit einer ‚ganzheitlichen‘ Betrachtung des Filmes gemeint ist, eine Betrachtung auf der Grundlage von (vielfältigen) Relationen.<sup>99</sup> Denn dabei geht es nicht darum, jedes Detail des

<sup>99</sup> So heißt es auch bei Deleuze (1997a: 24): „Wäre das Ganze zu definieren, dann durch die Relation.“

Bildes zu erfassen, sondern zentrale Elemente des Bildes in ihren zentralen Relationen (wie bspw. die dominanten Linien der planimetrischen Komposition).

Nach Imdahl (1996a: 23) ist das Bild eine „in seiner Ganzheitlichkeit invariable und notwendige, das heißt alles auf alles und alles aufs Ganze beziehende Simultanstruktur. Diese erzwingt keineswegs ein starres Hinsehen auf das Bildganze, wohl aber eröffnet sie ein Sehen simul und singulariter: Das Ganze ist von vornherein in Totalpräsenz gegeben und als das sinnfällige Bezugssystem in jedem einzelnen kopräsent, wann immer jedes Einzelne in den Blick genommen wird.“ (Imdahl 1996a: 23).

Die Bedeutung einer Bewegung oder Äußerung bestimmt sich im Sinne der dokumentarischen Methode von der Relation zu jenem *Kontext* her, wie er von den Akteuren in ihrer Handlungspraxis selbst hergestellt wird. Dies ist im Falle der *Text*-Interpretation die sequentielle Relation der jeweils zu interpretierenden Äußerung oder Geste zu den ihr nachfolgenden. Auf diese Weise verleihen die Äußerungen oder Gesten – durchaus im Sinne von George Herbert Mead – einander wechselseitig ihre Signifikanz.

Im Falle der *Bild*-Interpretation ist dies nicht eine *sequentielle*, sondern eine *simultane* Relation von Einzelelement und Gesamtkontext des Bildes, im Film sind diese Relationen zugleich simultaner und sequenzieller Art (siehe auch Wagner-Willi 2005: 269ff. u. 2007). Die Relation von Kontext und Einzeläußerung resp. Einzelelement ist eine ‚*reflexive*‘, wie die Ethnomethodologen dies bezeichnet haben (vgl. Garfinkel 1961 u. 1967). Das zirkelhafte Oszillieren zwischen den Einzelelementen und der Sinnstruktur des Gesamtkontextes stellt eine der Ausprägungen des klassischen hermeneutischen Zirkels im Sinne von Dilthey (1924: 330) dar: „Aus den einzelnen Worten und deren Verbindungen soll das Ganze eines Werkes verstanden werden, und doch setzt das volle Verständnis des einzelnen schon das Ganze voraus.“

Diese Relation von Einzelelement und Gesamtkontext hat einerseits eine formale Seite: Im Bereich von Gesprächen ist dies die „Diskursorganisation“ (vgl. u.a. Bohnsack/Przyborski 2006), im Bereich des Bildes ist es die formale Komposition, im Bereich des Films sowohl die formale Komposition des Fotogramms als auch die formale Komposition der einzelnen Fotogramme und Einstellungen in ihrer Relationierung, wie sie durch Schnitt und Montage hergestellt wird.

Der Blick auf die formale Struktur der Relationierung, also darauf, *wie* (im Bereich der Gesprächsanalyse) auf eine Äußerung durch andere Bezug genommen wird, *wie* ein Element des Bildes dem anderen und wie eine Einstellungssequenz der vorhergehenden formal zugeordnet ist, weist uns dann zugleich den Weg zur tiefer liegenden Semantik.

An dem im obigen Exkurs angeführten Beispiel kann auch deutlich werden, dass die beiden interpretierten Videosequenzen (der Blick auf die Füße der Lehrerin und der Blick auf die Schüler/innen) ihren Bezug zueinander erst durch ein *Kontext-Wissen* auf der *ikonografischen* Ebene erhalten, ein (stereotypes) Wissen um die Institution Schule. Sobald wir wissen, dass wir uns in einer Schule befinden und dem Weg der Lehrerin folgen, erscheinen die Gruppen von jungen Menschen als Schüler/innen. Somit ist, wie mit Bezug auf die Einzelbildinterpretation mehrfach dargelegt (vgl. Kap. 3), ein Wissen um Institutionen und Rollenbeziehungen,

ein sprachlich-textliches Wissen auf der *kommunikativ-generalisierten* Ebene Voraussetzung für die dokumentarische Bildinterpretation, nicht aber ein Wissen auf der *konjunktiven* Ebene. Denn letztere Art von Kontext-Wissen – bspw. ein Wissen um die Probleme der Lehrer/innen auf der Grundlage von Gesprächen mit ihnen – würde uns in der Interpretation den Blick verstellen für die im Bild selbst vermittelten konjunktiven Wissensbestände.

## 5.7 Zur Transkription

Wenn wir Interpretation im Sinne von Mannheim (1980: 272) als „begrifflich-theoretische Explikation“ definieren (vgl. auch Bohnsack: 2001: 332f.), so können wir zwei *Arten der Explikation* unterscheiden: eine Explikation des *Was*, des thematischen Gehalts, der thematischen (Formal-) Struktur, also die Thematisierung von Themen, wie sie durch die *formulierende Interpretation* geleistet wird, und die Explikation des *Wie*, des modus operandi oder der Orientierungsstruktur, wie sie durch die *reflektierende Interpretation* vorgenommen wird.

### 5.7.1 Interpretation, Transkription und Protokoll

Entsprechend der Definition von Interpretation als einer Leistung der *Explikation* ist die Wiedergabe von Datenmaterialien – auch wenn sie eine Übersetzungsleistung, eine Transformation in ein anderes Zeichensystem (bzw. von akustischen in schriftliche Zeichen) darstellt – noch so lange nicht als Interpretation zu bezeichnen, wie sie keine Explikation impliziter Bedeutungsgehalte vornimmt.

Die Transkriptionsverfahren, wie sie in der qualitativen Sozialforschung, vor allem in der Konversationsanalyse, entwickelt worden sind und auch unser eigenes Transkriptionssystem, welches wir in gewisser Anknüpfung an dasjenige der Konversationsanalyse entwickelt (siehe ursprünglich Bohnsack 1989: 387 f.) und neuerdings als „TiQ“ (Talk in Qualitative Social Research) bezeichnet haben (vgl. Bohnsack 2008a: Kap. 12.3), sind entsprechend auf dieser vor-interpretativen Ebene angesiedelt. Damit soll nicht gesagt sein, dass es sich beim Transkript nicht bereits um einen Teil der Analyse handeln würde. Denn, wie auch Angela Keppler (2006: 107) bemerkt, „ist die Verschriftlichung des sprachlichen und nicht-sprachlichen Materials selbst als Teil der Analyse zu betrachten und nicht lediglich als ein Akt der Übertragung“.

Da eine Explikation impliziter Bedeutungsgehalte grundsätzlich bereits gegeben ist, wenn der semantische Gehalt eines Bildes durch einen Text ersetzt wird, wie dies etwa in den „Transkripten“ bei Keppler geschieht, haben wir es in diesem Fall (unserer Definition entsprechend) nicht mehr mit einem Transkript, sondern bereits mit einer Interpretation zu tun. Eine derartige Interpretation ist jedoch adäquater als „Protokoll“ zu bezeichnen. Ein Protokoll ist auf der Ebene einer formulierenden Interpretation angesiedelt (vgl. Bohnsack et al. 1995: Kap. 7). Während Keppler (vgl. 2006: 107) zunächst von „Transkript“ spricht, überschreibt sie das

entsprechende Kapitel (3.3) dann jedoch schließlich mit „Filmprotokoll“. Um ein *Transkript* eines Filmes handelt es sich nach unserem Verständnis dann, wenn die visuelle Ebene durch Visuelles und die verbale Ebene durch Textförmiges repräsentiert wird, wenn wir also weitest möglich *innerhalb* der jeweiligen Medien verbleiben.

Keppler (2006: 106) bezieht sich auf das von Thomas Kuchenbuch (1978) vorgelegte Transkriptionssystem, welches, wie sie formuliert, einen „Anknüpfungspunkt für meinen eigenen filmanalytischen Ansatz“ darstellt, weicht dann aber von Kuchenbuch insofern ab, als bei diesem Bilder, Fotogramme Bestandteil des Transkripts sind.<sup>100</sup> Auch Hickethier (1996: 37) differenziert nicht zwischen der Transkription und der Protokollierung von Filmen. Beides versteht er als *vollständige Verschriftlichung* und „Literarisierung“ des Films“ (1996: 38), welcher er kritisch gegenübersteht: „Der *Einsatz und Gebrauch des Videorecorders* in der Analyse schafft einen direkteren Zugang zum Film als das von allen sinnlichen Konkretionen abstrahierende Protokoll.“ Dem ist zuzustimmen. Allerdings können wir auf diese Weise keinen Zugang zum Fotogramm und damit auch nicht zu der an die Simultaneität gebundenen Sinnstruktur des Bildes gewinnen. Auch die Relation von Bild- und Textsequentialität, also die Synchronizität von Text(-transkript) und Bild kann in präziser Weise nur dann rekonstruiert werden, wenn die Synchronizität von Texttranskript und Bild (und hier vor allem: den Gebärden der abgebildeten Bildproduzent(inn)en) in detaillierter Weise im Transkript dargestellt wird.

### 5.7.2 Das Transkriptionssystem MoViQ

Das von Stefan Hampl und Aglaja Przyborski für die dokumentarische Film- und Videointerpretation entwickelte Transkriptionssystem „MoViQ“ („Movies and Videos in Qualitative Social Research“) leistet die Repräsentation des Films durch Bildsequenzen (siehe Kap. 7.1 sowie auch: Przyborski/Wohlrab-Sahar 2008: 169-172). Bild und Ton werden hier synchron und in einem *konstanten* Zeitrhythmus transkribiert, dessen Intervalle nicht länger als 1 Sek. sein sollten (Die genaue Dauer der Intervalle ist aber abhängig von der Bewegungsintensität). Dies vermittelt dem Betrachter einen unmittelbaren Eindruck von der Relation von Bild- und Textsequentialität im Hinblick auf deren Synchronizität, also der Synchronizität von Text und Bild. Mit einem Filmprotokoll im Sinne von Keppler ist dies in präziser Weise nicht möglich.

Das Transkriptionssystem MoViQ ist darüber hinaus so angelegt, dass das von mir in Anknüpfung an die Konversationsanalyse für Gesprächs-, also Textanalysen und deren Sequenzialität entwickelte Transkriptionssystem „TiQ“ („Talk in Qualitative Social Research“; siehe Bohnsack 2008a: Kap. 12.3; ursprünglich Bohnsack 1989: 387f.) mit der Darstellung der Bildsequenzialität verbunden werden kann. Auch das erwähnte „Storyboard“-System von Kuchenbuch (1978) vermag einen

<sup>100</sup> In dieser Hinsicht können wir Reichertz (1990: 48) zustimmen, wenn er sich gegen die „Forderung zur vollständigen Verschriftlichung“ wendet u.a. mit der Begründung, „dass die Besonderheit des Audio-visuellen unwiederbringlich verloren geht.“

adäquaten Eindruck von der Bildsequenzialität nicht zu vermitteln, und es müssen sehr frühzeitig Entscheidungen über die Auswahl von Einstellungsvarianten getroffen werden.

Die mit dem Transkriptionssystem MoViQ getroffene Entscheidung dafür, Einzelbilder, also Fotogramme, als Grundeinheiten zu nehmen, ist nicht nur darin begründet, dass sich technisch gesehen der Film aus diesen zusammensetzt, sondern auch darin, dass uns dies den Zugang zu den Bewegungen der abgebildeten und des abbildenden Bildproduzent(inn)en eröffnet. Im Bereich der Gebärden der abgebildeten Bildproduzent(inn)en sind es, wie dargelegt, die Teilelemente dieser Gebärden, die Kineme, zu denen uns das Fotogramm den Zugang eröffnet. Im Bereich der Bewegungen der abbildenden Bildproduzent(inn)en sind es die jeweiligen *Relationen* der (in den Fotogrammen sich dokumentierenden) Einstellungen zueinander, die „Wechsel der Einstellungen innerhalb einer Einstellungsfolge“ (Hickethier 1996: 60), einer *Sequenz*, die uns den Zugang zu den Strukturen der Montage eröffnen, in denen sich der *modus operandi* der abbildenden Bildproduzent(inn)en dokumentiert.

Um die *Relationen zwischen* den Einstellungen, also die Leistung des Wechsels von Einstellungsgrößen, Kamerastandorten und der Kamerabewegung sowie der Montage, genauer interpretieren zu können, müssen wir notwendiger Weise zunächst die in den Fotogrammen sich dokumentierenden Einstellungen selbst identifizieren, sodass wir auf der Ebene der reflektierenden Interpretation mit den Fotogrammen beginnen. An ausgewählten Fotogrammen können wir eine tiefer gehende Betrachtung der (Einzel-) Einstellungen vornehmen – im Sinne einer genauen Rekonstruktion der *formalen Komposition* des Bildes, wie wir sie in der Einzelbildanalyse ausgearbeitet haben (siehe Kap. 3.6), die es aber hinsichtlich der technischen, vor allem elektronischen Möglichkeiten im Sinne einer „Montage im Bild“ (Keppler 2006: 118) oder „inneren Montage“ (Hickethier 1996: 153) zu differenzieren und zu konkretisieren gilt. Von hier aus können wir dann zur Rekonstruktion der Montage im eigentlichen Sinne, also zur Rekonstruktion von Schnitten sowie von Einstellungswechsel und Bildmischung, fortschreiten.

Die Transkription im Sinne von MoViQ eröffnet darüber hinaus die Möglichkeit, die in der Frequenz von weniger als 1 Sekunde ausgewählten Fotogramme in mehr oder weniger schnellem Rhythmus (nach Art einer Diashow) derart abzuspielen, dass der Film im Zeitlupenmodus wiedergegeben wird. Nach aller Erfahrung ermöglicht dies – zusammen mit der Analyse der Fotogramme selbst – eine genaue Rekonstruktion der Bewegungsabläufe sowohl der abbildenden wie der abgebildeten Bildproduzent(inn)en.

## 5.8 Arbeitsschritte dokumentarischer Film- und Videointerpretation

Vor dem Hintergrund meiner Ausführungen in Kapitel 3 habe ich mit Bezug auf den gegenwärtigen Stand qualitativer Methoden, aber auch mit Bezug auf Semiotik, Philosophie und Kunstgeschichte zu zeigen versucht, dass es bei der Entwicklung von sozialwissenschaftlichen Methoden der Bildinterpretation um das gehen muss,

was Hans Belting (2001: 15) aus eher philosophischer und kunsthistorischer Perspektive gefordert hat, nämlich „Bilder nicht mehr mit Texten zu erklären, sondern von Texten zu unterscheiden“. Insgesamt geht es darum, den Bildern das zu zuerkennen, was im Bereich von Texten schon geleistet ist, nämlich den Status selbstreferentieller Systeme. Das bedeutet, dass wir Texte wie auch Bilder primär auf der Grundlage des in ihnen selbst vermittelten Wissens und erst sekundär auf der Grundlage von Kontext-Wissen zu interpretieren suchen.

### 5.8.1 Zum Verhältnis der Interpretationen auf der Bild- und Textebene

Sofern man nicht, wie dies für weite Bereiche der sozialwissenschaftlichen Videointerpretation zutrifft (vgl. 5.3.3), das Bild lediglich in seiner ergänzenden Funktion zum Text in die Analyse einbezieht, hat dies die Konsequenz, dass im Falle der Interpretation von Videos und Filmen Bild und Text zunächst weitgehend unabhängig voneinander interpretiert werden. Erst auf dieser Grundlage werden dann im nächsten Schritt die Ergebnisse aufeinander bezogen, ggf. wechselseitig validiert, ergänzt oder auch differenziert.

Weitergehend habe ich in meinen Überlegungen in Kapitel 3 insbesondere mit Bezug auf Imdahl, Foucault und Barthes deutlich zu machen versucht, dass wir im Common Sense weniger dazu neigen, Texte mit Bildern ‚zu erklären‘ als vielmehr Bilder mit Texten. Wir müssen uns also auf der Grundlage einer systematischen methodischen Kontrolle vor allem die Bilder in ihrer Eigensinnigkeit und Eigenlogik in Abgrenzung gegenüber interpretativen Einflüssen unseres textlichen Vorwissens erschließen. Von daher erscheint es angebracht, der Interpretation in der Bilddimension im zeitlichen Ablauf der Arbeitsschritte Vorrang vor der Interpretation in der Textdimension einzuräumen.

Dies gilt insbesondere für die reflektierende Interpretation. Hier sollte Sorge getragen werden, dass der am Text rekonstruierte Orientierungsrahmen dem Bild nicht aufgelegt und die Rekonstruktion des Bildes in seiner Eigenlogik somit beeinträchtigt wird. Auf den wörtlichen (nicht aber den dokumentarischen) Gehalt des Textes, also auf das Texttranskript als solches, nehmen wir im Zuge der Bildinterpretation bereits Bezug – allerdings nur insoweit, als es unumgänglich ist. In der reflektierenden Interpretation des Textes werden dann jeweils bereits Bezüge zur Bildinterpretation hergestellt und ggf. Homologien zu den dort erarbeiteten ikonologischen Rahmenkomponenten hergestellt.

Bei der Textinterpretation folgen wir dem inzwischen insbesondere im Bereich der Gesprächsanalyse vielfach erprobten Verfahren der dokumentarischen Methode (siehe u.a. Bohnsack 2001, 2008a: Kap. 8 u. 7.3; Bohnsack/Przyborski 2006; Bohnsack/Schäffer 2007; Przyborski 2004).

Der zeitliche Vorrang der Interpretation in der Bilddimension vor derjenigen in der Textdimension kann ja nach Art und Genre des Films oder Videos im Hinblick auf die Arbeitsschritte unterschiedlich realisiert werden. So kann die Interpretation der dazugehörigen Textsequenz jeweils nach der Bildinterpretation einer Untersequenz, einer Hauptsequenz oder der gesamten Passage erfolgen. Letzterer Weg wird im Folgenden in der exemplarischen Interpretation in Kapitel 6 realisiert.

### 5.8.2 Zur Auswahl der für die Interpretation relevanten Sequenzen und Fotogramme

Die grundlegenden Analyseeinheiten dokumentarischer Videointerpretation stellen, wie in 5.6.1 dargelegt, die Sequenzen und (innerhalb dieser) die Fotogramme dar. „Eine detaillierte Analyse (...) wird man immer nur auf einzelne ausgewählte Sequenzen oder Shots anwenden, die unter der gewählten Fragestellung von besonderem Interesse sind“, wie Ehrenspeck/Lenzen (2003: 446) bemerken. Je weniger wir allerdings die Analyse und deren Fragestellung durch konkrete inhaltliche Fragestellungen vorstrukturieren wollen, desto mehr stellt die Auswahl der für die Interpretation herangezogenen Analyseeinheiten – hier: der Sequenzen – ein Problem qualitativer Forschung im allgemeinen und, wie Knoblauch/ Schnettler/Raab (2006: 14) ausführen, der Videoanalyse im besonderen dar: „(...) das methodologische Problem der Bestimmung der Analyseeinheit und der Balance zwischen einer zeitaufwendigen Mikroanalyse und einem Überblick über den gesamten Daten-Korpus bleiben offene Fragen für zukünftige methodologische Debatten.“

Im Sinne der dokumentarischen Methode wird die Problematik der Auswahl hinsichtlich ihrer Gültigkeit allerdings dadurch relativiert, dass sich die Grundstruktur des Falles – der *modus operandi* oder Habitus der Bild- oder Filmproduzent(inn)en – grundsätzlich in allen Sequenzen des Films oder Videos dokumentiert – allerdings mehr oder weniger deutlich, d.h. mehr oder weniger leicht zugänglich. Somit geht es eher darum, einen möglichst leichten oder direkten und in diesem Sinne forschungsökonomischen Zugang zu finden. Dabei folgen wir – wie auch im Falle von Passagen im Bereich der Textinterpretation, also u.a. der Gesprächs- und Interviewanalyse – dem Auswahlkriterium der *Fokussierung* (siehe auch Bohnsack 2006c).

Allerdings ist zur Bestimmung derartiger fokussierter Sequenzen die in Kap 5.1 angesprochene Unterscheidung von Bedeutung. Ich hatte unterschieden zwischen solchen Videos und Filmen, die zu Forschungszwecken produziert werden, die also als *Erhebungsinstrument* der (sozial-) wissenschaftlichen Forschung eingesetzt werden und solchen Filmdokumenten, die in den jeweiligen kulturellen Lebenszusammenhängen oder Milieus, welche Gegenstand der sozialwissenschaftlichen Analyse sind, *selbst* produziert worden sind. Dies gilt einmal für den *privaten* Bereich – beispielsweise für Familienvideos, die von Angehörigen der Familie aufgenommen worden sind, aber auch für Filme und aus dem *öffentlichen* Bereich, für Produkte der *Massenmedien*, wo es sich sowohl bei den abbildenden wie den abgebildeten Bildproduzent(inn)en überwiegend um die Gestaltungsleistungen professioneller Akteure handelt.

Im ersteren Fall, in demjenigen von Filmen und Videos als Erhebungsinstrument, wird das Dokument in Interaktion von Forschern und Erforschten produziert. Wir haben es – ebenso wie etwa im Falle von Gruppendiskussionen (vgl. u.a. Bohnsack 2008: Kap. 12.1 u. 2004a u. b) – mit zwei ineinander verschränkten Diskursen oder Gestaltungsleistungen zu tun: mit denjenigen zwischen Forschern und Erforschten einerseits und den denjenigen der Erforschten untereinander andererseits. Die genaue Rekonstruktion des Verlaufs der Videoaufzeichnungen ermöglicht (neben der Selbst-Reflexion der Forschenden hinsichtlich ihrer Interventionen) eine

Differenzierung zwischen jenen Sequenzen, die primär durch die Reaktion auf die Interventionen der Forschenden strukturiert sind, und jenen, in denen (primär) die Erforschten (die abgebildeten Bildproduzenten und -produzentinnen) wechselseitig aufeinander reagieren („Selbstläufigkeit“) und sich untereinander wechselseitig steigern, sodass wir es mit „dramaturgischen Höhepunkten“ zu tun haben. Die Erforschten ‚lösen‘ sich hier zumeist von den Interventionen der Forscher/innen und steigern sich erst allmählich zu derartigen dramaturgischen Höhepunkten, deren besondere kompositorische und interaktive Dichte (oder auch andere kompositorische Auffälligkeiten) diese Passagen oder Sequenzen als *Fokussierungsmetaphern* ausweist. Diese stellen Indikatoren für besondere Fokussierungen – und das bedeutet: für eigene Relevanzsetzungen – seitens der Erforschten dar.

Auch im Falle der von den Erforschten insgesamt – im Bereich der abgebildeten wie der abbildenden Bildproduzent(inn)en – selbst hergestellten und gestalteten Videos ist die Identifikation von dramaturgischen Höhenpunkten von Bedeutung. Hier kommen allerdings noch *weitere* Indikatoren für Fokussierungen bzw. Relevanzsetzungen hinzu: Da sich die Erforschten hier im Zuge der Eröffnung der visuellen Präsentation nicht erst von den Interventionen der Forscher/innen lösen müssen und die Strukturierung der Präsentation mehr oder weniger planbar ist, wird es hier aufschlussreich sein, *wie* die Erforschten selbst ihre Beiträge *eröffnen* und damit die Weichen für ihre visuelle Präsentation stellen. Somit gewinnen hier auch die *Eingangssequenzen* den Charakter von Indikatoren für Fokussierungen und Relevanzsetzungen.

So können wir beispielsweise in der in Kapitel 6 exemplarisch analysierten Fernsehshow den *am Anfang* dieser Show – und noch mehr: am Anfang der gesamten *Showserie* – stehenden Sequenzen einen Fokussierungscharakter zuschreiben. Diese Art der Fokussierung gewinnt insbesondere dann noch an Relevanz für die Auswahl von Sequenzen, wenn andere Kriterien für Fokussierungen – wie eine besondere kompositorische und interaktive Dichte und andere kompositorische Auffälligkeiten (siehe bspw.: 6.3.5) – kaum ausgeprägt sind.

Da auch die fokussierten Sequenzen – zumindest, was die Interpretation der Fotogramme anbetrifft – zumeist nicht insgesamt ausgewertet werden können, wählen wir *innerhalb* des Kriteriums der Fokussierung bzw. innerhalb der fokussierten Sequenzen als eine weitere Strategie die für die Auswertung relevanten Fotogramme und Untersequenzen nach dem Kriterium der *Repräsentanz* aus, d.h. danach, inwieweit sie geeignet sind, die umfangreichsten (Haupt-) Sequenzen eines Films oder Videos zu repräsentieren (vgl.: 6.3.2).

Dabei lassen sich alle drei genannten Kriterien miteinander kombinieren, wie dies in der empirischen Analyse in Kapitel 6 exemplarisch dargelegt ist: Hier wurden zunächst aus dem ersten Abend der (mehrere Abende umfassenden) Showserie die ersten Minuten ausgewählt und (da die Anzahl der Fotogramme, die einer Interpretation unterzogen werden können begrenzt ist) innerhalb dieser Eingangssequenz einerseits solche Fotogramme, welche die umfangreichsten (Haupt-) Sequenzen repräsentieren, und andererseits solche, die sich durch kompositorische Auffälligkeiten auszeichnen.

### 5.8.3 Überblick über die Arbeitsschritte im Ablauf

#### 1. Interpretation in der Bilddimension

##### 1.1 Auswahl der Sequenzen

nach dem Kriterium der Fokussierung

- aufgrund von Steigerungen in der Dichte der Komposition, der Gebärden und der Interaktion
- aufgrund von Diskontinuitäten u. anderen Auffälligkeiten in Szenerie, Einstellung u. Montage
- Im Falle des Videos als Alltagsdokument: aufgrund ihrer Positionierung im Gesamtprodukt (u.a. Einleitungssequenzen)

##### 1.2 Formulierende Interpretation von Sequenzen und von Einstellungswechsel und Montage

- Identifikation und Benennung von Hauptsequenzen (HS), Untersequenzen (US) und eingelagerten Sequenzen (ES) anhand vor-ikonografischer und ikonografischer Elemente der Szenerien
- Beschreibung von Kameraführung sowie Schnitt und Montage (Bildmischung)

##### 1.3 Auswahl der Fotogramme

- nach dem Kriterium der Fokussierung (siehe 1.1) sowie
- hinsichtlich ihrer Repräsentanz für die umfangreichsten Hauptsequenzen und die häufigsten Einstellungsvarianten

##### 1.4 Formulierende und Reflektierende Interpretation der Fotogramme

##### 1.5 Reflektierende Interpretation von Einstellungswechsel und Montage

#### 2. Interpretation in der Dimension von Text und Ton

##### 2.1 Formulierende Interpretation des Textes: Thematische Gliederung

##### 2.2 Reflektierende Interpretation in der Dimension von Text und Ton

#### 3. Reflektierende Gesamtinterpretation

## 6. Zur Forschungspraxis dokumentarischer Videointerpretation

### Exemplarische Analyse einer Fernsehshow: „Istanbul Total“

Zu Beginn des Kapitels 5 hatte ich zwischen solchen Filmdokumenten unterschieden, die in den jeweiligen kulturellen Lebenszusammenhängen oder Milieus, welche Gegenstand der sozialwissenschaftlichen Analyse sind, *selbst* produziert worden sind und somit Alltagsdokumente darstellen, und jener anderen Art von Filmdokumenten, die zu Forschungszwecken erstellt werden und somit den Charakter von Erhebungsinstrumenten annehmen. Ich werde die hier entwickelte Methode im Folgenden in exemplarischer Weise an ersterer Art von Dokumenten demonstrieren, da wir es dort mit dem komplexen Zusammenspiel von abgebildeten und abbildenden Bildproduzent(inn)en zu tun haben, welches methodisch eine Herausforderung darstellt.

Innerhalb dieser Art von Dokumenten, also derjenigen, die von Erforschten selbst produziert worden sind, lässt sich noch einmal zwischen solchen aus dem öffentlichen und denen aus dem privaten Bereich unterscheiden. Für die exemplarische Analyse wähle ich ein Dokument aus ersterem Bereich, ein öffentliches und professionell hergestelltes Video, da diese in der Regel technisch komplexer sind und somit erhöhte Anforderungen an die Analyse im Bereich der Leistungen der abbildenden Bildproduzent(inn)en, also hinsichtlich des Einstellungswechsels, der Kameraführung und der Montage (Bildmischung), stellen. Wie sich allerdings zeigen wird, gestalten sich in der hier ausgewählten Fernseh-Show die Bewegungen der abgebildeten Bildproduzent(inn)en, ihre Gebärden und deren interaktiver Bezug aufeinander nur wenig komplex.

Für den Bereich, in dem das Video als Erhebungsinstrument fungiert und sich die Analyse auf die abgebildeten Bildproduzenten und -produzentinnen konzentriert, liegen bereits seit einigen Jahren Analysen auf der Grundlage der dokumentarischen Methode in ersten Ansätzen vor (Wagner-Willi 2004, 2005, 2006 u. 2007; Klambeck 2007).

### 6.1. Zur Auswahl des Genres und des Videos

Die dokumentarische Interpretation von Videos vermag uns Aufschluss zu geben über den Erfahrungsraum, den Orientierungsrahmen und den Habitus derjenigen, die dieses Video (im weitesten Sinne) produziert haben: also über die abgebildeten und die abbildenden Bildproduzent(inn)en. Dies gilt sowohl für den privaten wie für den öffentlichen Bereich. Beispielsweise bieten uns Familienvideos, die von



Angehörigen der Familie aufgenommen worden sind, Einblicke in den Orientierungsrahmen, das Milieu oder den Habitus einer Familie. Ähnliches gilt aber auch für Filme und Bilder aus dem öffentlichen Bereich: Wollen wir etwas über den Erfahrungsraum und den Orientierungsrahmen einer deutschen Fernsehanstalt erfahren, so erscheint es sinnvoll, jene Gattung von Sendungen, also jene Genres auszuwählen, die insgesamt von dieser Anstalt (bspw. Pro 7) selbst produziert worden sind oder das Programm dieser Anstalt nachhaltig bestimmen. Dazu gehören u.a. die Shows, die, da sie zumeist über einen längeren Zeitraum gesendet werden und somit sozusagen den professionellen Alltag einer Anstalt repräsentieren.

Ausgehend von einer Idee von Stefan Hampl und Aglaja Przyborski,<sup>101</sup> die ich dann im Rahmen eines Lehrforschungsprojekts an der Freien Universität Berlin aufgegriffen habe, ging es um die Frage, welche Konstruktionen von Interkulturalität und Fremdheit wir in Produkten des deutschen Fernsehens finden. Zu dieser Thematik bot sich die Sendung „Istanbul Total“ des Moderators Stefan Raab beim Fernsehsender Pro 7 an. Und zwar nicht zuletzt deshalb, weil „Istanbul Total“ laut Pro 7 eine der erfolgreichsten Sendungen des Jahres 2004 war (Quelle: [www.quotenmeter.de](http://www.quotenmeter.de)) und noch dazu in der Presse als interkulturelles Ereignis gefeiert wurde. In einem Bericht des Berliner Tagesspiegels (vom 10.05. 2004) stellt „Stefan für mich besonders nach dieser Sendung den Beweis dafür dar, dass die sogenannte Integrationspolitik nicht gänzlich gescheitert ist. (...) Eine solche Sendung bringt Einschaltquoten; da bin ich mir sicher, liebe TV-Sender, denn hier findet sich nicht nur der deutsche Türke, sondern auch der türkische Deutsche wieder!“

Neben der für uns relevanten Thematik der Konstruktion von Interkulturalität und Fremdheit ließ also insbesondere auch der Erfolg und die Prominenz dieser Sendung sie interessant erscheinen für eine exemplarische Analyse von Fernsehshows. „Istanbul Total“ gehört zur Late Night-Show „TV Total“ von Stefan Raab und wurde an vier aufeinander folgenden Abenden von einem Istanbul Studio aus gesendet. Mit diesen Sendungen hat Stefan Raab zugleich den deutschen Wettbewerber des Eurovision Song-Contest 2004 mit dem Namen „MAX“ (Max Mutzke) promotet und in diesem Zusammenhang den Adolf-Grimme-Preis erhalten.<sup>102</sup>

Ausgewählt für die Interpretation wurde der erste der vier aufeinander folgenden Abende. Die ursprüngliche Abendsendung hat eine Dauer von 55 Min. bzw. von 35 Min. ohne Werbeblöcke und Einspielungen. Aus der dem Transkript zugrunde gelegten Internet-Fassung<sup>103</sup> sind diese herausgeschnitten. Von diesem ersten Abend wurde eine längere *Eingangsequenz*, d.h. die ersten beiden zentralen Szenerien im Umfang von 6 Min. und 21 Sek., ausgewählt (genauer zur Auswahl: 6.3.1 u. 6.3.3 sowie auch 5.8.3).

<sup>101</sup> Mit Stefan Hampl und Aglaja Przyborski, von denen das Videotranskriptionssystem MoViQ (siehe 5.7.2 sowie Anhang 7.1) entwickelt worden ist, hat insgesamt ein reger Austausch über die Methodenentwicklung stattgefunden. Ich danke den beiden für die spannenden und anregenden Diskussionen. Stefan Hampl (2008) stellt in seiner Dissertation die hier interpretierte Abendsendung in den Kontext weiterer Sendungen von „Istanbul Total“ und der gesamten Show „TV Total“ von Stefan Raab (siehe auch: 6.5.1).

<sup>102</sup> Den Preis erhielt Stefan Raab für die Entdeckung und Förderung von Musiktalenten durch die Casting-Show „SSDSGPS – Ein Lied für Istanbul“ (ProSieben), die Max Mutzke gewann. Die Begründung der Jury findet sich unter: <http://www.grimme-institut.de>

<sup>103</sup> <http://tvtotal.prosieben.de>

## 6.2 Videotranskript

Das Transkriptionssystem MoViQ („Movies and Videos in Qualitative Social Research“) wurde von Stefan Hampl und Aglaja Przyborski für die dokumentarische Film- und Videointerpretation entwickelt (genauer dazu: 5.7.2 sowie Anhang: Kap. 7.1).

