

Wolfram Fischer

Rekonstruktive Videoanalyse.

Wahrnehmungs- und interaktionstheoretische Grundlagen, Methoden

„Ich seh’ etwas, was du nicht siehst ...“ Das Kinderspiel lebt von der Vieldeutigkeit der individuellen visuellen Wahrnehmung und dem Wunsch, sie in ein Gemeinsames zu überführen, etwas, worüber man eine klare Beschreibung ‚von etwas‘ in der Welt und somit das Gefühl, in einer gemeinsamen Wirklichkeit zu leben, erzielen kann. Es lebt von der Differenz der Perspektivität der Mitspieler und der Multivaldität visueller Phänomene und ihrer Beschreibbarkeit nach Qualitäten, Formen und Prozessen. Es lebt von den schier unendlichen Möglichkeiten, Bedeutung zu erkennen, zu erteilen und gemeinsam zu verifizieren oder zu modifizieren. Dass diese Eigenarten besonders dem Visuellen eignen, kann man erkennen, wenn man das Spiel auditiv variiert: „Ich hör’ etwas, was du nicht hörst ...“, das macht viel weniger Sinn, weil Hören von vorneherein selektiver und perspektivfrei funktioniert. Das Kinderspiel kann als symbolische Praxis, als Einübung in soziales Handeln und somit Herstellung geteilter Welt gesehen werden, die beim Visuellen ansetzt, aber dort nicht stehen bleibt. Damit sind schon zentrale Probleme visuellen Verhaltens, seiner Analyse und schließlich der instrumentellen Nutzung des Bildhaften, die in diesem Beitrag als Videoanalyse im Mittelpunkt steht, thematisiert.

Das Sehen ist also anthropologisch tief und vielleicht trotz der Einheit aller Sinne *prävalent* (Plessner 1980: 275) als eine hervorragende Steuerungsleistung des Handelns und der Wahrheitsgarantie mundaner Existenz in die sozial-leibliche Kompetenz und alltägliche Praxis des Handelns und der Wahrnehmung eingelassen. Seine Bedeutung wird im Laufe der Kulturgeschichte gesteigert durch visuelle Praxen, die *mehr* zeigen, sie wollen beschwören, festhalten, rückmelden, chiffrieren, Gegebenes transzendieren, zur Besinnung oder Irritation des Alltäglichen dienen, Stimmungen zu diesem oder jenem Zweck erzeugen. Man bemalt Felswände, benutzt Spiegel, erzeugt Bilder, entwickelt optische Abbildungsapparate wie Kameras und produziert damit Fotos, schließlich Filme, zunächst ohne, dann mit Ton, audio-visuelle Aufzeichnungsgeräte werden entwickelt. Man nutzt

elektronische Verbreitungsmedien wie das Fernsehen und das Internet. Das letzte ‚Medium‘ zeichnet sich mittlerweile dadurch aus, dass es eine große Mitbeteiligung der Betrachter an der Herstellung der visuellen Produkte erlaubt, somit Konsumenten zu Produzenten, Betrachter (und Zuhörer) zu Zeigenden (und Sprechenden) werden lässt und damit auch medial eine soziale Reziprozität erlaubt, die bislang dem Alltagsleben in direkter Interaktion und Kommunikation unter Anwesenden vorbehalten war.

Alle diese visuellen Techniken haben je eigene kulturelle Darstellungs-, Seh- und Ausführungspraxen des Alltäglichen, der besonderen Sphären der Religion, Kunst und Wissenschaften, der kulturellen Chiffrierung des Selbstverständnisses von Milieus, Gruppen und Gesellschaften hervorgebracht; sie haben somit die (visuelle) Komplexität des Zusammenlebens gesteigert, indem sie dessen Symbolisierungen stark erweitert haben. In dem Maße, in dem das gesellschaftliche Leben Bilder und Filme sowie Audi-Video-Aufzeichnungen hervorbringt, re-produzieren sie als Symbole des Sozialen nicht nur die Ordnung des Sichtbaren, sondern ko-produzieren überhaupt die gemeinsame Kultur. Dazu gehören die öffentlichen Diskurse und symbolischen Ordnungen ebenso wie das Handeln und Erleben bis in die Mikrobereiche der Interaktion.

Das besondere Interesse, *Bewegungsabläufe zu filmen* und somit zeitlich abfließende und damit flüchtige Prozesse zu fixieren und jenseits der Handlungssituation wiederholbar zugänglich zu machen, hat sich in den letzten Jahrzehnten enorm gesteigert. Hier spielt die Entwicklung und massenhafte Vermarktung von leistungsfähigen, handlichen und preiswerten Videokameras (nicht nur bild-, sondern auch tonfähig), digitaler Technik, Aufzeichnungsträgern und Bearbeitungsinstrumenten, die allesamt auch von nicht ausgebildeten Fachleuten bedienbar sind, eine entscheidende Rolle. In keinem anderen ‚Medium‘ steigern sich die alltäglichen Leistungen von *Sehen und Hören, Sagen und Zeigen* so nahe an den lebensweltlichen Handlungen und Erfahrungen der Lebenswelt wie der Videotechnik. Es verwundert daher nicht, dass sich AV-Aufzeichnungen nicht nur im Alltag zur Selbstpräsentation und Gruppen- oder Familienkonsolidierung finden, sondern auch verstärkt in beruflichen Ausbildungssituationen und beim Einüben von professionellen Praxen, bei denen es um Erlernen und Kontrolle von professionsgerechten Handlungsketten geht, eingesetzt werden.

Da besonders auch im Bereich der Sozialen Arbeit solche Videoverfahren Eingang gefunden haben, soll zunächst kurz auf die *Videoanalyse als Feedbackinstrument in der Praxis*

eingegangen werden (1). Dem massenhaften Videogebrauch korrespondiert bislang noch keine entsprechende wissenschaftliche Aufklärung über Funktionen, Möglichkeiten und Grenzen dieser visuellen Praxis. Der vorliegende Artikel lotet die Voraussetzungen solcher Forschung aus. Er befasst sich mit Videoverfahren als Methoden einer qualitativen und rekonstruktiven Sozialforschung und steht damit im Kontext neuerer Entwicklungen visueller Soziologie. Nach grundlegenden *wahrnehmungs- und interaktionstheoretischen Überlegungen* (2), die mit der Frage *nach nicht nur bewegten Bildern und Ikonik* (3) weiter auf *Videoanalyse als Interaktionsanalyse und forschungspraktische Verfahren* zugespitzt werden (4), kommen abschließend *videoanalytische Forschungsdesiderata für die Felder der Sozialen Arbeit* zur Sprache (5).

1. Videoanalyse als Feedbackinstrument in der Praxis

Im Bereich der Praxis der Sozialen Arbeit und der verschiedenen professionellen Beratungsformate bezeichnet der Begriff ‚Videoanalyse‘ meist weniger ein wissenschaftliches Forschungsverfahren, sondern eine *Praxismethode*, die dazu dient, den alltäglichen Blick der Klienten und Professionellen zu schärfen. Die allgemeine Funktion der Videoaufzeichnung besteht hier darin, zunächst den Akteuren nach erfolgter Handlung einen Blick von außen auf ihre eigene Praxis zu geben. Die sich durch die Akteure wechselseitig verschränkende komplexe intrasituative Handlungsperspektivität wird dabei durch eine einfache gleichgerichtete Quasi-Außenperspektive, die das Ablaufdokument der Handlung ermöglicht, konterkariert. Diese veränderte Wahrnehmung, die immer auch das Hören des Gesprochenen beinhaltet (und bei dem sich die Sprecher selbst besonders fremd erleben), soll den Akteuren einen anderen Zugang auf ihr Handeln erlauben, ‚blinde Flecken‘ identifizieren und somit schließlich auch alternative Handlungsoptionen eröffnen. Die Befremdung der eigenen Praxis durch ein scheinbar objektives Dokument erlaubt probeweise Rollenübernahme wie alternative Handlungsentwürfe und erzeugt im günstigsten Fall bislang nicht verfügbares Wissen, indem unter der Oberfläche der beobachtbaren Erscheinungen liegende allgemeine Erzeugungsstrukturen des Handelns erkannt werden. Entscheidend für diese Praxis ist dabei die Einbindung des Gesehenen in ein Gespräch unter den Aktanten, also etwa zwischen Klient und professionellem Berater, Student und Dozent, Coachee und Coach, oder auch zwischen Gleichgestellten, wie etwa im Peer-Audit unter Experten. Es wird dabei immer versucht, von individuellen Wahrnehmungsvarianten