Abbildung 27: Videotranskript Istanbul Total: 0:00 – 6.21. Entnommen aus: [http://tvtotal.prosieben.de/show/letzte\\_sendung/2004/05/10/00257.html#](http://tvtotal.prosieben.de/show/letzte_sendung/2004/05/10/00257.html#)

Passage (oder Sequenz):	Eingangspassage
Film (oder Video):	Istanbul Total
Datum der Endfassung:	25.09.2008
Time Code:	0:00:00 – 0:06:21
Dauer:	6:21 min.
Transkription:	Stefan Hampl
Korrektur:	Ralf Bohnsack

TC:	0:00	0:01	0:02	0:03	0:04
Am:	Ich befinde mich hier ähm auf dem Balkon unseres Istanbul Studios und wie				
Musik:					
Geräusch:					

TC:	0:05	0:06	0:07	0:08	0:09
Am:	man sehen kann hier hat man einen fantastischen Blick über die ganze Stadt hier ist der <u>Bosp-</u>				
Musik:					
Geräusch:					

TC:	0:10	0:11	0:12	0:13	0:14
Am:	orus hier geht's ins goldene Horn ja? Die Schiffe fahren rein sie fahren raus ein				
Musik:					
Geräusch:					

TC:	0:15	0:16	0:17	0:18	0:19
Am:	ständiges Heckmeck ja? Und äh da drüben äh kann man fantastisch äh ähm die Moscheen				
Musik:					
Geräusch:					

TC:	0:20	0:21	0:22	0:23	0:24
Am:	sehen da vorne ist die Hagia Sofia eine der ältesten				
Musik:					
Geräusch:					

TC:	0:25	0:26	0:27	0:28	0:29
Am:	Kirchen äh hier in Istanbul ühm glaub über tausend Jahre alt und				
Musik:					
Geräusch:					

TC:	0:30	0:31	0:32	0:33	0:34
Am:	äh daneben ist der Sultanspalast dort vorne auf dieser Anhöhe ja?				
Musik:					
Geräusch:					

TC:	0:35	0:36	0:37	0:38	0:39
Am:	und wenn man mal auf die andere Seite rüberschaut dann kann man den Bosphorus entlang-				
Musik:					
Geräusch:					

TC:	0:40	0:41	0:42	0:43	0:44
Am:	kucken und sieht dort hinten die Bosphorusbrücke die Asien und Europa miteinander				
Musik:					
Geräusch:					

TC:	0:45	0:46	0:47	0:48	0:49
Am:	verbindet ja? die einzige Stadt der Welt auf zwei Kontinenten				
Musik:					
Geräusch:					

TC:	0:50	0:51	0:52	0:53	0:54
Am:	u n d	da vorne kommt grade ein Schlepper an ja?			
Musik:					
Geräusch:					

TC:	0:55	0:56	0:57	0:58	0:59
Am:	das sieht man hier unten das sieht man sind immer die Boote mit den kann man das mal zeigen? die Boote mit diesen ä hm mit diesen Reifen				
Musik:					
Geräusch:					

TC:	1:00	1:01	1:02	1:03	1:04
Am:	die sch -	die schleppen immer die großen äh Schiffe deswegen gibt's auch			
Musik:					
Geräusch:					

TC:	1:05	1:06	1:07	1:08	1:09
Am:	hier den äh Spruch Nepper-Schlepper-Bauernfänger @@@			@ j a	und ich hier war vielleicht
Musik:					
Geräusch:			[mehrstimmiges Gelächter]		

TC:	1:10	1:11	1:12	1:13	1:14
Am:	was los ich weiß nicht wie das bei ihnen war ein fantastisches Fußballwochenende oder? ein großartiger Aus-				
Musik:					
Geräusch:					

TC:	1:15	1:16	1:17	1:18	1:19
Am:	wärtssieg und damit vorzeitig Meister	Herzlichen Glückwunsch Fenerbahçe Istanbul			
Musik:					
Geräusch:					[Applaus]

TC:	1:20	1:21	1:22	1:23	1:24
Am:			"übrigens"		
Musik:					
Geräusch:	[Applaus, Pfeifen und Johlen]				

TC:	1:25	1:26	1:27	1:28	1:29
Am:			Die Mannschaft von Christoph Daum		
Musik:					
Geräusch:	[Applaus, Pfeifen und Johlen]				

TC:	1:30	1:31	1:32	1:33	1:34
Am:		ja	die Mannschaft von Christof Daum		
Musik:					
Geräusch:					

TC:	1:35	1:36	1:37	1:38	1:39
Am:	ist vorzeitig Meister geworden hier in der Türkei Fenerbahçe liegt übrigens drüben auf der asia-				
Musik:					
Geräusch:					

TC:	1:40	1:41	1:42	1:43	1:44
Am:	tischen Seite ähh	da sind diese Anhöhen zu sehen dahinter dem ersten Hügel ist das Fener-			
Musik:					
Geräusch:					

TC:	1:45	1:46	1:47	1:48	1:49
Am:	bahçe-Stadion	und ähh	da war natü rlich gestern was los hier in der Stadt das können Sie sich gar nicht vorstellen überall		
Musik:					
Geräusch:					

TC:	1:50	1:51	1:52	1:53	1:54
Am:	ü ü ü ü ü	es wurde gehupt und gefeiert ich dachte im ersten Moment ich bin doch schon seit			
Musik:					
Geräusch:					

TC:	1:55	1:56	1:57	1:58	1:59
Am:	ein paar Tagen da warum erst heute die Freude ja? aber es handelte sich äh um die Fans von Fenerbahçe				
Musik:					
Geräusch:					

TC:	2:00	2:01	2:02	2:03	2:04
Am:	die dort gefeiert haben und äh hier ist einiges los ich weiß nicht wer schon mal in Istanbul war von Ihnen ich kann Ihnen				
Musik:					
Geräusch:					

TC:	2:05	2:06	2:07	2:08	2:09
Am:	nur empfohlen fahren Sie hier hin wenn sie hier kriegen sie einfach alles überall Händler auf der Straße die die				
Musik:					
Geräusch:					

TC:	2:10	2:11	2:12	2:13	2:14
Am:	Jacke auf machen und so sagen [sag mal brauchst du 12 Punkte für Max ja? isch habe dabei hier]				
Musik:					
Geräusch:	[Am spricht mit türkischem Akzent]				

TC:	2:15	2:16	2:17	2:18	2:19
Am:	und ähh	fantastisch wir ham ne Schifffahrt gemacht entlang des Bosphorus eine kleine Tour hin-			
Musik:					
Geräusch:					

TC:	2:20	2:21	2:22	2:23	2:24
Am:	ten wollt ma durch die Meerenge und sind leider stecken		geblieben äh weil Elton		
Musik:					
Geräusch:					

TC:	2:25	2:26	2:27	2:28	2:29
Am:	an Bord war	@ja@	aber wirklich man kann hier die tolle Sachen erleben fantastisches Wetter		
Musik:					
Geräusch:					

TC:	2:30	2:31	2:32	2:33	2:34
Am:	und ähh	viele glauben ja		dass Istanbul die Hauptstadt der Türkei	
Musik:					
Geräusch:					

TC:	2:35	2:36	2:37	2:38	2:39
Am:	ist das stimmt aber nicht die Hauptstadt der Türkei ist Ankara ja das ist ähn-				
Musik:					
Geräusch:					

TC:	2:40	2:41	2:42	2:43	2:44
Am:	ich wie in Deutschland wo viele ja glauben Berlin ist die Hauptstadt is aber Köln ja?				
Musik:					
Geräusch:	[lauter werdendes Motorengeräusch]				

TC:	2:45	2:46	2:47	2:48	2:49
Am:	Hallo!	ich glaube der machtn bisschen Stenz hier unten			
Musik:					
Geräusch:	[lauter werdendes Motorengeräusch]			[Motorengeräusch]	

TC:	2:50	2:51	2:52	2:53	2:54
Am:	das kennt man ja oder einfach mal Vollgas geben bei uns mitm 3er-BMW hier einfach mal mitm Schlepper ja?				
Musik:					
Geräusch:	[Motorengeräusch]				

TC:	2:55	2:56	2:57	2:58	2:59
Am:	siehste		hat ihn schwer motiviert	fantastisch ich bin froh	
Musik:					
Geräusch:	[schnellerer Motor]	[Motorengeräusch]			

TC:	3:00	3:01	3:02	3:03	3:04
Am:	dass wir hier mit TV Total es geschafft haben in Istanbul zu sein bevor der Musikantenstadt kommt ja				
Musik:					
Geräusch:	[Motorengeräusch]				

TC:	3:05	3:06	3:07	3:08	3:09
Am:	und äh ich	und wir ham im ähh	im was ist das	Norden	
Musik:					
Geräusch:					

TC:	3:10	3:11	3:12	3:13	3:14
Am:	Äh mal kucken was das ist (Fener) Ost-Nord-Ost in			Ost-Nord-Ost ist unser Hotel und ich	
Musik:					
Geräusch:					

TC:	3:15	3:16	3:17	3:18	3:19
Am:	kann Ihnen sagen wir ham ein fantastisches Hotel ja unsere komplette Crew hat einen herrlichen Blick				
Musik:					
Geräusch:					

TC:	3:20	3:21	3:22	3:23	3:24
Am:	auf meinen Pool ja das können Sie sich gar nicht vorstellen			wunderschön ja	
Musik:					
Geräusch:					

TC:	3:25	3:26	3:27	3:28	3:29
Am:	ssss ühmm	toll oder	Genießen Sie den Ausblick wir werden jetzt jeden Abend		
Musik:					
Geräusch:					

TC:	3:30	3:31	3:32	3:33	3:34
Am:	Montag, Dienstag, Mittwoch und	Donnerstag hier aus Istanbul senden		und ähh	viele von Ihnen wissen ja
Musik:					
Geräusch:					

TC:	3:35	3:36	3:37	3:38	3:39
Am:	auch in der deutschen Sprache kommt das Wort Istanbul oft vor zum Beispiel		ä h m C l a u d i a	S c h i f f e r i s t = a n =	
Musik:					
Geräusch:					

TC:	3:40	3:41	3:42	3:43	3:44
Am:	B u l i m i e e r k r a n k t j a ? o d e r		E l t o n i s t = a n = B u l e t t e n i n t e r e s s i e r t j a ?		
Musik:					
Geräusch:					

TC:	3:45	3:46	3:47	3:48	3:49
Am:	solche Sachen und	ich nehm mal Platz hier in unserer Ecke	fantastische	schöne	
Musik:					
Geräusch:					

TC:	3:50	3:51	3:52	3:53	3:54
Am:	wie setzt man sich hin im Schneidersitz?				Ährg
Musik:					
Geräusch:					

TC:	3:55	3:56	3:57	3:58	3:59
Am:	so	das hat ja schon mal ganz gut geklappt		Fantastisch, oder? Das ist ein	
Musik:					
Geräusch:		[Ping-Geräusch]		[Applaus]	

TC:	4:00	4:01	4:02	4:03	4:04
Am:	Studio so wie man sich das vorstellt ähh hinter uns ist eine Moschee neben der anderen und was wir da eben gesehen haben da				
Bf:				"ja"	
Musik:					
Geräusch:					

TC:	4:05	4:06	4:07	4:08	4:09
Am:	drüben hab ich eh ich hab eben gesagt das ist eine Kirche das war früher eine Kirche die Hagia Sofia das war äh				
Musik:					
Geräusch:					

TC:	4:10	4:11	4:12	4:13	4:14
Am:	lange Zeit eine Kirche und wurde dann zur Moschee umgewandelt falls der eine oder andere Klugscheißer am				
Musik:					
Geräusch:					

TC:	4:15	4:16	4:17	4:18	4:19
Am:	Fernseher sitzt und sagt das stimmt ja nit dass das ne Kirch is ja? sooo ihr wart ja äh schon				
Musik:					
Geräusch:		[Am spricht mit tiefer Stimme und kölnischem Akzent]			

TC:	4:20	4:21	4:22	4:23	4:24
Am:	auch hier äh ein paar Tage unterwegs und habt ein bisschen gedreht. Gülcan du sprichst ja die Sprache perfekt				
Bf:					Ja ich versuch's
Musik:					
Geräusch:					

TC:	4:25	4:26	4:27	4:28	4:29
Am:		Du bist aber in Deutschland geboren oder?			
Bf:	immer wieder			Ja ich bin in Lübeck geboren aber ähm äh Türkisch verlernt man ja	
Musik:					
Geräusch:					

TC:	4:30	4:31	4:32	4:33	4:34
Am:			Ach ok.		
Bf:	nicht mein ich konnte bis ich fünf war überhaupt kein Deutsch und äh wir ham zuhause nur Türkisch gesprochen und ähm im Kindergarten				
Geräusch:					

TC:	4:35	4:36	4:37	4:38	4:39
Am:		Wo kommen deine Eltern her aus welchem Teil der Türkei?			
Bf:	hab ich dann Deutsch gelernt so			Die kommen also genau hier wo wir gerade sind	
Geräusch:					

TC:	4:40	4:41	4:42	4:43	4:44
Am:	Hier aus?				
Bf:	Die sind hier aufgewachsen und ich kenn das alles hier auch die Straßen weil als ich mit meinen Eltern hier war die ham mir				
Geräusch:					

TC:	4:45	4:46	4:47	4:48	4:49
Am:					
Bf:	dann so die Straßen gezeigt und ja mein Papa so ja kuck mal da hat Mama gewohnt und da hab ich mal Steine geschmissen und				
Geräusch:					

TC:	4:50	4:51	4:52	4:53	4:54
Am:				acha	
Bf:	da hab ich mit Schokolade unten gewartet und so (ese) ja Papa schön schön schön aber das ham sie mir mal erzählt das is hier irgendwo				
Geräusch:					

TC:	4:55	4:56	4:57	4:58	4:59
Am:		acha und ihr wart zusammen ein bisschen einkaufen Elton?		Ähh habt auch mal euch umgeschaut ist ja	
Bf:	an den ecken			Ja	
Cm:				Ja	

TC:	5:00	5:01	5:02	5:03	5:04
Am:	ne sehr ähh moderne Stadt ja?		also eine eine wirkliche Metropole es gibt hier ein		
Bf:					
Cm:	Tatsächlich ja				

TC:	5:05	5:06	5:07	5:08	5:09
Am:	Viertel das sieht aus wie New York mit mit Hochhäusern dann gibt's hier wieder				
Musik:					
Geräusch:					

TC:	5:10	5:11	5:12	5:13	5:14
Am:	sehr traditionelle Viertel die Altstadt und so weiter äh und ihr habt euch mal ein bisschen eingekleidet oder				
Cm:					ja das könn ma uns
Geräusch:					

TC:	5:15	5:16	5:17	5:18	5:19
Am:	habt ihr denn was gefunden		also erst möchte ich nochmal das zeigen was ich auf dem Basar		
Cm:	auch gerne mal ansehen	auch für mich ja			
Geräusch:					

TC:	5:20	5:21	5:22	5:23	5:24
Am:	das müßma bevor wir diese MAZ zeigen wo ihr unterwegs wart äh was ich auf dem Basar gefunden				
Cm:		ja			
Geräusch:					

TC:	5:25	5:26	5:27	5:28	5:29
Am:	hab äh äh ja es is ja so hier findet man ja äh @viele@ Sachen Taschen von Dolce und Gabana, Gucci und alles				
Musik:					
Geräusch:					

TC:	5:30	5:31	5:32	5:33	5:34
Am:	ganz billig hab mich gefragt warum ist das so ja	Rolex Uhren	ja	Für Bruch-	
Cm:		@			
Geräusch:					

TC:	5:35	5:36	5:37	5:38	5:39
Am:	-teil für einen Bruchteil des Preises ja und auch äh äh äh Raubkopien von CDs gibt's hier einige				
Bf:		(Preis)			
Musik:					
Geräusch:					

TC:	5:40	5:41	5:42	5:43	5:44
Am:	ich hab hier einige gekauft schaunse mal	äh zum Beispiel die CD	@ja@	Sie kenn ja die CD	
Bf:		@@			
Musik:					
Geräusch:					

TC:	5:45	5:46	5:47	5:48	5:49
Am:	das ist se original	Original-CD von Max	ja kennse vielleicht	und dann hab ich	
Musik:					
Geräusch:					



TC:	5:50	5:51	5:52	5:53	5:54
Am:	hier aufm Basar mich mal umgeschaut und hab mal gefragt ob die das auch ham ja und ich hab zwei verschiedene				
Musik:					
Geräusch:					

TC:	5:55	5:56	5:57	5:58	5:59
Am:	gefunden einmal hab ich die hier gefunden schaunse mal		ja		
Musik:					
Geräusch:			[mehrstimmiges Gelächter]		

TC:	6:00	6:01	6:02	6:03	6:04
Am:	und		ja	hab ich gesagt is is it real äh	
Musik:					
Geräusch:	[mehrstimmiges Gelächter]				

TC:	6:05	6:06	6:07	6:08	6:09
Am:	I mean is it a legal copy yees of course und dann bin ich in nächsten Laden gegangen und da				
Bf:					@@@
Geräusch:					

TC:	6:10	6:11	6:12	6:13	6:14
Am:	wollten se ma das verkaufen ja				
Bf:				* @ C a n ' t w a i t ü n t y l	
Musik:					
Geräusch:		[mehrstimmiges Gelächter]		[mehrstimmiges Gelächter und Applaus]	

TC:	6:15	6:16	6:17	6:18	6:19
Am:			Vorsicht		ja
Bf:	t o n i g h t @ °				
Musik:					
Geräusch:	[Gelächter, Applaus und Gejöhle]			[Applaus]	

TC:	6:20	6:21	6:22	6:23	6:24
Am:	ansonsten aber alles korrekt				
Musik:					
Geräusch:	[Applaus]				

### 6.3 Interpretation in der Bilddimension

Die Interpretation in der Bilddimension erfolgt, wie in Kapitel 3 und 5.8.1 begründet, nach Möglichkeit zeitlich vor der Interpretation in der Dimension des Textes. Dabei kann die Interpretation der zur jeweiligen Bildsequenz dazugehörigen Textsequenz je nach Art und Genre des Films entweder nach der Bildinterpretation jeweils einer Untersequenz oder einer Hauptsequenz oder auch der gesamten Passage erfolgen. In der hier vorgelegten exemplarischen Interpretation habe ich letzteren Weg gewählt. Es wird also zunächst die gesamte Passage in der Bild- und erst dann in der Text- und Tondimension interpretiert.

Dies stellt sicherlich eine extreme Variante dar, welche lediglich für spezifische Genres von Videos und Filmen geeignet ist. Mit der hier realisierten extremen Variante vermag die Eigenart dieser Vorgehensweise allerdings besonders deutlich hervortreten. In der auf die Interpretation in der Bilddimension folgenden Textinterpretation wird dann bereits auf Homologien zwischen den beiden Dimensionen Bezug genommen, um schließlich in der reflektierenden Gesamtinterpretation (6.5) auf der Grundlage einer umfassenden Rekonstruktion dieser Homologien und deren Relationierungen den Film bzw. das Video in seiner Gesamtgestalt zu erschließen.

### 6.3.1 Zur Auswahl der Sequenzen

Wie in Kapitel 5.6.1 dargelegt, definiere ich als *Sequenz* jenen Ausschnitt eines Films oder Videos, der sich durch eine Kontinuität oder Identität im Bereich der abgelenkten *Szenen* oder *Szenarien* oder deren Teilen auszeichnet, also eine Kontinuität im Bereich dessen, was abgelenkt wird. Diese Identität betrifft also die Bewegungen oder Handlungen der *abgebildeten* Bildproduzent(inn)en bzw. Kontinuitäten im Bereich der abgebildeten Objekte. Die Art und Weise, wie diese Bewegungen oder Handlungen abgelenkt werden, betrifft im Wesentlichen *Einstellung* und *Montage* als das Produkt der *abbildenden* Bildproduzent(inn)en.

Solange wir jenseits des Wechsels in der Art und Weise, wie etwas abgelenkt wird, also durch den Einstellungswechsel und durch Veränderungen aufgrund von Montageleistungen hindurch, noch eine Kontinuität oder Identität im Bereich dessen, was abgelenkt wird, also eine Identität in der *Szenarie* oder Teilen der *Szenarie*, feststellen können (bspw.: „Stefan Raab auf dem Balkon eines Istanbuler Studios und der Blick auf den Bosphorus“), spreche ich von einer *Hauptsequenz (HS)*. Jene Ausschnitte, die – aufgrund von Einstellungswechsel und Montageleistungen – als spezifische Modifikationen der Hauptsequenz identifiziert werden können (bspw.: „Stefan Raab auf dem Balkon in halbnaher Einstellung versus Naheinstellung“), nenne ich *Untersequenz (US)*. Sobald fremde Szenarien in die Hauptsequenz eingelagert sind (bspw.: „Die Hagia Sophia am Bosphorus“), nach denen immer wieder zur Hauptszenarie zurückgekehrt wird, spreche ich von *eingelagerten Sequenzen (ES)*.

Sofern notwendig, wird zur Benennung der Sequenzen auf die Begriffe zurückgegriffen, die wir dem Textteil des Videotranskripts entnehmen (bspw. „Hagia Sophia“, „Istanbuler Studio“ etc.). Dazu reicht zumeist das Transkript selbst, ohne dass bereits auf die formulierende Interpretation zurückgegriffen werden muss (auch nicht auf der Ebene der thematischen Gliederung), wie sie erst später in 6.2.1 erstellt wird.

Die Auswahl der für die Interpretation relevanten Sequenzen und Fotogramme folgt, wie dies ganz allgemein bei der Auswertung auf der Grundlage der dokumentarischen Methode üblich ist, dem Kriterium der *Fokussierung* (genauer dazu auch: 5.8.2; vgl. auch: 4.3.1). Die Fokussierungen indizieren die herausgehobene Bedeutung dieser Sequenzen für die Produzent(inn)en, geben also Aufschluss über deren besondere (wie auch immer geartete) Involviertheit in die performative Präsentation und damit über deren Relevanzsetzungen und ermöglichen somit einen umfassenderen und in der Regel direkteren und somit ökonomischeren Zugang zu deren Orientierungsrahmen oder Habitus.<sup>104</sup>

Im Bereich der Filminterpretation betrifft dies sowohl die formalen Strukturen der Performanz der abbildenden Bildproduzent(inn)en (Einstellungsgröße, Kamerastandpunkt, Kamerabewegung) (vgl. Kap. 5.6.) wie auch der Performanz der abgebildeten Bildproduzent(inn)en (interaktiver Bezug aufeinander, Struktur der Gebärden) (vgl. Kap. 5.4).

<sup>104</sup> Fokussierungen oder fokussierte Passagen bieten also keineswegs den einzigen, wohl aber ‚besten‘ Zugang zu zentralen Komponenten des Orientierungsrahmen.

Fokussierungen finden ihren Ausdruck *zum einen* in *dramaturgischen Höhepunkten* im Bereich der kompositorischen und interaktiven Dichte. Derartige Fokussierungsmetaphern sind dann das Produkt (kontinuierlicher) dramaturgischer Steigerungen (bspw. zunehmender Steigerung der Einstellungsgröße durch Zoom oder zunehmender interaktiver Dichte der Bewegungen der abgebildeten Bildproduzent(inn)en).<sup>105</sup> *Zum anderen* finden Fokussierungen ihren Ausdruck in kompositorischen Auffälligkeiten. Sie sind dann das Produkt von Diskontinuitäten und Brüchen (bspw. überraschende Detailaufnahmen oder Bewegungsabbrüche), in denen ebenfalls eine (positive oder negative) Involviertheit der Bildproduzent(inn)en ihren Ausdruck finden kann (dazu: 6.3.5).

Für die hier vorliegende Art von Videodokument kann noch eine *dritte Ausprägung von Fokussierungen* geltend gemacht werden. Wir haben es hier mit einem Video zu tun, welches – als Alltagsdokument – in den jeweiligen kulturellen und organisatorischen Lebenszusammenhängen oder Milieus, die Gegenstand der sozialwissenschaftlichen Analyse sind, hier: der Fernsehanstalt Pro 7, insgesamt *selbst* produziert worden und als professionelles Produkt in deren organisatorische Planungszusammenhänge eingebunden ist.

Da die Erforschten hier in ihren Gestaltungsleistungen nicht durch die Interventionen der Forscher gerahmt sind (von denen sie sich allmählich lösen müssten), können ihre Relevanzsetzungen bereits darin zum Ausdruck kommen, *wie* sie selbst ihre Beiträge *eröffnen* und damit die Weichen für die gesamte visuelle Präsentation stellen (siehe dazu auch: 5.8.2). Somit gewinnen hier auch die *Eingangssequenzen* den Charakter von Indikatoren für Fokussierungen und Relevanzsetzungen.

Wir haben deswegen zunächst eine derartige Eingangssequenz für die Interpretation ausgewählt – sozusagen die Eingangssequenz der Eingangssequenz, d.h. die Eingangssequenz des ersten Abends von „Istanbul Total“. Eine weitere sinnvolle und begründete Auswahl aus der Fülle des Datenmaterials eröffnet uns jene gekürzte Version der Sendung, wie sie im Internet zu finden ist. Da die gekürzte Internet-Fassung vom Sender Pro 7 stammt,<sup>106</sup> bietet sich – analog zur Auswahl bei den Familienfotos (vgl. 4.3.1) – die Möglichkeit, den ersten Auswahlsschritt aus der Fülle des Datenmaterials den Bild- resp. Film-Produzent(inn)en selbst zu überlassen. Diese Auswahlen oder Kürzungen sind somit von den Produzent(inn)en authentisiert. Sie lassen sich in diesem Sinne sogar, wie am Beispiel der Familienfotos erläutert, als zunehmende Verdichtung der stilistischen oder orientierungsmäßigen Präferenzen der Produzent(inn)en – hier also des Senders – verstehen.<sup>107</sup>

Die ursprüngliche Sendung des ersten Abends hat, wie gesagt, eine Dauer von 55 Minuten und die hier zugrunde gelegte Internet-Fassung ohne Werbeblöcke und Einspielungen 35 Minuten. Von diesem ersten Abend wurde die Eingangspassage

<sup>105</sup> Für letzteren Bereich hat Iris Nentwig-Gesemann (2006a u. 2007a) im Sinne einer Differenzierung des Begriffs der Fokussierungsmetapher denjenigen der „Fokussierungsakte“ geprägt.

<sup>106</sup> [http://tvtotal.prosieben.de/show/letzte\\_sendung/2004/05/10/00257.html#](http://tvtotal.prosieben.de/show/letzte_sendung/2004/05/10/00257.html#)

<sup>107</sup> Ein bei der Produktionsgesellschaft (Brainpool TV GmbH) erworbener Mitschnitt auf DVD der *vollständigen* ursprünglichen Sendung gibt uns Aufschluss darüber, dass in der Internet-Fassung neben dem Werbeblock die im Studio stattfindende Begrüßungszeremonie zwischen St. R. und den Studiogästen sowie eine Einspielung herausgeschnitten ist, welche St. R. selbst als „Exkurs auf den Bazar in Istanbul“ bezeichnet.

von 6 Min. und 21 Sek. für die Interpretation ausgewählt. Innerhalb dieses Ausschnitts lassen sich drei Hauptsequenzen (HS) identifizieren:

- *Hauptsequenz I: 0:00-3:33*  
„Stefan Raab auf dem Balkon des Istanbul Studios am Bosphorus und der Blick auf den Bosphorus“
- *Hauptsequenz II: 03:34-03:57*  
„Stefan Raab stehend und sich setzend vor der Sitzecke des Studios am Fenster mit Blick auf den Bosphorus“
- *Hauptsequenz III: 03:58-06:21*  
„Stefan Raab sitzt mit Gülçan und Elton auf Kissen in der Sitzecke des Studios“

### 6.3.2 Formulierende Interpretation von Sequenzen und von Einstellungswechsel und Montage

#### Methodische Vorbemerkung

Analog zum thematischen Verlauf als elementarem Schritt der formulierenden Interpretation von Texten geht es hier bei der Videotranskription um die Beschreibung des Verlaufs von Sequenzen als elementarem Schritt der formulierenden Interpretation in diesem Bereich. Dazu gehört auch die Beschreibung der Zuordnung, der Relationen der Sequenzen untereinander, wie sie durch Einstellungswechsel und Schnitt (Bildmischung) hergestellt werden. Erst durch die Beschreibung dieser Relationen wird die Identifikation von Hauptsequenzen (HS) und Untersequenzen (US) und auch eingelagerten Sequenzen (ES) möglich. Bei der Beschreibung der Relationen der Sequenzen untereinander bedienen wir uns einer Terminologie, mit der wir weitgehend an die Filmwissenschaft anzuschließen suchen (siehe Kap. 5.6 sowie Anhang 7.3).

Allerdings sind das Erkenntnisinteresse und die Gegenstandsbereiche der Filmwissenschaften andere als die einer sozialwissenschaftlichen Analyse. Wenn diese, wie dies bei der dokumentarischen Interpretation der Fall ist, auf den Habitus oder Stil der Produzent(inn)en zielt, ist es nicht notwendig, tiefer gehend in die *technische* Herstellung im Bereich der Kameraführung, Einstellung und Montage einzudringen. Wie bereits gesagt (vgl. Kap. 5), ist für die Interpretation nicht ausschlaggebend, welcher technischer Instrumente sich der Gestaltungswille der abbildenden Bildproduzent(inn)en (während und nach der filmischen Aufzeichnung) bedient hat. Entscheidend sind das gestaltete Produkt und der sich in ihm dokumentierende Habitus.<sup>108</sup>

Zugleich mit der Identifizierung von Sequenzen wird auch die Gestaltung des Filmausschnitts durch die abbildenden Bildproduzent(inn)en, wie sie durch die *Kameraführung* und durch *Schnitt und Montage* bzw. (im Falle des Fernsehens) *Bildmi-*

<sup>108</sup> Dabei können sich allerdings in der *spezifischen Selektivität* der technischen Herstellung (*dass* bspw. ein Zoom gewählt wurde und nicht etwa eine Überblendung) Komponenten des Habitus dokumentieren – getreu der Mannheimschen Maxime: „Nicht das ‚Was‘ eines objektiven Sinns, sondern das ‚Daß‘ und das ‚Wie‘ wird von dominierender Wichtigkeit.“ (Mannheim 1964a: 134).

*schung* hergestellt wird, (auf der Ebene der formulierenden Interpretation in der Bilddimension) beschrieben. Zur Gestaltung durch die Kamera gehört zum einen die *Einstellungsgröße*, also die Detail-, Groß- und Naheinstellung, die halbnah, amerikanische, halbtotale, totale und die Weit- oder Panorama-Einstellung (siehe dazu u.a.: Hickethier 1996: 58f.; Borstnar/Pabst/Wulff 2002: 91; Kuchenbuch 2005: 44 sowie im Anhang: 7.3). Zum anderen gehört die Wahl des *Kamerastandpunkts* dazu (Differenzierung von Normal-, Auf- und Untersicht; vgl. u.a. Hickethier 1996: 61f.) und zum dritten die Art der *Kamerabewegung* (Kameraschwenk, Fahraufnahme, Zoom).

Sofern notwendig, wird zur Identifikation der Kamerabewegung (bzw. zur Unterscheidung von Kamerabewegung und Schnitt) auch der Modus der Zeitlupenwiedergabe herangezogen oder andere Techniken, mit denen die Wiedergabe in Zeitverzögerung möglich ist.

#### Ad Hauptsequenz I: 0:00-3:33

„Stefan Raab auf dem Balkon des Istanbul Studios am Bosphorus und der Blick auf den Bosphorus“

Diese Hauptsequenz ist die umfangreichste. Sie wird in unterschiedlichen Einstellungsvarianten – hier: Variationen der *Einstellungsgröße* – realisiert, sodass sich folgende *Untersequenzen* (US) identifizieren lassen:

US 0:00 u. 0:16-0:23(24); 0:49-0:56; 1:05-1:07; 1:09-1:21; 1:28-1:43; 1:48-2:18; 2:23-2:24; 2:31-2:32; 2:35-2:45; 2:49-3:08 als häufigste Einstellungsvariante: Amerikanische Einstellung bzw. halbnah Einstellung (Oberkörper ohne Beine)

US 0:01-0:08 u. 2:19-2:22; 2:33-2:34 als zweithäufigste Einstellungsvariante: Totale

US 2:25-2:26; 2:28-2:30; 2:18 als dritthäufigste Einstellungsvariante: Übergang von der halbnahen Einstellung zur Naheinstellung (Brustbild).

*In die Hauptsequenz eingelagerte Sequenzen: ES*  
(Umfang insgesamt: 33 Sek.)

ES „Brücke am Bosphorus“ 0:12 -0:15 Weit (Panorama) // 0:11-12 Kameraschwenk bis 0:14-15 Überblendung //

ES „Hagia Sophia“ 0:24-0:28 Totale // 0:23-0:24 Überblendung bis 0:28-29 harter Schnitt //

ES „Sultanspalast“ 0:34-0:38 Totale // 0:34 Überblendung bis 0:38/39 harter Schnitt //

ES „Bosphorus“ 0:44-0:48; 1:44-47 Weit (Panorama) // 0:43-44 Überblendung bis 0:48-49 harter Schnitt // 1:43-44 harter Schnitt bis 1:47-48 harter Schnitt. – Der Wegführung von der Hauptszenarie geht im letzteren Fall ein Zeigen (und die verbale Bezugnahme) von St. R. voran//

ES „Schlepper“ 0:57-1:04; 2:46-2:48 Totale bzw. Halbtotale // 0:56-57 harter Schnitt bis 1:04-05 harter Schnitt. – Der Wegführung von der Hauptszenarie geht ein Zeigen von Stefan Raab in 0:00:56 voran // 2:45-46: harter Schnitt bis 2:48: harter Schnitt. – Der Wegführung von der Hauptszenarie geht der Wechsel der Blickrichtung (Drehen des Kopfes und Oberkörpers von Stefan Raab) in 2:45 vor-

an, ebenso wie der Rückführung zur Hauptszenerie //– Da der Schlepper vorher und hinterher im Hintergrund im Bild ist, könnte dies auch als Untersequenz (US) kategorisiert werden.

ES „Das im Studio sitzende Publikum“ 1:08 Halbnah; 1:22-1:27 Totale; 2:27 Nah // jeweils harte Schnitte hin und zurück. – Der Wegführung von der Hauptszenerie geht der Off-Ton des applaudierenden Publikums voran. //

#### **Ad Hauptsequenz II: 03:34-57**

„Stefan Raab stehend und sich setzend vor der Sitzecke des Studios am Fenster mit Blick auf den Bosphorus“

Diese Hauptsequenz ist die am wenigsten umfangreiche, und hier finden sich nur wenige Einstellungsvarianten. Die Einstellungsgröße variiert lediglich zwischen der Halbtotale und der Totalen. Diese Sequenz (neue Szenerie) wird durch einen Kameraschwenk vom Balkon in den Innenraum eingeführt (03:33-36), bei dem die Kamera der Bewegung von Stefan Raab folgt, und durch einen harten Schnitt (3:57) beendet.

*In die Hauptsequenz eingelagerte Sequenzen ES: keine*

(3:57 ist schwarz. An dieser Stelle sind in der hier zugrunde liegenden Internet-Version die Einspielung und der Werbeblock herausgeschnitten)

#### **Ad Hauptsequenz III: 03:58-06:21**

„Stefan Raab sitzt mit Gülçan und Elton auf Kissen in der Sitzecke des Studios“

Diese Hauptsequenz ist die zweitumfangreichste. Die häufigste Einstellungsvariante ist die Totale: 3:58-4:05; 4:09-4:12; 4:19-4:21; 4:26 u. 27; 4:34-4:37; 4:44-4:47; 4:55-5:00; 5:03-5:09; 5:12-5:16; 5:19-5:47; 5:50-5:56; 6:00; 6:03-6:10; 6:18-6:21

Die anderen Einstellungsvarianten stellen erkennbar Ausschnitte dieser Totalen dar und haben somit den Charakter von Untersequenzen (US). Bei dieser Einstellung in der Totale handelt sich um eine „Einstellung zu Beginn einer Sequenz, die einen allgemeinen Überblick über Lokalität, Personal und Situation gibt“ (Monaco 2003: 58). Sie wird als „Establishing Shot“ oder „Master Shot“ bezeichnet.<sup>109</sup>

*Folgende Untersequenzen (US) lassen sich identifizieren:*

US 3:58: „Gülçan und Elton“: Halbtotale // harte Schnitte hin und zurück //

US 4:22-4:33; 4:38-4:43; 4:48-4:54: „Gülçan“: Groß // Jeweils harter Schnitt von der Totalen auf die Gruppe und zurück //

<sup>109</sup> Den Master Shot definiert James Monaco (2003: 103) in ähnlicher oder identischer Weise wie den Establishing Shot als „meist in der Totale oder Halbtotale durchgedrehte Einstellung, die beim Schnitt als Grundlage für eine Szene genommen werden kann“. -- Indem hier die Begriffe „Sequenz“ und „Szene“ nahezu synonym verwendet werden, haben wir hier im übrigen wieder ein Beispiel für eine Unklarheit in der Verwendung dieser beiden Begriffe in der filmwissenschaftlichen Literatur (siehe auch: 6.2.1), auf die Monaco (2003: 159) allerdings selbst kritisch hinweist.

US 4:51 Stefan Raab: Groß // jeweils harte Schnitte von der Totalen auf die Gruppe und zurück // Es handelt sich um die einzige Groß-Ansicht von Stefan Raab. Sie steht in einer Reihe von Groß-Ansichten von Gülçan.

US 5:01-5:02; 5:10-5:11; 5:17-5:18: „Elton“: Groß // Jeweils harter Schnitt von der Totalen auf die Gruppe und zurück //

US 5:48-5:49; 5:57-5:59; 6:11-6:14 „CD-Hülle ‚MAX‘“: Detail (Es handelt sich um die einzige Detail-Ansicht der gesamten Passage) // Jeweils harter Schnitt von der Totalen (auf die Gruppe) und zurück //

*In die Hauptsequenz eingelagerte Sequenzen: ES*

ES 4:06-07; 4:13-4:15: „Hagia Sofia“: Totale // 4:06-07: harter Schnitt von der Totalen auf die Gruppe hin und Zoom auf die Gruppe zurück. Der Zoom dauert bis 0:04:21//

ES 6:01-02; 6:15-17: „Zuschauer/innen“: Groß // jeweils harte Schnitte hin und zurück //

### *6.3.3 Zur Auswahl der Fotogramme*

Innerhalb dieser durch ihren Eröffnungscharakter fokussierten Sequenzen galt es nun, eine begründete Auswahl von Fotogrammen zu treffen, um den Aufwand bei deren Interpretation in Grenzen zu halten. Hier folge ich *zum einen* der Auswahlstrategie der *Repräsentanz*: Die für die Auswertung relevanten Fotogramme werden danach ausgewählt, inwieweit sie geeignet sind, die umfangreichsten (Haupt-) Sequenzen der Eröffnung zu repräsentieren (vgl.: 6.3.2). Das Auswahlkriterium der Fotogramme, die für eine intensive Einzelbildinterpretation herangezogen werden, orientiert sich somit daran, dass diese Fotogramme zumindest die beiden umfangreichsten Hauptsequenzen (HS) in deren umfangreichsten Einstellungsvarianten, also in den Untersequenzen (US), repräsentieren sollten.

*Zum anderen* und zusätzlich können und sollten innerhalb dieser Eröffnungssequenzen Fotogramme bzw. Untersequenzen nach dem Prinzip der Fokussierungen aufgrund *dramaturgischer Steigerungen* wie auch aufgrund von *Diskontinuitäten* und *Brüchen* ausgewählt werden. Da der vorliegende Videoausschnitt relativ wenig ausgeprägte dramaturgische Steigerungen aufweist, orientierte sich die Auswahl vor allem an der Fokussierung auf der Grundlage von Diskontinuitäten (siehe Kap. 6.3.5).

Ausgewählt für eine genaue ikonologisch-ikonische Interpretation wurden also innerhalb der Eröffnungssequenzen solche Fotogramme, welche die umfangreichste Hauptsequenz in ihrer häufigsten Einstellungsvariante (HS IA) und in ihrer zweithäufigsten Einstellungsvariante (HS IB) sowie die vom Umfang her an zweiter Stelle stehende Hauptsequenz in ihrer häufigsten Einstellungsvariante (HS III) zu repräsentieren vermögen.