ausgehend gemeinsame Deutungsvarianten der videographisch dokumentierten Handlungsabläufe zu produzieren und so potentielle Handlungsvarianten vorzubereiten. Das Verfahren bewährt sich besonders dort, wo es um eingeschliffene Handlungsformen geht, die als selbstverständlicher Bestandteil kontinuierlicher Praktiken von den Akteuren selbst nicht explizit gewusst und nicht hinterfragt werden. Es ist weiterhin besonders geeignet, Einübungen in noch nicht beherrschte Verhaltensweisen oder Kommunikationstechniken (etwa der Beratung oder des Coaching) zu erleichtern, indem mit den Akteuren die Wirkung des Geschehens auf die Beteiligten sowohl *in situ* wie *post et extra situm* verbalisiert werden kann. Es geht bei dieser weit verbreiteten Art der Videoanalyse darum, etwas zu *sehen*, was man sonst nicht sehen kann, etwas zu *zeigen*, was wichtig erscheint, etwas gemeinsam zu *besprechen* und zu entwickeln, was die *Optionen des Erlebens und Handelns* erweitert. Die angewendete Methode ist immer ähnlich: Eine natürliche Handlungs-Situation wird per Videokamera gefilmt und danach gemeinsam betrachtet - was das gemeinsame Hinhören einschließt - und diskursiv nach bestimmten Hinsichten ausgewertet. Teilweise sind hieraus in beruflichen Settings Qualitätszirkel entstanden, die sich im Peer-Audit die Sicherung und Pflege der Qualität des eigenen beruflichen Handelns zur Aufgabe machen (Bahrs u. a. 1996; Bahrs u. a. 1998; Fischer-Rosenthal 1996). Auf dieser Linie liegen auch die seit Anfang der 1980er Jahre aus den Niederlanden importierten Modelle des so genannten Video-Home-Trainings (Kreuzer/ Räder 1999; Schepers/ König 2000) in der Sozialen Arbeit, vor allem in der Familien-, Kinder- und Jugendhilfe (Leist 1998; ter Horst 2008), die es mittlerweile in verschiedenen Formaten (Cordes 2000) gibt und die auch kommerziell als ‚Markenprodukte‘, wie z.B. SPIN (SPIN/Deutschland 1994), Orion (Schepers 1996) oder Marte Meo (Aarts 2002; Aarts u. a. 2002; Bündler u. a. 2009), auf dem - teilweise trägerinternen (Brümmer/ ter Horst 1997; Gens 2002) - Fortbildungsmarkt auch für die Förderung von Leitungskompetenzen (Gens 2002) angeboten werden. Die (lizenzierte) Videotrainerin erstellt und bereitet die Videoaufnahme – etwa von einer Familieninteraktion in der Klientenfamilie – vor, indem sie bestimmte ‚Szenen‘, bzw. Handlungssequenzen auswählt, die im Sinne einer sozialpädagogischen Förderung positive Handlungsansätze erkennen lassen oder in einem schwierigen Verhaltensmuster die Diskussion von alternativen Handlungsmöglichkeiten zulassen. Diese Ausschnitte werden in einem entsprechend hergestellten Klientensetting dann mit den Klienten zusammen betrachtet und besprochen. Der Schwerpunkt ist hierbei im Sinne

der humanistischen Psychologie primär das fördernde und unterstützende sozialpädagogische Gespräch, weniger die mikroanalytische Betrachtung von Fehlverhalten.

Dieser pädagogischen Zielrichtung ordnet sich auch die mittlerweile recht verbreitete Videoarbeit ein, die innerhalb einer aktiven Medienpädagogik die Produktion von Videos (Aufnahme, Schnitt und Vertonung) durch Jugendliche in den Mittelpunkt stellt. Die Jugendlichen werden in pädagogischer und medientechnischer Begleitung in die Lage versetzt, dokumentarische oder in Spielszenen umgesetzte Videos zu produzieren, die von sich selbst und ihrer Lebenswelt oder einem bevorzugten jugendkulturellen Milieu (z.B. Hip Hop oder Rap) handeln. Es sollen dabei pädagogisch erwünschte Bildungsprozesse, selbstkritische Auseinandersetzung und Gruppenprozesse befördert werden (vgl. das Projekt in einer hess. JVA Nolle/ Bauer 2008; Nolle/ Hildebrandt 2006). Vielfach wird dabei das Web als Träger- und Verbreitungsmedium genutzt, das auch kürzere ‚Video-Clips‘ und unter Einschluss von komplementärem Text und Tonereignissen collagenhafte (‚multimediale‘) Produkte hervorbringt, die insgesamt stark graphisch gestaltet werden müssen. Neben kurzfristigen Projekten entstehen hier auch innerhalb institutionalisierter Jugendarbeit nachhaltige Teileinrichtungen (vgl. für viele: in Kiel www.videomobil.de ; in Wuppertal www.medieprojekt-wuppertal.de). Zur Unterstützung von AV- und Web-Projekten entstehen - oft als Spin-Offs sozialpädagogischer Fachbereiche - auf dem Bildungsmarkt pädagogische Fortbildungsangebote und Netzwerke zur Herstellung multimedialer Produkte, die der Selbstreflexion und Identitätsformation bestimmter Zielgruppen dienen (z.B. in Leipzig www.medienblau.de oder in Hessen <http://medienpaedagogik-hessen.de>).

Es fällt auf, dass die genannten professionellen Praxen und die hohe Verbreitung von Videoverfahren kaum wissenschaftliche empirische Begleitstudien inspiriert haben. Entsprechend findet auch eine Auseinandersetzung mit diesen Produktions- und Gebrauchstechniken fast ausschließlich in pädagogisch motivierten interaktiven Lernsettings und weniger im wissenschaftlichen Diskurs statt. Über die möglicherweise auch durch die Praxen, Erzeugungstechniken, Abbildungs- und Darstellungsapparaturen unerkannten Vorannahmen, Selbstverständlichkeiten und ihre Wirkungen ist wenig bekannt. Zu einem professionellen Selbstverständnis in den so genannten helfenden Berufen gehört jedoch unabdingbar dazu, dass die eigenen Praxen wissenschaftlich reflektiert und formiert werden. Hier liegt ein großer Nachholbedarf an Forschung und Verbesserung wissenschaftlicher Professionalität (s. u. Abschnitt 5). Von diesen feedback-orientierten

Verfahren einer Videoanalyse *als* pädagogischer, sozialpädagogischer oder ähnlicher *Praxis* ist mithin die Videoanalyse als *Instrument der Sozialforschung* zu unterscheiden. Darum wird es im Folgenden gehen.

2. Videoanalyse in der qualitativen Sozialforschung

– wahrnehmungs- und interaktionstheoretische Grundlagen

2.1 Wahrnehmen – oder wie sich Innen und Außen verschränken

Die Beschäftigung mit einer seit wenigen Jahrzehnten gut verfügbaren einzelnen Forschungstechnik wie der Aufzeichnung von Ereignissen mithilfe audiovisueller Verfahren, kurz Video genannt, wird in jüngerer Zeit meist mit einem gesellschaftlichen Wandel in Zusammenhang gebracht. Er wird dadurch rhetorisch belegt, dass er durch die Entwicklung visueller Medien begleitet oder gar getragen sei. Die Videotechnik wird so in den größeren Zusammenhang gesellschaftlicher und visueller Evolution gestellt. Das vollmundige Postulat einer ‚ikonischen Wendung‘ zugunsten der Wissenschaften vom Bild – so anfangs der 1990er Jahre Boehm oder Mitchell (Boehm 2006: 13; Mitchell 1994) – wird als Markierung der eigenen Aktualität mit dem inflationären Gebrauch zunehmend wertloser.