Die beiden ausgewählten Fotogramme der umfangreichsten Hauptsequenz in ihrer häufigsten und in ihrer zweithäufigsten Einstellungsvariante (HS IA u. HS IB) folgen direkt aufeinander, (0:00 zu 0:01) und ihr Wechsel bietet sich somit zugleich

für eine genauere reflektierende Interpretation von Einstellungswechsel und Montage an (vgl. 6.3.4.).<sup>110</sup>

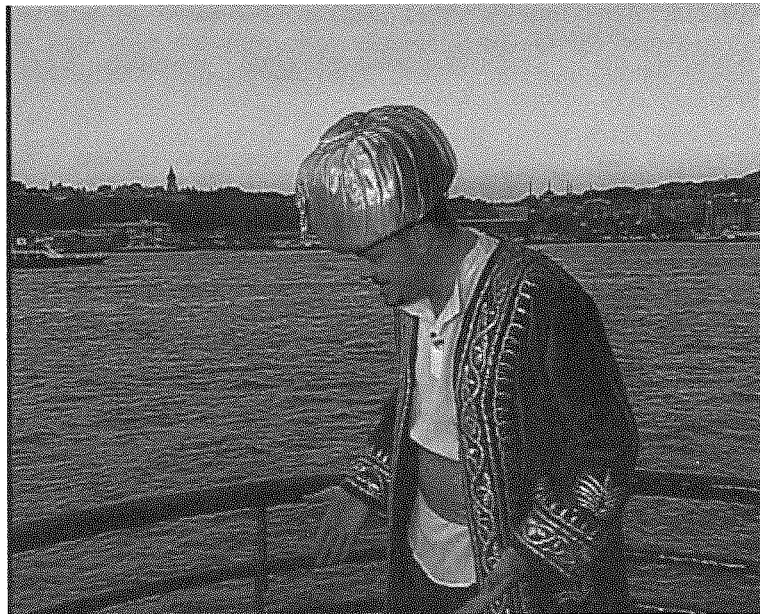
#### 6.3.4. Formulierende und Reflektierende Interpretation der Fotogramme

##### Fotogramm IA: 0:00

Hauptsequenz: „Stefan Raab auf dem Balkon“

in der häufigsten Einstellungsvariante: Amerikanisch bzw. Halbnahe

Abbildung 28: Istanbul Total: Fotogramm IA. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD



##### Formulierende Interpretation

##### Vor-ikonografische Interpretation

Zu Beginn der gesamten Abendsendung ist der Oberkörper eines auf einem Balkon stehenden, dem Zuschauer seitlich zugewandten, Mannes bis zur Hüfte zu sehen. Er deutet mit seinen geöffneten Händen auf den Boden und senkt auch seinen Kopf und Blick in diese Richtung. Hinter dem Mann befindet sich eine nach außen ge-

<sup>110</sup> Während das Videotranskript auf der Grundlage der Internet-Version erstellt wurde, liegt den Fotogrammen ein bei der Produktionsgesellschaft (Brainpool TV GmbH) erworbener Mitschnitt auf DVD der vollständigen ursprünglichen Sendung zugrunde, da dessen höhere Bildauflösung eine bessere Qualität der vergrößerten Fotogramme liefert).

wölbte und mit mehreren massiven senkrecht verlaufenden Rohren gitterartig konstruierte Brüstung. Dahinter sehen wir eine ca. zwei Drittel des Bildes umfassende große blaue Wasserfläche mit Wellen. Ungefähr auf Kopfhöhe des Mannes ist in der Ferne am Horizont ein schmaler Küstenstreifen mit Häusern zu erkennen und darüber hellblauer Himmel.

Der Mann trägt eine silberne glänzende topfförmige Kopfbedeckung aus gewickeltem Tuch und einen langen weiten Umhang aus braunem Samt, der entlang seiner vorderen Öffnung und der Manschetten mit goldenen Stickereien verziert ist. In der Öffnung ist ein weißes Hemd zu sehen, um welches eine rosafarbene Bauchbinde geschlungen ist. Unterhalb dieser fällt das Hemd über eine violett schimmernde Hose.

##### Ikonografische Interpretation

Stefan Raab (im Folgenden: St. R.) befindet sich auf einem Balkon des Istanbul Studios von Pro 7. Er ist bekleidet mit einem „Sarik“ (im deutschen „Turban“ genannt), welcher in öffentlichen Institutionen verboten ist und lediglich bei folkloristischen Veranstaltungen getragen wird, einem „Mintan“, einem Hemd, das aus der osmanischen Periode stammt, der „Şalvar“, einer Hose, welche heute noch auf dem Dorf zu finden ist, und dem „Kaftan“, der Weste, die heute ebenfalls bei Folkloreveranstaltungen getragen wird.<sup>111</sup>

Im Hintergrund ist der Bosphorus zu sehen.

##### Reflektierende Interpretation

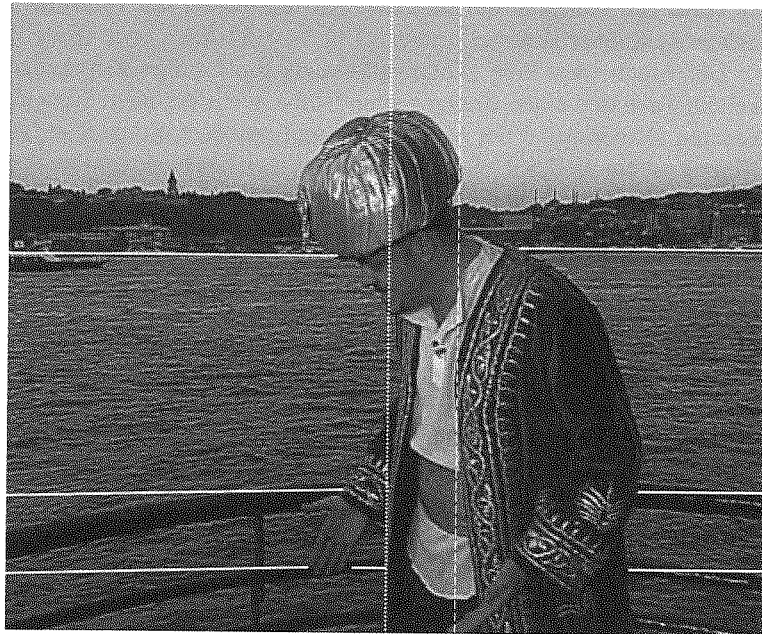
##### Formale Komposition

##### Perspektivität:

Das *perspektivische Zentrum* des Bildes ist kaum zu erschließen. Und da St. R. die einzige auf dem Bild abgebildete Person ist, entfallen auch die Ausführungen zur *szenischen Choreografie*. Es ist allerdings für sich genommen bereits aufschlussreich, dass St. R. in der umfangreichsten Hauptsequenz (sowie auch im ersten Fotogramm) allein abgebildet ist, da sich allein hierin eine Zentrierung auf seine Person dokumentiert.

<sup>111</sup> Der Sarik ist seit dem Gesetz zur Kleidung aus der Frühzeit der Republik in öffentlichen Institutionen verboten. Mintan ist in letzter Zeit – vor allem bei Islamisten – wieder ein wenig Mode geworden. Auch Şalvar ist neuerdings bei modernen, sich alternativ kleidenden jungen Frauen zu finden. – Diese Informationen verdanken wir Kolleg(inn)en, die aus der Türkei stammen und/oder sich regelmäßig dort aufhalten.

Abbildung 29: Istanbul Total: Fotogramm IA: Planimetrie. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.)



#### Planimetrie:

St. R. ist auch planimetrisch insofern fokussiert – und das gilt für alle Bilder in dieser häufigsten Einstellungsvariante der umfangreichsten Hauptsequenz – als seine Körperachse auf dem Goldenen Schnitt<sup>112</sup> liegt (gestrichelte senkrechte Linie) und sein Kopf auf der senkrechten Bildmittelachse (gepunktete senkrechte Linie).

Auffällig ist, dass das gesamte Fotogramm durch das Balkongitter planimetrisch zerteilt ist. Unterstrichen wird diese Teilung durch zwei bzw. drei weitere waagerechte Linien, die durch den Landstreifen am Horizont gebildet werden. Das Bild ist somit in vier waagerechte Blöcke oder Streifen zerschnitten: Balkonbrüstung, Meer, Landstreifen am Horizont, Himmel. Man kann sagen: das weist eine planimetrische Struktur auf, die es in ästhetischer Hinsicht zerfallen lässt. – Einen gewissen Zusammenhalt erhält das Bild lediglich durch St. R., also dadurch, dass St. R. mit seinem Oberkörper alle drei Streifen verbindet. Sein Sarik thront sozusagen auf dem Horizont. Das Bild ist also vollkommen auf ihn ausgerichtet.

#### Ikonologisch-ikonische Interpretation

St. R. steht im Vordergrund und im planimetrischen Zentrum. Die Umgebung, die Landschaft, bildet lediglich den Hintergrund. In der Einstellung wurde Wert darauf

<sup>112</sup> Zum Goldenen Schnitt siehe: Kapitel 4.2 (Anmerkungen).

gelegt bzw. wurde in Kauf genommen, dass das Gitter der Balkonbrüstung im Bild ist, obschon damit in ästhetischer Hinsicht ein Bruch verbunden ist.<sup>113</sup> St. R. ist von dieser Umwelt durch das Gitter, welche das Bild dominiert, getrennt. Man gewinnt den Eindruck, dass ihn dieses Gitter vor der Umwelt schützt und/oder ihn hindert oder davor bewahrt, sich so ohne Weiteres dieser Umwelt zu nähern. Da diese Szenerie in dieser Einstellungsvariante das häufigste Bild darstellt und das Gitter auch in der zweithäufigsten Einstellungsvariante abgebildet ist, bleibt dieses insgesamt 5 Min. 57 Sek., also im weitaus überwiegenden Teil des interpretierten Ausschnitts, im Bild.

Es wird zwischen St. R. und der Umwelt aber nicht nur eine *Grenzlinie* gezogen, sondern es wird auch markiert, dass St. R. eine Position *oberhalb* dieser Umgebung einnimmt, von oben herab blickt. Denn auch wenn das Gitter nicht vorhanden wäre, befände er sich nicht in der Welt des Bosphorus, d.h. nicht auf Augenhöhe mit ihren Bewohnern. Das Gitter und die Position von St. R. zusammen mit dem Hintergrund des Meeres erwecken auch den Eindruck, dass er an der Reling oder auf der Brücke eines Schiffes steht, sich also gar nicht in Istanbul befindet.

Insgesamt dokumentiert sich hier also, dass St. R. weniger in der Türkei bzw. in Istanbul sich befindet, als vielmehr aus der Ferne und von oben auf die Türkei bzw. Istanbul blickt und zugleich aber auch von dieser Umwelt abgegrenzt bzw. vor ihr geschützt wird. Genauer betrachtet, schaut er aber in dieser allerersten Szene nicht einmal von oben auf die Türkei, sondern, da der Blick nach unten gerichtet ist, auf seinen eigenen Standort. Wir haben es auf diesem Bild also mit einer *gegenläufigen Bewegung* bzw. mit einer *Übergegensätzlichkeit* zu tun: St. R. begibt sich in die Welt von Istanbul bzw. umgibt er sich mit dieser, wird aber zugleich von dieser getrennt oder distanziert sich von ihr bzw. bleibt er auf sich selbst bezogen. Diese Welt ist somit nicht etwas, mit dem man in Interaktion und Kommunikation tritt. In diesem Sinne bleibt die Umwelt in gewisser Weise Kulisse und Dekoration.

Hier zeigen sich auch Homologien zur Bekleidung von St. R.. Denn diese repräsentiert keineswegs den Alltag in der Türkei. Sarik und Kaftan finden sich auch in der Türkei regelmäßig lediglich auf der Theaterbühne oder ggf. bei Folkloreveranstaltungen. Somit ist auch der mit dieser Verkleidung hergestellte Bezug zur Welt Istanbul oder der Türkei als Dekoration markiert. Man kann auch sagen, dass eine *Annäherung* an die Welt der Türkei durch Kleidung bzw. Kostümierung mit kulturellem bzw. regionalem Bezug auf dem Wege der ironisierenden Übertreibung und der Exotisierung den Charakter einer Distanzierung annimmt.

#### Fotogramm IB: 0:01

Hauptsequenz: „Stefan Raab auf dem Balkon“  
in der zweithäufigsten Einstellungsvariante: Totale

<sup>113</sup> Es soll hier natürlich nicht übersehen werden, dass der Einsatz von Kameras, die für den Studiobetrieb eingerichtet sind, erhebliche technische Restriktionen mit sich bringt, sodass es auch möglich ist, dass hier etwas „in Kauf“ genommen wird. Die Tatsache aber, dass es in Kauf genommen wurde – und zwar durchgängig – und nicht eine ganz andere Konzeption gewählt wurde, ist aufschlussreich. Bourdieu (1983a: 17) hat mit Bezug auf die Amateurfotografie bemerkt, „daß noch die unbedeutendste Photographie neben den expliziten Intentionen ihres Produzenten das System der Schemata des Denkens, der Wahrnehmung und der Vorlieben zum Ausdruck bringt, die einer Gruppe gemeinsam sind.“

Abbildung 30: Istanbul Total: Fotogramm IB. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD



### Formulierende Interpretation

#### Vor-ikonografische Interpretation

Dieser Schritt ist hier aus Raumgründen nicht abgedruckt.

#### Ikonografische Interpretation

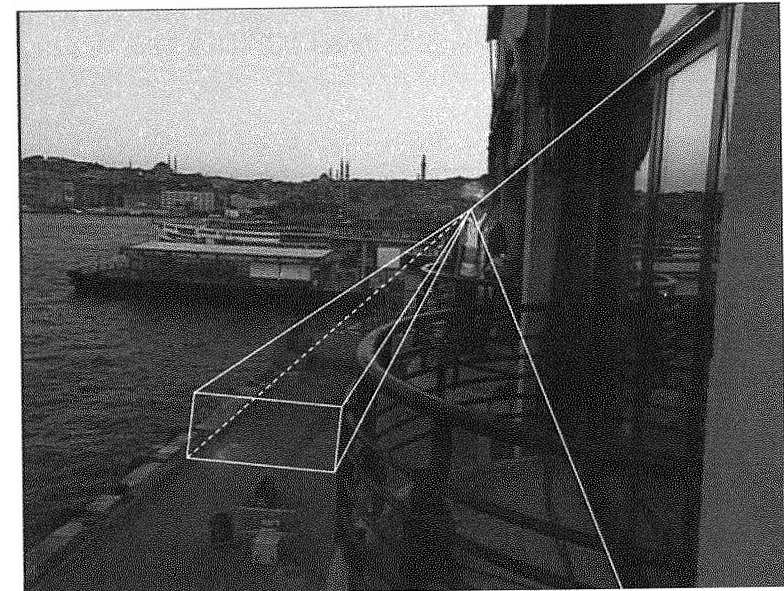
St. R. befindet sich auf einem Balkon des Istanbul Studios von Pro 7. Er ist bekleidet mit einem Sarik und einem traditionellen Gewand offensichtlich türkischen Ursprungs. Links unterhalb vom Balkon befindet sich ein Hafengelände mit einem Quai im Vordergrund und im Mittelgrund. Dort haben Schiffe festgemacht. Im Hintergrund sieht man ein dicht besiedeltes Ufer des Bosphorus.

### Reflektierende Interpretation

#### Formale Komposition

Perspektivität:

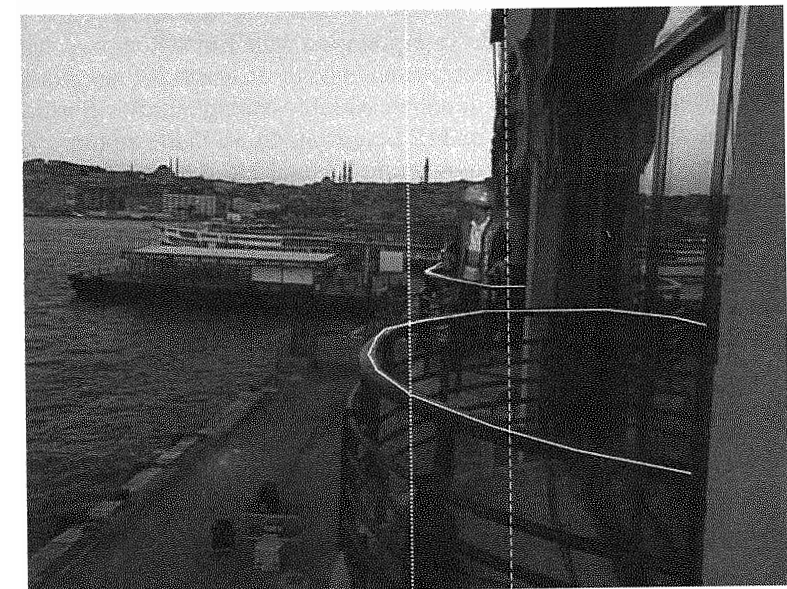
Abbildung 31: Istanbul Total: Fotogramm IB: Perspektivität. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.)



St. R. bzw. genauer: sein Kopf ist im perspektivischen Zentrum positioniert.

#### Planimetrie:

Abbildung 32: Istanbul Total: Fotogramm IB: Planimetrie (Linien von R.B.)



St. R. steht auch im planimetrischen Zentrum: sein Körper befindet sich im Goldenen Schnitt (gestrichelte Linie) bzw. zwischen Goldenem Schnitt und senkrechter Bildmittellachse (gepunktete Linie). Ansonsten wird auch dieses Bild durch die Balkongitter in gewisser Weise zerschnitten. Diese drängen in massiver Weise ins Bild, zumal sie sich im Fensterglas spiegeln. Es findet sich somit keine in sich geschlossene planimetrische Komposition, sodass auch hier das Bild in ästhetischer Hinsicht zerfällt.

#### *Ikonologisch-ikonische Interpretation*

Die gegenläufige Bewegung, die Übergegensätzlichkeit (die Hinwendung zur Umwelt und zugleich die Abgrenzung ihr gegenüber), die bereits in Bild 0:00 herausgearbeitet werden konnte, dokumentiert sich deutlicher noch in diesem Fotogramm: St. R. befindet sich auch hier am Bosphorus und mehr noch relativ dicht am Hafengeschehen, also fast im türkischen Alltag von Istanbul, ist aber zugleich durch die das Bild beherrschenden Gitter, die hier noch massiver ins Bild gerückt werden, noch deutlicher von diesem getrennt als in der vorherigen Einstellung. Die Balkonbrüstung im Vordergrund nimmt mehr als  $\frac{1}{4}$  des gesamten Bildes ein. Sie erscheint wie eine Art Abschirmung, Rüstung oder Bollwerk und das Studio somit wie eine Bastion. Zugleich wird noch deutlicher erkennbar, dass St. R. *von oben herab* auf den Bosphorus und das Geschehen am Hafen blickt. Auch wird hier die Distanz zwischen ihm und dem Bosphorus und Istanbul dadurch markiert, dass kein Weg nach unten führt. Es gibt lediglich den Weg zurück ins Studio.

#### *Zur Reflektierenden Interpretation des Einstellungswechsels von Fotogramm IA zu Fotogramm IB*

Obschon die reflektierende Interpretation von Einstellungswechsel und Montage systematisch und im Gesamtzusammenhang erst in 6.3.4 ausgearbeitet wird, empfiehlt es sich aufgrund der direkten Anschlussfähigkeit dies für die spezifische Relation der Fotogramme IA und IB bereits an dieser Stelle zu tun.

Innerhalb der beiden Fotogramme IA und IB, d.h. in ihren *internen Relationen*, dokumentiert sich auf der Grundlage der Einzelbildinterpretation eine gegenläufige Bewegung, eine Übergegensätzlichkeit, bei der eine Hinwendung zur Umwelt zugleich mit einer Abgrenzung ihr gegenüber einhergeht. Diese gegenläufige Bewegung oder Übergegensätzlichkeit wird (wenn wir die *externe Relationierung* der Fotogramme betrachten) durch den Einstellungswechsel von Fotogramm IA zu IB noch einmal verstärkt: Denn mit diesem Wechsel von der amerikanischen Einstellung zur Totale vollzieht sich (mit der Technik des Zooms) einerseits eine weitergehende Öffnung hin zur Umwelt, die sich von St. R. löst bzw. über ihn hinaus zur Umwelt des türkischen Bosphorus führt, die aber konterkariert bzw. geradezu ins Gegenteil verkehrt wird dadurch, dass damit zugleich das Gitter sich noch massiver in das Bild schiebt.

#### **Fotogramm III: 0:05:28**

Hauptsequenz mit dem zweitgrößten Umfang: „Stefan Raab sitzt mit Gülçan und Elton auf den Kissen der Ostecke“

Einstellungsgröße: Totale (Establishing Shot oder Master Shot)

Abbildung 33: Istanbul Total: Fotogramm III. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD



#### *Formulierende Interpretation*

#### *Vor-ikonografische Interpretation*

Dieser Schritt ist hier aus Raumgründen nicht abgedruckt.

#### *Ikonografische Interpretation*

St. R. sitzt in der bereits beschriebenen Bekleidung in einer Ecke, die mit Sitzkissen ausgelegt ist. Diese befinden sich auf einem ca. 40 cm hohen Podest im Istanbul Studio von Pro 7 vor großen Fenstern, aus denen sich der Blick auf den Bosphorus öffnet. Es handelt sich um eine für das traditionelle türkische Haus typische Sitzkecke, die auch als „orientalische“ bezeichnet wird, da sie sich im Osten des Hauses befindet. Direkt neben ihm, zu seiner Linken, sitzt Gülçan (Gülçan Karahançi) und links neben dieser mit etwas Abstand Elton (Alexander Duszat). Gülçan ist 1982 in Lübeck geboren. Ihre Eltern sind türkischer Herkunft und stammen aus Istanbul. Sie ist als Moderatorin beim Musiksender VIVA tätig. Elton ist in Berlin geboren, war Show-Praktikant bei TV Total und hat eine eigene Sendung auf Pro7 moderiert (siehe auch Kap. 6.5.2).

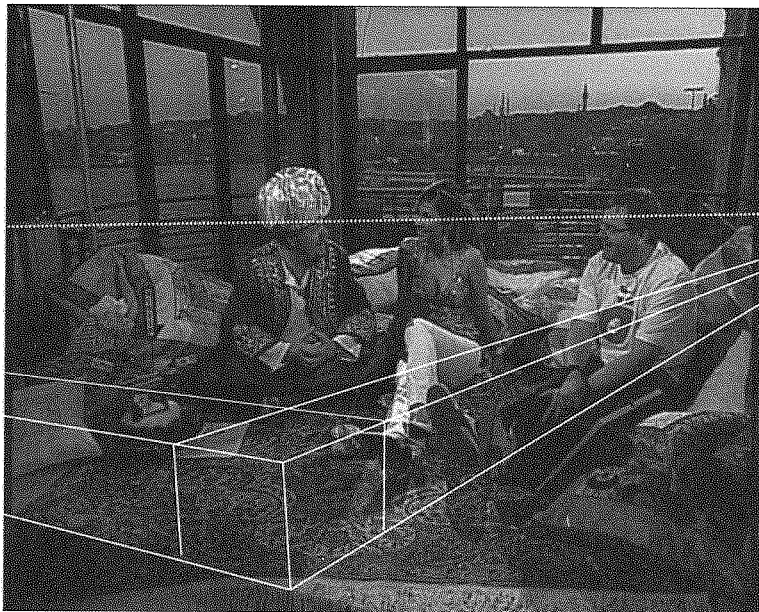


## Reflektierende Interpretation

### Formale Komposition

#### Perspektivität:

Abbildung 34: Istanbul Total: Fotogramm III: Perspektivität. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.)



Wir haben eine Übereckperspektive oder auch Schrägperspektive<sup>114</sup> mit zwei Fluchtpunkten, die sich außerhalb des Bildes befinden. Die Horizontlinie (gepunktet) verläuft auf der Höhe der Köpfe der drei Personen und auf Augenhöhe mit Gülçan und St. R. Die Kamera ist also Auge in Auge mit diesen positioniert. Zugleich geht damit der Blick durch das Fenster auf den Bosphorus. Unterhalb der Horizontlinie befinden sich also die Körper der Bildakteure und die Sitzfläche der „orientalischen Ecke“. Der Blick des Betrachters wird somit zu den Köpfen und zugleich zum Bosphorus gelenkt.

Durch die Perspektivität kommt darüber hinaus ein doppeltes Gefälle in das Bild hinein: Dadurch, dass die Kamera von oben blickt (somit Elemente einer „Vogelperspektive“ aufweist, was filmtechnisch als „Aufsicht“ vgl. Hickethier 1996:

<sup>114</sup> siehe dazu genauer: Kapitel 7.4: Anhang. Diese Art der Perspektivität, die als „Übereckperspektive“ oder als „Zweipunktperspektive“ (Der Brockhaus Kunst 2006: 692) oder auch als „Schrägperspektive“ (Paramón/Calbó 1999: 51) bezeichnet wird, unterscheidet sich von der „Frontalperspektive“ (Der Brockhaus Kunst 2006: 692), wie wir sie in Fotogramm I B finden. Wir finden hier in Fotogramm III aber auch Elemente einer „Vogelperspektive“ (ebenda), also einer Aufsicht, die sich noch einmal von der Untersicht, der „Froschperspektive“ unterscheiden lässt (ebenda).

61f. bezeichnet werden kann) haben wir ein Gefälle nach vorn. Zum anderen haben wir – wie auch die Horizontlinie markiert – ein Gefälle nach links. Die Personen scheinen nach links, insbesondere aber nach vorne weg zu rutschen. Dadurch gelangen erhebliche Spannungen in das Bild.

Abbildung 35: Istanbul Total: Fotogramm III: Planimetrie. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.)



#### Planimetrie:

Planimetrisch zerfällt das Bild in zwei Bereiche, die völlig unterschiedlich strukturiert sind. Während der obere Bereich durch die Fensterbalken geometrisch in klare Rechtecke strukturiert ist, erscheint der untere Bereich in runderen Formen und eher ungeordnet. Er ist allenfalls – in starkem Kontrast zum oberen – durch eine Ellipse strukturiert, die der Positionierung der aufgetürmten Sitzkissen und der Positionierung der Personen folgt.

Gestützt und verstärkt wird diese Teilung des Bildes bzw. dessen Auseinanderfallen in einen oberen und einen unteren Teil durch die (gestrichelte) Horizontlinie als die zentrale Linie im Bereich der Perspektivität, sodass es im Bereich der formalen Komposition eine eindeutige Struktur aufweist.

#### Szenische Choreografie:

Im Bildmittelpunkt (markiert durch das weiße Kreuz) befinden sich Gülçan und St. R., bzw. befindet sich dieser Mittelpunkt genau zwischen ihnen. Hier ist es Gülçan, deren Körpermittelachse sich in etwa auf dem Goldenen Schnitt befindet, welcher auf dem senkrechten Balken des Fensterrahmens liegt, vor dem der Oberkörper von Gülçan positioniert ist.

Gülçan hat (im übrigen die ganze Zeit) ihre Beine zu St. R. hin geöffnet, den Oberkörper ihm zugewandt. Sie schaut hier – und auf den meisten Einstellungen – zu ihm hin. Elton sitzt in wesentlich größerem Abstand zu Gülçan als diese zu St. R.

In Relation zu St. R. und Gülçan befindet sich Elton auf diesem Fotogramm – und das bestätigt sich an anderen Fotogrammen – in einer Zuschauerposition, in der Position des abgebildeten Zuschauers. Dies dokumentiert sich auch in seiner Kleidung: Er ist gekleidet wie ein Tourist. In Relation zum Publikum und zu den eigentlichen Zuschauern vor den Bildschirmen befindet er sich demgegenüber in der Rolle oder Funktion des Darstellers oder Bühnenakteurs. Elton nimmt somit eine Zwischenstellung ein zwischen Bühnenakteur und abgebildetem Zuschauer und vermittelt somit in gewisser Weise zwischen den Zuschauern und der Bühnenszenerie.

### *Ikonologisch-ikonische Interpretation*

Die Planimetrie bietet uns (wie dies auch bei Max Imdahl betont wird; vgl. Kap. 3.6) den Einstieg in die ikonologisch-ikonische Interpretation, da sie den Blick auf dessen Gesamtkomposition lenkt. Auch dieses Bild zerfällt – wie die Fotogramme IA u. IB – sozusagen in zwei Teile, die sowohl in planimetrischer Hinsicht, nämlich durch die horizontale Bildmittelachse, als auch im Hinblick auf die Perspektivität, nämlich durch die Horizontlinie, voneinander getrennt sind und dem Bild somit eine eindeutige Struktur geben. Durch diese strukturelle Verdoppelung im Bereich der formalen Komposition bricht das Bild mehr oder wenig auseinander, sodass zwei Bereiche der türkischen oder Istanbuler Welt (so wie sich aus touristischer Perspektive darbieten) von einander getrennt werden:

*Oben* haben wir den Blick auf den Bosphorus und auf das blaue sonnige Meer, die Moscheen und die Minarette als Metapher für die türkische Natur und die begehrten Objekte des Sightseeing. Dies ist sozusagen die *äußere* Seite der türkischen Welt. Indem der Beobachter durch das Fenster blickt, vor allem dadurch, dass auch hier die allgegenwärtigen Gitter im unteren Bereich der Fenster sichtbar sind, wird auch hier – wie bereits auf den Balkon-Bildern – der Beobachter auf Distanz gehalten. Die türkische oder Istanbuler Welt erscheint eher als Objekt der Betrachtung denn als etwas, mit dem man in Kommunikation bzw. in näheren Kontakt tritt.

*Unten* im Bild haben wir die orientalische Ecke als Element des Inneren des traditionellen türkischen Hauses (zumindest wie Touristen sich dieses vorstellen). Sie steht als Metapher für die innere Seite der türkischen Welt: als Ort der Kommunikation und der sozialen Beziehungen innerhalb des Hauses, in Familie und Verwandtschaft.

Während der erstere Bereich, die *äußere* Seite der türkischen Welt: die Natur, das Meer und das Sightseeing, zwar auf Distanz gehalten wird, aber hell, offen und klar strukturiert ist, erscheint die andere – die innere – Seite der türkischen Welt eher unstrukturiert, ungeordnet und opak, wirkt also undurchschaubar. Dieser Eindruck wird dadurch erzeugt, dass die Kissen kaum Konturen aufweisen, also keine scharfen Umrisse erkennen lassen, aber auch die Muster wenig konturiert sind sowie die Farben mit den ineinander fließenden grauen, graubraunen und grau-beigen Tönen. Auch ist die Kissenlandschaft insgesamt unstrukturiert und ein wenig unterbelichtet. Dies steht im starken Kontrast zum oberen Bereich, also der äußeren

Welt. Im Studio ist hier – im Sinne einer ‚*Exotisierung*‘ – ein Symbol für den türkischen Alltag aufgebaut worden, welche diesen fremd und undurchschaubar macht (ohne ihn wirklich zu repräsentieren).

Durch die weiche, konturlose Figuration und das dämmrige Licht könnte der untere Bereich gemütlich oder heimelig wirken. Dies wird allerdings nicht nur durch den opaken Charakter beeinträchtigt, sondern auch durch Elemente der Spannung: Durch die Wahl der Perspektivität weist das Bild, wie dargelegt, ein Gefälle nach links und insbesondere nach vorne auf. Die Personen, vor allem Gülçan und Elton, scheinen nach vorne zu rutschen. Dieser Eindruck wird durch die Spannung verstärkt, welche durch die Sitzhaltung von Gülçan hier hineinkommt, die sich mit den Armen abstützt und das linke Bein in einer angestregten Spannung hält. Dies vermittelt einen Eindruck, als müsse sie sich gegen das Wegrutschen abstützen. All dies: die angespannte Sitzposition der Personen wie der opake Charakter des Gesamtarrangements der Sitzecke laden nicht dazu ein, sich hier niederzulassen.

### *6.3.5 Reflektierende Interpretation von Einstellungswechsel und Montage*

Es ist wesentliches Prinzip des Wechsels der Sequenzen, des Wechsels von Szenerie und Einstellung, dass dieser durch St. R. ‚dirigiert‘ wird. Der Wechsel folgt der Blick- bzw. Zeigerichtung und/oder der Bewegung von St. R. – entweder durch harten Schnitt oder durch Kameraschwenk oder Zoom. In 0:56 gibt St. R. auch explizite Anweisungen zum Wechsel der Szenerie.

Indem wir das zu sehen bekommen, was St. R. sieht bzw. wohin er blickt oder worauf er zeigt, wird hier Räumlichkeit strikt vom Standort von St. R. konstruiert. Dies geschieht einerseits auf dem Wege des durch Montage bzw. Bildmischung hergestellten Raumes, des ‚*montierten Raumes*‘, des ‚*editing space*‘, den Bordwell (1985: 117) von dem ‚*shot space*‘ unterscheidet, welcher durch das Arrangement der Einstellungen komponiert wird, und den Hickethier (1996: 82f.) auch als den ‚*fotografierten optischen Raum*‘ bezeichnet, der sich vom eigentlichen ‚*filmischen Raum*‘, also dem montierten Raum, unterscheiden lässt (genauer dazu: Kap. 5.6.3). Der fotografierte Raum erschließt sich in seiner formalen Komposition im Wesentlichen durch das Fotogramm.

Zum ‚*fotografierten optischen Raum*‘ ist aber auch der durch die Kamerabewegung, insbesondere den Kameraschwenk oder die Kamerafahrt, hergestellte Raum zu zählen (der allerdings mit nur einem Fotogramm nicht zu erschließen ist), wie er uns in diesem Video dort begegnet, wo die Kamera uns vom Balkon in den Innenraum des Studios zur Sitzecke (Ostecke) führt (3:33-3:36). Auch hier folgt die Kamera St. R., weniger seinem Blick oder seiner Zeigerichtung, als seiner Bewegung, seinem Ortswechsel insgesamt.

Eine gewisse Ausnahme bildet die Ablichtung des ‚*Schleppers, der Stenz macht*‘: St. R. sieht sich – auf dem Balkon stehend – durch das immer lauter werdende Motorgeräusch des Schleppers (vgl. Video-Transkript 2:42-3:09) genötigt, die Störungsquelle zu thematisieren und (negativ) zu kommentieren (vgl. Textinterpretation von UT 2:43-0:02:58: ‚*Nepper, Schlepper, Bauernfänger*‘), sodass der

Schlepper (von 2:46-2:48) durch einen Wechsel von der Totalen in die Halbtotale allein ins Bild gebracht wird. Der Wechsel der Einstellung wird hier also durch die Störquelle bestimmt (vgl. auch: *ES* „Schlepper“ 0:57-1:04; 2:46-2:48). Allerdings bleibt das übergreifende Prinzip dennoch insofern erhalten, als der Wechsel, die Wegführung von der dominanten Einstellung, zwar nicht primär, aber sekundär durch St. R. (in 2:45) dirigiert wird: durch die Blickrichtung (Drehen des Kopfes und Oberkörpers), die Zurückführung durch Änderung seiner Blickrichtung.

Hier greift also der türkische Alltag, das menschliche Leben am Bosphorus, gleichsam unkontrolliert in die Aufnahme ein. Es stört. Interessant ist hier, dass dies durch den Ton, nicht durch das Bild, geschieht, da sich Tonaufnahmen weniger fokussieren und somit auch weniger kontrollieren lassen als Bildaufnahmen (der Schlepper kann optisch, kaum aber akustisch, marginalisiert werden). Somit tritt hier ein Ton-Raum (das sich nähernde und dann wieder entfernende Motorgeschall schafft ein Raumerlebnis), welcher den türkischen Alltag am Bosphorus repräsentiert, gleichsam in Konkurrenz zu dem montierten und fotografierten optischen Raum, welcher auf St. R. und die Sightseeing-Kulisse des Bosphorus fokussiert ist.

Wie in 5.6.3 dargelegt, setzt sich nach Bordwell (1985: 113) der filmische Raum in seiner Gesamtheit als „szenographischer Raum“ aus dem Arrangements der drei unterschiedlichen Komponenten zusammen: montierter Raum, fotografierter optischer Raum oder Einstellungsraum sowie Ton-Raum („sonic space“).

Insgesamt dokumentiert sich also in dieser *Formalstruktur der Konstruktion von Räumlichkeit* auf der Basis des fotografierten optischen Raumes wie auch des montierten Raumes und schließlich des Ton-Raumes, welche spezifische Perspektive uns hier auf die Türkei vermittelt wird, von welchem Standort und aus welcher Perspektive in diesem Video Räumlichkeit konstruiert wird – nämlich aus der Perspektive von St. R. und von dessen Standort aus.

Hinsichtlich dieser *Zentrierung auf St. R.* zeigen sich Homologien auch zur Interpretation der Fotogramme, die uns Aufschluss über die *Formalstruktur des fotografierten optischen Raumes* geben. St. R. ist in der umfangreichsten Hauptsequenz allein abgebildet, steht überwiegend im perspektivischen Zentrum und ist planimetrisch fokussiert. Letzteres wird dadurch hergestellt, dass – und das gilt für alle Bilder der häufigsten und zweithäufigsten Einstellungsvariante der umfangreichsten Hauptsequenz – seine Körperachse mehr oder weniger auf dem Goldenen Schnitt liegt. Damit wird mittels dieser spezifischen Art der Zentrierung auch Kontinuität hergestellt.