Fest steht, es gibt in der gegenwärtigen globalen Kultur neben bekannten eine Fülle neuer und expandierender visueller Objekte und Praktiken, die offenbar gesellschaftliche Funktionen erfüllen. Diese lassen sich wissenschaftlich untersuchen und auch instrumentell in einem definierten Interesse nutzen und verändern. Die sozialwissenschaftliche Beschäftigung mit visueller Kultur in den letzten Jahren, die mittlerweile auch in Deutschland eine Reihe von gewichtigen Methodenhandbüchern (Bohnsack 2009; Ehrenspeck/ Schäffer 2003; Friebertshäuser u. a. 2007; Knoblauch u. a. 2006; Raab 2008) hervorgebracht hat, zeigt allerdings, wie komplex und quer durch die Einzelwissenschaften verlaufend die theoretischen Diskurse (s. beeindruckende Aufreihung Sachs-Hombach 2005) und die auf sie bezogenen methodologischen Grundfragen sind. Ebenso ist nach mehr als einem Jahrzehnt nach der Ausrufung des ‚iconic turn‘ deutlich geworden, dass die scheinbare Aktualität und Modernitätsbehauptung eine rhetorisch stark zugespitzte Epochalisierung im eigenen Interesse darstellte. Sie war vielleicht wissenschaftspolitisch kurzfristig nützlich, verdeckt jedoch zu leicht, dass die Auseinandersetzung mit dem Visuellen im Konzert aller Sinne und Handlungsweisen des Menschen, einschließlich der Thematik von Schrift und

Text, schon lange in verschiedenen Einzelwissenschaften ihren Platz hat. Das anspruchsvolle aktuelle Programm einer wissenschaftlichen Disziplin namens ‚Visual Culture‘ wird also entwertet, wenn sie zu eng auf gegenwärtigen Bildgebrauch westlicher Gesellschaften eingegrenzt wird. „If visual culture is to mean anything, it has to be generalized as the study of all the social practices of human visibility, and not confined to modernity or the West.“ (Mitchell 2005: 349)

Die entscheidenden Fragen sind einfach zu stellen und schwierig zu beantworten: Was ist Sehen? Was ist ein Bild? Was ist ein visuelles Image? Was ist ein Medium? Was ist der Unterschied zwischen einem (Stand-) Bild und einem bewegten Bild? Welche Beziehung zu anderen Sinnen liegt vor? Welche zur Sprache? Warum ist visuelle Erfahrung so mit Angst und Phantasie befrachtet? Hat das Sehen eine Geschichte? Welche Rolle spielen visuelle Begegnungen mit anderen Menschen (Bildern, Objekten) bei der Konstruktion der sozialen Wirklichkeit? (Mitchell 2005: 355)

Es handelt sich beim Visuellen um Fähigkeiten, Praktiken, Produkte und deren gesellschaftlichen Gebrauch, die einer Vielzahl von theoretischen und einzelwissenschaftlichen Kontexten zuzuordnen sind. Kontexte philosophischer Ästhetik (meist im Schatten der Erkenntnistheorie), der Semiotik, der Anthropologie der Sinne, der Wahrnehmungspsychologie, der Kognitionspsychologie, der Ethnologie, der Kunsttheorie, der Kunstgeschichte, der Filmtheorie, der Kulturwissenschaften und nicht zuletzt der Soziologie - sie alle tragen zur Wissenschaft vom Sehen und den Bildern etwas bei. Dabei werden je nach einzeldisziplinärer Zuordnung auch recht verschiedene Methodiken der Bild- und Videoanalyse begründet. Es wird kaum gelingen, Fragen und Verfahren visueller Analyse zu entwickeln, die in allen diesen Kontexten bestehen können. Gleichwohl erscheint es nützlich, wenn man sich nicht damit begnügt, Videoanalyse als eine ‚methodologiefreie‘ Forschungstechnik im Rahmen einer Einzeldisziplin zu betreiben, sondern sich zu vergegenwärtigen, welche Grundfragen quer durch die Disziplinen an dieser Stelle hervorgebracht wurden. So wird man der Komplexität des Gegenstandes gerechter und vermeidet naive Forschungspraktiken, deren Vorannahmen – z.B. das einer alltagsweltlichen Abbildungsvorstellung - ungeklärt einfließen. Diese interdisziplinäre Aufgabe und forschungspraktische Kooperation sind weithin noch zu leisten. Gleichwohl gehen etwa die Versuche innerhalb der Wissenssoziologie (Bohnsack 2003, 2007; Raab 2008: 20-95; Raab/ Tänzler 2006; Soeffner 2004) und Ethnologie (Amann 1997; Doering/ Hirschauer 1997; Mohn

2002), philosophische, anthropologische und kunsttheoretische Diskussionen aufzugreifen, genau in diese Richtung und belegen die Fruchtbarkeit eines solchen multidisziplinären Weges. Im Rahmen des vorliegenden Beitrags können lediglich einige der Grundfragen, die aus den verschiedenen theoretischen Kontexten inspiriert sind, skizziert werden.

Einen allgemeinen und fundierenden Zugang sucht bereits im ersten Drittel des 20. Jahrhunderts eine *Theorie der Wahrnehmung*, die sich einerseits in der frühen phänomenologischen Philosophie (seit der Vorlesung vom WS 1904/05 Husserl 1980; Husserl 2004; vgl. allgemein Orth 1994-95), andererseits in der Wahrnehmungspsychologie, besonders auch in der Gestalttheorie (Wertheimer, Köhler, Koffka, Lewin), entwickelt. Diese theoretischen wie empirisch gestützten Versuche waren mittelfristig in der Sozialphilosophie (Gurwitsch 1974; Merleau-Ponty 1966, 1986; Schütz 2003b, 2003f, 2003c, 2003a) und langfristig in der neueren Wissenssoziologie sehr einflussreich. Sie haben auch über den Gestaltbegriff in anderen Disziplinen und in therapeutischen Praktiken (Gestalttherapie, Humanistische Psychologie) Wirkungen entfaltet und einen ganzheitlichen Wahrnehmungs- und Interaktionsbegriff entwickelt, der auch für visuelle Phänomene offen ist. Die differenzierten und verschiedenen Konzepte, die sich aus der Frage nach Wahrnehmung, Realitätskonstitution, der Rolle des Handelns und Erlebens, der Rolle von Erinnerung und Imagination, oder nach sprachlichen und nicht sprachlichen Symbolisierungen ergeben, umfassen ein ausgesprochen großes Feld und können hier nur beispielhaft ange deutet werden.

Edmund Husserl (1859-1938) geht von der alltäglichen Annahme und Erfahrung aus, dass die Wahrnehmung in der ‚natürlichen Einstellung‘ zur selbstverständlichen Konstitution einer existierenden Welt führt, die sich im Handeln zu bewähren hat. Die Frage, aus der er das große Programm der phänomenologischen Philosophie entwickelt, lautet: *Wie* geschieht das? So nimmt die Aufklärung der Wahrnehmung – und dazu gehört auch die visuelle Wahrnehmung – einen zentralen Platz ein. Er versteht sie als eine komplexe (aktive und passive) Bewusstseinsleistung (cogitatio), bei der in der relativ natürlichen Einstellung etwas aus dem größeren Wahrnehmungsfeld aktuell gewahr wird. Unter Wahrnehmungsfeld versteht er nicht den sich dem Blick bietenden Sektor der ‚äußeren‘ Welt, sondern es geht ihm um das *Bewusstseinsfeld* mit seinen Leistungen, internen Organisation und Logik. Er verdeutlicht den Vorgang an der Wahrnehmung eines Blatt Papiers. „Im eigentlichen

Wahrnehmen, als einem Gewahren, bin ich dem Gegenstande, z.B. dem Papier zugewendet, ich erfasse es als dieses hier und jetzt Seiende. Das Erfassen ist Herausfassen, jedes Wahrgenommene hat einen Erfahrungshintergrund. Rings um das Papier liegen Bücher, Stifte, [...] in gewisser Weise auch ‚wahrgenommen‘, perzeptiv da, im Anschauungsfelde [...] Jede Dingwahrnehmung hat so einen Hof von Hintergrundsanschauungen“.