Indem St. R. die Konstruktion des filmischen Raumes und damit die Richtung und Selektivität der Performanz der abbildenden Bildproduzent(inn)en dirigiert und zugleich in planimetrischer wie perspektivischer Hinsicht im Zentrum des fotografierten optischen Raumes steht, also auch als abgebildeter Bildproduzent die Performanz des Videos bestimmt, haben wir es hier mit einer Personalunion von abbildendem und abgebildetem Bildproduzenten zu tun, sodass ich hier von einer *Monostrukturierung durch den Showmaster* und einer *Hyperzentrierung* auf seine Person sprechen möchte. Der Orientierungsrahmen der Sendung und der Habitus des Moderators sind somit nicht mehr zu trennen.

Während es ein zentrales Prinzip des Wechsels von Szenerie und Einstellung ist, dass dieser der Blick- bzw. der Zeigerichtung und der Bewegung von St. R. fol-

gen, strukturiert sich die Einblendung des applaudierenden oder lachenden Publikums nach einem anderen Prinzip. Die Kameraführung wendet sich dem Publikum dann zu, wenn der Off-Ton des Applauses sehr stark wird und verwandelt den Off-Ton dann in einen On-Ton (Ausnahme: 0:02:27). – Darin dokumentiert sich, dass das Publikum weitgehend die Funktion eines Applausegebers hat. Indem, wie sich zeigt, die autochthone Bevölkerung nur im Publikum – und dort offensichtlich auch nur in der Minderzahl – in Erscheinung tritt, wird auch hierin noch einmal die marginale Bedeutung sichtbar, welche dieser zukommt.

Eine Variante dieses Prinzips, den Off-Ton zum On-Ton zu machen, ist der oben dargelegte Modus, dass sich St. R. einem stärker werdenden Off-Ton bzw. Hintergrund-Ton zuwendet und die Kamera dieser Hinwendung folgt (Beispiel: der nahende Schlepper 0:50-1:07). Ein dritter Weg des Wechsels von Szenerie und Einstellung ist der, dass diejenigen in Groß- oder Nahaufnahme eingeblendet werden, denen St. R. sich kommunikativ zuwendet (4:22-5:18), der er verbal adressiert. Hier sind es Gülçan und Elton.

*Generalisierend* lässt sich zur Struktur des Wechsels von Szenerie und Einstellung sagen, dass dieser Wechsel den Objekten oder Personen folgt, denen St. R. sich zuwendet, sodass er selbst den Wechsel dirigiert, oder dass er den Personen folgt, die sich St. R. zuwenden. In beiden Fällen haben wir es mit einer Zentrierung auf seine Person zu tun.

### 6.3.6 *Fokussierungen aufgrund von Steigerungen, Verdichtungen und Diskontinuitäten*

Für die Auswahl von Fotogrammen oder auch von Ausschnitten von Sequenzen oder Teil-Sequenzen ziehe ich, wie bereits dargelegt, neben dem Kriterium der Fokussierung durch die Einleitungssequenzen und demjenigen, dass die ausgewählten Fotogramme die umfangreichsten Einleitungssequenzen repräsentieren sollten, ein weiteres – ergänzendes oder alternatives – Auswahlkriterium heran: dasjenige der Fokussierung aufgrund von (dramaturgischen) Steigerungen, Verdichtungen und/oder Diskontinuitäten.

Wie auch im Bereich der Textinterpretation (vgl. dazu u.a. Bohnsack 2008a: 138 sowie 2006: 67) zeigen sich Fokussierungen hier an Auffälligkeiten hinsichtlich der Formalstruktur der Performanz seitens der Text- resp. Bildproduzent(inn)en, die auf die herausgehobene Bedeutung der Teilsequenzen für die Produzent(inn)en verweisen. Im Bereich der Filminterpretation sind dies sowohl die formalen Strukturen der Performanz der abbildenden Bildproduzent(inn)en (Einstellungsgröße, Kamerastandpunkt, Kamerabewegung) (vgl. Kap. 5.6) wie auch der abgebildeten Bildproduzent(inn)en (interaktiver Bezug aufeinander, Struktur der Gebärden) (vgl. Kap. 5.4.). In diesen Sequenzen des Videos finden sich, wie gesagt, weniger dramaturgische Steigerungen und Verdichtungen als vielmehr Diskontinuitäten als Indikatoren für Fokussierungen.

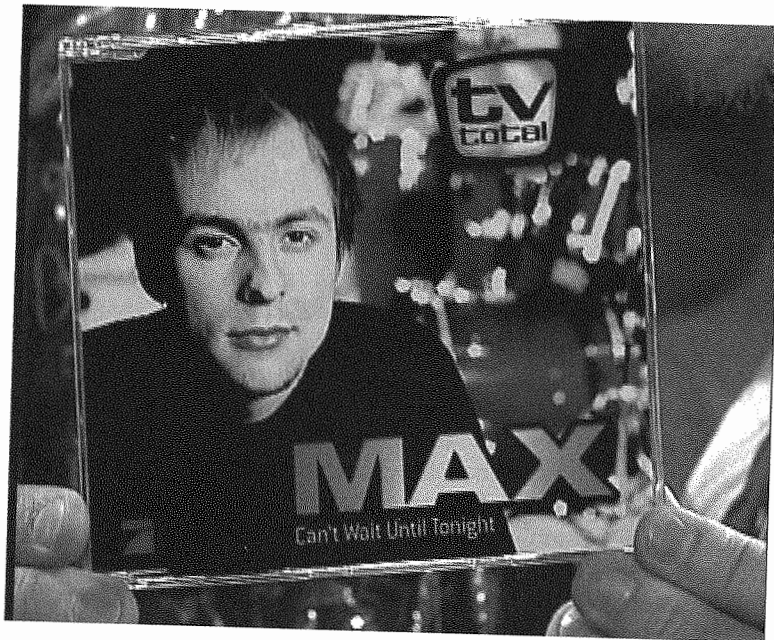
*Fokussierungen im Bereich der Performanz der abbildenden Bildproduzent(inn)en*

In diesem Bereich geht es um Auffälligkeiten, d.h. Diskontinuitäten und Steigerungen in Schnitt bzw. Bildmischung, Kameraführung und Einstellungsgröße.

**Fokussierungssequenz I: 5:48-49; 5:57-59; 6:11-14**

Auffälligkeit in der Einstellungsgröße: die einzige Detailansicht des gesamten Ausschnitts

Abbildung 36: Istanbul Total: Fotogramm IV (5:48). Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD



US „CD-Hülle ‚MAX‘“: 5:48-49; 5:57-59; 6:11-14.

Detailansicht // jeweils harter Schnitt von der Totalen, durch welche die HS („Stefan Raab sitzt mit Gülçan und Elton auf den Kissens der Ostecke“) konstituiert wird, und zurück //

*Formulierende Interpretation*

*Vor-ikonografische Interpretation:*

Auf der Hülle einer CD, welche ins Bild gehalten wird, ist ein ca. 30-jähriger Mann mit schwarzen Haaren und hohem Haaransatz abgebildet, der ein schwarzes T-Shirt oder Polohemd trägt.

Die CD-Hülle wird in drei unterschiedlichen Varianten präsentiert:

5:48-49: mit dem Schriftzug: „MAX“ und darunter: „Can't Wait Until Tonight“ sowie oben rechts „TV Total“ in einem mit Gold unterlegten Rahmen im Fernseh-Bildschirm-Format und unten links eine stilisierte rote „7“

5:57-59: mit dem Schriftzug: „MAKS“ und darunter: „Çan't Wayt Üntyl Tonight“ sowie ohne „TV Total“

6:11-6:14: mit dem Schriftzug: „MÜKS“, darunter: „Çan't Wayt Üntyl Tonight“ ohne „TV Total“. Der abgebildete Mann trägt hier einen schwarzen Schnauzer.

Diese drei Varianten (der Szenerie) werden ihrerseits noch einmal in leicht variierten Einstellungen (alle: Detail-Größe) präsentiert.

Oben abgebildet ist die erste Variante in der ersten Einstellung.

*Ikonografische Elemente:*

Der Mann auf der Hülle der CD, der mit bürgerlichem Namen Max Mutzke heißt, ist Teilnehmer des Eurovision Song Contest 2004, der in Istanbul stattfindet. Unter dem Schriftzug „MAX“ befindet sich (in der ersten Variante) der Titel der CD: „Can't Wait Until Tonight“, mit dem Max Mutzke beim Song Contest antritt, oben rechts das Logo der Show „TV Total“, unten links das Logo der Fernsehanstalt Pro 7. Mit dieser Single hat Max Mutzke es auf Platz 1 der deutschen Charts geschafft. Sie wurde von St. R. geschrieben und produziert.

*Reflektierende Interpretation*

*Formale Komposition:*

Da die formale Komposition der CD-Hülle nicht von den abbildenden Bildproduzent(inn)en der Sendung „Istanbul Total“ hergestellt wurde, deren Orientierungsrahmen hier Gegenstand der Forschung ist, ist sie nicht von direkter Relevanz für die Auswertung.

*Ikonologisch-ikonische Interpretation:*

Da ein Bild ins Bild gehalten wird, haben wir es hier mit einem Metabild, einem Bild im Bilde zu tun. Auch ohne Einbeziehung des Textes wird erkennbar, dass es um die Promotion von „MAX“ und seines Beitrages für den Song-Contest geht. Durch das Logo von „TV-Total“ und Pro 7 wird das kommerzielle Interesse an dieser Promotion (auch ohne Kontextwissen) erkennbar. Es liegt somit nahe, dass die Promotion von „MAX“ einer der Anlässe dafür ist, die Sendung TV-Total nach Istanbul zu verlagern. Somit geht es darum, der Präsentation der „MAX“-CD innerhalb der Sendung eine besondere Aufmerksamkeit zu verschaffen. Formal findet dies – von der Performanz der abbildenden Bildproduzent(inn)en her betrachtet – seinen Ausdruck darin, dass wir es hier mit der einzigen Detail-Einstellung des gesamten Ausschnitts zu tun haben. (Homolog dazu finden wir in der Dimension des Textes hierzu eine detaillierte Erzählung mit der Wiedergabe wörtlicher Rede; vgl. Kap. 6.4.3). Dass St. R. die Präsentation und damit die Promotion im wahrsten Sinne des Wortes ‚fest im Griff‘ hat, findet seinen Ausdruck darin, die seine die CD umgreifenden Finger überwiegend mit ins Bild genommen werden.

Von der Performanz des abgebildeten (und hier wiederum zugleich: abbildenden) Bildproduzenten her wird die Präsentation als Witz inszeniert. Dabei lassen sich im Wesentlichen drei Komponenten des Orientierungsrahmens identifizieren:

- Die CD wird hier als eine Raubkopie und somit ein Stück (Klein-) Kriminalität präsentiert. Durch den Bezug auf den Transkript-Text wird dies ganz deutlich. Die Inszenierung ist aber so aufgebaut, dass dies auch ohne den gesprochenen Text, allein durch den Bild-Text, bereits erkennbar ist.
- Vorgeführt wird hier aber nicht nur die Illegalität der Aktivitäten, sondern auch die Unbeholfenheit, Ungeschicklichkeit der Akteure. Die Präsentation lebt von der Belustigung über die angeblich fehlenden Sprachkenntnisse der türkischen Händler. Ausgearbeitet wird damit das Stereotyp der ‚kleinen Gauner‘ (vgl. auch Text-Interpretation 2:01-2:14 sowie 5:17-5:39). Die türkischen Händler sind nicht geschickt genug, um große Gauner sein zu können.
- Zugleich geht es darum, dass den türkischen Kleinhändlern unterstellt wird, das Fremde – hier die englische Sprache und die deutschen Namen, aber auch das Portrait von „MAX“ – kulturell zu vereinnahmen, zu ‚nostrifizieren‘, indem es umstandslos dem eigenen Erfahrungsraum angepasst wird. Damit wird Naivität unterstellt, was seinerseits den Gaunereien eine gewisse Bedrohlichkeit nimmt und das Bild vom ‚kleinen Gauner‘ stützt.

Die Sequenz zeichnet sich nicht allein in der Bilddimension, sondern auch in derjenigen des Textes durch ein hohes Detaillierungsniveau aus (detaillierte Erzählung mit der Wiedergabe wörtlicher Rede; vgl. 6.4.3) und ist somit in beiden Dimensionen (formal) fokussiert. Doppelt fokussiert wird damit zugleich die (unterstellte) Kleinkriminalität im türkischen Alltag wie auch die (durch diesen Aufmerksamkeitsgewinn noch einmal unterstützte) Promotion der CD von MAX (an der – dem Kontextwissen zufolge – St. R. wie auch der Sender Pro 7 ein unmittelbares kommerzielles Interesse haben).

#### **Fokussierungssequenz II: 4:51**

*Fotogramm V: (Abbildung siehe Videotranskript Istanbul Total: Abbildung 27)*  
Das einzige Fotogramm mit St. R. in der Nah- oder Groß-Einstellung

Insgesamt ist St. R. in dem analysierten Ausschnitt lediglich 1 Sek. (nur in Fotogramm 4:51) in Groß-Einstellung zu sehen, und dies ist offensichtlich Ästhetik und Aufnahmetechnik geschuldet: Denn würde St. R. hier nicht in dieser Einstellung gezeigt, hätten wir, da vorher und nachher je 3 Sek. lang Gülçan in dieser Einstellung zu sehen ist, einen Bruch in der Bildfolge, der nicht nur ästhetisch problematisch ist, sondern wohl auch kamera- und mischtechnische Probleme mit sich bringen würde. Bei den anderen Akteurinnen und Akteuren auf dem Podium und auch bei Personen aus dem Publikum wird die Groß-Einstellung häufiger gewählt.

Gülçan ist allerdings die Einzige, die in umfassender Weise in dieser Einstellung abgelichtet wird. Dies geschieht dort, wo sie (als Einzige) über ihr Leben berichtet, also über etwas, was unmittelbar ihre persönliche Identität betrifft (vgl. Goffman 1963: 62ff.). Dies wird vor dem Hintergrund des engen Zusammenhangs von Gesicht, von Physiognomie und persönlicher Identität plausibel: Die Physiognomie gilt als der wohl wichtigste Träger persönlicher Identifizierung (vgl. Goff-

man 1963: 57ff.). Da eine Ablichtung der Person von St. R. in Nah- oder Groß-Einstellung, also mit Fokussierung auf seine Physiognomie, weitgehend vermieden wird, geht es – trotz der hier zu beobachtenden Hyperzentrierung auf seine Person (vgl. 6.3.4) – offensichtlich nicht um seine persönliche (individuelle) Identität (Was ja auch, wenn wir Kontextwissen heranziehen, darin zum Ausdruck kommt, dass seine Privatsphäre, ebenso wie diejenige vergleichbar prominenter Moderatoren, in der Regel aus der medialen Präsentation herausgehalten wird).

Das legt nahe, dass St. R. – und allgemeiner Moderator(inn)en mit einer derartigen Hyperzentrierung auf ihre Person – eher als Repräsentanten einer *sozialen* Identität, eines kollektiven (milieuspezifischen) Habitus bzw. als Repräsentanten von Stereotypen in diesem Bereich von Bedeutung sind. Ein derartiger ‚Kult‘ um die Person hat dann die Funktion der Konstitution von Milieus, als deren Repräsentantin und Identifikationsfigur diese Person erscheint, die Funktion der Konstitution von rudimentären Erfahrungsräumen, von Gemeinsamkeiten und Zugehörigkeiten im Sinne habituelier Übereinstimmungen.

#### *Fokussierungen im Bereich der Performanz der abgebildeten Bildproduzent(inn)en*

Auffälligkeiten oder Diskontinuitäten in diesem Bereich, die Fokussierungscharakter gewinnen können, betreffen insbesondere die Abläufe der interaktiven und individuellen Bewegungen. Da ihr Fokussierungscharakter besondere Aufschlüsse über den Habitus oder Stil der Bildproduzent(inn)en verspricht, werden sie einer intensiven ‚mikroskopischen‘ Analyse unterzogen.

#### **Fokussierungssequenz III: 3.53-3.56**

Eine Auffälligkeit des Bewegungsablaufs von St. R. lässt sich innerhalb der Hauptsequenz II (03:34-03:57: „Stefan Raab stehend und sich setzend im Studio vor der Sitzecke am Fenster mit Blick auf den Bosphorus“) identifizieren. Diese Hauptsequenz ist die am wenigsten umfangreiche, und hier finden sich nur wenige Einstellungsvarianten. Die Szenerie ist hinsichtlich der *formalen Komposition* ähnlich strukturiert wie die in der Hauptsequenz III (3:58-6:21), und wir können zu deren Rekonstruktion und Interpretation somit auf die Ausführungen zu Fotogramm III zurückgreifen.

Die Auffälligkeit des Bewegungsablaufs selbst befindet sich innerhalb der Teilsequenz, in welcher der Ablauf des Sich-Setzens abgelichtet ist: 3.53-3.56. Diese Teilsequenz ist nicht nur durch den auffälligen Bewegungsablauf, sondern auch durch die dramaturgische Stellung exponiert, da an ihrem Ende der Übergang zur Einspielung und zum Werbeblock sich befindet (die in der Internet-Version herausgeschnitten sind).

#### *Methodischer Exkurs*

Ich erinnere an die Differenzierung unterschiedlicher Dimensionen des Bewegungsablaufs in Kapitel 5.5.1: Wenn ich davon spreche, dass St. R. „sich setzt“, be-

wege ich mich auf der Ebene „operativer Handlungen“, wie ich es dort genannt habe. Die Elemente dieser operativen Handlung, welche ich „Gebärden“ nenne (bspw. das ‚Beugen des Rumpfes‘) stehen dann (im Sinne von Alfred Schütz 1971) in einer Beziehung des „Um-zu“ zu dieser operativen Handlung (St.R. beugt den Rumpf, *um sich zu setzen*). Die Elemente dieser Gebärden (‚Kopf und Schulter gehen nach vorn, das Becken nach hinten, die Knie werden gebeugt‘) habe ich mit Birdwhistell (1952) als „Kineme“ bezeichnet.

Dieselbe Gebärde oder Handlung kann immer auf zwei Sinnebenen zugleich interpretiert werden: zum einen – zweckrational – nach Art der Konstruktion eines Um-zu-Motivs (die Frage nach dem, *was* jemand tut), zum anderen als Dokument für das „Wesen“ oder den Habitus der AkteurInnen (bspw.: Unsicherheit oder Starrheit). Entscheidend ist dann, *wie* jemand die Gebärde oder operative Handlung ausführt. Während die Motivkonstruktion eine *sequenzielle* Betrachtung des Bewegungsverlaufs voraussetzt, setzt eine dokumentarische oder ikonologische Interpretation der Gebärde oder Handlung zusätzlich die Beobachtung des *simultanen* Kontexts voraus, in den die Bewegung eingebettet ist. Gemeint ist sowohl der Kontext der Kineme und Gebärden des einzelnen Akteurs wie auch der interaktive Kontext (szenische Choreografie) in seiner Simultaneität. Die Interpretation von Gebärden macht, wie Birdwhistell (1952: 15) zuerst betont hat, die Rekonstruktion von elementaren Bewegungseinheiten, den Kinen oder Kinemen, in ihrem simultanen und ganzheitlichen Zusammenspiel erforderlich (vgl. 5.5.1)

Hier geht es vor allem um die Interpretation der Gebärden und teilweise ihrer Elemente, der Kineme.<sup>115</sup> Diese werden sowohl in ihrer sequenziellen Struktur auf der Grundlage der Zeitlupenbetrachtung dieses Ausschnitts bzw. einer Betrachtung mit Hilfe einer Programmfunktion, welche die Sequenz in kleine Schritte (in Sekundenbruchteile) zerlegt (bspw. Programmfunktion „Schritt vorwärts“ in Power DVD) interpretiert als auch – in den besonders auffälligen Situationen – in ihrer Simultanstruktur auf der Grundlage eines auf diesem Wege ausgewählten Fotogramms.

#### Formulierende Interpretation

St. R. hat sich von einem Balkon des Istanbul Studios von Pro 7 in das Studio selbst begeben, wo die orientalische Ecke auf einem ca. 40 cm hohen Podest aufgebaut ist.

Nachdem St. R. den Zuschauern zunächst den Rücken zugewandt hat, um auf das Podest zu steigen, wendet er sich – oben angekommen – nun wieder den Zuschauern zu. Dabei ist die Haltung des Rumpfes und der Arme im Unterschied zur Sequenz vor dem Betreten des Podests vollkommen verändert. Vorher waren (wie in dem weitaus überwiegenden Teil der Szenerien, in denen er stehend auftritt) die Arme und auch der Oberkörper bzw. die Brust zu den Zuschauer(inne)n weit geöffnet, unterstrichen durch eine Mimik mit einem breiten Lächeln. Nachdem St. R. (als Vorbereitung für den ‚Schneidersitz‘) die Füße und Unterschenkel hintereinander positioniert hat und dabei ein wenig ins Ungleichgewicht geraten ist, rafft er nun – schwankend – die Beine der weiten Pluderhose mit beiden Händen nach oben.

<sup>115</sup> Die Interpretation dringt hier – nicht zuletzt aufgrund der niedrigen Bildauflösung – nicht sehr weit in die Dimension der Kineme ein. Für eine intensivere Analyse auf der Basis der dokumentarischen Methode siehe Klambeck 2007.

Er sackt nun regelrecht in sich zusammen und lässt die Hosenbeine wieder los. Dann kippt er nach hinten, trifft rücklings auf das undurchsichtige Gemenge von Kissen auf und rollt schwankend und von der Richtung her zunächst unkontrolliert nach hinten ab.

Dieses In-sich-zusammen-Sacken erscheint in seinem Bewegungsablauf besonders interessant und wird nun genauer – und das bedeutet: in seiner simultanen Bewegungsstruktur – beleuchtet. Dies geschieht auf der Grundlage eines Fotogramms (welches Sekundenbruchteile dem Fotogramm 3:53 im Transkript voraus geht):

Abbildung 37: Istanbul Total: Fotogramm V (3:52/3:53). Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD



St.R. knickt dann in der Hüfte und auch in den Knien leicht ein, sodass die Beine in den Knien nach außen angewinkelt sind. Er lässt dabei die Hose wieder los, sodass seine Hände sich funktions- und ziellos bewegen und halb geschlossen sind. Die Schultern sind nach vorne gezogen – die rechte mehr als die linke – und wirken verkrampt. Der Kopf ist leicht nach vorne gebeugt, das Gesicht nach unten gewendet und die Mimik wirkt (soweit aufgrund der niedrigen Bildauflösung diese Beobachtung valide ausfällt) angestrengt.

#### Reflektierende Interpretation

Insgesamt vermittelt dies den Eindruck eines Gebärdenablaufs, der in seiner Unkontrolliertheit am ehesten einer Filmszene vergleichbar ist, in der ein von einer Kugel getroffener Mensch zu Boden sinkt. Der Gebärdenablauf in seiner Ungeordnetheit und Undurchschaubarkeit fügt sich damit in den Charakter des szenischen Ambientes,

also der orientalischen Ecke, die ja, wie in 6.3.3 (Fotogramm III) bereits beschrieben, unstrukturiert, ungeordnet und opak, als undurchschaubar erscheint.

St. R. begibt sich hier nicht nur in das Studio, sondern damit auch in das Innere des türkischen Hauses, für welches die orientalische Ecke symbolisch steht. Er wechselt somit von der äußeren in die innere Welt der Türkei, von der auf dem Balkon eingenommenen Perspektive des Sightseeing und der Betrachtung aus der Ferne zur Innenwelt. St. R. fällt in dieses – die innere Welt der Türkei symbolisierende – undurchschaubare Durcheinander relativ unkontrolliert hinein bzw. lässt er sich mehr oder weniger unkontrolliert fallen.

Dieser Übergang bzw. diese Annäherung in die innere Welt ist offensichtlich mit Unsicherheiten und Unklarheiten verbunden. Und der Gebärdenablauf zeichnet sich durch einen gewissen Kontrollverlust aus, der nicht eigentlich komisch wirkt. Wenn wir hier bereits einen Bezug zur Dimension von Text und Ton (auf der Ebene der formulierenden Interpretation) herstellen (vgl. 6.4.2), so leitet St. R. das Betreten des Podests mit Witz-Wortspielen zu „Istanbul“ ein (3.34-3.46), über die aber nicht gelacht wird – ebenso wenig wie sein Fall in die Kissen der orientalischen Ecke für Heiterkeit sorgt, auch dann nicht, als er mit einem übersteigerten Stöhnen oder Ächzen („Ährg“) diesem eine witzige Note zu geben versucht.

Der symbolische Übergang in die innere Welt der Türkei misslingt technisch und, da der Witz fehlt, auch in dramaturgischer Hinsicht. Dies steht im Gegensatz zum verbalen Kommentar, den St. R. selbst gibt: „das hat ja schon mal ganz gut geklappt“ (3.56-3.57). Darin kommt zum Ausdruck, wie sehr er seine eigenen Erwartungen hinsichtlich der gelungenen Bewältigung dieses Übergangs heruntergeschraubt hat. Die expressive Distanz gegenüber dem, was hier als typische Welt des türkischen Hauses konstruiert wird, gerät zur Unbeholfenheit (was als solche allerdings der Identifikation mit dem Moderator und dessen Habitus keinen Abbruch tun muss, sondern diese möglicherweise sogar befördert).

## 6.4 Interpretation in der Dimension von Text und Ton

Bei der Interpretation des Textes orientieren wir uns an der seit über 20 Jahren erprobten Vorgehensweise der dokumentarischen Gesprächsanalyse (vgl. u.a. Bohnsack 2008: Kap. 7.3 u. 12.2; Bohnsack/Przyborski 2006; Bohnsack/Schäffer 2007; Przyborski 2004), wie sie ursprünglich im Bereich des Gruppendiskussionsverfahrens entwickelt worden ist (Bohnsack 1989).

### 6.4.1 Texttranskript

Um für die Interpretationen, die ausschließlich auf den Text bezogen sind, einen besseren Überblick über diesen zu gewinnen, empfiehlt es sich, das Texttranskript gesondert abzudrucken. Die Transkription folgt dem System TiQ („Talk in Qualitative Social Research“) (siehe Anhang 7.2 u. Bohnsack 2008a: Kap. 12.3). Allerdings treten im Unterschied zur Transkription von Gesprächen, Interviews und anderen Texten ohne Bildbezug hier an die Stelle der Zeilennummern die Laufzeiten

des Videos (in Sek.), die sich auch im Videotranskript finden. Diese werden jeweils am Beginn von Themen und Unterthemen in das Transkript hineingeschrieben und durch einen Kasten markiert.

### Transkript „Istanbul Total“ – von Anfang bis 6 Min.21 Sek.

Am = Stefan Raab (St. R.); Bf = Gülçan; Cm = Elton

Am: Ich befinde mich hier ähm auf dem Balkon unseres Istanbul Studios und wie man sehen kann hier hat man einen fantastischen Blick über die ganze Stadt hier ist der Bosporus hier geht's ins goldene Horn ja? Die Schiffe fahren rein sie fahren raus ein ständiges Heckmeck ja? 0:17 Und äh da drüben äh kann man fantastisch äh ähm die Moscheen sehen da vorne ist die Hagia Sofia eine der ältesten Kirchen äh hier in Istanbul ühm glaub über tausend Jahre alt und äh daneben ist der Sultanspalast dort vorne auf dieser Anhöhe ja? Und wenn man mal auf die andere Seite rüberschaut dann kann man den Bosporus entlang kucken und sieht dort hinten die Bosporusbrücke die Asien und Europa miteinander verbindet ja? Die einzige Stadt der Welt auf zwei Kontinenten 0:50 und da vorne kommt grade ein Schlepper an ja? Das sieht man hier unten das sieht man sind immer die Boote mit den kann man das mal zeigen? Die Boote mit diesen ähm mit diesen Reifen die sch- die schleppen immer die großen äh Schiffe deswegen gibt's auch hier den äh Spruch Nepper-Schlepper-Bauernfänger @ @ @ @ ja

L ((mehrstimmiges

Gelächter))

Am: und ich 1:09 hier war vielleicht was los ich weiß nicht wie das bei ihnen war ein fantastisches Fußballwochenende oder? Ein großartiger Auswärtssieg und damit vorzeitig Meister Herzlichen Glückwunsch

((Applaus))

Am: Fenerbahçe Istanbul °übrigens° Die Mannschaft von Christoph Daum ja

L ((Applaus, Pfeifen und Johlen))

Am: die Mannschaft von Christoph Daum ist vorzeitig Meister geworden hier in der Türkei Fenerbahçe liegt übrigens drüben auf der asiatischen Seite ähh da sind diese Anhöhen zu sehen dahinter dem ersten Hügel ist das Fenerbahçe-Stadion und ähh da war natürlich gestern was los hier in der Stadt das können Sie sich gar nicht vorstellen überall ü ü ü es wurde gehupt und gefeiert ich dachte im ersten Moment ich bin doch schon seit ein paar Tagen da warum erst heute die Freude ja? Aber es handelte sich äh um die Fans von Fenerbahçe die dort gefeiert haben und äh hier ist einiges los 2:02 ich weiß nicht wer schon mal in Istanbul war von Ihnen ich kann Ihnen nur empfehlen fahren Sie hier hin wenn sie hier kriegen sie einfach alles überall Händler auf der Straße die die Jacke auf machen und so sagen [sag mal brauchst du 12 Punkte

\*(Am spricht mit türkischem

für Max ja? Isch habe dabei hier] 2:15 und ähh fantastisch wir ham ne Schifffahrt Akzent))

gemacht entlang des Bosporus eine kleine Tour hinten wollt ma durch die Meerenge und sind leider stecken geblieben äh weil Elton an Bord war @ja@ aber wirklich man kann hier die tolle Sachen erleben fantastisches Wetter 2:30 und ähh viele glauben ja dass Istanbul die Hauptstadt der Türkei ist das stimmt aber nicht die Hauptstadt der Türkei ist Ankara ja das ist ähnlich wie in Deutschland wo viele ja glauben Berlin ist die Hauptstadt is aber Köln ja? 2:43 Hallo! Ich glaube der machtn bisschen Stenz hier unten das kennt

L ((lauter werdendes Motorengeräusch))

Am: man ja oder einfach mal Vollgas geben bei uns mitm 3er-BMW hier einfach mal mitm

Am: Schlepper ja? Siehste hat ihn schwer motiviert fantastisch 2:59 ich bin froh dass wir hier

Am: mit TV Total es geschafft haben in Istanbul zu sein bevor der Musikantenstadl kommt ja und äh ich 3:06 und wir ham im ähh im was ist das Norden Äh mal kucken was das ist

Am: (Fener) Ost-Nord-Ost in Ost-Nord-Ost ist unser Hotel und ich kann Ihnen sagen wir ham ein fantastisches Hotel ja unsere komplette Crew hat einen herrlichen Blick auf meinen Pool ja das können Sie sich gar nicht vorstellen wunderschön ja ssss ühmm 3:26 toll oder Genießen Sie den Aus-Blick wir werden jetzt jeden Abend Montag, Dienstag, Mittwoch und Donnerstag hier aus Istanbul senden und ähh 3:34 viele von Ihnen wissen ja auch in der deutschen Sprache kommt das Wort Istanbul oft vor zum Beispiel ähm Claudia Schiffer ist=an=Bulimie erkrankt ja? Oder Elton ist=an=Buletten interessiert ja? Solche Sachen und 3:46 ich nehm mal Platz hier in unserer Ecke fantastische schöne wie setzt man sich hin im Schneidersitz? Ährg so das hat ja schon mal ganz gut geklappt. Einspielung/Werbung Fantastisch, oder? Das ist ein Studio

Am: wie man sich das vorstellt ähh 3:58 hinter uns ist eine Moschee neben der anderen und was wir da eben gesehen haben da drüben hat ich eh

Bf: ja°

Am: ich hab eben gesagt das ist eine Kirche das war früher eine Kirche die Hagia Sofia das war äh lange Zeit eine Kirche und wurde dann zur Moschee umgewandelt falls der eine oder andere Klugscheißer am Fernseher sitzt und sagt das stimmt ja nicht dass

Am: und kölnischem Akzent))

Am: das ne Kirch is ja? 4:19 sooo ihr wart ja äh schon auch hier äh ein paar Tage unterwegs und habt ein bisschen gedreht. Gülçan du sprichst ja die Sprache perfekt

Bf: Ja ich versuchs immer wieder

Am: Du bist aber in Deutschland geboren oder?

Bf: Ja ich bin in Lübeck geboren aber ähm äh Türkisch verlernt ma ja nicht mein ich konnte bis ich fünf war überhaupt kein Deutsch

Am: Ach ok.

Bf: und äh wir ham zuhause nur Türkisch gesprochen und ähm im Kindergarten hab ich dann Deutsch gelernt so

Am: Wo kommen deine Eltern her aus welchem Teil der Türkei

Bf: Die kommen also genau hier wo wir gerade sind

Am: Hier aus?

Bf: Die sind hier aufgewachsen und ich kenn das alles hier auch die Strassen weil als ich mit meinen Eltern hier war die ham mir dann so die Strassen gezeigt und ja mein Papa so ja kuck mal da hat Mama gewohnt und da hab ich mal Steine geschmissen und da hab ich mit Schokolade unten gewartet und so (ese) ja Papa schön schön schön aber das ham sie mir mal erzählt das is hier irgendwo in den ecken acha 4:56 und ihr wart zusammen ein bisschen einkaufen Elton?

Am: Ja

Cm: Ja

Am: Ähh habt auch mal euch umgeschaut ist ja ne sehr ähh morderne Stadt Ja?

Cm: Tatsächlich ja

Am: also eine eine wirkliche Metropole es gibt hier ein Viertel das sieht aus wie New York mit mit Hochhäusern dann gibt's hier wieder sehr traditionelle Viertel die Altstadt uns so weiter ähh und ihr habt mal ein bisschen eingekleidet oder

Cm: ja das könn ma uns auch gerne mal ansehen

Am: habt ihr denn was gefunden

Cm: auch für mich ja

Am: 5:17 also erst möchte ich noch mal das zeigen was ich auf dem Basar das muss ma bevor wir diese MAZ zeigen wo ihr unterwegs wart äh was ich auf dem Basar gefunden hab

Cm: ja

Am: äh äh ja es is ja so hier findet man ja äh @ viele @ Sachen Taschen von Dolce und Gabana, Gucci und alles ganz billig hab mich gefragt warum ist das so ja Rolex Uhren

Cm: @

Am: ja für Bruchteil für einen Bruchteil des Preises ja und auch äh äh äh Raubkopien

Bf: (Preis

Am: von CDs gibt's hier einige 5:40 ich hab hier einige gekauft schause mal äh zum

Bf: @ @

Am: Beispiel die CD @ja@ sie kenn ja die CD das ist se original Original-CD von Max ja kenne vielleicht und dann hab ich hier aufm Basar mich mal umgeschaut und hab mal gefragt ob die das auch ham ja und ich hab zwei verschiedene gefunden einmal hab ich die hier gefunden schause mal ja

Am: und ja hab ich gesagt is is it real äh I mean is it a legal copy yees

Am: of course und dann bin ich in nächsten Laden gegangen und da wollten se ma das verkaufen ja

Bf: @ @ @

Am: ° @ Can't wait üntyl tonight@°

Am: Vorsicht ja ansonsten aber alles korrekt

#### 6.4.2 Formulierende Interpretation des Textes: Thematische Gliederung

OT: 0:00-01:08 Der Blick auf den Bosphorus  
 UT: 0:00-0:16 Der Blick über die ganze Stadt und der Schiffsverkehr  
 UT: 0:17-0:49 Moscheen, Paläste und die Bosphorusbrücke  
 UT: 0:50-1:08 Der Schlepper

OT: 1:09-2:01 Fenerbahçe ist „Meister“ geworden  
 OT: 2:02-2:14 Die Straßenhändler von Istanbul und „MAX“  
 OT: 2:15-2:29 Die „fantastische“ Schifffahrt am Bosphorus und Eltons Körperfülle  
 OT: 2:30-2:42 Istanbul und Ankara, Berlin und Köln: Wer ist Hauptstadt?  
 OT: 2:43-2:58 Der Schlepper „macht Stenz“  
 OT: 2:59-3:05 Der „Musikantenstadl“  
 OT: 3:06-3:25 Das „fantastische“ Hotel  
 OT: 3:26-3:33 Der „tolle“ Ausblick und die Sendezeit  
 OT: 3:34-3:45 Wortspiele mit „Istanbul“  
 OT: 3:46-3:57 Der „Schneidersitz“ in der Ostecke  
 OT: 3:58-4:18 Der Blick auf die Moschee oder die Kirche  
 OT: 4:19-4:55 Gülçan und Elemente ihrer Lebensgeschichte  
 OT: 4:56-5:16 Der Einkauf von Gülçan und Elton in der „modernen“ „Metropole“ Istanbul  
 OT: 5:17-6:21 „Raubkopien“  
 UT: 5:17-5:39 „Billige“ Taschen und Uhren und „Raubkopien“ von CDs  
 UT: 5:40-6:21 „Raubkopien“ der CD von „MAX“



### 6.4.3 Reflektierende Interpretation in der Dimension von Text und Ton

#### 0:00-1:08 Der Blick auf den Bosporus

- *Proposition durch St. R. im Modus des Beschreibens*

In seinen ersten verbalen Äußerungen bezieht sich St. R. auf seinen eigenen Standort: „Ich befinde mich hier (...)“ und somit auf seine Person. Er lenkt dadurch nicht nur die Aufmerksamkeit des Zuschauers, sondern auch die Kameraführung zunächst auf sich selbst, und erst dann – sekundär – auf das, was er sieht, was in seinem Blickfeld liegt: „die ganze Stadt“, den „Bosporus“ und das „Goldene Horn“, also auf die bekanntesten Objekte des Sightseeing. Insbesondere an dem nun folgenden Themenwechsel von den Sehenswürdigkeiten zum „Schlepper“ wird deutlich, dass der Text von St. R. dem folgt, was ihm in den Blick gerät.<sup>116</sup>

Hier zeigen sich Homologien zur Bilddimension und der dort rekonstruierten Blickrichtung und der Gebärden von St. R., mit denen er, wie in der Interpretation des ersten Fotogramms (IA) herausgearbeitet ist, ebenfalls zunächst auf seinen eigenen Standort verweist und dann auf die Objekte des Sightseeing (die Kameraführung hat allerdings etwas Mühe, dem zu folgen und zeigt zunächst nicht Letzteres, sondern die Hafenmole bzw. den Anleger direkt unterhalb des Balkons).

Die Textinterpretation stützt somit in homologer Weise die Selbstfokussierung oder Selbstbezogenheit von St. R., wie sie auch im Bild bereits in der allerersten Einstellung der Internetversion zum Ausdruck kommt und für das gesamte Video konstitutiv ist. Dieser Selbstbezogenheit entspricht auch der monologische Charakter seiner verbalen Äußerungen. Interaktive Elemente bzw. Elemente einer reziproken Kommunikation zwischen Moderator und Publikum finden sich – wenn auch nur rudimentär – lediglich dort, wo St. R. Applaus erhält (was in dieser Sequenz nicht der Fall ist) oder wo gelacht wird (in: 1:08).

Als eine (wohl nicht selten angewandte) Strategie, die fehlende Reziprozität nicht allzu deutlich werden zu lassen, lassen sich die nachgestellten „Ja“ mit steigender Intonation, also mit Frageintonation, am Ende von Sätzen ( u.a. 0:12; 0:16; 0:34; 0:45/46; 0:53; 1:08) interpretieren, die in der Konversationsanalyse „tag-questions“ genannt werden (Sacks/Schegloff/Jefferson 1978). Sie geben seinen Äußerungen den Anschein einer Frage. St. R. suggeriert damit eine dialogische Struktur, also dass die Zustimmung des Publikums von Bedeutung für seine Äußerungen sei, ohne dass das Publikum allerdings eine wirkliche Chance hätte, eine derartige Zustimmung zu geben oder zu verweigern.

Die erste Bemerkung, in der überhaupt die Menschen in Istanbul Erwähnung finden, rahmt diese als „Nepper, Schlepper, Bauernfänger“ (1:06-07), also als kleine Gauner, als Kleinkriminelle. Dieser Witz auf Kosten (der Kriminalisierung) anderer lässt die Welt, auf die er blickt, sogleich als nicht ganz geheuer, als undurchsichtig oder opak erscheinen – homolog zum opaken Charakter der inneren Seite

<sup>116</sup> Der von Stefan Raab angesprochene rege Schiffsverkehr, das „Heckmeck“, steht im Widerspruch zu den Bildern, auf denen pro Einstellung nur je ein Schiff zu sehen ist. „Heckmeck“ (türkischer Ursprung) meint ja nicht nur eine rege Betriebsamkeit, sondern auch eine solche relativ unkoordinierter Art. Beides ist hier nicht zu beobachten. Entweder produziert Stefan Raab Stereotype bzw. Klischees oder die Kameraführung ist nicht in der Lage, diese Beobachtungen einzufangen.

der türkischen Welt in Fotogramm III. Zugleich verweist dieser Spruch auf ein *mediales Ereignis*.<sup>117</sup> Damit ist zum einen der Wiedererkennungswert und somit ein Aufmerksamkeitsgewinn gesichert oder zumindest sehr wahrscheinlich, und zum anderen wird an gemeinsame Erfahrungen und somit an elementare oder rudimentäre Gemeinsamkeiten zwischen Moderator und Zuschauern sowie der Zuschauer untereinander appelliert.

#### 1:09-2:00 Fenerbahçe ist „Meister“ geworden

- *Proposition durch St. R. im Modus des Erzählens (mit eingelagerter Beschreibung: 1:37-1:45)*

St. R. sucht auch hier den Anschluss an ein mediales Großereignis, genauer: ein Großereignis des Fernsehens. Es ist dieser Bezug auf mediale Ereignisse, der überhaupt eine Verbindung zwischen dem vorherigen und dem jetzigen Thema schafft. St. R. kann damit von dem Interesse für das mediale Großereignis profitieren. Das mediale Ereignis St. R. zieht (wie auch schon im vorherigen Kommentar zum Schlepper) Aufmerksamkeitsgewinne aus anderen Ereignissen der Medien. Ich möchte deshalb von einem *parasitären Aufmerksamkeitsgewinn* sprechen. An dem scherzhaften Hinweis von St. R., dass er die Begeisterung für Fenerbahçe im ersten Moment auf sich selbst bezogen habe, wird die Orientierung an einem parasitären medialen Aufmerksamkeitsgewinn noch einmal in spezifischer Weise zum Ausdruck gebracht.

Zugleich gelingt es ihm auch, dieses türkische Fußballereignis in einen deutschen Rahmen zu stellen, es gleichsam zu nostrifizieren: Wenn die Mannschaft von Fenerbahçe Istanbul als die „Mannschaft von Christoph Daum“, also eines deutschen Trainers, bezeichnet wird, dann kann sie nahezu als deutsche Mannschaft betrachtet werden. Zumindest wird herausgehoben, dass der Verdienst am Sieg einem deutschen Trainer zukommt.

#### 2:01-2:14 Die Straßenhändler von Istanbul

- *Proposition durch St. R. im Modus globalen Beschreibens*

St. R. nimmt nun Bezug auf Straßenhändler, die – versteckt unter der Jacke, also in illegaler Weise – Waren anbieten. Es wird damit zum zweiten Mal (zuerst mit der Schlepperbesatzung) überhaupt Bezug genommen auf Menschen, die hier in Istanbul leben (ohne dass St. R. allerdings hier in eine Kommunikation mit ihnen eintritt). Auch beim zweiten Mal werden sie im Rahmen der Illegalität, der Kleinkriminalität, der kleinen Gaunereien, präsentiert. Indem St. R. seine „Empfehlung“ nach Istanbul zu fahren, mit der Attraktivität der Schwarzware begründet, sind es nunmehr zwei Bereiche, die Istanbul attraktiv machen: die interessante Kulisse des

<sup>117</sup> Es handelt sich um den Untertitel einer Sendung mit dem Titel „Vorsicht Falle!“, welche zuerst am 24. März 1964 und bis 2001 ausgestrahlt wurde, zunächst von Eduard Zimmermann selbst und dann von seiner Adoptivtochter Sabine Zimmermann moderiert. In der Sendung wurden zahlreiche im Alltag und bei Geschäften an der Haustür begangene Betrügereien aufgedeckt. Meist wurden dabei kurze Episoden nachgespielt und eine Off-Stimme kommentierte das Geschehen. Der jeweils letzte Beitrag der 45-minütigen Sendung bestand aus einem vom ZDF selbst inszenierten Betrugsversuch, bei dem mit versteckter Kamera Leute auf ihre Leichtgläubigkeit getestet wurden.

Sightseeing (Blick auf den Bosphorus) und die Waren der schlitzohrigen Straßenhändler.

Hier zeigen sich Homologien zur Bildebene hinsichtlich der Diskrepanz zwischen der heilen Welt des Sightseeing als der äußeren Seite der türkischen Kultur und dem undurchschaubaren, dem opaken Charakter der inneren Seite der türkischen Kultur in der Interpretation des Fotogramms III. Damit wird Fremdheit im Sinne einer undurchschaubaren, nicht ganz geheueren, Andersartigkeit konstruiert bzw. verfestigt.

St.R. wählt hier nicht – wie in der Propositionen vorher – den Modus einer (relativ) konkretisierenden Erzählung, sondern den einer (verallgemeinernden) Beschreibung („überall“; 2.08); wodurch der stereotypisierende Charakter der Darstellung in formaler Weise hergestellt wird.

#### 2:15-2:29 Die „fantastische“ Schiffahrt am Bosphorus und Eltons Körperfülle

- *Proposition durch St. R. im Modus globalen Erzählens*

Der Versuch, sich in die Welt der Türkei – wenn auch nur in die touristische – hinein zu begeben, wird dann doch wieder zu einer Selbst-Fokussierung der deutschen Protagonisten hingeführt. Dieses Unterthema wird mit der Charakterisierung „fantastisch“ gerahmt, d.h. sowohl eingeleitet als auch beendet. Allerdings bleibt dies auf das Sightseeing bezogene Adjektiv ebenso wie die gesamte Erzählung unkonkret bzw. unausgeführt, hat eher den Charakter einer Beschwörung. Der Bosphorus bleibt auch verbal Kulisse. Die Fokussierung liegt auf den deutschen Protagonisten und deren weniger ästhetischen Eigenarten: auf der Körperfülle von Elton.

#### 2:30-2:42 Istanbul und Ankara, Berlin und Köln: Wer ist Hauptstadt?

- *Proposition durch St. R. im Modus eines fiktiven Dialogs*

Indem St. R. betont, dass für ihn nicht die rechtlich-politische Definition gilt, nach der Berlin Hauptstadt Deutschlands ist, sondern die Heimatverbundenheit (er stammt aus Köln und wohnt dort), hebt er die Bedeutung dieser *regionalen Bindung* hervor. – In gewisser Weise fügt sich dies in eine der zentralen Botschaften dieser Sendung (vgl. 6.5), derzufolge alles Ferne und Unbekannte von geringer Relevanz ist.

Von der Formalstruktur des Textes her gesehen, inszeniert St. R. hier einen fiktiven Dialog mit denen, die Istanbul für die Hauptstadt der Türkei halten, was sich u.a. wiederum (wie in 0:00-1:08) in die Strategie einer Kompensation der fehlenden Reziprozität im Bereich der Diskursorganisation interpretieren lässt.

#### 2:43-2:58 Der Schlepper „macht Stenz“

- *Proposition durch St. R. im Modus des Evaluierens resp. Theoretisierens*

Die dritte Bezugnahme auf Menschen, die hier in Istanbul wohnen, bzw. auf den Alltag hier, wird durch die Störung aufgrund des stärker werdenden Motorengeräusches sozusagen erzwungen (vgl. auch 6.3.4). Der türkische Alltag, das menschliche Leben am Bosphorus, greift gleichsam unkontrolliert in die Aufnahme ein. Wiederum nimmt der Kommentar die Form einer ironischen Diskriminierung an: Nachdem die Leute hier als Nepper, Schlepper und Bauernfänger sowie als Kleinkrimi-

nelle kategorisiert worden sind, machen sie, die hier am Bosphorus ihre Arbeit tun, nun auch „Stenz“, sind also kleine Wichtigtuer.

Im Fall der „Nepper, Schlepper, Bauernfänger“ („auch hier“) wie auch des „Stenz“ („bei uns mitm 3er BMW“) werden Parallelen zu Deutschland gezogen, in dem Sinne: kleine Gaunereien finden sich hier wie dort. Dadurch scheint die Diskriminierung zunächst wieder aufgehoben im Sinne einer Nostrifizierung („Verunsicherung“): die Menschen sind hier genauso wie bei uns zu Hause. Allerdings wird dort, wo überhaupt auf die Bevölkerung von Istanbul Bezug genommen wird, diese *ausschließlich* in diese Kategorien eingeordnet: Die Istanbuler sind genauso wie die kleinen Gauner und Angeber, die wir von uns zu Hause kennen. Man kann von einer Normalisierung der Ent-Normalisierung sprechen, die es ermöglicht, Andersartigkeit zu konstruieren, ohne dadurch in die Verlegenheit von Prozessen der Reflexion und des Fremdverstehens zu geraten, da diese Andersartigkeit ja schon bekannt ist. Wir haben es also mit der Doppelstruktur zu tun, die zugleich eine *Nostrifizierung* und *Diskriminierung* umfasst.

Wie bereits in 6.3.4 dargelegt, tritt hier ein Ton-Raum (das sich nähernde und dann wieder entfernende Motorgeräusch schafft ein Raumerlebnis), welcher den türkischen Alltag am Bosphorus repräsentiert, gleichsam in Konkurrenz zu dem montierten und dem fotografierten optischen Raum, welcher auf St. R. und die Sightseeing-Kulisse des Bosphorus fokussiert ist. Der Ton-Raum verschafft dem türkischen Alltag, der ansonsten (mit Ausnahme des „Bazars“) aus der Welt von „Istanbul Total“ herausgehalten wird, gleichsam in unkontrollierter Weise Zugang zu dieser Welt.

#### 2:59-3:05 Der „Musikantenstadl“

- *Proposition durch St. R. im Modus des Evaluierens*

Indem St. R. betont, dass er „froh“ ist, noch vor dem „Musikantenstadl“ in Istanbul zu sein, haben wir es hier mit einer Distinktion gegenüber dem Milieu derjenigen zu tun, die den Musikantenstadl präferieren, einer Distinktion der eigenen (wie auch immer gearteten) medial konstituierten Gemeinschaft gegenüber anderen – sich ebenfalls über das Medium Fernsehen konstituierenden – Gemeinschaften (vgl. 6.5). Er ist „froh“, diesen anderen medialen Protagonisten und deren Publikum nicht zu begegnen.

#### 3:06-3:26 Das „fantastische“ Hotel

- *Proposition durch St. R. im Modus des globalen Beschreibens*

Das Hotel ist „fantastisch“, weil es einen „herrlichen Blick“ auf den Pool von St. R. („meinen Pool“) eröffnet (nicht etwa wegen der Aussicht auf Istanbul oder den Bosphorus). Analog zu 2:15-2:29 wird auch hier – im Sinne einer permanenten Selbstbezogenheit – die Aufmerksamkeit von der Welt der Türkei weggeführt und wieder zurück zu St. R., der ja auch auf der Bildebene immer wieder den Fokus bildet.

Zugleich wird (auch durch die stereotype Wahl der Adjektive: „fantastisch“ und „herrlich“) der Blick auf den Pool auf eine Stufe gestellt mit dem Blick auf den Bosphorus und letzterer somit diskreditiert.

### 3:26-3:33 *Der „tolle“ Ausblick und die Sendezeit*

- *Proposition durch St. R. im Modus einer Handlungsempfehlung*  
Indem die Programmankündigung (3.28-3.32) mit der Empfehlung („Genießen Sie...“) zusammengezogen wird, wird sie dieser Empfehlung integriert.

Grammatikalisch bezieht sich auch hier der „Ausblick“ auf den Pool von St. R., so dass der „Ausblick“ auf Istanbul und den Bosphorus auf der verbalen Ebene wiederum zurückgelenkt wird in die Selbstbezogenheit.

Auf der Bildebene wird „Ausblick“ in Bezug zum Bosphorus gesetzt, weil St. R. sich diesem durch die Kopf- und Körperdrehung um 90° nach rechts zuwendet.

### 3:34-3:45 *Wortspiele mit „Istanbul“*

- *Proposition durch St. R. im Modus des Beschreibens*  
Die Wortspiele mit „Istanbul“ beziehen (wie dasjenige zur Schlepperbesatzung) ihren ‚Witz‘ aus der ironischen Diskriminierung oder Degradierung von Personen: Claudia Schiffer ist magersüchtig und Elton isst zuviel.

Es gelingt auch hier wiederum – in einer anderen Variante von Selbstbezogenheit – der Sprung von der Welt in Istanbul nach Deutschland, d.h. zur deutschen Sprache und zur deutschen Medienlandschaft auf dem Wege über zwei ihrer mehr oder weniger bekannten Figuren.

### 3:46-3:56 *Der „Schneidersitz“ in der orientalischen Ecke*

- *Proposition durch St. R. im Modus des Fragens*  
Mit der in der Frage: „wie setzt man sich hin im Schneidersitz?“ implizierten Proposition bringt St. R. seine Unbeholfenheit im Hinblick auf den Umgang mit dem zum Ausdruck, was hier – auf dem Wege einer ‚Exotisierung‘ sozusagen – als Element, als Symbol des türkischen Alltags von ihm selbst in seiner Fremdheit konstruiert worden ist: die orientalische Ecke, welche den Schneidersitz notwendig macht (die aber für den türkischen Alltag nicht eigentlich von Bedeutung ist).

[ Einspielung und Werbeblock herausgeschnitten ]

### 3:58-4:18 *Der Blick auf die Moschee oder die Kirche*

- *Proposition durch St. R. im Modus des Beschreibens (3:58-4:12) und des fiktiven Dialogs (4.13-18)*

Es ist für St. R. von sekundärem Interesse, dass es sich bei der Hagia Sofia in der Gegenwart um eine Moschee handelt. Er rückt jene historische Phase ins Zentrum, in der sie eine Kirche war. Denen, die behaupten, dass es sich nicht um eine Kirche handele, bescheinigt er, „Klugscheißer“ zu sein. Auch hier haben wir es (wie in 2.43-2.85) mit einer Spielart von „Nostrifizierung“ zu tun: die Hagia Sofia gilt ihm in erster Linie oder ursprünglich als eine Kirche, also als ein Denkmal westlich-christlicher Kultur, und erst sekundär als eine Moschee.

St. R. führt wiederum einen fiktiven Dialog wie in 2:30-2:42. Da sich nun zeigt, dass St. R. an dieser Strategie (bzw. diesem Stilmittel) auch dann festhält, wenn Studiogäste neben ihm sitzen, mit denen ein realer Dialog sich anbietet, wird deutlich, wie stark St. R. an der monologischen Struktur orientiert ist.

Zudem wird aufgrund der Homologien mit 2.30-2.42 (Istanbul ist nicht die Hauptstadt der Türkei) nunmehr deutlich, dass St. R. fiktive Dialoge in einer Weise inszeniert, dass er selbst als der Belehrende auftritt (wodurch die monologische Struktur noch deutlicher hervortritt).

### 4:19-4:55 *Gülçan und Elemente ihrer Lebensgeschichte*

- *Propositionen durch St. R. im Modus des Fragens (4:44-23; 4:26-27; 4:36-37)*  
*Anschlusspropositionen durch Gülçan im Modus des Beschreibens (4:24-4-2) und des Erzählens (4.28-36; 4.38-55)*

Da Gülçan nicht in der Türkei lebt und auch dort nicht geboren ist, sondern lediglich ihre Eltern, zeigt es sich somit, dass St. R. für das Studiogespräch dieser ersten Folge von „Istanbul Total“ keine Gesprächspartner/innen aus Istanbul oder der Türkei eingeladen hat (wie auch der Blick auf den weiteren Verlauf dieses Abends bestätigt).<sup>118</sup> Ein Zugang zum Erfahrungsraum, zur Perspektive und somit zur Lebenspraxis der autochthonen Bevölkerung wird also auch auf diesem Wege nicht gesucht. Auch hier tritt St. R. nicht in eine Kommunikation mit der Welt der Türkei.

St. R. unterstellt Gülçan, dass sie die Sprache „perfekt“ spricht. Er suggeriert es ihr. Ihre Antwort zeigt aber, dass von Perfektion wohl nicht die Rede sein kann. Indem St. R. dies aber für wesentlich hält, dokumentiert sich hier ein (vergebliches) Bemühen, einen Bezug zur türkischen Kultur herauszustreichen.

### 4:56-5:16 *Der Einkauf von Gülçan und Elton in der „modernen“ „Metropole“ Istanbul*

- *Propositionen durch St. R. im Modus des Fragens (4:56-5.01; 5:12-13; 5:15) und des Beschreibens (5.02-5.11)*  
*Antwort durch Elton (4.57; 5:01-02; 5:14-16) und Gülçan (4:57)*

Der nun folgende Bezug zur türkischen Alltagskultur bleibt wiederum, wie in 2:01-2:14, auf den Bereich des Shopping beschränkt. Als ob St. R. diese Selektivität nicht ganz so stehen lassen möchte, wird in einem Satz der Charakter von Istanbul als „moderne Stadt“ und „Metropole“ erwähnt (analog zur „fantastischen“ Schifffahrt auf dem Bosphorus), um dann wieder auf das Shopping zurückzukommen. St. R. richtet Fragen an Gülçan und Elton, erteilt ihnen somit das Wort, um dies aber, als Elton es ergreift (4:53-54) und über seine und Gülçans Erlebnisse berichten will (5:14-16), sogleich wieder zurückzunehmen und sich selbst, d.h., seine eigenen Einkäufe, in den Fokus zu rücken.

Die monologische Struktur (und auch die „Monostrukturierung“ durch St. R.; siehe: 6.3.4) wird hierdurch bestätigt.

<sup>118</sup> Eine Ausnahme bildet der Auftritt des – die internationale Medienwelt repräsentierenden – türkischen Schlagerstars Sertab (siehe auch: Kap. 6.5).

#### 5:17-5:41 Billige Taschen und Uhren und „Raubkopien“ von CDs

- *Proposition durch St. R. im Modus der Erzählung*

St. R. erzählt, was er „auf dem Basar gefunden“ hat. Dies sind nun nicht etwa Waren, die aufgrund ihrer Originalität oder ihres Bezuges zur türkischen Kultur oder zum türkischen Alltag interessant sind, sondern globale Konsumartikel, die hier besonders „billig“ erworben werden können. Vor allem aber wird mit den „Raubkopien“ wieder der Bezug zur Kleinkriminalität und damit zum opaken, nicht ganz geheueren, türkischen Alltag hergestellt.

#### 5:42-6:21 „Raubkopien“ der CD von „MAX“

- *Anschlussproposition durch St. R. im Modus der detaillierten Erzählung*

Die Thematisierung von „Raubkopien“ im Allgemeinen wird nun hingeführt hin zur speziellen Thematik der Raubkopien von CDs und schließlich zur CD von „MAX“, die (angeblich) ebenfalls als „Raubkopie“ zu finden war. Die gegenüber dem Original veränderten Titel sind – wie das Zitat des Straßenhändlers belegen soll – nicht beabsichtigt, sondern im guten Glauben entstanden, eine möglichst getreue Kopie des Originals zu produzieren.

Es bestätigt sich damit, das, was bereits in der Bilddimension herausgearbeitet werden konnte (6.3.5): Vorgeführt wird nicht nur die Illegalität der Aktivitäten, sondern auch die Unbeholfenheit, die Ungeschicklichkeit der Akteure: Ausgearbeitet wird damit das Stereotyp der kleinen Gauner.

Ob die vor der Kamera präsentierten Fälschung nun ‚echte Fälschungen‘ oder gefälschte (von St. R. zu Zwecken eines Gags produzierte) Fälschungen sind, spielt für die sich hierin dokumentierende doppelte Diskriminierung der Repräsentanten des türkischen Alltags – als Gauner und zugleich ungeschickte (kleine) Gauner – keine entscheidende Rolle.

Diese Sequenz hat formal gesehen auch hier in der Textdimension eine Fokussierungsqualität, da es sich (wie auch die Wiedergabe wörtlicher Rede indiziert) um die detaillierteste Darstellung (Erzählung) der hier untersuchten Eingangspassage handelt. Die Sequenz zeichnet sich somit sowohl in der Text- als auch in der Bilddimension (vgl. 6.3.5: Fokussierungssequenz I: Detaileinstellung) durch ein hohes Detaillierungsniveau aus (sowohl erzähltheoretisch als auch bildtheoretisch steht ein hohes Detaillierungsniveau für Fokussierung). Doppelt fokussiert wird damit zugleich die (unterstellte) Kleinkriminalität im türkischen Alltag wie auch (durch diesen Aufmerksamkeitsgewinn unterstützt) die Promotion der CD von „MAX“ (an dessen Erfolg als Teilnehmer des Eurovision Song Contest 2004 St. R. wie auch der Sender Pro 7 ein unmittelbares kommerzielles Interesse haben).

#### *Schluss*

Durchgehend rahmt St. R. jene Darstellungen, in denen die Landschaft und die architektonischen Sehenswürdigkeiten der Türkei, also Objekte des Sightseeing, sowie spektakuläre mediale Ereignisse (0:00-0:16; 0:17-0:49; 1:09-2:00; 2:15-2:29; 2:59-3:05 und 3:06-3:26; 3:46-3:57) thematisiert werden, als „fantastisch“. Genauer betrachtet, bildet „fantastisch“ erst die Überleitung, die Verbindung zwischen den z. T. thematisch vollkommen diskrepanten Sequenzen.

Die Kategorisierung der touristischen Attraktionen als „fantastisch“ bleibt dabei seltsam vage oder unausgefüllt. So beispielsweise, wenn das Thema der Schifffahrt auf dem Bosphorus als „fantastisch“ eingeleitet und beendet wird, ohne dass überhaupt eine Schilderung dieser Sehenswürdigkeiten eingebracht würde. Die Etikettierung als „fantastisch“ gewinnt damit vielmehr die Funktion, das Fehlen eines derartigen Sich-Einlassens zu überdecken. Der Bosphorus bleibt somit Kulisse.

Jene Äußerungen, welche die Verhaltensweisen von Menschen im türkischen Alltag thematisieren, lassen ein derartiges Sich-Einlassen gleichermaßen vermissen, stehen aber dennoch im Kontrast zu dieser positiven Stereotypisierung im Bereich der Attraktionen des Sightseeing. Sie haben die Form einer Karikierung, einer tendenziell negativen Stereotypisierung, die mit Elementen der Diskriminierung verbunden ist.

Insgesamt ist der gesprochene Text auf der thematischen Ebene – zumindest, was die Oberthemen anbetrifft – zusammenhanglos. Die Oberthemen stehen (wie bereits der Blick auf die thematische Gliederung in 6.4.2 zeigt) unverbunden nebeneinander (dies findet in der Diskursorganisation darin seinen Ausdruck, dass wir es durchgängig mit Propositionen und nur in einem Fall mit einer Anschlusspropositionen zu tun haben). Der Zusammenhang wird auf der Ebene des *Bildes* hergestellt, genauer: der abgebildeten Szenerien, deren Struktur aber wiederum, wie dargelegt, durch St. R. dirigiert wird: Ich hatte dargelegt, dass der Wechsel von Szenerie und Einstellung den Objekten oder Personen folgt, denen St. R. sich zuwendet, oder dass er den Personen folgt, die sich ihm zuwenden. Hier zeigen sich Homologien zu dem, was den *Text* überhaupt noch in gewisser Weise zusammenhält. Dies findet sich unterhalb der thematischen Ebene, also auf derjenigen des Orientierungsrahmens, in dessen Fokus die Selbstbezogenheit auf St. R. und die anderen Protagonisten steht sowie der selbstreferentielle Bezug dieses Fernsehereignisses auf andere Fernsehereignisse mit der Funktion des parasitären Aufmerksamkeitsgewinns.

## 6.5 Reflektierende Gesamtinterpretation

Der Ort, an dem die Sendung live produziert wird, vor allem aber der Titel der Sendung („Istanbul Total“) suggerieren, dass St. R. sich in die Welt von Istanbul begibt bzw. sich sogar umfassend („total“) auf diese einlässt. In Medienberichten und -kommentaren *über* diese Sendung wird ein derartiger Anspruch explizit formuliert, wie ich einleitend (vgl. 6.1) dargelegt habe.

Eine genauere Betrachtung zeigt, dass dieser Anspruch nicht eingelöst wird: Diese Welt von Istanbul und der Türkei ist nicht etwas, mit dem St. R. in Interaktion und Kommunikation tritt, sondern Gegenstand entfernter und ironisierender Betrachtung. Nicht allein, dass St. R. im Studio bzw. auf dessen Balkon verbleibt. Wichtiger ist noch, wie er sich dort präsentiert. Sowohl dort, wo er auf dem Balkon sich befindet, als auch dort, wo er im Studio sitzt, ist er in der weitaus überwiegenden Zahl der Bilder von der Welt Istanbul bzw. der Türkei durch ein Gitter getrennt ist, welches das Bild dominiert. Es wird dabei zwischen St. R. und der Umwelt aber nicht nur eine *Grenzlinie* gezogen. Es wird ebenso markiert, dass Stefan Raab eine Position *oberhalb* dieser Umgebung einnimmt, von oben herab blickt.

Das Gitter und die Position von St. R. zusammen mit dem Hintergrund des Meeres erwecken auch den Eindruck, dass St. R. an der Reling oder auf der Brücke eines Schiffes steht und sich somit gar nicht in Istanbul befindet.

- *Die Übergegensätzlichkeit von Anwesenheit und Abwesenheit*

Insgesamt dokumentiert sich hier eine *gegenläufige Bewegung* oder eine *Übergegensätzlichkeit*: St. R. begibt sich in die Welt von Istanbul bzw. umgibt er sich mit dieser, hält sich aber zugleich von ihr und ihrem Alltag fern, wird von diesem getrennt und/oder vor ihr geschützt. Diese Welt ist somit nicht etwas, mit dem man in Interaktion und Kommunikation tritt. Der Blick auf diese Welt oder Umwelt hat den Charakter einer Aneinanderreihung von Objekten des touristischen Sightseeing wie auch des Fern-Sehens im wahrsten Sinne des Wortes. Diese Umwelt bleibt somit in gewisser Weise auch Kulisse und Dekoration. Hier zeigen sich Homologien zur Bekleidung von St. R., die als Kostümierung erscheint – insbesondere auch deshalb, weil diese dem türkischen Alltag weitgehend fremd ist, lediglich im Museum oder Theater regelmäßig vorkommt.

Die Anwesenheit in Istanbul bei gleichzeitiger Distanz gegenüber seinem Alltag, die als *Übergegensätzlichkeit von Anwesenheit und Abwesenheit* bezeichnet werden kann, dokumentiert sich sowohl innerhalb der Fotogramme als auch in ihrer Sequenzierung, d.h. in Einstellungswechsel und Montage resp. Bildmischung: Gleich im ersten Einstellungswechsel der Show „Istanbul Total“ vollzieht sich im Wechsel von der amerikanischen Einstellung zu derjenigen der Totale (mit der Technik des Zooms) einerseits eine weitergehende Öffnung hin zur Umwelt, zur Hafenanlage des türkischen Bosphorus. Andererseits wird dies aber konterkariert bzw. geradezu ins Gegenteil verkehrt dadurch, dass damit zugleich das Balkongitter, welches St. R. vom Bosphorus und der Türkei fernhält, sich noch massiver in das Bild schiebt.<sup>119</sup>

- *Die Konstruktion von Räumlichkeit und die Hyperzentrierung auf die Person des Moderators*

Wenn wir uns nun weitergehend der Dimension der Kameraführung und des Schnitts resp. der Bildmischung, also den Leistungen der abbildenden Bildproduzent(inn)en, zuwenden, so hatte ich als wesentliches Prinzip des Wechsels von Sequenzen (des Wechsels von Szenerie und/oder Einstellung) herausgearbeitet, dass dieser Wechsel wesentlich dem Blick- bzw. der Zeigerichtung und der Bewegung von St. R. folgt – entweder durch harten Schnitt oder durch Kameraschwenk oder Zoom. Auf diese Weise wird hier Räumlichkeit strikt vom Standort von St. R. konstruiert.

Dies geschieht einerseits auf dem Wege des durch Montage bzw. Bildmischung hergestellten Raumes, des „*montierten Raumes*“. Andererseits findet diese Zentrierung auf St. R. aber auch in der Dimension des „*fotografierten optischen Raumes*“

119 In jener Einspielung, die aus der hier zugrunde gelegten Internet-Version herausgeschnitten und von St. R. selbst als „Exkurs auf den Bazar in Istanbul“ bezeichnet wurde, begeben sich St. R. und MAX in der Rolle der Touristen in Begleitung eines Fremdenführers in den Bazar. Abgesehen davon, dass wir uns hier in der Welt des touristischen Sightseeing befinden, ließe sich (wenn der entsprechende Raum für die Analyse zur Verfügung stünde) auch hier zeigen, dass diese Situation durch den primären Rahmen der ‚Türkei als Welt der kleinen Gauner‘ (siehe unten) strukturiert ist und sich fast ausschließlich darum dreht, dass St.R. den Händlern unterstellt, ihn übers Ohr hauen zu wollen. Eingeleitet wird der Besuch mit der Frage an den Fremdenführer: „Wie groß ist die Wahrscheinlichkeit, dass man hier übers Ohr gehauen wird?“.

ihren Ausdruck, welcher einerseits durch die Kamerabewegung, insbesondere den Kameraschwenk, hergestellt wird. Auch hier folgt die Kamera St. R., also nicht allein seinem Blick oder seiner Zeigerichtung, sondern seiner Bewegung, seinem Ortswechsel insgesamt. Somit dokumentiert sich u.a. auch in dieser *Konstruktion von Räumlichkeit* auf der Basis des fotografierten optischen Raumes wie auch des montierten Raumes, dass diese in umfassender Weise aus der Perspektive von St. R. und von dessen Standort aus konstruiert werden. Hinsichtlich dieser Zentrierung auf seine Person zeigen sich Homologien auch zur Interpretation der Fotogramme, die uns Aufschluss über die Formalstruktur des *fotografierten optischen* Raumes geben. St. R. ist in der umfangreichsten Hauptsequenz allein abgebildet, steht überwiegend im perspektivischen Zentrum und ist planimetrisch fokussiert.

Dabei zeigt sich an dem vorliegenden Video sehr deutlich, dass dieses seine Struktur nicht durch die verbalen Äußerungen, also nicht durch die Textdimension erhält. Insgesamt ist der gesprochene Text auf der thematischen Ebene – zumindest, was die Oberthemen anbetrifft – zusammenhanglos. Die Oberthemen stehen unverbunden nebeneinander. Seine übergreifende Struktur erhält dieses Video allein in der Bilddimension, welche ihrerseits wesentlich durch St. R. strukturiert wird.

Indem St. R. die Konstruktion des filmischen Raumes und damit die Richtung und Selektivität der Performanz der *abbildenden* Bildproduzent(inn)en dirigiert und zugleich in planimetrischer wie perspektivischer Hinsicht im Zentrum des fotografierten optischen Raumes steht, also auch als *abgebildeter* Bildproduzent die Performanz des Videos bestimmt, haben wir es hier mit einer *Personalunion von abbildendem und abgebildetem Bildproduzenten* zu tun, sodass ich hier von einer *Monostrukturierung durch den Showmaster* und einer *Hyperzentrierung* auf seine Person sprechen möchte. Die Struktur der Sendung und der Habitus des Moderators sind somit nicht mehr zu trennen.

- *Die Identifikation mit dem Moderator und die Herstellung gemeinsamer Erfahrungsräume*

Die alles strukturierende und stilistisch prägende Person des Moderators ist eine ganz wesentliche Komponente der Herstellung von Gemeinsamkeit, Vertrautheit und Kontinuität in der Beziehung des Publikums zum Moderator. Allerdings gilt es genauer zu bestimmen, in welcher Weise seine Person hier Bedeutung gewinnt, welche Art von Habitus oder Identität des Moderators hier im Zentrum steht.

Wenn wir davon ausgehen, dass das Gesicht, die Physiognomie, einer der bedeutendsten Areale zur persönlichen, zur individuellen Identifikation eines Menschen ist, dann ist es von Bedeutung, dass St.R. in dem analysierten Ausschnitt lediglich 1 Sek. in Groß-Einstellung zu sehen und dies offensichtlich auch noch der Aufnahmetechnik und Ästhetik geschuldet ist. Da eine Ablichtung der Person von St. R. mit Fokussierung auf seine Physiognomie offensichtlich vermieden wird bzw. nicht von Bedeutung ist, geht es – trotz der hier zu beobachtenden Hyperzentrierung auf seine Person – also offensichtlich nicht um seine persönliche (individuelle) Identität.<sup>120</sup> Das legt nahe, dass St. R. – und allgemeiner Moderator/innen

120 Was ja auch – wenn wir Kontextwissen heranziehen – darin zum Ausdruck kommt, dass St. R. und vergleichbare prominente Moderatoren ihre Privatsphäre aus der medialen Präsentation weitest möglich herauszuhalten suchen.

mit einer derartigen Hyperzentrierung auf ihre Person – eher als Repräsentanten einer *sozialen* Identität, eines kollektiven (milieuspezifischen) Habitus bzw. als Repräsentanten von Stereotypen in diesem Bereich von Bedeutung sind.

Dies wiederum lässt sich in der Weise plausibilisieren, dass ein derartiger „Kult“ der Identifikation mit Stil und Habitus des Moderators wesentlich die Funktion der Konstitution oder Selbstvergewisserung von Gemeinsamkeiten zwischen Publikum und Moderator und auf diesem Wege auch des Publikums untereinander hat. Hier lässt sich vermuten, letztlich aber nur durch eine Rezeptionsanalyse klären (vgl. 5.2.2), dass es (mit bestimmten Bereichen des Publikums) zu einer Art Milieubildung, zur Konstitution von rudimentären konjunktiven Erfahrungsräumen (zum Begriff: Kap. 2.3), von Gemeinsamkeiten, Zugehörigkeiten im Sinne habitueller Übereinstimmungen kommt. Nahe gelegt wird dies auch in der Text-Dimension durch die von St. R. klar artikulierte *Distinktion* gegenüber anderen derartigen über das Medium Fernsehen konstituierten Milieus (so durch die Distinktion gegenüber denjenigen, die den „Musikantenstadl“ präferieren; vgl. 2:59-3:05).

Die Identifikation mit dem Moderator ist dabei nicht an die Perfektion von Inszenierungen gebunden. Vielmehr könnten (wie in einer Rezeptionsanalyse zu klären wäre) gerade ‚Schwächen‘ und Unbeholfenheiten Grundlage für Identifikationen sein (vgl. 6.3.5).