(Husserl 1976b: 71) Anders als die Gestaltpsychologie, die diese Relation von „Figur und Grund“ als Gestalt-Gegebenheiten der Objektwelt belegt, der die Organisation des Bewusstseinsfeldes (Gurwitsch 1974) folgt, kommt es Husserl nicht auf den realen Umkreis von wahrgenommenen Objekten an, sondern auf den Erfahrungshintergrund *im Bewusstsein*. Indem er so den Akt der Zuwendung in der Wahrnehmung akzentuiert, kommt er auf Modifikationen des ursprünglichen Wahrnehmungserlebnisses im ‚geistigen Blick‘. Er spricht hier von implizit und explizit bewussten ‚Gegenständen‘, die in die Wahrnehmung hineinspielen. Damit bekommen unmittelbar auch Erinnerung, erinnerungsähnliche Vergegenwärtigungen und Phantasie eine mögliche Rolle bei der Wahrnehmung. „Die Wahrnehmung charakterisierten wir als einen Akt, in dem uns das Gegenständliche als in eigener Person gleichsam, als selbst gegenwärtig erscheint. In der Phantasie erscheint der Gegenstand zwar insofern selbst, als eben er es ist, der da erscheint, aber er erscheint nicht als gegenwärtig, er ist nur vergegenwärtigt, es ist gleichsam so, als wäre er da, aber nur gleichsam, er erscheint uns im Bilde. Die Lateiner sagen *imaginatio*.“ (Husserl 1980: 16) Die aktuelle Wahrnehmung eines Dings lässt sich somit modifizieren, indem sie in den Modus der Inaktualität (Erinnerung oder Phantasie) überführt wird. Es kann eine Erinnerung oder Phantasie auftauchen und aktuell werden, die wir noch gar nicht fokussieren (a.a.O.: 72). Das Aktuelle im Bewusstsein ist von einem Hof von Inaktuellem umgeben; das Bewusstsein wandert zwischen diesen Sphären hin und her. Später führt Husserl hierfür den allgemeinen Begriff der *Appräsentation* ein. Die „kontinuierlich fortlaufende Kette von *cogitationes* [ist] beständig von einem Medium der Inaktualität umgeben, diese [ist] immer bereit, in den Modus der Aktualität überzugehen wie umgekehrt die Aktualität in die Inaktualität.“ (a.a.O.: 73) Selbst für die Wahrnehmung eines einfachen, scheinbar statischen Objekts in der Welt (z.B. eines konkreten Tisches, der vor mir als Betrachter steht und mir als real existierender Tisch erscheint) ergibt sich daraus die Frage, *wie* sich aus dem visuell immer perspektivisch reduzierten Teilbild, bzw. aus einem beständigen Fluss verschiedener perspektivischer Wahrnehmungen, wenn ich mich um den Tisch

herumbewege, automatisch die Vorstellung von einem soliden Tisch als wirklich da seiendem Gegenstand konstituiert. Es entsteht mir eben nicht ein *Bild* vom Tisch, sondern die Wahrnehmung eines *realen Tisches*, der jetzt vor mir steht und den ich so unmittelbar als existent erlebe.

Wenn man so fragt, stößt man einerseits auf den Zusammenhang von Bewusstsein, Zeit und Dingkonstitution, denn „immerfort wandelt sich das Wahrnehmungs-Jetzt in das sich anschließende Bewusstsein des Soeben-Vergangenen, und zugleich leuchtet ein neues Jetzt auf“ (a.a.O.: 84). Zum anderen wird eine genauere Bestimmung der Wahrnehmung (*noesis*) und des wahrgenommenen Gegenstandes (*noema*) - zu dem das tatsächliche Ding transzendent ist - notwendig. Beide thematischen Stränge führen schließlich zur Frage, was ein Bild sei und wie sich die Bildwahrnehmung zur direkten Wahrnehmung von Ereignissen und Objekten in der Lebenswelt verhält. Ich skizziere zunächst die Rolle der Phantasie und Bildwahrnehmung und komme dann auf die Zeitlichkeit der Wahrnehmung zu sprechen.

Husserl hat großen Wert darauf gelegt, dass die Wahrnehmung einer Phantasie, eines Bildes und die Wahrnehmung von Ereignissen oder Dingen in der Welt deutlich zu unterscheiden sind. Es hängt zunächst damit zusammen, dass er einen Unterschied zwischen *Bild als Imagination*, also Phantasiebild, und Sache macht. Mit Sache ist nicht der reale Gegenstand gemeint, sondern der immanente reelle Inhalt, also der in der Vorstellung gemeinte Gegenstand, „der für seiend gehaltene (z.B. erinnerte oder erwartete) oder der für unwirklich gehaltene, wie in der bewussten Fiktion, oder der bezweifelte, erwünschte, erfragte, erhoffte, befürchtete usw. [...] Wenn das Berliner Schloss uns im Phantasiebild vorschwebt, so ist eben das Schloss zu Berlin die gemeinte, die vorgestellte Sache. Davon unterscheiden wir aber das vorschwebende Bild, das natürlich kein wirkliches Ding und nicht zu Berlin ist. Das Bild macht die Sache vorstellig, ist aber nicht sie selbst.“ (Husserl 1980: 18)

Wie steht es nun weiter mit dem physischen Bild, das in irgendeiner Weise materialiter gegeben ist? Husserl unterscheidet hier begrifflich drei Ebenen: einmal das *Bild als physisches Objekt* (z.B. an der Wand oder im Photoalbum); dann das *Bildobjekt*, darunter versteht er das, was man auf dem Bild optisch sieht; schließlich das bildlich dargestellte „*Sujet*, dem wir in der normalen Einstellung des Abbildungsbewusstseins zugewendet sind und das je nach dem als wirklich seiend oder nicht seiend, ferner als gegenwärtig oder vergangen seiend charakterisiert ist.“ (Husserl 2004: 350f.) An anderer Stelle heißt es

knapp: „Das physische Bild weckt das geistige Bild, und dieses wieder stellt ein anderes: das Sujet vor.“ (Husserl 1980: 29)

Die Wahrnehmung des *Bildsujets*, das, was den *Sinn der Betrachtung*, die *Bedeutung*, ausmacht, ist möglicherweise nur locker an das Bildobjekt gebunden, sie ist eine Interpretation des Bewusstseins. Somit ist die Konstitution des Bildsujets als des Bildsinns zwar durch das optisch Sichtbare angeregt, geht aber durch das Bildobjekt hindurch. Das Bildobjekt ist nur ein nicht wahrgenommenes Durchgangsstadium auf das Sujet, den eigentlichen Wahrnehmungsinhalt des Bildes. Dieses stellt etwas vor, vergegenwärtigt etwas (eine Phantasie, eine Geschichte, einen Charakter, etc.) und ist deshalb nicht dasselbe wie die (auch visuelle) Wahrnehmung einer Situation oder eines Dinges in der Lebenswelt, die unmittelbar gegeben und evident ist. „Diese Art der Anschaulichkeit durch ein Bild (Bildobjekt) hindurch ist offenbar eine wesentlich andere und uneigentliche gegenüber der echten unmittelbaren Anschaulichkeit, wie sie in jeder Perzeption vorliegt, es ist eine Quasi-Leibhaftigkeit, die doch nicht wirkliche Leibhaftigkeit ist. [...] Es scheidet sich danach scharf dieses mittelbare Anschauen, das ein unmittelbares voraussetzt, vom unmittelbaren.“ (a.a.O.: 29) Auf diesen Unterschied von mittelbarer und unmittelbarer Wahrnehmung, etwa der Eigenwahrnehmung einer Interaktion durch die Interaktanten in der Situation und das anschließende Betrachten einer AV-Aufzeichnung dieser Interaktion, wird später eingegangen.

Die Differenz zwischen dem Bildobjekt, das mir sozusagen auf dem Bildträger optisch geliefert wird, und dem Bildsujet, das ich dadurch erkenne, macht vollends deutlich, dass letzteres als der Sinn des Wahrgenommenen eine Leistung des Betrachters, eine Interpretation ist. Selbst wenn ich nur ein ‚Abbild‘ (z.B. die Photographie von etwas) vor mir habe und kein Kunstobjekt - Boehm nennt das ein ‚schwaches Bild‘ im Unterschied zum ‚starken Bild‘ in der Kunst (Boehm 2008: 246 ff.) – macht dies für Husserl keinen prinzipiellen Unterschied, denn wahrnehmungstheoretisch geht die Bedeutung immer vom Betrachter aus. Sie ‚hängt‘ nicht am Bildobjekt und wird von ihm erzwungen, sondern entsteht beim Betrachter. Wie frei der Betrachter dabei ist, etwas so oder anders zu interpretieren, also verschieden ‚zu sehen‘, oder ob es nicht kollektiv übereinstimmende Sichtweisen gibt, ist die entscheidende Frage, wenn man die Konstitution einer gemeinsam geteilten Welt erklären will.

Methodologisch liegt in dieser Frage der Interpretierbarkeit der Ansatzpunkt für die

Thematisierung des Verstehens. Hermeneutik als methodische Kunst der Auslegung kommt erst auf, wenn man dauerhaft etwas Fremdes erlebt, etwas nicht versteht und in der Aneignung oder der Änderung seiner bisherigen Wahrnehmung einen Interpretationsspielraum findet. Man kann dann das Problem der Eindeutigkeit oder des Interpretationsspielraums der Wahrnehmung in Richtung Bewusstsein oder in Richtung Bildobjekt ausloten. Im ersten Fall geht es um das beim Betrachter vorhandene Wissen, seine Erinnerungen und Phantasien, allgemein seine sonstigen aktualisierbaren Appräsentationen. Im anderen Fall geht es um die Frage, wieweit die Bindung durch das im Bildobjekt visuell Gegebene geht. Die Frage lautet dann, inwieweit macht das Sichtbare in seiner Oberflächen- oder Formalstruktur Bedeutungs- und Sinnvorgaben, und wie ist es überhaupt der Wahrnehmung gegeben? Ebenso ist diese Frage für das Verstehen von Texten oder des Sprechens in einer Interaktionssituation hoch relevant: liest man, was da steht, hört man was gesagt wird oder liest man das, was man aus seiner eigenen Sicht versteht, und hört man nur das, was man selber zulässt.

Husserl beginnt seine Reflexion damit, dass er vom Blickfeld, also der visuellen Leistung eines Betrachters, ausgeht und dabei zunächst die Gleichzeitigkeit aller Gegenstandsmomente im Augen-Blick konstatiert. „Halten wir uns an das Blickfeld, in dem die Gegenständlichkeiten nicht nacheinander, sondern zugleich erscheinen. Dem entsprechen die mannigfaltigen Wahrnehmungen dieser Gegenständlichkeiten bzw. die Wahrnehmungserscheinungen. Diese sind ebenfalls zugleich und nicht nacheinander. Nun umfasst das Blickfeld seinem Begriff nach alle in der Form des Zugleich koexistierenden Erscheinungen, und diese all konstituieren einen einzigen Zusammenhang: d.h. in ihnen erscheint ein gegenständlicher Zusammenhang. Diese Koexistenz gehört zu einem Querschnitt des Wahrnehmungsbewusstseins. In Wahrheit geht der Zusammenhang kontinuierlich in der Sukzession fort: Die wahrgenommene Gegenständlichkeit erstreckt sich durch die Zeit hindurch. D.h. durch die Folge des Nacheinander, und zwar stetig, wobei sie schon in jedem Querschnitt des Zugleich, der Koexistenz, einen einheitlichen Zusammenhang bildet.“ (Husserl 1980: 72) Die Präsenz des Blickfeldes ist also zugleich in der Zeitlichkeit des Bewusstseinsstroms aufgehoben. Diese Zeitlichkeit hat Husserl in großen Studien (Husserl 1966, 2001) sein Leben lang beschäftigt. Die Wahrnehmung als zeitliche Dauer, ein Hauptthemenkreis, der Dauer des Bewusstseinsstroms entspricht, wird in Bezüge von Retention und Protention, also eine Einbettung des fortströmenden Jetzt in fortwährend

sich modifizierendes Rückschauen und Vorschauen aufgefächert. Rückschau und Vorschau sind dabei eng mit den bereits angeschnittenen Themen der Phantasie, Erinnerung, Erwartung und Bilderfassung verbunden. Kurz gesagt, die aktuelle Wahrnehmung von etwas Bestimmtem, die sich auf Empfindungsinhalte stützt, und die im Licht der Empfindungsinhalte modifizierte Wahrnehmung von Vergangenem und Zukünftigem verschmelzen miteinander und konstituieren, was jetzt wahrgenommen wird. Vereinfacht, die aktuellen Empfindungen werden angereichert oder überlagert durch das, was ich jetzt erinnere und jetzt erwarte. Erinnerung und Erwartung werden also nicht als feststehende mentale ‚Aufzeichnungen‘ gesehen, sondern werden durch die aktuelle Situation verändert. Doch gibt es so etwas wie einen ersten Impuls von außen, der dann weiter modifiziert wird? Das wäre eine „Ur-Impression“ (Husserl), die nur vom Ereignis angestoßen wird, bevor sie im Ichprozess der Wahrnehmung modifiziert wird. Hier entstünde dann auch eine rein perzeptive Evidenz, die keine weiteren, vor allem auch keine sprachlichen oder kommunikativen Einflüsse und Modifikationen braucht, um sich gewiss zu sein (vgl. dazu kritisch im Sinne einer Sprachpragmatik Apel 1990). Wenn es so etwas gibt, dann wäre auch eine ichliche, also auf das Bewusstsein bezogene Zeit von einer vor-ichlichen, quasi äußeren (Husserl sagt ‚hyletischen‘) Zeit zu unterscheiden. Getreu seinem transzendentalphilosophischen bewusstseinszentrierten Ansatz entscheidet sich Husserl hier eindeutig für die alleinige Geltung der ichlichen, phänomenalen Zeit und schließt eine bewusstseinsunabhängige ‚äußere Zeit‘ aus (vgl. genauer Husserl 2001: 129-163). Dennoch war er offenbar nicht ganz zufrieden damit, sonst hätte ihn die Frage nach intersubjektiver Zeit nicht sein Leben lang beschäftigt. Es muss hier offen bleiben, ob Husserls spätere Zuwendung zur „Lebenswelt‘ als dem allgemeinen ‚Boden‘ menschlichen Weltlebens“ und das Hauptaugenmerk „auf einen abstrakt herauszupräparierenden Weltkern: die Welt der schlichten intersubjektiven Erfahrungen“ (Husserl 1976a: 158, 136) die bewusstseinstheoretische Schlagseite seines Ansatzes überwunden haben.

Der Zusammenhang von *jetzt zu vorher und nachher*, von *nicht mehr und noch nicht* für die Erzeugung einer (visuellen) Wahrnehmung - sei sie auf Gegebenheiten und Abläufe in der Welt oder auf Bilder bezogen -, gilt analog auch für die anderen Sinne, die an der Wahrnehmung beteiligt sind. Dass die Wahrnehmung, in die sowohl inhaltlich wie zeitlich verschiedene Faktoren hineinspielen, eine einheitliche Gegenständlichkeit zur Erscheinung bringt, liegt „im Zusammenhang der Auffassungsinhalte, der Empfindungen und Emp-

findungsfelder. Die Empfindungen sind im Gesichtsfeld nicht isoliert, sondern kontinuierlich einheitlich zusammenhängend, miteinander verschmolzen.“ (Husserl 1980: 72) Das gilt auch für die Inhalte der anderen Sinne, denn immer liefert die sinnliche Wahrnehmung die Vorstellung eines realen ganzen Objekts, nicht nur Teilaspekte. Doch woher kommt die Einheit? Sie kommt durch nichts anderes als durch die Wahrnehmungsleistung selbst: Es „ist der apperzeptive Zusammenhang, der Einheit darstellt, im sinnlich erscheinenden Objekte, das zugleich gesehen und getastet ist, haben Gesichts- und Tastinhalte fühlbare Einheit, die Einheit der gegenständlichen Zusammengehörigkeit.“ (a.a.O.: 72) „Das perzeptive Gesichtsfeld und das perzeptive Tastfeld oder Gehörsfeld koexistieren, die Empfindungsgruppen sind gesondert, aber sie sind zusammen zu schauen, und sie schmelzen auch zusammen in intuitiv-apperzeptiven Einheiten, es erscheinen Gegenstände, die die zugehörigen Empfindungen, nur gedeutet, vereinigt enthalten mögen.“ (a.a.O.: 75) Husserl betont allerdings, dass sich diese gemeinsame Gegenstandskonstitution der verschiedenen Sinne in der Wahrnehmung nicht vollzieht, wenn das Feld der Phantasie etwa mit dem Gesichtssinn verglichen werde. Es entstehe entweder eine Phantasiewahrnehmung *oder* eine visuelle Wahrnehmung, sie können nicht in eins verschmelzen. Ich weiß zum Beispiel um die Unterschiede der in meinem Gesichtsfeld wahrgenommenen roten Rose, die ich als reale Rose erlebe, und der nur phantasierten gelben Rose, die eben keine Präsenz eines realen Gegenstands zeitigt. Geht es im ersten Fall um Präsenz des Gegenstands in seiner Leibhaftigkeit, handelt es sich im zweiten Fall um eine gedankliche Imagination. Der Fall dazwischen ist nun die Rose, die ich etwa in einem Stillleben sehe. Sie wird mir als Bildsujet anlässlich dessen, was ich optisch sehe, vorstellig, sie wird aber nicht real und ist dennoch keine Phantasie. Ihr kommt vielmehr eine Quasi-Leibhaftigkeit zu, für die möglicherweise das materielle Bild eine Stellvertreter- und das Bildobjekt eine Auslösefunktion haben. In jedem Fall hat die auf dem Bild wahrgenommene Rose eine *artifizielle Präsenz*. Es geht also nicht nur darum, dass das Bildobjekt ein *Zeichen* für etwas ist, sondern es *vergegenwärtigt* etwas, was *sich im Bildsujet zeigt* und benennbar ist. Hierin sieht Wiesing den Grundgedanken einer phänomenologischen Bildwissenschaft (Wiesing 2005: 31). Der Sinn des Bildes besteht also letztlich in einer Auslegungsregel, durch die das Bildobjekt zum Bildsujet wird. Diese Regel ist nicht eine einmalige individuelle Angelegenheit, sondern sie ist stark kontextbezogen, ist in ‚Sehgemeinschaften‘ entstanden (vgl. auch Wiesing 2005: 62ff.). „Der Sinn legt fest, *wie* ein Bildobjekt als Zeichen genutzt