- *Die monologische Struktur und die reduzierte Reziprozität mit Publikum und Studiogästen*

Diese Monostrukturierung und Selbstbezogenheit in der Bilddimension findet in homologer Weise auch in der monologischen Struktur *des Textes* ihren Ausdruck. So lauten die allerersten Worte (nach der in der Internet-Version herausgeschnittenen Begrüßungsformel), die St. R. in der Show spricht: „Ich befinde mich hier (...)“. St. R. initiiert im Folgenden Erzählungen seiner beiden Podiumsgäste hinsichtlich ihrer Erfahrungen beim Shopping, unterbricht diese aber sogleich wieder, um über seine eigenen derartigen Erfahrungen zu berichten.

Gestützt wird diese Rahmenkomponente der Monostrukturierung und Selbstfokussierung auch durch die Beobachtung der Funktion des Publikums, indem eine Kommunikation mit diesem nicht aufgenommen, sondern es lediglich in seiner Funktion als Applausgeber ins Bild gerückt wird (vgl. 6.3.4).

Dort, wo St. R. keinen Applaus erhält, wodurch der Eindruck einer zumindest rudimentären Reziprozität geschaffen wird, findet sich die (in der alltäglichen Kommunikation wohl nicht seltene) Strategie, die fehlende Reziprozität dadurch zu überspielen, dass er seinen Sätzen durch ein nachgestelltes „ja“ mit Frageintonation (u.a. 0:12; 0:16; 0:34; 0:45/46; 0:53; 1:08.) den Charakter einer Frage gibt („tag-question“) und somit eine dialogische Struktur suggeriert (vgl. 6.4.3). Eine ähnliche Funktion erhält die Inszenierung eines fiktiven Dialogs (mit dem Publikum) in 2.30-2.42 u. in 4.13-4.18.

Homolog zum fehlenden Dialog mit dem Publikum findet in dieser ersten Sendung auch auf dem Podium kein Dialog mit Gesprächspartner(inn)en aus Istanbul oder der Türkei statt. Denn auch die Gesprächspartnerin Gülçan hat zwar Eltern türkischer Herkunft, ist aber – ebenso wie der Gesprächspartner Elton – in Deutschland geboren und dort aufgewachsen, und auch die türkische Sprache ist ihr, wie sie selbst ausführt (4.24-4.25) keineswegs selbstverständlich. Über das Gespräch mit

den Gästen im Studio wird somit ein Zugang zum Erfahrungsraum, zur Perspektive und zur Lebenspraxis der autochthonen Bevölkerung, also zur Welt der Türkei nicht eröffnet.

- *Die Unzugänglichkeit der Innenwelt der Türkei*

In die im Studio aufgebaute „orientalische“ Ecke, welche paradigmatisch steht für das Innere des türkischen Hauses und somit in gewisser Weise für die innere Welt der Türkei, sind also Gesprächspartner/innen aus Istanbul oder aus der Türkei nicht eingeladen.<sup>121</sup> Dies verweist bereits darauf, dass diese orientalische Ecke, die hier – in einem Schritt der Exotisierung sozusagen – als Symbol für den türkischen Alltag konstruiert wird, Dekoration oder Kulisse bleibt. Sie ist als solche in ihrer Gestaltung besonders aufschlussreich: Diese Ecke wirkt vergleichsweise dunkel, die Gegenstände bleiben unkonturiert sowohl aufgrund geringer Helligkeit als auch aufgrund der Farbgebung. Während die *äußere* Welt Istanbuls und der Türkei: die Natur, das Meer, die Objekte des Sightseeing, zwar auf Distanz gehalten werden, aber hell, offen und klar strukturiert sind, erscheint die andere – die *innere* – Seite der türkischen Welt eher unstrukturiert, ungeordnet, undurchschaubar oder opak und somit nicht ganz geheuer. Durch die weiche, konturlose Figuration und das dämmrige Licht könnte der untere Bereich gemütlich oder heimelig wirken. Dem stehen allerdings Elemente der Spannung entgegen, wie sie im Wesentlichen durch die Positionierung der Personen ins Bild kommen: Durch die Wahl der Perspektivität weist dieses Bild ein starkes Gefälle nach links und vor allem nach vorne auf. Die Personen – insbesondere Gülçan und Elton – scheinen gleichsam wegzurutschen.

Die vermeintlichen Accessoires dieser Lebenspraxis bleiben – wie die Präsentation der Sitzecke, aber auch die Selbstpräsentation von St. R. in seiner Verkleidung zeigen – *Kulisse* und *Dekoration* und werden als solche auch karikiert. Dieses ließe sich durchaus als eine *Stilisierung* der eigenen Unbeholfenheit in der Begegnung mit der fremden Kultur und somit als eine (selbstreflexive) Pose interpretieren. Allerdings wird dieser kulissenhafte Zugang zur fremden Welt eben auch in anderen Dimensionen und dort auch ohne Selbststilisierung erkennbar. Denn auch die äußere Welt des Bosphorus und der Sightseeing-Objekte bleibt Kulisse.

Auf der Bildebene findet der *fehlende Zugang zum fremden Erfahrungsraum* bzw. die fehlende Orientierung an einem solchem Zugang ihren Ausdruck wesentlich im opaken Charakter, im Charakter der Undurchschaubarkeit, welcher durch die Komposition des Fotogramms im Studio mit der orientalischen Sitzecke von der inneren Welt der Türkei vermittelt wird.

- *Die Türkei als Welt der kleinen Gauner*

Dieser opake Charakter der inneren Welt, der diese ein wenig undurchschaubar und nicht ganz geheuer erscheinen lässt, dokumentiert sich auch in den – sowohl in der Bild- wie in der Textdimension fokussierten – „Raubkopien“. Auf der Textebene begründet St. R. seine „Empfehlung“ nach Istanbul zu fahren, einerseits mit der „fantastischen“ Kulisse des Sightseeing (des Blicks auf den Bosphorus) und zum an-

<sup>121</sup> Eine Ausnahme bildet der türkische Schlagerstar Sertab (Erener). Sie repräsentiert die Welt der (internationalen) medialen Größen und tritt neben dem deutschen Moderator (mit arabischen und türkischen Vorfahren) Kaya Yanar und dem zu promotenden Max im weiteren Verlauf dieses Sendeabends auf, um dann u.a. „Heideröslein“ zu singen.

deren mit den billigen Waren der Straßenhändler, die (seiner Darstellung zufolge) im Wesentlichen aus „Raubkopien“ bestehen: „ich kann Ihnen nur empfehlen fahren Sie hier hin denn Sie kriegen hier einfach alles überall Händler auf der Straße die die Jacke aufmachen“ (2.04-10).

Vorgeführt wird dabei nicht nur der illegale Charakter der Aktivitäten, die ‚kleineren Gaunereien‘ und undurchschaubaren Trickereien der Händler: „und alles ganz billig hab mich gefragt warum ist das so ja“ (5:30-31). Vorgeführt wird auch die Ungeschicklichkeit oder Unbeholfenheit der Akteure. Eine Fälschung, die vom Original so gut wie nicht zu unterscheiden ist, bringen diese nicht zustande. Ausgearbeitet wird damit das Stereotyp der ‚kleinen Gauner‘. Die Türken sind zu ungeschickt, um große Gauner sein zu können. Sie sind eben die „Nepper, Schlepper, Bauernfänger“, die wir schon von zu Hause (aus den deutschen Medien) kennen, wie St. R. an anderer Stelle mit Bezug auf die Schlepperbesatzung assoziiert. Die Welt des Alltags von Istanbul – also diejenige jenseits der „fantastischen“ Kulisse eines Sightseeing und der Welt der Hotels – ist so wie die uns von zu Hause bekannte Welt der kleinen Gauner und zwielfichtigen Gestalten. Es ist eine andere Welt, aber letztlich in einer Andersartigkeit, die wir in ihren negativen Aspekten von zu Hause schon kennen, um deren genauere Kenntnis wir uns also nicht mehr bemühen müssen.

- *Die Haltung des Fern-Sehens, die Gewissheiten der medialen Welt und der parasitäre Aufmerksamkeitsgewinn*

Es reicht also vollkommen – dies scheint ein Element der Botschaft dieser Sendung zu sein – wenn wir uns diese Welt vermittelt der Medien präsentieren lassen. Und dies ist auch besser so, denn diese Welt von Istanbul und der Türkei erscheint – als Produkt der Konstruktion von Undurchschaubarkeit und Exotisierung – nicht ganz geheuer und letztlich unzugänglich. Da ist es angebracht, sie aus der Ferne zu betrachten, zum Gegenstand des Fern-Sehens im wahrsten Sinne des Wortes zu machen und auch die kleinen Abenteuer, wie den Einkauf auf dem Basar, den Protagonisten auf der Fernseh-Bühne zu überlassen.<sup>122</sup> Vollständig vertraut erscheint dagegen eine andere Welt, deren Thematik die Kommentare von St. R. dominiert. Es sind dies die *medialen Großereignisse und Größen*, genauer: diejenigen des Fernsehens: der Fußballclub Fenerbahçe, der gerade Meister geworden ist, der Musikantenstadl, Claudia Schiffer, die Sendung „Nepper, Schlepper, Bauernfänger“.

Mit dem Bezug auf allgemein bekannte mediale Ereignisse ist *zum einen* – durch deren Wiedererkennungswert – ein gewisser Aufmerksamkeitsgewinn garantiert oder zumindest sehr wahrscheinlich. Und *zum anderen* wird – mit dem damit verbundenen Bezug auf gemeinsame Wissensbestände – an eine elementare oder rudimentäre Gemeinsamkeit und Vertrautheit mit dem Publikum und des Publikums untereinander appelliert. Ich möchte mit Bezug auf beide Komponenten, wie bereits angesprochen, von einem *parasitären Aufmerksamkeitsgewinn* sprechen: Das mediale Ereignis „Istanbul Total“ lebt nicht zuletzt oder sogar primär vom Bezug auf andere Ereignisse des Mediums Fernsehen.<sup>123</sup>

<sup>122</sup> Das Fernsehen wird in diesem Sinne, wie Hepp (2004: 126) im Kontext der Cultural Studies formuliert, zu einem „Kommunikator der kollektiven Selbstvergewisserung“ – mit Bezug auf spezifische Milieus, wie hier ergänzt werden müsste.

<sup>123</sup> In der Sprache der Cultural Studies stellt dies eine spezifische Ausprägung des Charakters der „Intertextualität“ von Medienprodukten dar (vgl. Mikos 2003a: 261 ff.).

Der Blick, den wir mit St. R. auf die Welt von Istanbul und der Türkei werfen, ist aber nicht nur in mehrfacher Weise über das *Medium Fernsehen* vermittelt. Er ist darüber hinaus und vor allem über eine (im ursprünglichen Sinne des Wortes zu verstehende) *Haltung des Fern-Sehens* vermittelt. Indem nämlich in der Haltung, die St. R. dieser Welt gegenüber einnimmt, mit ihr eine Kommunikation im eigentlichen Sinne, im Sinne von Reziprozität, nicht wirklich möglich wird. Diese Welt bleibt lediglich Objekt der Betrachtung, des Sightseeing. St. R., der ja selbst Produkt des Fernsehens ist, begibt sich gegenüber der Welt von Istanbul und der Türkei in eine analoge Rolle. Zudem promoted St. R. mit dieser Show ein weiteres mediales Ereignis, nämlich den Song-Contest bzw. dessen Teilnehmer „MAX“. Schließlich sind die Gäste auf der Bühne im Studio selbst auch Medienstars und Moderator(inn)en – wenn auch von bescheidenerer Größe. Der Titel seiner Show, also „TV-Total“, ist in diesem Sinne somit durchaus treffend, der Titel dieser Sendung „Istanbul Total“ allerdings überhaupt nicht. Einen Zugang zur Welt von Istanbul erhalten wir keineswegs – schon gar nicht in ihrer Totalität.

## 6.6 Komparative Analyse und Gattungsanalyse

Die hier vorgelegte Skizze einer Produktanalyse dient als exemplarische Fallanalyse primär der Demonstration eines empirisch-methodischen Zugangs im Bereich der Videointerpretation und seiner methodologischen Grundeinstellung. Dabei geht es insbesondere darum, der Bilddimension eine größere systematische Beachtung zu schenken als dies in anderen Methoden der Fall ist (vgl. 5.3.3). Es hat sich gerade an diesem Video gezeigt, dass es als ein zusammenhangloses Produkt erscheint, wenn man auf der Ebene des Textes verbleibt bzw. sich primär an dieser orientiert. Zugleich sind aber Fortschritte in methodischer Hinsicht davon abhängig, inwieweit es gelingt, sich an generellen methodischen und methodologischen Grundprinzipien und Standards zu orientieren, die gleichermaßen für die Text- wie für die Bildinterpretation Gültigkeit haben (vgl. auch Kap. 3.9 sowie Bohnsack 2005a).

Verallgemeinerbare Aussagen zu Strukturen und Typen des Mediums Fernsehen – insbesondere im Hinblick auf dessen unterschiedliche Gattungen – sind von einer derartigen isolierten Fallanalyse einer einzelnen Abendveranstaltung selbstverständlich nicht zu erwarten. Dazu bedarf es des Fallvergleichs, also einer systematischen komparativen Analyse, welche die Konstruktion einer (mehrdimensionalen) Typologie ermöglicht. Eine derartige Typenbildung ist im Bereich der dokumentarischen Methode auf der Grundlage von Textinterpretationen bereits ausgearbeitet worden und weit fortgeschritten (vgl. u.a. Bohnsack 2007c u. 2009a). Im Bereich der dokumentarischen Bild- und Videointerpretation steht sie noch am Anfang.

Auch hat die komparative Analyse die Funktion, die eigene Standortgebundenheit des Interpreten zumindest ansatzweise der methodischen Kontrolle zuzuführen, indem die für jegliche Interpretation konstitutiven Vergleichshorizonte zunehmend durch empirische (und somit intersubjektiv überprüfbare) Vergleichshorizonte (Vergleichsfälle) ersetzt werden (vgl. 2.4 sowie Bohnsack 2001). Auch in dieser Hinsicht stehe ich mit der hier vorgelegten Fallanalyse noch ganz Anfang.

Es ist Angela Keppler (2006: 139) zuzustimmen, wenn sie Fernsehanalyse ganz wesentlich als „Gattungsanalyse“ versteht und diese methodisch auf das „komparative Verfahren“ stützen will: „Erst durch dieses komparative Verfahren kann beides erreicht werden: eine genaue Interpretation einzelner Sendungen und eine am Material belegte Interpretation der übergreifenden Formen, in denen sich die TV-Kommunikation zu einer bestimmten Zeit vollzieht.“

Bezogen auf die hier begonnene Interpretation einer Fernsehshow würde eine derartige komparative Analyse sowohl Veranstaltungen derselben Show („Istanbul Total“) an verschiedenen Abenden als Vergleichshorizonte heranziehen als auch andere Veranstaltungen desselben Moderators („TV Total“). Eine komparative Analyse mit anderen Folgen von „Istanbul Total“ findet sich in der Dissertation von Stefan Hampl (2009). Auf der Grundlage der Untersuchung drei weiterer Folgen können dort wesentliche Komponenten des Orientierungsrahmens, welche hier auf der Grundlage der ersten Folge herausgearbeitet worden sind, überprüft und weitgehend bestätigt werden, wie u.a. die Konstruktion der Opakheit und Undurchschaubarkeit der Innenwelt der Türkei sowie deren Konstruktion als Welt der kleinen Gauner. Auch die Struktur-Komponente der Monostrukturierung durch den Moderator wird in der komparativen Analyse mit anderen Folgen von „Istanbul Total“ und „TV Total“ einer genaueren Überprüfung unterzogen.

Eine komparative Analyse sollte auch Veranstaltungen einbeziehen, die – zumindest auf den ersten Blick – Gattungsähnlichkeiten aufweisen, aber von anderen Moderatoren oder Moderatorinnen getragen werden (bspw. die „Harald Schmidt Show“ und die Show „Was guckst Du?“ von Kaya Yanar bieten sich hier an). Auf diesem Wege könnte überprüft werden, inwieweit einige der an der vorliegenden Fallanalyse aufgewiesenen Strukturelemente für die gesamte Gattung, das Genre oder Format Late Night Show insgesamt Gültigkeit hat und wo es sinnvoll und notwendig ist, Untergattungen auszudifferenzieren.

Auf ihre generelle Bedeutung für die Struktur von Shows zu überprüfen sind vor allem zwei am Fall „Istanbul Total“ herausgearbeitete Rahmenkomponenten: zum einen diejenige einer *Monostrukturierung durch den Showmaster*, einer Hyperzentrierung auf dessen Person, indem dieser zugleich die Performanz der abgebildeten wie auch der abbildenden Bildproduzent(inn)en durchgängig strukturiert. Die Struktur der Sendung insgesamt und der individuelle Habitus des Moderators sind somit nicht mehr zu trennen. Die andere Rahmenkomponente ist diejenige des *parasitären Aufmerksamkeitsgewinns*: Welche (Unter-) Gattungen von Shows sind dadurch charakterisiert und von anderen Gattungen unterschieden, dass sie primär vom Bezug auf andere Fernseh- oder Medienereignisse leben?<sup>124</sup>

124 Auch Formate aus anderen zeitgeschichtlichen Phasen der BRD und aus der Zeit der DDR stellen – hinsichtlich der Rahmenkomponenten der Monostrukturierung und des parasitären Aufmerksamkeitsgewinns – ebenso interessante Vergleichshorizonte dar wie internationale Vergleiche. – In die komparative Analyse einzubeziehen wären dann auch solche Sendungen, die zwar derselben Gattung, derjenigen der Shows, zuzurechnen sind, die aber vermutlich ganz andere Formen aufweisen (bspw. „Wer wird Millionär“, moderiert von Günther Jauch u. „Clever – Die Show, die Wissen schafft“, moderiert von Barbara Eligmann und Wigald Boning) und schließlich Fernsehereignisse, die vermutlich ganz anderen Gattungen zuzurechnen sind.

## 7. Anhang

### 7.1 Transkriptionssystem MoViQ: Movies and Videos in Qualitative Research

In Kapitel 5.7 habe ich die Anforderungen herausgearbeitet, die an ein Transkriptionssystem für Filme und Videos zu stellen sind, welches den Prinzipien und dem Erkenntnispotential der dokumentarischen Methode gerecht zu werden vermag. In diesem Zusammenhang ist das Transkriptionssystem „MoViQ“ von Stefan Hampl und Aglaja Przyborski für die dokumentarische Film- und Videointerpretation entwickelt worden (siehe auch: Hampl 2009 sowie Przyborski/Wohlrab-Sahar 2008: 169-172), unter Einbeziehung des von mir für die dokumentarische Textinterpretation ausgearbeiteten Transkriptionssystems „TiQ“ (siehe Kap. 7.2).

Mit dem Transkriptionssystem MoViQ werden Bild und Ton synchron und in einem *konstanten* Zeitrhythmus transkribiert, dessen Intervalle nicht länger als 1 Sek. sein sollten. Die entsprechenden Anleitungen und Arbeitsschritte zur Transkription sind auf der von Stefan Hampl eingerichteten und betreuten Website [www.moviscript.net](http://www.moviscript.net) nachzulesen. Da die Erstellung der Videotranskription nach MoViQ in einzelnen manuellen Arbeitsschritten ein relativ umständlicher, zeitintensiver und vergleichsweise fehleranfälliger Prozess ist, wird im Rahmen des Dissertationsvorhabens von Stefan Hampl (2009) derzeit die Software „*MoviScript*“ entwickelt, welche ebenfalls über die Website [www.moviscript.net](http://www.moviscript.net) ab Herbst 2008 zugänglich sein wird. Die erste Version der Software MoviScript soll für den Anwender insbesondere hinsichtlich des Einlesens von Videosequenzen und deren Übertragung in ein druckbares Format Erleichterungen bringen. Ziel ist die weitgehende Automatisierung dieses Prozesses, sodass letztlich nur noch das manuelle Eintippen des gesprochenen Textes sowie die Kennzeichnung anderer Tonereignisse (Musik und Geräusche) in den dafür vorgesehenen Feldern notwendig ist.

Zusätzlich wird auf der Seite [www.moviscript.net](http://www.moviscript.net) ständig eine Liste verfügbarer Publikationen zum Thema Videotranskription und -interpretation aktualisiert. Es soll somit eine zentrale Anlaufstelle für die methodische Auseinandersetzung mit Videotranskription und -interpretation geschaffen werden.

### 7.2 Transkriptionssystem TiQ: Talk in Qualitative Social Research

Dieses Transkriptionssystem ist in seiner ursprünglichen Fassung (Bohnsack 1989: 387f.), welche zunehmend verfeinert worden ist, seit nunmehr 20 Jahren im Ge-



brauch und insbesondere für die Gesprächsanalyse im Rahmen der dokumentarischen Methode entwickelt worden.

### Transkriptionszeichen

L	Beginn einer Überlappung bzw. direkter Anschluss beim Sprecherwechsel
J	Ende einer Überlappung
(.)	Pause bis zu einer Sekunde
(2)	Anzahl der Sekunden, die eine Pause dauert
nein	betont
nein	laut (in Relation zur üblichen Lautstärke des Sprechers/der Sprecherin)
°nee*	sehr leise (in Relation zur üblichen Lautstärke des Sprechers/der Sprecherin)
.	stark sinkende Intonation
;	schwach sinkende Intonation
?	stark steigende Intonation
,	schwach steigende Intonation
viellei-	Abbruch eines Wortes
oh=nee	Wortverschleifung
nei:n	Dehnung, die Häufigkeit vom : entspricht der Länge der Dehnung
(doch)	Unsicherheit bei der Transkription, schwer verständliche Äußerungen
( )	unverständliche Äußerungen, die Länge der Klammer entspricht etwa der Dauer der unverständlichen Äußerung
((stöhnt))	Kommentare bzw. Anmerkungen zu parasprachlichen, nicht-verbalen oder gesprächsexternen Ereignissen; die Länge der Klammer entspricht im Falle der Kommentierung parasprachlicher Äußerungen (z. B. Stöhnen) etwa der Dauer der Äußerung. In vereinfachten Versionen des Transkriptionssystems kann auch Lachen auf diese Weise symbolisiert werden. In komplexeren Versionen wird Lachen wie folgt symbolisiert:
@nein@	lachend gesprochen
@(.)@	kurzes Auflachen
@(3)@	3 Sek. Lachen

für biografische Interviews zusätzlich:

//mhml// Hörersignal des Interviewers, wenn das „mhml“ nicht überlappend ist

### Groß- und Kleinschreibung:

Hauptwörter werden groß geschrieben, und bei Neuansetzen eines Sprechers/einer Sprecherin am Beginn eines ‚Häkchens‘ wird das erste Wort mit Großbuchstaben begonnen. Nach Satzzeichen wird klein weitergeschrieben, um deutlich zu machen, dass Satzzeichen die Intonation anzeigen und nicht grammatikalisch gesetzt werden.

### Zeilennummerierung:

Zum Auffinden und Zitieren von Transkriptstellen ist es notwendig, eine durchlaufende Zeilennummerierung zu verwenden.

Bei allen Transkripten zu Beginn vermerken: Codename der Gruppe, Name der Passage, wo die Passage auf der Kassette beginnt (bspw.: „1/5“ oder: Zählwerkangabe der verwendeten Geräte – die Zählwerkangaben sind je nach Gerät unterschiedlich), Dauer der Passage (bspw. 10 Minuten), Kürzel für die Personen, die transkribiert und die Transkription korrigiert haben.

gabe der verwendeten Geräte – die Zählwerkangaben sind je nach Gerät unterschiedlich), Dauer der Passage (bspw. 10 Minuten), Kürzel für die Personen, die transkribiert und die Transkription korrigiert haben.

### Maskierung:

Allen Personen einer Gruppendiskussion wird ein Buchstabe zugewiesen. Diesem wird je nach Geschlecht „f“ (für weiblich) oder „m“ (für männlich) hinzugefügt. Die Zuweisung lautet bei einer Diskussion mit 2 Mädchen und 3 Jungen bspw.: Af, Bf, Cm, Dm, Em. Dieser Buchstabe bleibt auch bei allen etwaigen weiteren Erhebungen bzw. bei der teilnehmenden Beobachtung bestehen, bei denen die Person beteiligt ist. Ist eine Person neben der Gruppendiskussion auch an einem biographischen Interview beteiligt, so erhält sie einen erdachten Namen, der mit dem zugewiesenen Buchstaben beginnt (bspw.: Bm, Berthold).

Alle Ortsangaben (Straße, Plätze, Bezirke) werden maskiert.

Namen, die im Interview genannt werden, werden durch erdachte Namen ersetzt. Dabei versuchen wir, einen Namen aus dem entsprechenden Kulturkreis zu nehmen, bspw. könnte „Mehmet“ zu „Kamil“ werden.

## 7.3 Zur Rekonstruktion der Perspektivität

Panofsky (1964a) bezeichnet die Perspektive mit einem Begriff von Ernst Cassirer als „symbolische Form“. Damit ist gesagt, dass die Perspektive bzw. der Modus der Perspektivität nicht als Charakteristikum einer Realität anzusehen ist, welche unabhängig von unserem historischen Wissen und Denken existiert. Vielmehr ist die Perspektive und somit auch die Zentral- oder Linearperspektive, wie sie in der Renaissance ihren Anfang nahm, nicht nur Element, sondern Ausdruck einer spezifischen, d.h. auch historisch einzuordnenden Weltanschauung. In der Wahl der Perspektivität dokumentiert sich im Sinne von Panofsky der „Geist“ einer Epoche. Weiter gefasst können wir sagen, dass sich hierin auch Elemente des ‚Geistes‘ oder des Habitus eines Milieus (bspw. des Familienmilieus in einem Familienfoto) oder einer Person dokumentieren.

Die Zentralperspektive, wie wir sie in der Fotografie finden, wurde in der Renaissance entwickelt. Sie macht die gesamte Bild-Konstruktion vom subjektiven Blickpunkt des abbildenden Bildproduzenten (bzw. des Betrachters) abhängig. Dies steht im Zusammenhang mit der Entdeckung des Subjekts in der Renaissance. Dabei ist nicht nur entscheidend, dass der Blickpunkt ein strikt subjektabhängiger ist, sondern auch, dass das Subjekt den Blickpunkt frei wählen kann. Die gesamte Konstruktion wird – wie Panofsky (1964a: 123) formuliert – durch die „frei wählbare Lage eines subjektiven ‚Blickpunktes‘ bestimmt“.

Panofsky hebt den ambivalenten Charakter der Zentralperspektive hervor: Sie hat einerseits eine „objektivistische“ (Panofsky 1964a: 124) Bedeutung. Sie „bringt die künstlerische Erscheinung auf feste, ja mathematisch-exakte Regeln, aber sie macht sie auf der anderen Seite vom Menschen, ja vom Individuum abhängig“ (Pa-

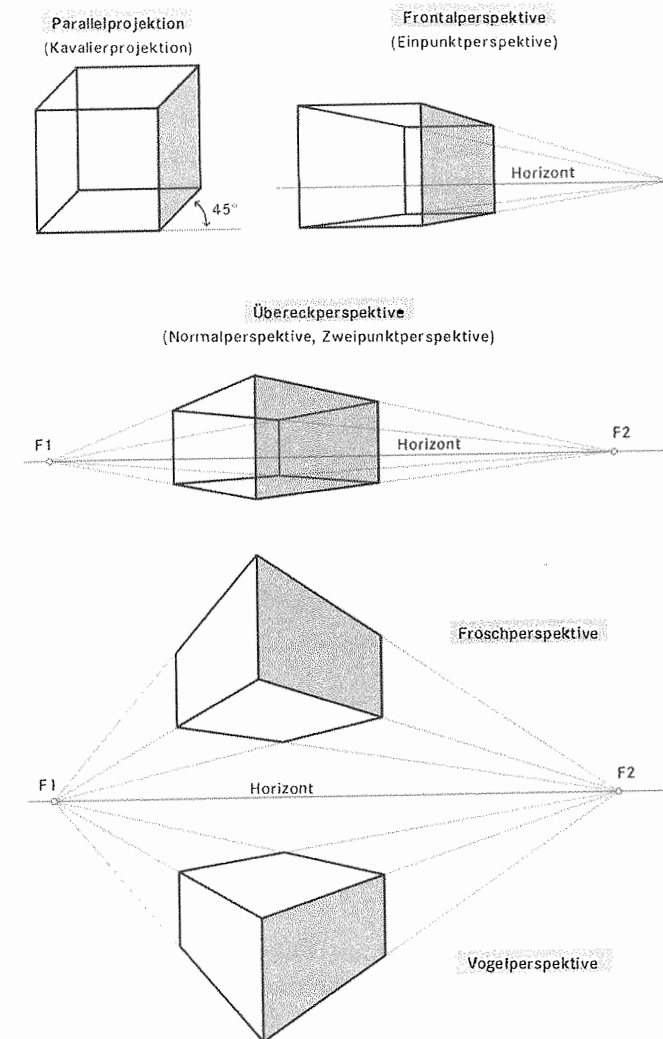
nofsky 1964a: 123) – eben aufgrund der Abhängigkeit der gesamten Konstruktion vom subjektiven Blickpunkt. Mit Max Imdahl (1996a: 19) lässt sich formulieren, „daß die Projektion, nämlich die Konvergenz der raumprojizierenden Fluchtlinien, als solche ein Ausdruck von Invariabilität, Gesetzmäßigkeit, Unzufälligkeit und Notwendigkeit ist, daß diese Invariabilität aber zugleich offen ist auf Variabilität, das heißt auf die freie Wahl der Lokalisierung des Fluchtpunkts oder der Fluchtachse“. Dadurch wird nach Imdahl der verbildlichte Raum zu einem „einfühlbaren Erlebnisraum“.

Die (im Kontext mathematisch exakter Regeln) „frei wählbare Lage des subjektiven, Blickpunktes“ (Panofsky 1964a: 124) im Rahmen der Zentralperspektive, wie sie in der Regel auch für die Fotografie von Bedeutung ist, kann unterschiedliche Ausprägungen oder Modi annehmen, die uns im wahrsten Sinne des Wortes Einblicke in die Perspektive eines Individuums oder eines Milieus zu vermitteln vermögen. Diese unterschiedlichen Modi der (Zentral-)Perspektivität werden in verschiedenen Veröffentlichungen in unterschiedlicher Weise schematisch dargestellt und bezeichnet. Ich beziehe mich hier auf die Schematisierung und grafische Darstellung in „Der Brockhaus Kunst“ (2006: 692) (s. S. 247).

Die Frontalperspektive habe ich in früheren Texten mit Parramón/Calbó (1999: 50) auch als „Parallelperspektive“ bezeichnet. Dies bringt allerdings mögliche Verwechslungen mit der „Parallelprojektion“ (Der Brockhaus Kunst: 692) mit sich, deren Fluchtpunkt sich im Unendlichen befindet. Eine derartige Parallelprojektion findet sich nicht in der Fotografie, sondern allenfalls in Malerei und Zeichnung.

In Anlehnung an die Terminologie, wie sie in der Filmwissenschaft im Zusammenhang mit der Bestimmung des „Kamerastandpunktes“ üblich ist (vgl. u.a. Hikethier 1996: 61f.), verwende ich anstelle des Begriffes der „Froschperspektive“ denjenigen der „Untersicht“ und anstelle des Begriffes der „Vogelperspektive“ denjenigen der „Aufsicht“. Gelegentlich wird in der Filmwissenschaft (vgl. etwa: Borstnar/Pabst/Wulf 2002: 92ff.) noch einmal differenziert zwischen Untersicht und (deren Steigerung durch die) Froschperspektive und der Aufsicht und (deren Steigerung durch die) Vogelperspektive. Dieser Differenzierung folge ich hier nicht.

Abbildung 38: Perspektive. Entnommen aus: Der Brockhaus Kunst 2006: 692

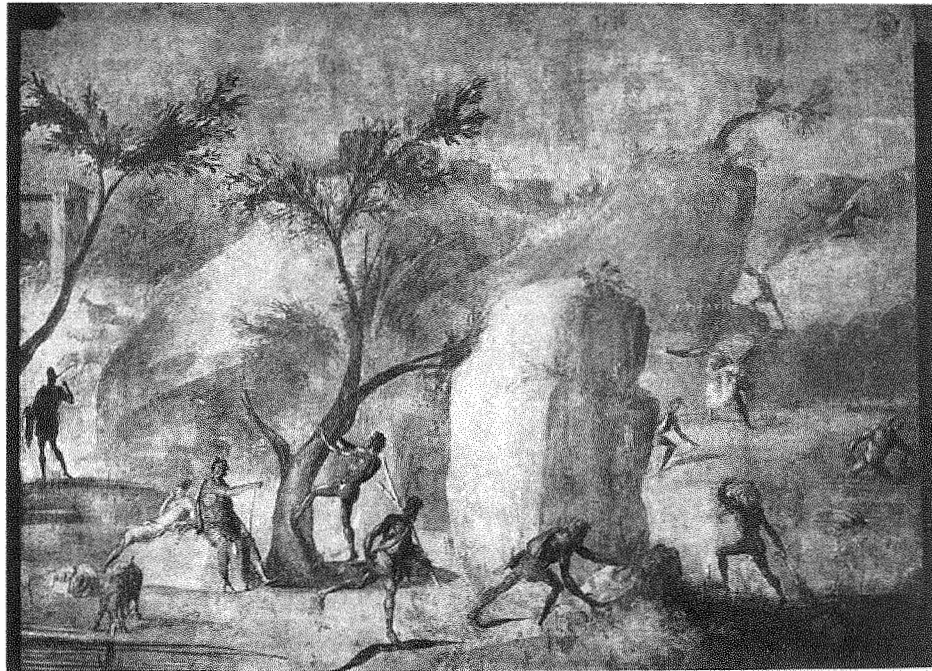


#### Alternativen zur Zentralperspektive

Im Unterschied zur Zentralperspektive blickte beispielsweise in der mittelalterlichen Malerei der Betrachter nicht wie durch ein Fenster *auf* das Dargestellte, sondern der Betrachter und der abbildende Bildproduzent, der Maler, waren gleichsam selbst im Bild mit anwesend. Während im Mittelalter also ohne Perspektive gearbeitet wurde, (s. die Abbildung in: Parramón/Calbó 1999: 12), war bereits die Antike zu einer perspektivischen Darstellung gelangt. Aber diesem Raum fehlt jenes „unendliche Kontinuum“ (Panofsky 2001a: 17), welches für die nachmittelalterliche Kunst, d.h. für die Zentralperspektive, charakteristisch ist. In den antiken Dar-

stellungen vermittelt dies den Eindruck einer „in sich selbst un stabilen und un zusammenhängenden Welt. Felsen, Bäume, Schiffe und kleine Figuren sind frei über weite Abschnitte von Land und Meer verteilt“. Als Beispiel hierzu das Bild „Odysseus im Land der Kannibalen“.

Abbildung 39: Odysseus im Land der Kannibalen. Rom: Vatikanische Bibliothek. Entnommen aus: Panofsky 2001b (Band 2): 290, Abb. 446



446. Rom, Vatikanische Bibliothek, Odysseus im Land der Kannibalen.

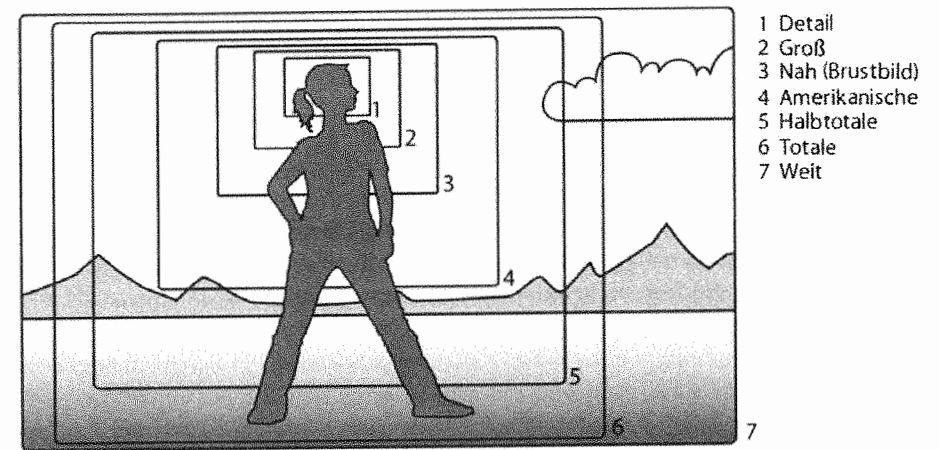
In der Frührenaissance, in der wir auch die Werke von Giotto einordnen können, begannen die Künstler bereits die Dinge darzustellen, als ob das Bild ein Fenster wäre. Aber in dieser Zeit finden wir noch nicht die Zentralperspektive, d.h. es gibt nicht den *einen*, sondern mehrere Fluchtpunkte, die auf einer *Fluchtachse* liegen.

Eine Kenntnis des Umgangs mit der Perspektivität in unterschiedlichen Epochen kann eine systematisierende und sensibilisierende Funktion bei der Interpretation von Zeichnungen – beispielsweise von Kinderzeichnungen (vgl. Wopfner 2008) – haben. Es erscheint allerdings problematisch, direkte Parallelen von Ontogenese und Phylogenese zu ziehen, wie wir dies im Ansatz bei Edgerton (2002: 25) finden, der „interessante Übereinstimmungen“ des Mittelalters mit ontogenetischen Entwicklungsphasen des Kindes herausarbeitet: „denn die Kinder begreifen sich selbst wie die Künstler des Mittelalters als ganz und gar in der Welt, die sie malen wollen, enthalten.“

## 7.4 Einstellungsgrößen: Übersicht über Begriffe

Die übersichtlichste Darstellung der Einstellungsgrößen (siehe dazu auch: Hicke thier 1996: 58f.; Borstnar/Pabst/Wulff 2002: 91; Faulstich 2002) findet sich nach meiner Einschätzung bei Kuchenbuch (2005: 44).