werden soll. Doch das Bildobjekt legt fest, *was* ein Sinn sein kann.“ (a.a.O: 67) Dieses ‚Wie‘ kommt aus den Gebrauchsweisen und Kontexten. Ich sehe auf dem Foto die Ähnlichkeit der Frau im Garten mit einer mir bekannten Person (meine Mutter beim Pflanzen) oder auf dem Gemälde erkenne ich in der Frau mit dem Schwan eine visuelle Vergegenwärtigung einer antiken Narration (Zeus schwängert Leda). Die Verstehensleistung, ihre Bedingungen und Grenzen werden allgemein unter dem Titel Hermeneutik bearbeitet; *Hermeneutik* wird zum allgemeinen *Verfahren für das Erkennen und Zuschreiben von Sinn* (vgl. Abschnitt 4).

Diese nach dem Programm der phänomenologischen Philosophie streng auf Wahrnehmung als Bewusstseinsleistung eines abstrakt gedachten Ich beschränkten Überlegungen, haben sich selber als Transzendentalphilosophie verstanden, die keine Aussagen über die reale Welt außerhalb der Wahrnehmungskonstitution machen will. Intersubjektivität, Kommunikation und Sprache – also das, was Sozialität ausmacht - werden so zum theoretischen Problem, weil sie vom transzendentalen Ego her nicht zu begründen sind. Gleichwohl waren diese Konzepte der Wahrnehmung auch stark wirksam in Richtung Soziologie, die sich als mundane Wissenschaft zur empirischen und theoretischen Aufklärung der sozialen Welt versteht, und haben vor allem über Alfred Schütz die moderne Wissenssoziologie bis heute inspiriert. Thomas Luckmann hat sich stets dagegen gewandt, dies als phänomenologische Soziologie zu bezeichnen, sondern sprach von *Proto-Soziologie* (Luckmann 1980; 2002: 45ff.; zugespitzt auf Symbolgebrauch Soeffner 2004), insofern sich phänomenologische Konzepte den empirischen Zugriff auf die Welt versagen, aber begriffliche Voraussetzungen dafür in einer Bestimmung der „Strukturen der Lebenswelt“ (Schütz/Luckmann 1979/1984) geschaffen haben. Bevor ich darauf im Zusammenhang mit Interaktion, Symbol und Sprache zu sprechen komme, sollen die für eine Bildwissenschaft wichtigen leibzentrierten Konzepte der Wahrnehmung bei Merleau-Ponty und Plessner vorgestellt werden. Das Konzept der Leiblichkeit kann eine Empirie des Visuellen fundieren, die jenseits der Dichotomie von Betrachter und objekthaftem Bild und jenseits einer stillschweigenden Abbildtheorie liegt, indem sie Betrachter und Bild einem gemeinsamen handelnd-sinnlich konstituierten Lebenszusammenhang zurechnet.

2.2 Leiblichkeit der Sinne

Für *Maurice Merleau-Ponty (1908 - 1961)* wird Leiblichkeit (Fischer 2003) geradezu das

zentrale begriffliche Mittel, sich aus philosophischen Sackgassen zu befreien, die wissenschaftsgeschichtlich höchst folgenreich waren, weil sie mit einer Spaltung von ‚innen‘ und ‚außen‘ auch diejenige zwischen ‚empirisch‘ operierenden Naturwissenschaften und Geisteswissenschaften fundieren. Es geht um den cartesianischen Dualismus von Subjekt und Objekt (Merleau-Ponty 1966), der die Philosophiegeschichte der Neuzeit prägt und sich bis in problematische dualistische wissenschaftliche Konzepte von ‚innen‘ und ‚außen‘, von geisteswissenschaftlicher Kulturanalyse einerseits und empirischen Naturwissenschaften andererseits fortsetzt. Auch die Gegenüberstellung einer naiv alltagsweltlichen und positivistischen Vorstellung der Abbildbarkeit, Beschreib- und Messbarkeit von Welt, *wie sie ist*, zu einer Welt, *wie sie Sinn macht* und daher notwendig hermeneutisch ausgelegt werden muss, bewegt sich in diesem Fahrwasser.

Zur Reflexion der Wahrnehmung, die für empirische Wissenschaften das zentrale Konzept ist, lässt sich Merleau-Ponty in erster Linie durch die Gestalttheorie (Merleau-Ponty 2003: 7, 10, 25) (und die neurologische Gehirnforschung nach dem 1. Weltkrieg) und nicht durch die Philosophie Husserls anregen. Letztere ist ihm zu bewusstseinszentriert, kann also den Dualismus von Subjekt und Objekt für ihn nicht befriedigend reflektieren und aufheben. Die cartesianisch nach Subjekt und Objekt strukturierte Wissenschaft setzt in seinem Urteil somit auch naiv Wahrnehmung als Abbildung voraus und klärt sie nicht auf. Die Berücksichtigung des Leibes führt zu einer Auflösung dieses Entweder-oder von Subjekt und Objekt, sie verschränkt beide Seiten vielmehr und realisiert so eine Wahrnehmungskonzeption des Sowohl-als-auch. Merleau-Ponty konstatiert, „dass unser Leib ein zweiblättriges Wesen ist, auf der einen Seite ist er Ding unter Dingen, und auf der anderen sieht und berührt er sie; und wir stellen fest, [...] dass er diese zwei Eigenschaften in sich vereinigt, und dass seine doppelte Zugehörigkeit zur Ordnung des ‚Objekts‘ und des ‚Subjekts‘ uns zur Entdeckung ganz unerwarteter Beziehungen zwischen diesen beiden Ordnungen führt.“ (Merleau-Ponty 1986: 180) Der Leib erscheint für Merleau-Ponty als der besondere ‚Gegenstand‘, der nicht mit anderen Gegenständen in der Welt gleichzusetzen ist, weil er die unumgehbare Basis aller Wahrnehmung ist, immer schon in der Welt und diese von seinem (nicht selber wahrgenommenen) Nullpunkt perspektivisch wahrnehmend. „Gegenwart und Abwesenheit äußerer Gegenstände sind nur Variationen innerhalb eines dem Vermögen meines Leibes zugeeigneten primordialen Gegenwartsfeldes und Wahrnehmungsbereiches. Nicht allein ist die Ständigkeit meines Leibes nicht lediglich ein Sonderfall der

Ständigkeit äußerer Gegenstände in der Welt, vielmehr versteht sich die perspektivische Darstellung der Gegenstände ihrerseits daraus allein, dass mein Leib sich jeder perspektivischen Variation widersetzt. Wenn Gegenstände mir notwendig stets nur eine ihrer Seiten zeigen, so weil ich selbst einen bestimmten Platz einnehme, von dem aus ich sie sehe, den ich selbst aber nicht sehen kann.“ (Merleau-Ponty 1966: 117) Dieser Leib ist aber nicht nur einfach ein Rezeptionsinstrument, eine Leinwand, auf der sich die Welt abbildet (das wäre eine cartesianische Konzeption), sondern er ‚zeigt sich‘ expressiv und impressiv auch anderen. Das heißt, er ist in einem lebensweltlichen kulturellen Kontext interaktiv, ‚gesehen‘ und handelnd, positioniert, aufgebaut und selber in die Welt eingreifend, was jegliche Wahrnehmung, auch die von Bildern mit bestimmt, ohne dass die Wahrnehmung das selbst weiß. Ebenso sind Objekte und Bilder nicht einfach Objekte, sondern auch sie haben Teil an der Leiblichkeit des Menschen und dem ‚Fleisch‘ seiner Welt. Sie sind somit auch soziale Gegebenheiten, die Gemeinsamkeit stiften.