Abbildung 40: Einstellungsgrößen. Entnommen aus Kuchenbuch 2005: 44 (Abb. 4)



- 1 Detail
- 2 Groß
- 3 Nah (Brustbild)
- 4 Amerikanische
- 5 Halbtotale
- 6 Totale
- 7 Weit

## Abbildungsverzeichnis

### *Kapitel 3:*

- Abb. 1: Diego Velázquez, Las Meninas, 1656. Madrid, Museo del Prado.  
Entnommen aus: Greub 2001: 295 ..... 34
- Abb. 2: Giotto, Gefangennahme, um 1305. Padua, Arena-Kapelle. Entnommen aus: Imdahl 1996a: Abbildungsverzeichnis S. 45 (die Schräge wurde eingezeichnet von R.B. nach Imdahl 1996a, Abb. 45) ..... 37

### *Kapitel 4.2:*

- Abb. 3: Burberry London. Entnommen aus: Vogue 2005 (russische Ausgabe) ..... 59
- Abb. 4: Burberry London. Entnommen aus: Vogue 2005 (russische Ausgabe): Planimetrie I (Linien von R.B.) ..... 61
- Abb. 5: Burberry London. Entnommen aus: Vogue 2005 (russische Ausgabe): Planimetrie II (Linien von R.B.) ..... 62
- Abb. 6: Burberry London. Entnommen aus: Vogue 2005 (russische Ausgabe): Perspektivität (Linien von R.B.) ..... 63
- Abb. 7: Burberry London. Entnommen aus: Vogue 2005 (deutsche Ausgabe) ..... 70
- Abb. 8: Burberry London. Entnommen aus: Vogue 2005 (deutsche Ausgabe): Planimetrie (Linien von R.B.) ..... 71

### *Kapitel 4.3:*

- Abb. 9: Foto der Familie Schiller „Museum“ ..... 78
- Abb. 10: Foto der Familie Schiller „Museum“: szenische Choreografie ..... 80
- Abb. 11: Foto der Familie Schiller: „Museum“: Perspektivität ..... 80
- Abb. 12: Foto der Familie Telchow: „Gartenfest“ ..... 85
- Abb. 13: Foto der Familie Telchow: „Gartenfest“: Planimetrie u. szenische Choreografie ..... 86
- Abb. 14: Foto der Familie Telchow: „Gartenfest“: Perspektivität ..... 87
- Abb. 15: Familie Schiller: Urlaubsfoto „Fahrradrast an der Straße“ ..... 93
- Abb. 16: Familie Telchow: Urlaubsfoto „In der Seilbahngondel“ ..... 94

Abb. 17: Foto zur Kommunion von Frau Schiller .....	95
Abb. 18: Foto zur Kommunion von Frau Schiller: Planimetrie .....	97
Abb. 19: Foto zur Kommunion von Frau Schiller: Perspektivität .....	98
Abb. 20: Foto zur Kommunion von Frau Schiller: szenische Choreografie ....	98
Abb. 21: Foto der Feier zur Jugendweihe der Schwester von Frau Telchow ...	101
Abb. 22: Foto zur Jugendweihe der Schwester von Frau Telchow: Planimetrie und szenische Choreografie .....	103
Abb. 23: Foto zur Kommunion der Kinder der Familie Schiller .....	106
Abb. 24: Foto zur Kommunion der Kinder der Familie Schiller: Planimetrie und szenische Choreografie .....	108
Abb. 25: Foto zur Kommunion der Kinder der Familie Schiller: szenische (Teil-)Choreografie .....	109

### *Kapitel 5:*

Abb. 26: Sequenz aus Videotranskript „Kammer des Schreckens“. Entnommen aus: Baltruschat 2008 .....	167
--	-----

### *Kapitel 6:*

Abb. 27: Videotranskript Istanbul Total: 0:00 – 6.21 auf der Basis des Videos <a href="http://tvtotal.prosieben.de/show/letzte_sendung/2004/05/10/00257.html#">http://tvtotal.prosieben.de/show/letzte_sendung/2004/05/10/00257.html#</a> tml# .....	179
Abb. 28: Istanbul Total: Fotogramm IA. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD .....	202
Abb. 29: Istanbul Total: Fotogramm IA: Planimetrie. Entnommen aus: Mit- schnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.) .....	204
Abb. 30: Istanbul Total: Fotogramm IB. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD .....	206
Abb. 31: Istanbul Total: Fotogramm IB: Perspektivität. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.) .....	207
Abb. 32: Istanbul Total: Fotogramm IB: Planimetrie. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.) .....	207
Abb. 33: Istanbul Total: Fotogramm III. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD .....	209
Abb. 34: Istanbul Total: Fotogramm III: Perspektivität. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.) .....	210
Abb. 35: Istanbul Total: Fotogramm III: Planimetrie. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD (Linien von R.B.) .....	211
Abb. 36: Istanbul Total: Fotogramm IV. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD .....	216

Abb. 37: Istanbul Total: Fotogramm V. Entnommen aus: Mitschnitt der Produktionsgesellschaft Brainpool TV GmbH auf DVD .....	221
--	-----

### *Kapitel 7:*

Abb. 38: Perspektive. Entnommen aus: Der Brockhaus Kunst 2006: 62 .....	245
Abb. 39: Odysseus im Land der Kannibalen. Rom: Vatikanische Bibliothek. Entnommen aus: Panofsky 2001b (Band 2): 289. Abb. 290 .....	246
Abb. 40: Einstellungsgrößen. Entnommen aus Kuchenbuch 2005: 44 (Abb. 4) .....	247

## Literaturverzeichnis

Dort, wo der nicht deutschsprachige Titel im Text zitiert wird, sind die Zitate von mir selbst ins Deutsche übertragen worden – mit Ausnahme der Zitate in den Anmerkungen, die ich in der Originalsprache belassen habe.

- Ang, Ien (1997): Radikaler Kontextualismus und Ethnographie in der Rezeptionsforschung. In: Andreas Hepp/Rainer Winter (Hg.): Kultur-Medien-Macht. Cultural Studies und Medienanalyse. Opladen. S. 85-102.
- Asbrand, Barbara (2009): Wissen und Handlungskompetenz in der Weltgesellschaft. Münster.
- Assmann, Jan (1997): Das kulturelle Gedächtnis. Schrift, Erinnerung und politische Identität in frühen Hochkulturen. München.
- Aufenanger, Stefan (1995): Qualitative Forschung in der Medienpädagogik. In: König, Eckard/Peter Zedler (Hg.): Bilanz qualitativer Forschung. Bd. I: Grundlagen qualitativer Forschung. Weinheim. S. 221-239.
- Aufschnaiter, Stefan von/Welzel, Manuela (Hg.) (2001): Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen. Aktuelle Methoden empirischer pädagogischer Forschung. Münster et al.
- Bachmeier, Ben (1984): Symbolische Verarbeitung von Fernseherlebnissen in assoziativen Freiräumen. Eine Bestandsaufnahme in der Grundschule. 1. Teil: Fernsehspuren im Handeln von Kindern. Kassel.
- Balázs, Belá (1976): Der Film. Werden und Wesen der neuen Kunst. Wien (5. Auflage).
- Balázs, Belá (2001a): Der sichtbare Mensch oder die Kultur des Films. Frankfurt a.M. (ursprüngl.: 1924).
- Balázs, Belá (2001b): Der Geist des Films. Frankfurt a.M. (ursprüngl.: 1930).
- Baltruschat, Astrid (2008): Die Dekoration der Institution Schule. Filminterpretationen nach der Dokumentarischen Methode von Beiträgen zu einem Ideenwettbewerb zur Schulgestaltung. Dissertationsprojekt an der Freien Universität Berlin. Berlin.
- Baltruschat, Astrid (2009): Film Interpretation According to the Documentary Method. In: Ralf Bohnsack/Nicolle Pfaff/Wivian Weller (Eds.): Qualitative Analysis and Documentary Method in International Educational Research. Opladen & Farmington Hills.
- Barboza, Amalia (2005): Kunst und Wissen. Die Stilanalyse in der Soziologie Karl Mannheims. Konstanz.
- Barthes, Roland (1983): Elemente der Semiologie. Frankfurt a.M. (Original: 1964).
- Barthes, Roland (1985): Die helle Kammer. Bemerkung zur Photographie. Frankfurt a.M. (Original: 1980).
- Barthes, Roland (1990): Der entgegenkommende und der stumpfe Sinn. Kritische Essays III. Frankfurt a.M. (Original: 1982).
- Becker, Howard S. (1986): Photography and Sociology. In: Howard S. Becker: Doing Things Together. Selected Papers. Evanston/Ill. S. 223-271.
- Beller, Hans (2005): Aspekte der Filmmontage. Eine Art Einführung. In: Ders. (Hg.): Handbuch der Filmmontage. München. (5. Auflage).

- Belting, Hans (2001): *Bild-Anthropologie. Entwürfe für eine Bildwissenschaft*. München.
- Berger, Peter L./Luckmann, Thomas (1969): *Die gesellschaftliche Konstruktion der Wirklichkeit*. Frankfurt a.M.
- Bergmann, Jörg/Thomas Luckmann/Hans-Georg, Soeffner (1993): *Erscheinungsformen von Charisma – Zwei Päpste*. In: Winfried Gebhardt/Arnold Zingerle/Michael N. Ebertz (Hg.): *Charisma – Theorie – Religion, Politik*. Berlin/New York. S. 121-155.
- Birdwhistell, Ray L. (1952): *Introduction to Kinesics (An annotation system for analysis of body motion and gesture)*. Louisville.
- Birdwhistell, Ray L. (1960): *Kinesics and Communication*. In: E. Carpenter/M. Mc Luhan (Hg.): *Exploration in Communication*. Boston. S. 54-64.
- Birdwhistell, Ray L. (1968): *Kinesics*. In: David L. Sills (Hg.): *International Encyclopedia of Social Sciences*. New York. Vol. 8, S. 379.
- Birdwhistell, Ray L. (1970): *Kinesics and Context. Essays on Body Motion Communication*. Philadelphia.
- Boehm, Gottfried (1978): *Zu einer Hermeneutik des Bildes*. In: Hans-Georg Gadamer/Gottfried Boehm (Hg.): (Prozess) Seminar: *Die Hermeneutik und die Wissenschaften*. Frankfurt a.M. S. 444-471.
- Boehm, Gottfried (1985): *Die Krise der Repräsentation. Die Kunstgeschichte und die moderne Kunst*. In: Volker Bätschmann/Lorenz Dittmann (Hg.): *Kategorien und Methoden der Kunstgeschichte 1900-1930*. Stuttgart. S. 113-128.
- Böhme, Gernot (1999): *Theorie des Bildes*. München.
- Bogen, Steffen (2005): *Kunstgeschichte/Kunstwissenschaft*. In: Klaus Sachs-Hombach (Hg.): *Bildwissenschaft. Disziplinen, Themen, Methoden*. Frankfurt a.M. S. 52-67.
- Bohnsack, Ralf (1983): *Alltagsinterpretation und soziologische Rekonstruktion*. Opladen.
- Bohnsack, Ralf (1989): *Generation, Milieu und Geschlecht – Ergebnisse aus Gruppendiskussionen mit Jugendlichen*. Opladen.
- Bohnsack, Ralf (1998a): *Rekonstruktive Sozialforschung und der Grundbegriff des Orientierungsmusters*. In: Dirk Siefkes/Peter Eulenhöfer/ Heike Stach/Klaus Städtler (Hg.): *Sozialgeschichte der Informatik*. Wiesbaden. S. 105-121.
- Bohnsack, Ralf (2000): *Jugendliche als Opfer und Täter. Das Fehlen der Jugend in der Forschung zur Jugendkriminalität*. In: Uwe Sander/Ralf Vollbrecht (Hg.): *Jugend im 20. Jahrhundert. Sichtweisen – Orientierungen – Risiken*. Neuwied.
- Bohnsack, Ralf (2001): *Dokumentarische Methode. Theorie und Praxis wissenssoziologischer Interpretation*. In: Theo Hug (Hg.): *Wie kommt Wissenschaft zu Wissen? Bd. 3: Einführung in die Methodologie der Kultur- und Sozialwissenschaften*. Baltmannsweiler. S. 326-345.
- Bohnsack, Ralf (2003a): *Qualitative Methoden der Bildinterpretation*. In: *Zeitschrift für Erziehungswissenschaft (ZfE)*, Jg. 6, Heft 2/2003, S. 239-256.
- Bohnsack, Ralf (2003b): *Dokumentarische Methode und sozialwissenschaftliche Hermeneutik*. In: *Zeitschrift für Erziehungswissenschaft (ZfE)*, Jg. 6, Heft 4/2003, S. 480-504.
- Bohnsack, Ralf (2003c): *Differenzerfahrungen der Identität und des Habitus. Eine empirische Untersuchung auf der Basis der dokumentarischen Methode*. In: Burkhard Liebsch/Jürgen Straub (Hg.): *Lebensformen im Widerstreit. Integrations- und Identitätskonflikte in pluralen Gesellschaften*. Frankfurt a.M. S. 136-160.
- Bohnsack, Ralf (2004a): *Gruppendiskussion*. In: Uwe Flick/Ernst v. Kardorff/Ines Steinke (Hg.): *Qualitative Forschung. Ein Handbuch*. Reinbek b. Hamburg (3. Aufl.). S. 369-384.
- Bohnsack, Ralf (2004b): *Group Discussion*. In: Uwe Flick/Ernst v. Kardorff/Ines Steinke (Eds.): *A Companion to Qualitative Research* (pp. 214-220). London.
- Bohnsack, Ralf (2005a): *Standards nicht-standardisierter Forschung in den Erziehungs- und Sozialwissenschaften*. In: *Zeitschrift für Erziehungswissenschaft (ZfE)*, 7. Jg., Beiheft Nr. 3 (Standards und Standardisierung in der Erziehungswissenschaft), Hg. von Ingrid Gogolin/Heinz-Hermann Krüger/Dieter Lenzen/Thomas Rauschenbach. S. 65-83.
- Bohnsack, Ralf (2005b): *Bildinterpretation und Dokumentarische Methode*. In: Christoph Wulf/Jörg Zirfas (Hg.): *Ikonologie des Performativen*. München. S. 246-262.
- Bohnsack, Ralf (2006a): *Mannheims Wissenssoziologie als Methode*. In: Dirk Tänzler/Hubert Knoblauch/Hans-Georg Soeffner (Hg.): *Neue Perspektiven der Wissenssoziologie*. Konstanz. S. 271-291.
- Bohnsack, Ralf (2006b): *Fokussierungsmetapher*. In: Ralf Bohnsack/Winfried Marotzki/Michael Meuser (Hg.): *Hauptbegriffe Qualitativer Sozialforschung*. Opladen & Farmington Hills. S. 67.
- Bohnsack, Ralf (2006c): *Qualitative Evaluation und Handlungspraxis. Grundlagen dokumentarischer Evaluationsforschung*. In: Uwe Flick (Hg.): *Qualitative Evaluationsforschung*. Reinbek b. Hamburg. S. 135-155.
- Bohnsack, Ralf (2007a): *Die dokumentarische Methode in der Bild- und Fotointerpretation*. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): *Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung*. Wiesbaden (2. Aufl.; ursprüngl.: 2001). S. 67-89. (Wieder abgedruckt in: Yvonne Ehrenspeck/Burkhard Schäffer (Hg.) (2003): *Film- und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch*. Opladen. S. 87-107).
- Bohnsack, Ralf (2007b): *„Heidi“: Eine exemplarische Bildinterpretation auf der Basis der dokumentarischen Methode*. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): *Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung*. Wiesbaden (2. Aufl. ; ursprüngl.: 2001). S. 323-337. (Wieder abgedruckt in: Yvonne Ehrenspeck/Burkhard Schäffer (2003) (Hg.): *Film- und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch*. Opladen. S. 109-120).
- Bohnsack, Ralf (2007c): *Typenbildung, Generalisierung und komparative Analyse. Grundprinzipien dokumentarischer Interpretation*. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): *Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung*. Wiesbaden (2. Aufl. ; ursprüngl.: 2001). S. 225-252.
- Bohnsack, Ralf (2007d): *Dokumentarische Bildinterpretation. Am exemplarischen Fall eines Werbefotos*. In: Renate Buber/Hartmut Holzmüller (Hg.): *Qualitative Marktforschung. Konzepte – Methoden – Analysen*. Wiesbaden. S. 951-979.
- Bohnsack, Ralf (2007e): *Zum Verhältnis von Bild- und Textinterpretation in der qualitativen Sozialforschung*. In: Barbara Friebertshäuser/Heide von Felden/Burkhard Schäffer (Hg.): *Bild und Text – Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft*. Opladen & Farmington Hills. S. 21-24.
- Bohnsack, Ralf (2007f): *Performativität, Performanz und dokumentarische Methode*. In: Christoph Wulf/Jörg Zirfas (Hg.): *Pädagogik des Performativen. Theorien, Methoden, Perspektiven*. Weinheim. S. 200-212.
- Bohnsack, Ralf (2007g): *Gruppendiskussionsverfahren und Focus Groups*. In: Renate Buber/Hartmut Holzmüller (Hg.): *Qualitative Marktforschung. Konzepte – Methoden – Analysen*. Wiesbaden. S. 491-506.
- Bohnsack, Ralf (2008a): *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden*. Opladen & Farmington Hills (7. Auflage).
- Bohnsack, Ralf (2008b): *Gruppendiskussionsverfahren und Gesprächsanalyse*. In: Ders.: *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden*. Opladen & Farmington Hills (7. Auflage). S. 187-206.
- Bohnsack, Ralf (2008c): *Exemplarische Bildinterpretationen*. In: Ders.: *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden*. Opladen & Farmington Hills (7. Aufl.). S. 236-257.
- Bohnsack, Ralf (2008d): *Praxeologische Methodologie*. In: Ders.: *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden*. Opladen & Farmington Hills (7. Auflage). S. 187-206.
- Bohnsack, Ralf (2009a): *Die Mehrdimensionalität der Typenbildung und ihre Aspekthaftigkeit*. In: Jutta Ecarius/Burkhard Schäffer (Hg.): *Typenbildung und Theoriegenerierung. Perspektiven qualitativer Biographie- und Bildungsforschung*. Opladen & Farmington Hills. S. 47-72.
- Bohnsack, Ralf (2009b): *The Interpretation of Pictures and the Documentary Method*. In: Ralf Bohnsack/Nicolle Pfaff/Wivian Weller (eds.): *Qualitative Analysis and Documentary Method*

- in International Educational Research. Opladen & Farmington Hills. – Wieder abgedruckt in: Hubert Knoblauch/Alejandro Baer/Eric Laurier/Sabine Petschke/Bernt Schnettler (eds.): Visual Methods. In: FQS (Forum Qualitative Social Research). Vol 9. No 3 (2008).
- Bohnsack, Ralf/Gebhardt, Winfried/Kraul, Margret/Wulf, Christoph (2001): Erziehung und Tradition. Tradierungsprozesse in Familien. Antrag für ein Forschungsprojekt bei der Deutschen Forschungsgemeinschaft. Berlin/Koblenz.
- Bohnsack, Ralf/Loos, Peter/Schäffer, Burkhard/Städtler, Klaus/Wild, Bodo (1995): Die Suche nach Gemeinsamkeit und die Gewalt der Gruppe. Hooligans, Musikgruppen und andere Jugendcliquen. Opladen.
- Bohnsack, Ralf/Nentwig-Gesemann, Iris (Hg.) (2008): Dokumentarische Evaluationsforschung. Theoretische Grundlagen und Beispiele aus der Praxis. Opladen u. Farmington Hills.
- Bohnsack, Ralf/Nentwig-Gesemann, Iris/Nohl, Arnd-Michael (2007): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Wiesbaden (2. Aufl.).
- Bohnsack, Ralf/Nohl, Arnd-Michael (2001a): Ethnisierung und Differenzierung. Fremdheit als alltägliches und als methodologisches Problem. In: ZBBS (Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung). Heft 3/2001. S. 15-36.
- Bohnsack, Ralf/Nohl, Arnd-Michael (2001b): Jugendkulturen und Aktionismus. Eine rekonstruktive empirische Analyse am Beispiel des Breakdance. In: Jürgen Zinnecker/Hans Merckens (Hg.): Jahrbuch Jugendforschung. Folge 1. Opladen. S. 17-37.
- Bohnsack, Ralf/Nohl, Arnd-Michael (2007): Exemplarische Textinterpretation. Die Sequenzanalyse der dokumentarischen Methode. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.) Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Wiesbaden (2. Aufl.). S. 303-308.
- Bohnsack, Ralf/Pfaff, Nicole/Weller, Wivian (2010) (eds.): Qualitative Analysis and Documentary Method in International Educational Research. Opladen & Farmington Hills.
- Bohnsack, Ralf/Przyborski, Aglaja (2006): Diskursorganisation, Gesprächsanalyse und die Methode der Gruppendiskussion. In: Ralf Bohnsack/Aglaja Przyborski/Burkhard Schäffer (Hg.): Das Gruppendiskussionsverfahren in der Forschungspraxis. Opladen (2. Auflage), S. 233-248.
- Bohnsack, Ralf/Schäffer, Burkhard (2002): Generation als konjunktiver Erfahrungsraum. Eine empirische Analyse generationspezifischer Medienpraxiskulturen. In: Günter Burkart/Jürgen Wolf (Hg.): Lebenszeiten. Erkundungen zur Soziologie der Generationen (Martin Kohli zum 60. Geburtstag). Opladen. S. 249-273
- Bohnsack, Ralf/Schäffer, Burkhard (2007): Exemplarische Textinterpretation: Diskursorganisation und dokumentarische Methode. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Wiesbaden (2. Aufl. ; ursprüngl.: 2001). S. 309-323.
- Bordwell, David (1985): Narration in the Fiction Film. London
- Bordwell, David (2006): Visual Style in Cinema. Vier Kapitel Filmgeschichte. Frankfurt a.M.
- Borstnar, Niels/Pabst, Eckhard/Wulff; Hans J. (2002): Einführung in die Film- und Fernsehwissenschaft. Konstanz.
- Bourdieu, Pierre (1970): Der Habitus als Vermittlung zwischen Struktur und Praxis. In: Ders.: Zur Soziologie symbolischer Formen. Frankfurt a.M.. S. 125-158.
- Bourdieu, Pierre (1976): Entwurf einer Theorie der Praxis. Frankfurt a.M..
- Bourdieu, Pierre (1982): Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft. Frankfurt a.M.
- Bourdieu, Pierre (1983a): Einleitung. In: Ders. et al.: Eine illegitime Kunst. Die sozialen Gebrauchsweisen der Photographie. Frankfurt a.M. S. 11-21.
- Bourdieu, Pierre (1983b): 2. Kap.: Die gesellschaftliche Definition der Photographie. In: Ders. et al.: Eine illegitime Kunst. Die sozialen Gebrauchsweisen der Photographie. Frankfurt a.M. S. 85-110.
- Bourdieu, Pierre (1992): Rede und Antwort. Frankfurt a.M.
- Bourdieu, Pierre (1996): Die Praxis der reflexiven Anthropologie. In: Ders./Loïc J.D. Wacquant (Hg.): Reflexive Anthropologie, Frankfurt a.M. S. 251-294.
- Bourdieu, Pierre/Boltanski, Luc/Castel, Robert/Chamboredon, Jean-Claude/Lagneau, Gerard/Schnapper, Dominique (1965): Un art moyen. Essais sur les usages sociaux de la photographie. Paris.
- Bourdieu, Pierre/Boltanski, Luc/Castel, Robert/Chamboredon, Jean-Claude/Lagneau, Gerard/Schnapper, Dominique (1983): Eine illegitime Kunst. Die sozialen Gebrauchsweisen der Photographie. Frankfurt a.M.
- Brandt, Birgit/Krummheuer, Götz/Naujok, Natascha (2001): Zur Methodologie kontextbezogener Theoriebildung im Rahmen interpretativer Grundschulforschung. In: Stefan von Aufschnaiter/Manuela Welzel (Hg.): Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen. Aktuelle Methoden empirischer pädagogischer Forschung. Münster et al. S. 17-40.
- Bredenkamp, Horst (2004): Drehmomente – Merkmale und Ansprüche des Iconic Turn. In: Christa Maar/Herbert Burda (Hg.): Iconic Turn. Die neue Macht der Bilder. S. 9-26.
- Breidenstein, Georg (2006): Teilnahme am Unterricht. Ethnographische Studien zum Schülerjob. Wiesbaden.
- Cappai, Gabriele (2008): Der empirische Zugang zum kulturell Fremden am Beispiel Zeit. Ein rekonstruktiver Ansatz. In: Ders. (Hg.): Forschen unter Bedingungen kultureller Fremdheit. Wiesbaden
- Castoriadis, Cornelius (1984): Gesellschaft als imaginäre Institution. Entwurf einer politischen Philosophie. Frankfurt a.M.
- Clarke, John et.al. (Hg.) (1979): Jugendkultur als Widerstand. Frankfurt a.M.
- Collier John, Jr./Collier, Malcolm (1986) :Visual Anthropology: Photography as a Research Method. Albuquerque (University of New Mexico Press).
- Deleuze, Gilles (1997a): Das Bewegungs-Bild. Kino 1. Frankfurt a.M.
- Deleuze, Gilles (1997b): Das Zeit-Bild. Kino 2. Frankfurt a.M.
- Denzin, Norman K. (1991): Hollywood shot by shot: Alcoholism in American cinema. New York.
- Der Brockhaus Kunst (2006). Hg. von der Lexikonredaktion des Verlags F.A. Brockhaus. Mannheim/Leipzig (3.Aufl.).
- Dilthey, Wilhelm (1924): Die Entstehung der Hermeneutik. In: Ders.: Gesammelte Schriften. 5. Band: Die geistige Welt. Leipzig/Berlin. S. 317-338.
- Dörner, Olaf (2007): Comics als Gegenstand qualitativer Forschung. Zu analytischen Möglichkeiten der dokumentarischen Methode am Beispiel der „Abrafaxe“. In: Barbara Friebertshäuser/Heide von Felden/Burkhard Schäffer (Hg.): Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen & Farmington Hills. S. 179-196.
- Duden „Fremdwörterbuch“ (1982). Mannheim/Wien/Zürich (4. Auflage).
- Eco, Umberto (1994): Einführung in die Semiotik. München (8. Auflage).
- Eco, Umberto (1977): Das offene Kunstwerk. Frankfurt a.M.
- Edgerton, Samuel Y. (2002): Die Entdeckung der Perspektive. München.
- Ehrenspeck, Yvonne/Lenzen, Dieter (2003): Sozialwissenschaftliche Filmanalyse. Ein Werkstattbericht. In: Yvonne Ehrenspeck/Burkhard Schäffer (Hg.): Film und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch. Opladen. S. 439-450.
- Ehrenspeck, Yvonne/Schäffer, Burkhard (Hg.) (2003): Film und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch. Opladen.
- Englisch, Felicitas (1991): Bildanalyse in strukturhermeneutischer Absicht. Methodische Überlegungen und Analysebeispiele. In: Dieter Garz/Klaus Kraimer (Hg.): Qualitativ-Empirische Sozialforschung. Konzepte, Methoden, Analysen. Opladen. S. 133-176.
- Erickson, Frederick (1992): Ethnographic Microanalysis of Interaction. In: Margret Le Compte/Wendy Milroy/Judith Preisle (Hg.): The Handbook of Qualitative Research in Education. New York et al. (pp. 201-225).
- Faulstich, Werner (2002): Grundkurs Filmanalyse. München.
- Fetzner, Daniel/Selke, Stefan (2005) (Hg.): selling politics. Bildinhalt und Bildwirkung. Ergebnisse des Forschungsprojekts selling politics zu Plakatmotiven der agenda 2010. Furtwangen 2005 (Schriftenreihe Fakultät Digitale Medien).



- Fiske, John (1989): Popular television and commercial structure: Beyond political economy. In: Gary Bums/Robert I. Thompson (Eds.): *Television Studies: Textual analyses*, Praeter: New York/London. S. 21-37.
- Flick, Uwe (2002): *Qualitative Sozialforschung. Eine Einführung*. Reinbek bei Hamburg (6. Aufl.).
- Flick, Uwe (2008): *Triangulation. Eine Einführung*. Wiesbaden (2. Aufl.).
- Foltys, Julia (2006): *Familiale Aushandlungsprozesse. Eine dokumentarische Interpretation von Tischgesprächen*. Diplomarbeit an der Freien Universität Berlin. Berlin.
- Foucault, Michel (1971): *Die Ordnung der Dinge. Eine Archäologie der Humanwissenschaften*. Frankfurt a.M. (Original: 1966).
- Friebertshäuser, Barbara/von Felden, Heide/Schäffer, Burkhard (Hg.) (2006): *Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft*. Opladen & Farmington Hills.
- Fritzsche, Bettina (2003): *Pop-Fans. Studie einer Mädchenkultur*. Opladen.
- Fritzsche, Bettina (2007): *Mediennutzung im Kontext kultureller Praktiken als Herausforderung für die qualitative Forschung*. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): *Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung*. Wiesbaden (2. Aufl.). S. 29-44.
- Fuhs, Burkhard (1997): *Fotografie und qualitative Forschung. Zur Verwendung fotografischer Quellen in den Erziehungswissenschaften*. In: Barbara Friebertshäuser/Annedore Prengel (Hg.): *Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft*. Weinheim u. München. S. 265-285.
- Fuhs, Burkhard (2006): *Narratives Bildverstehen. Plädoyer für die erzählende Dimension der Fotografie*. In: Winfried Marotzki/Horst Niesyto (Hg.): *Bildinterpretation und Bildverstehen. Methodische Ansätze aus sozialwissenschaftlicher, kunst- und medienpädagogischer Perspektive*. Wiesbaden. S. 207-225.
- Garfinkel, Harold (1961): *Aspects of Common Sense Knowledge of Social Structures*. In: *Transactions of the Fourth World Congress of Sociology*. Vol. IV. S. 51-65 (deutsch (1973): *Das Alltagswissen über soziale und innerhalb sozialer Strukturen*. In: Arbeitsgruppe Bielefelder Soziologen (Hg.): *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit*. Reinbek b. Hamburg. S. 189-260.
- Garfinkel, Harold (1967) *What is Ethnomethodology?* In: Ders.: *Studies in Ethnomethodology*. Englewood Cliffs/New Jersey. S. 1-34.
- Garz, Detlef/Kraimer, Klaus (1994): *Die Welt als Text. Theorie, Kritik und Praxis der objektiven Hermeneutik*. Frankfurt a.M.
- Gehlen, Arnold (1986): *Zeit-Bilder. Zur Soziologie und Ästhetik der modernen Malerei*. Frankfurt a.M.
- Geimer, Alexander (2008): *Kulturelle Praktiken der Aneignung von Filmen. Eine qualitativ-rekonstruktive Studie zur Rezeption bei Jugendlichen*. Dissertationsprojekt an der Freien Universität Berlin. Berlin.
- Geimer, Alexander (2009): *Cultural practices of the reception and appropriation of films. The analytical stance of the praxeological sociology of knowledge*. In: Ralf Bohnsack/Nicolle Pfaff/Wivian Weller (Eds.): *Qualitative Analysis and Documentary Method in International Educational Research*. Opladen & Farmington Hills.
- Geimer, Alexander/Ehrenspeck, Yvonne (2009): *Qualitative Filmanalyse in den Sozial- und Erziehungswissenschaften*. In: Barbara Friebertshäuser/Annedore Prengel (Hg.): *Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft (Neuausgabe)*. Weinheim u. München.
- Glaser, Barney G./Strauss, Anselm (1967): *The Discovery of Grounded Theory. Strategies for Qualitative Research*. Chicago.
- Goffman, Erving (1963): *Stigma. Notes on the Management of Spoiled Identity*. Englewood Cliffs/N. J.
- Goffman, Erving (1975): *Frame Analysis*. Harmondsworth/Middlesex
- Goffman, Erving (1979): *Gender Advertisements*. New York et al.
- Goffman, Erving (1981): *Geschlecht und Werbung*. Frankfurt a.M.
- Goffman, Erving (1990): *Some Context for Content Analysis: A View on the Origins of Structural Studies of Face to Face Interaction*. In: Erving Goffman: *Conducting Interaction. Patterns of Behavior in Focused Encounters* (pp. 15-49). Cambridge.
- Gombrich, Ernst H. (1978): *Kunst und Illusion. Zur Psychologie bildlicher Darstellung*. Stuttgart/Zürich.
- Goodwin, Charles (2001): *Practices of Seeing Visual Analysis: An Ethnomethodological Approach*. In: Theo van Leeuwen/Carey Jewitt (eds.): *Handbook of Visual analysis* (pp. 157-182). Los Angeles et al.
- Greub, Thierry (Hg.) (2001): *Las Meninas im Spiegel der Deutungen. Eine Einführung in die Methoden der Kunstgeschichte*. Berlin.
- Guba, Egon G./Lincoln, Yvonna S. (1989): *Fourth Generation Evaluation*. Newbury Park, CA.
- Habermas, Jürgen (1971): *Vorbereitende Bemerkungen zu einer Theorie der kommunikativen Kompetenz*. In: Jürgen Habermas/Niklas Luhmann (Hg.): *Theorie der Gesellschaft oder Sozialtechnologie – Was leistet die Systemforschung?* Frankfurt a.M. S. 101-141.
- Habermas, Jürgen (1973): *Notizen zum Begriff der Rollenkompetenz*. In: Ders.: *Kultur und Kritik*. Frankfurt a.M. S. 195-231.
- Habermas, Jürgen (1981): *Theorie des kommunikativen Handelns*. Bd. 1. Frankfurt a.M.
- Halbwachs, Maurice (1985): *Das kollektive Gedächtnis*. Frankfurt a.M.
- Hall, Stuart (1980): *Encoding/decoding*. In: Stuart Hall/D. Hobson/A. Lowe/Paul Willis (Eds.): *Culture, Media, Language* (pp. 128-138). London/New York.
- Hampl, Stefan (2009): *Die Artikulation des Türkischen in der deutschen Fernsehshow Istanbul Total. Rekonstruktive Videointerpretation nach der dokumentarischen Methode*. Dissertation an der Universität Wien. Wien.
- Hart, Joan (1993): *Erwin Panofsky and Karl Mannheim: A Dialogue on Interpretation*. In: *Critical Inquiry* 19: 3. (pp. 534-566).
- Heath, Christian (1997): *The Analysis of Activities in Face to Face Interaction Using Video*. In: Silverman, David (Hg.): *Qualitative Research. Theory, Method and Practice* (pp. 183-200). London.
- Heidegger, Martin (1986): *Sein und Zeit*. Tübingen (Original: 1927).
- Hepp, Andreas (2004): *Cultural Studies und Medienanalyse. Eine Einführung*. Wiesbaden (2. Aufl.).
- Herrle, Matthias (2007): *Selektive Kontextvariation. Die Rekonstruktion von Interaktionen in Kursen der Erwachsenenbildung auf der Basis visueller Daten*. Frankfurt a.M.
- Herrle, Matthias/Kade, Jochen/Nolda, Sigrid (2009): *Erziehungswissenschaftliche Videografie*. In: Barbara Friebertshäuser/Antje Langer/Annedore Prengel (Hg.): *Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft (Neuausgabe)*. Weinheim u. München.
- Herrlitz, H.-G./Rittelmeyer, C. (Hg.): *Exakte Phantasie. Pädagogische Erkundungen bildender Wirkungen in Kunst und Kultur*. S. 105-121.
- Hickethier, Knut (1996): *Film- und Fernsehanalyse*. Stuttgart/Weimar (2. Auflage; Erstaussgabe 1993).
- Hüppauf, Bernd/Wulf, Christoph (2007): *Einleitung: Warum Bilder die Einbildungskraft brauchen*. In: Dies. (Hg.): *Bild und Einbildungskraft*. München. S. 9-44.
- Imdahl, Max (1979): *Überlegungen zur Identität des Bildes*. In: Odo Marquard/Karlheinz Stierle (Hg.): *Identität (Reihe: Poetik und Hermeneutik, Bd. VII)*. München. S. 187-211.
- Imdahl, Max (1994): *Ikonik. Bilder und ihre Anschauung*. In: Gottfried Boehm (Hg.): *Was ist ein Bild?* München. S. 300-324.
- Imdahl, Max (1996a): *Arenafresken. Ikonographie – Ikonologie – Ikonik*. München.
- Imdahl, Max (1996b): *Wandel durch Nachahmung. Rembrandts Zeichnung nach Lastmanns „Susanna im Bade“*. In: Ders.: *Zur Kunst der Tradition. Gesammelte Schriften*. Bd. 2. Frankfurt a.M. S. 431-456.
- Imdahl, Max (1996c): *Bildsyntax und Bildsemantik. Zum Centurioblat im Codex Egberti*. In: Ders. *Zur Kunst der Tradition. Gesammelte Schriften* Bd. 2. Frankfurt a.M. S. 78-79.
- Imdahl, Max (1996d): *Pose und Indoktrination. Zu Werken der Plastik und Malerei im Dritten Reich*. In: Ders.: *Reflexion – Theorie – Methode. Gesammelte Schriften* Band 3. Frankfurt a.M. S. 575-590.
- Imdahl, Max (2003): *Farbe. Kunsttheoretische Reflexionen in Frankreich*. München.