Das heißt in unserem Kontext für eine Untersuchung der visuellen Wahrnehmung, dass sie mit zu rekonstruieren hat, welche Konstruktionsleistungen die Betrachter und schließlich die Produzenten von (bewegten) Bildern erbringen, von deren Logik sie nichts wissen, weil sie sozial-leiblich in die konkreten Bildpraktiken der Rezeption und Produktion und der Bedeutung dieser Praktiken für ihre konkrete Lebenswelt verwickelt sind. Dieser Zugang erlaubt weiter eine Analyse, die danach fragt, in welchem Sinne ein Bild teilhaben könnte an einer sehenden und zeigenden Kultur, die sich dieser Praxen zur Ausbildung einer gemeinsam geteilten Welt blind bedient.

Für den Philosophen und Biologen *Helmuth Plessner (1892 -1985)* war das Projekt einer philosophischen Anthropologie in erkenntnistheoretischer wie empirisch-praktischer Hinsicht nicht zu trennen von der Frage nach der Erforschung des Lebendigen. Die tief greifende Kritik am cartesianischen Subjekt-Objekt-Schema wird in einer Reflexion auf die Stellung des Menschen in der Welt des Organischen geleistet (Plessner 1981: 78-126). Seine These von der exzentrischen Positionalität des Menschen (a.a.O.: 360ff.) hat er seit Mitte der 1920er Jahre in einer Reihe von großen Untersuchungen zur Leiblichkeit ausgearbeitet, die leibliche Akte wie Lachen und Weinen und besonders die Rolle der Sinne umfasst. Es gelingt ihm, sowohl der Differenziertheit der Sinne in einer Ästhesiologie des Gehörs, des Gesichts und der Zustandssinne (Riechen, Schmecken, Tasten, Temperatursinn, Organ- und Lagegefühl), gerecht zu werden, als auch ihre Einheit zu erfassen. Der

Zustandssinn, dessen Sinn in nichts anderem als „der Vergegenwärtigung des körperlichen Seins“ (Plessner 1980: 286) liegt, steht konträr und scheinbar unverbunden den visuellen und auditiven Sinnen gegenüber, die gegenständliche Sinngehalte (z.B. Bildwahrnehmung und -herstellung, Geometrie als theoretische Leistung, Musik), im weitesten Sinn gegenständliche und soziale Kulturinhalte, begründen. Dennoch belegt Plessner die These ihrer *Einheit* im Prinzip des Verhältnisses von „Leib und Geist“. Doch wie ist dieses Einheit stiftende Verhältnis zu denken? „Pure Vergegenwärtigung der Körper im Erleben ist die Funktion der Zustandssinne. Eine zentrale Stellung nimmt in dieser Vergegenwärtigung der Leib des Individuums, der ‚eigene‘ Körper, ein. Denn er ist Reizfeld, Ausdrucksfeld, Reaktionsschema und Ansatzgebiet beziehungsweise Instrument der Impulse und des Willens, die konkrete Figur, in der die seelische Wirklichkeit ihre Vertretung [...] besitzt. Da die seelische Realität unmittelbar den Artikulationsgesetzen der Sprache folgt, [...] da wir das Seelische in den Weisen möglicher Ausdrückbarkeit durch Sprache [...] bei uns und anderen wahrnehmen, ergeben sich zwei Möglichkeiten der Vergegenwärtigung körperlichen Seins. Einmal in der Wahrnehmung und Vorstellung, zweitens als ‚Hintergrund‘ der Seele“. (a.a.O: 286) Es geht also um die (vorwiegend) visuelle Fremdwahrnehmung der anderen körperlichen Dinge und die Selbstwahrnehmung des eigenen Körpers, weiterhin um die Vergegenwärtigung des eigenen Leibes im Zusammenhang mit dem eigenen Seelenzustand. Freude, Lust, Trauer etc. sind nicht einfach aktuelle körperliche Sensationen anlässlich irgendeines Auslösers, sondern untrennbar mit seelischen Inhalten verbunden, die durch Erfahrungen, Erwartungen und biographische Aufschichtungen konstituiert sind. Schließlich manifestieren sich der Geist und die Seele im *leiblichen Ausdruck*, der sich quasi nicht intentional herstellt, und in *zielgerichteten Handlungen* der Körper, die alles an menschlichen Aktivitäten abdecken, nicht zuletzt die Kunst. „Ursprüngliche Gegenwart des Geistes ist nur an Leibern in ihrer Haltung ablesbar, während das künstliche Verfahren jeder Kunst darin besteht, Körper wie: Leinwand, Farbe, Stein zu Ausdrucksfeldern zu machen, mit ihnen einen Sinn zu verleiblichen. Die unmittelbare Ausdrucksfähigkeit des Leibes, welcher eine gleich unmittelbare und ursprüngliche Auffassungsgabe des Geistes für den Ausdruckssinn seiner Gestik, Mimik, Physiognomik notwendig entspricht, hat Geltung für jeden Inhalt, mag er seelischer Natur sein, woran wir zunächst denken, oder selbst auch geistigen, gedanklichen, sinnhaften Wesens.“ (a.a.O.: 288)

Zweierlei scheint mir an dieser Konzeption der Einheit der Sinne durch die Wechselbeziehung von Leib und Geist für unseren Themenzusammenhang von besonderer Bedeutung. Zum einen richtet sich visuelle Aufmerksamkeit in Interaktionen unter Anwesenden offenbar in ständiger visueller Wahrnehmungsleistung auf das *Ausdrucksfeld des Leibes*, um für die Zwecke der Interaktion das Gegenüber zu lesen und das eigene Verhalten daran auszurichten. Etwas am Körper zu *zeigen* (z.B. Aufmerksamkeit, Gefühle, Zustimmung oder Ablehnung) und *sich zu zeigen* (z.B. als bedeutsame, witzige oder fügsame Person) fallen im Ausdrucksfeld des Leibes in eins. Dieses wird *sichtbar* in die Interaktion eingebracht, und es entstehen mitschwingende oder gegenläufige Expressionen im Ausdrucksfeld des Leibes der anderen Interaktionsteilnehmer (sie gleichen ihre Haltung an oder verändern sie) und so fort. Hier erfährt das Thema der so genannten *nonverbalen Kommunikation* eine begriffliche und beobachtungsfähige Zuspitzung. Über diese Darstellungs- und Wahrnehmungsleistung ist trotz der Prominenz des Terminus noch zu wenig geforscht worden. Die Videoanalyse von Interaktionen hat hier eine wichtige Aufgabe.

Zum anderen wird von Plessner – wie nebenbei - eine interessante Kunsttheorie formuliert. Kunst verwirklicht sich, indem sich das Ausdrucksfeld des Leibes quasi auf das Ausdrucksfeld von Material ausweitet. Dort werden zielgerichtet Objekte hergestellt, die in der Schaffung eines *Sujets* (im Sinne Husserls) das visualisieren, was im Zusammenspiel von Eigenleiblichkeit und Kultur wurzelt, am eigenen Leib aber nur unzureichend dargestellt werden kann. Die Betrachtung, Herstellung und der Gebrauch von Bildern und Filmen in der Kunst, in der Wissenschaft und im Alltag erfährt hier eine interessante Ursprungs-erklärung, die vom Leiblich-Geistigen ausgeht und im Medium des Visuellen als artifizielle Präsenz zu ihm zurückkehrt. Eine rekonstruktive Analyse von visuellen Objekten und Prozessen kann versuchen, diesen Doppelbogen vom Leib zum visuellen Produkt und von diesem zurück zum wahrnehmenden Leib in seinen Strukturmerkmalen nachzuzeichnen.