- Kade, Jochen/Nolda, Sigrid (2007): Das Bild als Kommentar und Irritation. Zur Analyse von Kursen der Erwachsenenbildung/Weiterbildung auf der Basis von Videodokumentationen. In: Barbara Friebertshäuser/Heide von Felden/Burkhard Schäffer (Hg.): Bild und Text – Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen & Farmington Hills. S. 159-177.
- Kendon, Adam (1990): Spatial Organization in Social Encounters: The F-formation System. In: Des.: Conducting Interaction. Patterns of Behavior in Focused Encounters. Cambridge. S. 209-262.
- Keppeler, Angela (2006): Mediale Gegenwart. Eine Theorie des Fernsehens am Beispiel der Darstellung von Gewalt. Frankfurt a.M.
- Klambeck, Amelie (2007): „Das hysterische Theater unter der Lupe“. Klinische Zeichen psychogener Gangstörungen. Wege der dokumentarischen Rekonstruktion von Körperbewegungen auf der Grundlage von Videografien. Göttingen.
- Klar, Isabelle (2007): Tischgespräche in Familien. Eine qualitative Analyse. Magisterarbeit an der Freien Universität Berlin. Berlin.
- Klika, Dorle/Kleynen, Thomas (2007): Adoleszente Selbstinszenierung in Text und Bild. In: Barbara Friebertshäuser/Heide von Felden/Burkhard Schäffer (Hg.): Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen & Farmington Hills. S. 121-140.
- Knoblauch, Hubert (1998): Pragmatische Ästhetik. Inszenierung, Performance und die Kunstfertigkeit alltäglichen kommunikativen Handelns. In: Herbert Willems/Martin Jurga (Hg.): Inszenierungsgesellschaft. Ein einführendes Handbuch. Opladen. S. 305-324.
- Knoblauch, Hubert (2004): Die Video-Interaktions-Analyse. In: Sozialer Sinn. Zeitschrift für hermeneutische Sozialforschung. Heft 1. S. 123-138.
- Knoblauch, Hubert (2005): Video-Interaktions-Analyse. In: Christoph Wulf/Jörg Zirfas (Hg.): Ikonologie des Performativen. München. S. 263-275.
- Knoblauch, Hubert/Schnettler, Bernd/Raab, Jürgen (2006): Video-Analysis. Methodological Aspects of Interpretive Audiovisual Analysis in Social Research. In: Hubert Knoblauch/Bernd Schnettler/Jürgen Raab/Hans-Georg Soeffner (Eds.): Video Analysis. Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology (pp. 9-26). Frankfurt a.M.
- Koller, Hans Christoph (1999): Bildung und Widerstreit. Zur Struktur biographischer Bildungsprozesse in der (Post-) Moderne. München.
- Kracauer, Siegfried (1964): Theorie des Films. Die Errettung der äußeren Wirklichkeit. Frankfurt a. M (Original: 1960).
- Kraul, Margret/Schmidtke, Adrian (2007): Mädchen und Jungen in der Eliteerziehung des Nationalsozialismus Eine Annäherung über Fotografien. In: Barbara Friebertshäuser/Heide von Felden/Burkhard Schäffer (Hg.): Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen & Farmington Hills. S. 249-259.
- Krüger, Heinz-Hermann (2000): Stichwort: Qualitative Forschung in der Erziehungswissenschaft. In: Zeitschrift für Erziehungswissenschaft, 3. Jg., Heft 3/2000. S. 323-342.
- Krüger, Heinz-Hermann/Marotzki, Winfried (Hg.) (1999): Handbuch erziehungswissenschaftliche Biographieforschung. (2. Auflage). Opladen.
- Krummheuer, Götz/Naujok, Natalie (1999): Grundlagen und Beispiele Interpretativer Unterrichtsforschung. Leske + Budrich.
- Kuchenbuch, Thomas (1978): Filmanalyse. Theorien – Modelle – Kritik. Köln.
- Kuchenbuch, Thomas (2005): Filmanalyse. Theorien – Methoden – Kritik. Wien/Köln/Weimar (2. Auflage)
- Langer, Susanne (1984): Philosophie auf neuen Wegen. Das Symbol im Denken, im Ritus und in der Kunst. Frankfurt a.M.
- Latour, Bruno (1998): Über technische Vermittlung. Philosophie, Soziologie, Genealogie. In: Werner Rammert (Hg.): Technik und Sozialtheorie. Frankfurt a.M. S. 29-81.
- Leeuwen, Theo van (2001): Semiotics and iconography. In: Theo van Leeuwen/Carey Jewitt (eds.): Handbook of Visual analysis (pp. 92-118). Los Angeles et al.
- Lenzen, Dieter (1993): Heiliges Bild oder Kreatur? Anmerkungen zum Kinderbild bei Otto Dix. In: Hans-Georg Herrlitz/Christian Rittelmeyer (Hg.): Exakte Phantasie. Pädagogische Erkundungen bildender Wirkungen in Kunst und Kultur. Weinheim u. München. S. 55-67.
- Lexikon der Kunst. Leipzig 2004.
- Loer, Thomas (1994): Werkgestalt und Erfahrungskonstitution. Exemplarische Analyse von Paul Cézannes ‚Montaigne Sainte-Victoire‘ (1904/06) unter Anwendung der Methode der objektiven Hermeneutik und Ausblick auf eine soziologische Theorie der Aesthetik im Hinblick auf eine Theorie der Erfahrung. In: Detlef Garz/Klaus Kraimer (Hg.): Die Welt als Text. Theorie, Kritik und Praxis der objektiven Hermeneutik. Frankfurt a.M. S. 341-382.
- Loer, Thomas (1996): Halbbildung und Autonomie. Über Struktureigenschaften der Rezeption bildender Kunst. Opladen.
- Luhmann, Niklas (1975): Macht. Stuttgart.
- Luhmann, Niklas (1990): Die Wissenschaft der Gesellschaft. Frankfurt a.M.
- Mannheim, Karl (1952a): Ideologie und Utopie. Frankfurt a.M. (Original: 1929. Bonn).
- Mannheim, Karl (1952b): Wissenssoziologie. In: Ders.: Ideologie und Utopie. Frankfurt a.M. S. 227-267 (Original: 1931. In: A. Vierkandt (Hg.): Handwörterbuch der Soziologie. Stuttgart. S. 659-680).
- Mannheim, Karl (1964a): Beiträge zur Theorie der Weltanschauungsinterpretation. In: Ders.: Wissenssoziologie. Neuwied S. 91-154 (Original: 1921-1922 in: Jahrbuch für Kunstgeschichte XV, 4).
- Mannheim, Karl (1964b): Das Problem der Generationen. In: Ders.: Wissenssoziologie. Neuwied. S. 509-565 (Original 1928 in: Kölner Vierteljahreshefte für Soziologie, 7. Jg. Heft 2, S. 157-185; Heft 3, S. 309-330).
- Mannheim, Karl (1980): Strukturen des Denkens. Frankfurt a.M.
- Marotzki, Winfried (1990): Entwurf einer strukturalen Bildungstheorie. Weinheim.
- Marotzki, Winfried/Schäfer, Eva (1999): Forschungsmethoden und Methodologie der Erziehungswissenschaftlichen Biographieforschung. In: Heinz-Hermann Krüger/Winfried Marotzki (Hg.): Handbuch erziehungswissenschaftliche Biographieforschung. (2. Auflage). Opladen. S. 258-275.
- Marotzki, Winfried (2006): Film- und Videoarbeit. In: Ralf Bohnsack/Winfried Marotzki/Michael Meuser (Hg.): Hauptbegriffe Qualitativer Sozialforschung. Opladen & Farmington Hills (2. Aufl.). S. 62-67.
- Marotzki, Winfried/Stoetzer, Katja (2007): Die Geschichten hinter den Bildern. Annäherungen an eine Methode und Methodologie der Bildinterpretation in biographie- und bildungstheoretischer Absicht. In: Winfried Marotzki/Horst Niesytho (Hg.): Bildinterpretation und Bildverstehen. Methodische Ansätze aus sozialwissenschaftlicher, kunst- und medienpädagogischer Perspektive. Wiesbaden. S. 15-44.
- Matthes, Joachim (1992): The Operation Called „Vergleichen“. In: Ders. (Hg.): Zwischen den Kulturen? (Sonderband 8 der Sozialen Welt). Göttingen. S. 75-99.
- Mead, George Herbert (1968): Geist, Identität und Gesellschaft. Frankfurt a.M. (Original 1934: Mind, Self and Society, Chicago).
- Mead, Margaret/Gregory Bateson (2002): On the Use of the Camera in Anthropology. In: Kelly Askew/Richard R. Wilk (Eds.): The Anthropology of Media (pp. 41-46). Malden/Oxford.
- Metz, Christian (1972): Semiologie des Films. München.
- Meuser, Michael (1998): Geschlecht und Männlichkeit. Soziologische Theorie und kulturelle Deutungsmuster. Opladen.
- Meuser, Michael (2007): Repräsentation sozialer Strukturen im Wissen. Dokumentarische Methode und Habitusrekonstruktion. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnold-Michael Nohl (Hg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Wiesbaden (2. Aufl.). S. 209-224.
- Michel, Burkard (2003): Dimensionen der Offenheit. Kollektive Sinnbildungsprozesse bei der Rezeption von Fotografien. In: Yvonne Ehrenspeck/Burkhard Schäffer: Film- und Photoanalyse in der Erziehungswissenschaft. Ein Handbuch. Opladen.

- Michel, Burkard (2004): Bildrezeption als Praxis. Dokumentarische Analyse von Sinnbildungsprozessen bei der Rezeption von Fotografien. In: ZBBS (Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung). Heft 1. S. 67-86.
- Michel, Burkard (2005): Kommunikation vs. Konjunktion. Zwei Modi der Medienrezeption. In: Volker Gehrau/Helma Bilandzik/Jens Woelke (Hg.): Rezeptionsstrategien und Rezeptionsmodalitäten. München.
- Michel, Burkard (2006a): Bild und Habitus. Sinnbildungsprozesse bei der Rezeption von Fotografien. Wiesbaden.
- Michel, Burkard (2006b): Das Gruppendiskussionsverfahren in der (Bild-) Rezeptionsforschung. In: Ralf Bohnsack/Aglaja Przyborski/Burkhard Schäffer (Hg.): Das Gruppendiskussionsverfahren in der Forschungspraxis. Opladen. S. 219-231.
- Michel, Burkard (2007): Fotografien und ihre Lesarten. Dokumentarische Interpretation von Bildrezeptionsprozessen. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Wiesbaden (2.Aufl.; ursprüngl.: 2001). S. 93-123.
- Mikos, Lothar (2003a): Film- und Fernsehanalyse. Konstanz.
- Mikos, Lothar (2003b): Zur Rolle ästhetischer Strukturen in der Filmanalyse. In: Yvonne Ehrenspeck/Burkhard Schäffer (Hg.): Film- und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch. Opladen. S. 135-149.
- Mitchell, William J.T. (1994): Picture Theory. Essays on Verbal and Visual Representation. Chicago & London.
- Mitchell, William J.T. (1997): Der Pictorial Turn. In: Christian Kravagna (Hg.): Privileg Bild. Kritik der visuellen Kultur. Berlin. S. 15-40.
- Mitchell, William J.T. (2002): Showing Seeing: A Critique of Visual Culture. In: Journal of Visual Culture Vol. I, Heft 2. S. 165-181.
- Mollenhauer, Klaus (1983): Streifzug durch fremdes Terrain. Interpretation eines Bildes aus dem Quattrocento in bildungstheoretischer Absicht. In: Zeitschrift für Pädagogik. 30.Jg., Heft 2. S. 173-194.
- Mollenhauer, Klaus (1996): Grundfragen ästhetischer Bildung. Theoretische und empirische Befunde zur ästhetischen Erfahrung von Kindern. Weinheim u. München.
- Mollenhauer, Klaus (1997): Methoden erziehungswissenschaftlicher Bildinterpretation. In: Barbara Friebertshäuser/Annedore Prengel (Hg.): Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim u. München. S. 247-264.
- Mollenhauer, Klaus (2003): Methoden erziehungswissenschaftlicher Bildinterpretation. In: Barbara Friebertshäuser/Annedore Prengel (Hg.): Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim. S. 247-264.
- Monaco, James (2003): Film und neue Medien. Lexikon der Fachbegriffe. Reinbek b. Hamburg (2. Aufl.).
- Morley, David (1999): Bemerkungen zur Ethnographie des Fernsehpublikums. In: Roger Bromley/Udo Göttlich/Carsten Winter (Hg.): Cultural Studies. Grundlagentexte zur Einführung. Lüneburg. S. 281-316.
- Müller-Doohm, Stefan (1993): Visuelles Verstehen. Konzepte kultursoziologischer Bildhermeneutik. In: Thomas Jung/Stefan Müller-Doohm (Hg.): „Wirklichkeit“ im Deutungsprozess. Verstehen und Methoden in den Kultur- und Sozialwissenschaften. Frankfurt a.M. S. 434-475.
- Müller-Doohm, Stefan (1997): Bildinterpretation als struktural-hermeneutische Bildanalyse. In: Ronald Hitzler/Anne Honer (Hg.): Sozialwissenschaftliche Hermeneutik. Opladen. S. 81-109.
- Nentwig-Gesemann, Iris (1999): Krippenerziehung in der DDR. Alltagspraxis und Orientierungen von Erzieherinnen im Wandel. Opladen.
- Nentwig-Gesemann, Iris (2001): Die Typenbildung der dokumentarischen Methode. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Opladen. S. 275-302.
- Nentwig-Gesemann, Iris (2006a): Regelgeleitete, habituelle und interaktionistische Spielpraxis. Die Analyse von Kinderspielkultur mit Hilfe videogestützter Gruppendiskussionen. In: Ralf Bohnsack/Aglaja Przyborski/Burkhard Schäffer (Hg.): Das Gruppendiskussionsverfahren in der Praxis. Opladen. S. 25-44.
- Nentwig-Gesemann, Iris (2006b): The ritual culture of learning in the context of family vacation: A qualitative analysis of vacation pictures. In: Tobias Werler/Christoph Wulf (Eds.): Hidden Dimensions of Education. Rhetoric, Rituals and Anthropology (pp. 135-148). Münster et al.
- Nentwig-Gesemann, Iris (2006c): Dokumentarische Evaluationsforschung. In: Uwe Flick (Hg.): Qualitative Evaluationsforschung. Reinbek b. Hamburg, S. 159-182.
- Nentwig-Gesemann, Iris (2007a): Sprach- und Körperdiskurse von Kindern. Verstehen und Verständigung zwischen Textförmigkeit und Ikonizität. In: Barbara Friebertshäuser/Heide von Felden/Burkhard Schäffer (Hg.): Bild und Text – Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen & Farmington Hills. S. 105-120.
- Nentwig-Gesemann, Iris (2007b): Der Familienurlaub; Rituelle Praxis; Differenzbearbeitung und Lernprozesse. In: Christoph Wulf et al. (Hg.): Lernkulturen im Umbruch. Rituelle Praktiken in Schule, Medien, Familie und Jugend. Wiesbaden. S. 220-252.
- Neumann-Braun, Klaus/Charlton, Michael (1989): Strukturanalytische Rezeptionsforschung. Theorie, Methode und Anwendungsbeispiele. In: Dieter Baacke/Hans-Dieter Kübler (Hg.): Qualitative Medienforschung. Konzepte und Erprobungen. Tübingen. S. 177-193.
- Neumann-Braun, Klaus/Schneider, Sylvia (1993): Biographische Dimensionen in der Medienaneignung. In: Werner Holly/Ulrich Püschel (Hg.): Medienrezeption als Aneignung. Methoden und Perspektiven qualitativer Medienforschung. Opladen. S. 193-208.
- Nohl, Arnd-Michael (2001): Migrationslagerung und Differenzzerfahrung. Junge Einheimische und Migranten im Milieuvergleich. Opladen.
- Nohl, Arnd-Michael (2002): Personale und soziotechnische Bildungsprozesse im Internet. In: Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung (ZBBS), Heft 2. S. 215-240.
- Nohl, Arnd-Michael (2006a): Bildung und Spontaneität. Phasen biographischer Wandlungsprozesse in drei Lebensaltern. Empirische Rekonstruktionen und pragmatistische Reflexionen. Opladen.
- Nohl, Arnd-Michael (2007): Komparative Analyse: Forschungspraxis und Methodologie dokumentarischer Interpretation. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Wiesbaden. S. 253-273 (2. Aufl.).
- Nohl, Arnd-Michael (2009): Interview und dokumentarische Methode.
- Nolda, Sigrid (2007): Videobasierte Kursforschung. Mögliche Erträge von interpretativen Videoanalysen für die Erforschung des organisierten Lernens Erwachsener. In: Zeitschrift für Erziehungswissenschaft. 10. Jg., Heft 4/2007. S. 478-492.
- Ondaatje, Michael (2002): Die Kunst des Filmschnitts. Gespräche mit Walter Murch. München/Wien.
- Panofsky, Erwin (1932): Zum Problem der Beschreibung und Inhaltsdeutung von Werken der Bildenden Kunst. In: Logos, XXI, S. 103-119. (Wieder abgedruckt in: Ders. (1964): Aufsätze zu Grundfragen der Kunstwissenschaft. Berlin. S. 85-97).
- Panofsky, Erwin (1964a): Die Perspektive als „symbolische Form“. In: Ders.: Aufsätze zu Grundfragen der Kunstwissenschaft. Berlin. S. 99-167.
- Panofsky, Erwin (1964b): Der Begriff des Kunstwillens. In: Ders.: Aufsätze zu Grundfragen der Kunstwissenschaft. Berlin. S. 29-43.
- Panofsky, Erwin (1975): Ikonographie und Ikonologie. Eine Einführung in die Kunst der Renaissance. In: Ders.: Sinn und Deutung in der bildenden Kunst. Köln. S. 36-67 (Original: 1955: Meaning in the Visual Arts. New York).
- Panofsky, Erwin (1989): Gotische Architektur und Scholastik. Zur Analogie von Kunst, Philosophie und Theologie im Mittelalter. Köln.
- Panofsky, Erwin (1999): Stil und Medium im Film. In: Ders.: Stil und Medium im Film & Die ideologischen Vorläufer des Rolls-Royce-Kühlers. Frankfurt a.M. S. 19-57.
- Panofsky, Erwin (2001a): Die altniederländische Malerei. Ihr Ursprung und Wesen. Band I. Köln (Ursprüngl: Early Netherlandish Painting. Harvard University Press. Cambridge/Mass. 1953).

- Panofsky, Erwin (2001b): Die altniederländische Malerei. Ihr Ursprung und Wesen. Band II. Köln (Ursprüngl: Early Netherlandish Painting. Harvard University Press. Cambridge/Mass. 1953).
- Panofsky, Erwin (2003): Korrespondenz 1910-1968. Eine kommentierte Auswahl in fünf Bänden. Hg. von Dieter Wuttke. Band II. Wiesbaden.
- Parmentier, Michael (1993): Sehen sehen. Ein bildungstheoretischer Versuch über Chardins ‚L'enfant au toton‘. In: Hans-Georg Herrlitz/Christian Rittelmeyer (Hg.): Exakte Phantasie. Pädagogische Erkundungen bildender Wirkungen in Kunst und Kultur. S. 105-121.
- Parmentier, Michael (2001): Jenseits von Idylle und Allegorie: Die Konstruktion des ästhetischen Subjekts in Bruegels "Kinderspielen". In: Christian Rittelmeyer/Michael Parmentier (Hg.): Einführung in die pädagogische Hermeneutik. Darmstadt. S. 89-104.
- Parramón Vilasaló, José M./Calbró Angrill, Muntsa (1999): Das große Buch vom Zeichnen und Malen in der Perspektive. München (2. Aufl.).
- Pasolini, Pier Paolo (1976): L'expérience herétique. Langue et cinema. Paris.
- Pfaff, Nicole (2006): Jugendkultur und Politisierung. Eine multimethodische Studie zur Entwicklung politischer Orientierungen im Jugendalter. Wiesbaden.
- Pilarczyk, Ulrike/Mietzner, Ulrike (2000): Bildwissenschaftliche Methoden in der erziehungs- und sozialwissenschaftlichen Forschung. In: Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung (ZBBS). Heft 2. S. 343-364.
- Pilarczyk, Ulrike/Mietzner, Ulrike (2005): Das reflektierte Bild. Die seriell-ikonografische Fotoanalyse in den Erziehungs- und Sozialwissenschaften. Bad Heilbrunn.
- Polanyi, Michael (1985): Implizites Wissen. Frankfurt a.M. (Original: The Tacit Dimension. Garden City/New York).
- Popper, Karl (1971): Logik der Forschung. Tübingen.
- Prange, Regina (1994): Stil und Medium im Film. Panofsky ‚On Movies‘. In: Bruno Reudenbach (Hg.): Erwin Panofsky: Beiträge des Symposiums. Hamburg. S. 171-191.
- Preuss-Lausitz, Ulf et al. (1983) (Hg.): Kriegskinder, Konsumkinder, Krisenkinder. Zur Sozialisationsgeschichte seit dem zweiten Weltkrieg. Weinheim.
- Przyborski, Aglaja (2004): Gesprächsanalyse und dokumentarische Methode. Qualitative Auswertung von Gesprächen, Gruppendiskussionen und anderen Diskursen. Wiesbaden
- Przyborski, Aglaja/Wohlrab-Sahr, Monika (2008): Qualitative Sozialforschung. Ein Arbeitsbuch. München.
- Pudowkin, Wsewolod I. (2003a): Filmregie und Filmmanuskript. In: Franz-Josef Albertsmeier (Hg.): Texte zur Theorie des Films. Stuttgart. S. 70-73.
- Pudowkin, Wsewolod I. (2003b): Über die Montage. In: Franz-Josef Albertsmeier (Hg.): Texte zur Theorie des Films. Stuttgart. S. 74-96.
- Raab, Jürgen (2002): „Der schönste Tag des Lebens“ und seine Überhöhung in einem eigenwilligen Medium. Videoanalyse und sozialwissenschaftliche Hermeneutik am Beispiel eines professionellen Hochzeitsvideofilms. In: Sozialer Sinn. Zeitschrift für hermeneutische Sozialforschung. Heft 3. S. 469-493.
- Raab, Jürgen (2008): Visuelle Wissenssoziologie. Theoretische Konzeptionen und materiale Analysen. Konstanz.
- Raab, Jürgen/Tänzler, Dirk (2002): Politik im/als Clip. Zur soziokulturellen Funktion politischer Werbespot. In: Herbert Willems (Hg.): Die Gesellschaft der Werbung. Wiesbaden. S. 217-245.
- Reckwitz, Andreas (2003): Grundelemente einer Theorie sozialer Praktiken. Eine sozialtheoretische Perspektive. In: Zeitschrift für Soziologie 32, Heft 4. S. 282-301.
- Reichertz, Jo (1992): Der Morgen danach. Hermeneutische Auslegung einer Werbefotografie in zwölf Einstellungen. In: Hans A. Hartmann/Rolf Haubl (Hg.): Bilderflut und Sprachmagie. Fallstudien zur Kultur der Werbung. Opladen. S. 141-163.
- Reichertz, Jo (2000): Die Frohe Botschaft des Fernsehens. Kulturwissenschaftliche Untersuchung medialer Diesseitsreligion. Konstanz.
- Renggli, Cornelia (2007): Selbstverständlichkeiten zum Ereignis machen: Eine Analyse von Sag- und Sichtbarkeitsverhältnissen nach Foucault [38 Absätze]. In: Forum Qualitative Sozialforschung/Forum Qualitative Social Research, 8 (2), Art. 23. <http://www.qualitative-research.net/fqs-texte/2-07/07-2-23-d.htm>.
- Ricoeur, Paul (1972): Der Text als Modell: Hermeneutisches Verstehen. In: Walter L. Brühl (Hg.): Verstehende Soziologie. Grundzüge und Entwicklungstendenzen. München. S. 252-283 (Wieder abgedruckt in: Gadamer, Hans-Georg/Boehm, Gottfried (1978): Seminar: Die Hermeneutik und die Wissenschaften. Frankfurt a.M. S. 83-117).
- Rittelmeyer, Christian/Parmentier, Michael (2001): Einführung in die pädagogische Hermeneutik. Darmstadt. S. 89-104.
- Rorty, Richard (1967): The Linguistic Turn: Recent Essays in Philosophical Method. Chicago.
- Rosenberg, Florian von (2008): Relationale Bildungsprozesse. Zur Rekonstruktion einer relationalen Bildungstheorie und Bildungsforschung. Dissertationsprojekt an der Helmut Schmidt-Universität Hamburg. Hamburg.
- Sabisch, Andrea (2007): Inszenierung der Suche. Vom Sichtbarwerden ästhetischer Erfahrung im Tagebuch. Entwurf einer wissenschaftskritischen Grafieforschung. Bielefeld.
- Sabisch, Andrea (2009): Recording and Representing Aesthetic Experience in Diaries. In: Ralf Bohnsack/Nicolle Pfaff/Wivian Weller (Eds.): Qualitative Analysis and Documentary Method in International Educational Research. Opladen & Farmington Hills.
- Sachs-Hombach, Klaus (2003): Vom Bild zum Film. Zur begrifflichen Analyse wahrnehmungsnaher Kommunikation. In: Yvonne Ehrenspeck/Burkhard Schäffer (Hg.): Film und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch. Opladen. S. 119-134.
- Sacks, Harvey (1995): Lectures on Conversation. Volumes I & II. Oxford (UK) & Cambridge (USA).
- Sacks, Harvey/Schegloff, Emanuel/Jefferson, Gail (1978): A Simplest Systematics for the Organization Turn Taking for Conversation. In: Jim Schenkein (Ed.): Studies in the Organization of Conversational Interaction (pp 7-55). New York.
- Schäfer, G./Wulf, Christoph (1999): Bild – Bilder – Bildung. Weinheim.
- Schäffer, Burkhard (1996): Die Band. Stil und ästhetische Praxis im Jugendalter. Opladen.
- Schäffer, Burkhard (1998): Generation, Mediennutzungskultur und (Weiter)Bildung. Zur empirischen Rekonstruktion medial vermittelter Generationenverhältnisse. In: Ralf Bohnsack/Winfried Marotzki (Hg.): Biographieforschung und Kulturanalyse. Transdisziplinäre Zugänge qualitativer Forschung. Opladen, S. 21-50.
- Schäffer, Burkhard (2003): Generation – Medien – Bildung. Medienpraxiskulturen im Generationenvergleich. Opladen.
- Schäffer, Burkhard (2005): Erziehungswissenschaft. In: Klaus Sachs-Hombach (Hg.): Bildwissenschaft. Disziplinen, Themen, Methoden. Frankfurt a.M. S. 213-225.
- Schäffer, Burkhard (2008): Bilder lebenslangen Lernens. Anmerkungen zu einem eigentümlichen Diskurs. In: Christiane Hof/Joachim Ludwig/Christine Zeuner (Hg.): Strukturen lebenslangen Lernens. Hohengehren.
- Schäffer, Burkhard (2009): Abbild – Denkbild – Erfahrungsbild. Methodisch-methodologische Anmerkungen zur Analyse von Altersbildern. In: Jutta Ecarius/Burkhard Schäffer (Hg.): Typenbildung und Theoriegenerierung. Perspektiven qualitativer Biographie- und Bildungsforschung. Opladen & Farmington Hills.
- Schierl, Thomas (2005): Werbungsforschung. In: Klaus Sachs-Hombach (Hg.): Bildwissenschaft. Disziplinen, Themen, Methoden. Frankfurt a.M. S. 309-319.
- Schulze, T. (1999): Bilder zur Erziehung. Annäherungen an eine „Pädagogische Ikonologie“. In: Gerd Schäfer/Christoph Wulf (Hg.): Bild – Bilder – Bildung. Weinheim. S. 59-88.
- Schumm, Gerhard (2004): Montage. In: Horst Schättle/Dieter Wiedemann (Hg.): Bewegte Bilder – Bewegte Zeit. Berlin. S. 167-169.
- Schütz, Alfred (1971): Gesammelte Aufsätze, Bd. 1: Das Problem der sozialen Wirklichkeit. Den Haag (Original 1962: Collected Papers, Vol. 1, The Problem of Social Reality. – Den Haag).
- Schütz, Alfred (1974): Der sinnhafte Aufbau der sozialen Welt. Eine Einleitung in die verstehende Soziologie. Frankfurt a.M. (zuerst: 1932: Wien).
- Schütze, Fritz (1987): Das narrative Interview in Interaktionsfeldstudien: Erzähltheoretische Grundlagen. Studienbrief der Fernuniversität Hagen. Teil I: Merkmale von Alltagserzählungen und was wir mit ihrer Hilfe erkennen können. Hagen.
- Selke, Stefan (2005a): Von der Formstruktur zum Sinngehalt. Kritische Würdigung der dokumentarischen Bildinterpretationsmethode am Beispiel der agenda-2010 Plakate. In: Daniel

- Fetzner/Stefan Selke (Hg.): selling politics. Bildinhalt und Bildwirkung. Ergebnisse des Forschungsprojekts selling politics zu Plakatmotiven der agenda 2010. Furtwangen 2005 (Schriftenreihe Fakultät Digitale Medien). S. 19-53.
- Selke, Stefan (2005b): Symbolische Politik oder Politik als Ware? Netzbasierte Bildwirkungsanalyse der agenda 2010-Plakatkampagne durch virtuelle Gruppendiskussionen. In: Daniel Fetzner/ Stefan Selke (Hg.): selling politics. Bildinhalt und Bildwirkung. Ergebnisse des Forschungsprojekts selling politics zu Plakatmotiven der agenda 2010. Furtwangen 2005 (Schriftenreihe Fakultät Digitale Medien). S. 125-175.
- Selke, Stefan (2007): Rekonstruktive Sozialforschung online. Qualitative Bildanalyse-Chats mit der Open Source Software VeraICON. In: Daniel Fetzner/Stefan Selke (Hg.): Bild – Raum – Interaktion. Angewandte empirische Wirkungsforschung. Ergebnisse interdisziplinärer Zusammenarbeit. Furtwangen 2005 (Schriftenreihe Fakultät Digitale Medien). S. 35-42.
- Sluneko, Thomas (2002): Von der Konstruktion zur dynamischen Konstitution. Beobachtungen auf der eigenen Spur. Wien.
- Soeffner, Hans-Georg/Raab, Jürgen (1998): Die Medialisierung des Sehens: Schnitt und Montage als Ästhetisierungsmittel medialer Kommunikation. In: Werner Rammert (Hg.): Technik und Sozialtheorie. Frankfurt a.M./New York. S. 121-148.
- Süddeutsche Zeitung (2007): Hollywood ist höllisch. Ein Gespräch mit David Lynch über seinen Film „Inland Empire“, Schönheit und Meditation, digitales Kino und Dinosaurier. 25. April. S. 15.
- Vogd, Werner (2004): Ärztliche Entscheidungsprozesse des Krankenhauses im Spannungsfeld von System- und Zweckrationalität. Eine qualitativ-rekonstruktive Studie. Berlin.
- Vogd, Werner (2005a): Systemtheorie und rekonstruktive Sozialforschung. Eine empirische Ver-söhnung unterschiedlicher theoretischer Perspektiven. Opladen.
- Vogd, Werner (2005b): Teilnehmende Beobachtung. In: Sven-Uwe Schmitz/Klaus Schubert (Hg.): Einführung in die Politische Theorie und Methodenlehre. Opladen. S. 89-109.
- Wacquant, Loïc D. (1996): Auf dem Weg zu einer Sozialpraxeologie. Struktur und Logik der Soziologie Pierre Bourdieus. In: Pierre Bourdieu/ Loïc D. Wacquant: reflexive Anthropologie. Frankfurt a.M. S. 17-93.
- Wagner-Willi, Monika (2004): Videointerpretation als mehrdimensionale Mikroanalyse am Beispiel schulischer Alltagsszenen. In ZBBS (Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung). Heft 1. S. 49-66.
- Wagner-Willi, Monika (2005): Zwischen Vorder- und Hinterbühne. Rituelle Übergangspraxen bei Kindern von der Hofpause zum Unterricht. Eine empirische Analyse in einer Berliner Grundschule. Wiesbaden.
- Wagner-Willi, Monika (2006): On the Multidimensional Analysis of Video-Data. Documentary Interpretation of Interaction in Schools. In: Hubert Knoblauch/Bernt Schnettler/Jürgen Raab/Hans-Georg Soeffner (Eds.): Video Analysis: Methodology and Methods (pp. 143-153). Frankfurt a.M. et al.
- Wagner-Willi, Monika (2007): Videoanalysen des Schulalltags. Die dokumentarische Interpretation schulischer Übergangsrituale. In: Ralf Bohnsack/Iris Nentwig-Gesemann/Arnd-Michael Nohl (Hg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. Wiesbaden (2. Aufl.). S. 121-140.
- Wehner, Christa (1996): Überzeugungsstrategien in der Werbung. Eine Längsschnittanalyse von Zeitschriftenanzeigen des 20. Jahrhunderts (Studien zur Kommunikationswissenschaft. Band 14). Opladen.
- Weller, Wivian (2003): Hip Hop in Berlin und São Paulo. Ästhetische Praxis und Ausgrenzungserfahrungen junger Schwarzer und Migranten. Opladen.
- Wiedemann, Dieter (2005): Film und Fernsehen. In: Klaus Sachs-Hombach (Hg.): Bildwissenschaft. Disziplinen, Themen, Methoden. Frankfurt a.M. S. 365-380.
- Willis, Paul (1979): Spaß am Widerstand. Gegenkultur in der Arbeiterschule Frankfurt a.M..
- Willis, Paul (1981): „Profane Culture“ – Rocker, Hippies: Subversive Stile der Jugendkultur. Frankfurt a.M..
- Winter, Rainer (1995): Der produktive Zuschauer: Medienaneignung als kultureller und ästhetischer Prozess. München.

- Winter, Rainer (1997a): Cultural Studies als kritische Medienanalyse. Vom „ecoding/decoding“-Modell zur Diskursanalyse. In: Andreas Hepp/Rainer Winter (Hg.): Kultur-Medien-Macht. Cultural Studies und Medienanalyse. Opladen. S. 47-63.
- Winter, Rainer (1997b): Reflexivität, Interpretation und Ethnografie: Zur kritischen Methodologie von Cultural Studies. In: Andreas Hepp/Rainer Winter (Hg.): Kultur-Medien-Macht. Cultural Studies und Medienanalyse. Opladen. S. 81-92.
- Winter, Rainer (2003a): Filmanalyse in der Perspektive der Cultural Studies. In: Yvonne Ehrenspeck/Burkhard Schäffer (Hg.): Film- und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch. Opladen. S. 152-164.
- Winter, Rainer (2003b): Polysemie, Rezeption und Handlungsmöglichkeit. Zur Konstitution von Bedeutung im Rahmen der Cultural Studies. In: F. Jannidis/G. Lauër/M. Martinez/S. Winko (Hg.): Regeln der Bedeutung. Zur Theorie der Bedeutung historischer Texte. Tübingen. S. 431-453.
- Winter, Rainer (2004): Cultural Studies. In: Uwe Flick/Ernst v. Kardorff/Ines Steinke (Hg.): Qualitative Forschung. Ein Handbuch. Reinbek b. Hamburg (3. Aufl.). S. 204-213.
- Wittpoth, Jürgen (2003): Fotografische Bilder und ästhetische Reflexivität. In: Yvonne Ehrenspeck/Burkhard Schäffer (Hg.): Film- und Fotoanalyse in der Erziehungswissenschaft. Ein Handbuch. Opladen. S. 73-86.
- Wopfner, Gabriele (2011): Kindliche Vorstellungen von Geschlechterbeziehungen. Dokumentarische Interpretation von Zeichnungen und Gruppendiskussionen. Dissertationsprojekt an der Freien Universität Berlin. Berlin.
- Wulf, Christoph (1998): Mimesis in Gesten und Ritualen. In: Paragrana. Internationale Zeitschrift für Historische Anthropologie. Bd. 7, Heft 1. S. 241-263.
- Wulf, Christoph (1999): Bild und Phantasie. Zur historischen Anthropologie des Bildes. In: Gerd Schäfer/Christoph Wulf: Bild – Bilder – Bildung. Weinheim. S. 331-344.
- Wulf, Christoph (2006): Mimesis. In: Ralf Bohnsack/Winfried Marotzki/Michael Meuser (Hg.): Hauptbegriffe Qualitativer Sozialforschung. Opladen & Farmington Hills (2. Aufl.). S. 117-119.
- Wulf, Christoph/Zirfas, Jörg (2001): Das Soziale als Ritual. Perspektiven des Performativen. In: Christoph Wulf et al. (Hg.): Das Soziale als Ritual. Zur performativen Bildung von Gemeinschaften. Opladen. S. 339-354.
- Wulf, Christoph/Zirfas, Jörg (Hg.) (2005): Ikonologie des Performativen. München.
- Wünsche, Konrad (1991): Das Wissen im Bild. Zur Ikonographie des Pädagogischen. In: Zeitschrift für Pädagogik. 27. Beiheft: Pädagogisches Wissen (hg. von Jürgen Oelkers u. Heinz-Elmar Tenorth). Weinheim u. Basel. S. 273-290.