Mit diesen Betrachtungen bin ich bei dem Verhältnis von Wahrnehmung und Sprache oder dem von Bild und Text angekommen. Wie kann man die Wendung Plessners, dass das Seelische den Ausdrucksmöglichkeiten der Sprache folgt, verstehen? Die nahe liegende Vermutung, dass auch miteinander Sprechen ohne leibliche Präsenz und Aktivität nicht möglich ist, wird meist zugunsten eines Sprachverständnisses, in dem sich Symbolsystem

und Symbolisierungsleistung des Menschen als geistige Kompetenz wechselseitig aufbauen, verdeckt. Während ersteres die Frage aufwirft, was Menschen miteinander *tun*, wenn sie sich unterhalten und interagieren, also einen interaktionslogischen, *sprachpragmatischen Fokus* hat, beschäftigt sich letzteres mit der Bedeutungskonstitution und inner-sprachlichen Ordnung, hat also einen semiotischen (semantischen und syntaktischen) Fokus. Ich gehe zunächst auf den semiotischen Aspekt ein, komme aber noch auf eine pragmatische, aktuallsprachliche Sicht von Interaktion zurück.

2.3 Symbolgebrauch, Interaktion und die Konstruktion einer gemeinsamen Wirklichkeit

Alfred Schütz (1899 - 1959) verbindet Konzepte zur Interaktion und zum Entstehen von Wissen zu einer Theorie der Lebenswelt (Schütz 2003d, 2003e; Schütz/ Luckmann 1979/1984), dabei spielen Symbole und Sprache eine zentrale Rolle.

Interaktion als gemeinsames Handeln geschieht immer in einem bereits kulturell, das heißt vor allem auch sprachlich und symbolisch geordneten Kontext. Dieser Kontext legt fest, was im Vollzug einer Handlung oder eines Erlebens sinnvoll ist, was als nächstes kommen kann, aufeinander beziehbar ist und welche Anschlüsse an Vergangenes oder Zukünftiges nahe liegen. Dabei wird Intersubjektivität sowohl vorausgesetzt als auch für die aktuellen Zwecke konstituiert. Dies funktioniert nur, wenn es möglich ist, in einer Situation etwas für die Anwesenden zu vergegenwärtigen, das nicht auf Antrieb zu erkennen oder jetzt abwesend ist. Diese Funktion der Vergegenwärtigung wird durch die *Appräsentationskraft von Symbolen* ermöglicht, weil die menschliche Wahrnehmung zur Transzendierung des unmittelbar sinnlich Gegebenen in der Lage ist. Wie hängen nun Intersubjektivität und die Vergegenwärtigung von sozialen Gruppen, von ihrem Wissen, ihrer Geschichte und ihrer Handlungspraxis mit dem Gebrauch von Symbolen zusammen?

Schütz unterscheidet zwei Ebenen der Appräsentationsbeziehungen. Einmal geht es beim Verstehen des anderen in der Reziprozität der Perspektiven um das Transzendieren des dem Ich jetzt unmittelbar Gegebenen. Vor allem die Wahrnehmung des anderen als alter ego (und nicht nur als Objekt) transzendiert egos Perspektive. Es geht um „immanente Transzendenz“ innerhalb der Alltagswelt. Die andere Ebene, der weitere Bezugskreis, der sich nicht mehr in meiner aktuellen Reichweite befindet, besteht aus sozialen Kollektiven und institutionalisierten Beziehungen. Sie überschreiten Egos unmittelbare Wirkmöglichkeit und erfordern so auch keine interaktionelle Rollenübernahme und Empathie und lassen

keine unmittelbare Erfahrung entstehen. Sie bilden ein vermitteltes, Interaktionen übersteigendes durch Symbole geprägtes Subuniversum. Solche Symbole stabilisieren und institutionalisieren soziale Beziehungen, werden somit alltagswirksam. Dazu gehören vor allem auch Symbole und Repräsentationen der politischen Ordnung (Schütz 2003b: 167-197). Schließlich lässt sich der Kreis noch erweitern auf abstraktere und komplexere höhersymbolische Welten wie Religion, Wissenschaft und Kunst. Die Unterscheidung von Sinn (oder Unsinn), der Verweis auf das, was im Meer des Unbedeutsamen bedeutsam werden kann, erfolgt in den Symbolisierungsapparaturen von *Sprache und Sprechen* (a.a.O.: 37-115), von *Bild und Bildgebrauch* und schließlich durch *Rituale*. Bei diesen drei Klassen von Symbolsystemen handelt es sich gleichermaßen um Verweis- und Herstellungsprozeduren von Bedeutungen, die gemeinsame und mannigfaltige Wirklichkeiten erzeugen, in denen sich Menschen handelnd orientieren und zuordnen können. Auf dem Boden der Alltagswelt entwickeln sich mannigfaltige Welten, die den Wirklichkeitsakzent der alles fundierenden Lebenswelt nicht erreichen, die man gleichwohl ‚erwandern‘ kann, ohne dass man die Alltagswelt je dauerhaft verlassen könnte (Schütz 2003d: 181-247).

Soeffner übernimmt von Schütz/Luckmann die dreiteilige räumliche und soziale Differenzierung der alltäglichen Lebenswelt und die damit verbundene Entstehung verschiedenen Wissensarten (Schütz/ Luckmann 1979/1984: I, 63-124, 314ff.). Die Unterscheidung der Welt in meiner aktuellen Reichweite (Wirkwelt, Bereich direkter Interaktion) von der Welt in meiner potentiellen Reichweite (institutionell vermitteltes gesellschaftliches Handeln) und dem Reich des symbolisch vermittelten Kosmos der Weltbilder (Soeffner 2004: 46ff.) schafft unterschiedliche Symbolisierungsanforderungen, die sich aber im Jetzt der Appräsentation treffen. Soeffner streicht die vermittelnde Funktion von Symbolen heraus. Sie sind sozusagen Grenzmarkierungen und verschaffen Zugang zu dem, wo ich jetzt (noch) nicht bin (a.a.O.: 51, 54). Gleichzeitig betont er eine paradoxe Struktur der Symbole, sie sind immer etwas, was sie nicht sind. Sie machen etwas vorstellig, vergegenwärtigen aber auch gleichzeitig dessen Abwesenheit. „Die Formulierung von Paradoxa ist Teil der symbolischen Arbeit, durch die wir unsere sichtbaren und unsichtbaren Konstruktionen der Welt mit einem Netz von Bedeutungen überziehen. Und es sind nicht zuletzt die Symbole, die diese Konstruktionen sichern. Diese Fähigkeit verdanken Symbole der Eigenschaft, Paradox und Ambivalenz zu betonen, aber zugleich auch auszuhalten und die Dissonanten des Gegensätzlichen in ästhetische Konsonanzen umzuformen.“ (a.a.O.: 55f.)

Noch deutlicher als der (1937) von Blumer so genannte ‚Symbolische Interaktionismus‘ (Blumer 1969) konzipiert Schütz in großer begrifflicher Tiefenschärfe die Herstellung einer gemeinsam geteilten Welt in einer Symbol- und Handlungstheorie. Dies heißt aber auch, dass über Symbolgebrauch Abgrenzung und Exklusion erfolgen können. Verschiedene Sprachen stehen für verschiedene kulturelle Welten, bevor Übersetzungen möglich oder notwendig werden. Bildsymbole gelten nur für die, die in ihnen etwas Bedeutsames erkennen, sie bleiben rätselhaft für den leeren Blick. Rituale vergegenwärtigen etwas nur, wenn ein praktischer Glaubenszusammenhang gestiftet wird, dem Zweifler und Außenstehenden vertiefen sie nur ihre Befremdung und die Gewissheit, nicht dazu zu gehören. Dabei funktioniert die jeweilige Bedeutungsherstellung nicht automatisch, indem sie sich in einen Hörer, Leser, Betrachter oder Teilnehmer eines Rituals einschreibt, sich ihm zeigt oder von ihm Besitz ergreift. Vielmehr ist der Sprechende, Hörende, Sehende, Handelnde jeweils selber gefordert, jetzt im Augenblick des Ausdrucks eine Agency zu verwirklichen, die diese Bedeutung auch erneut schafft und womöglich modifiziert. Die unterschiedlichen Symbolklassen (Sprache, Bild, Ritual) und ihr Gebrauch überschneiden sich, sie ergänzen oder unterlaufen sich, sie haben jeweils auch ihre nicht substituierbaren exklusiven Möglichkeiten. Sie wurzeln, das sollte die bisherige Darstellung vermittelt haben, *in der Einheit der Sinne, in der Wahrnehmung und der Expressivität der Leiblichkeit, im Sprechen und Hören, Sehen und Zeigen und einem auf die Welt und andere bezogenen Handeln.*

Die in jüngster Zeit zu beobachtende Konkurrenz zwischen Vertretern eines aktuellen ‚iconic turn‘ und denen des noch gestern angesagten ‚linguistic turn‘ bei der Beschreibung von Welt verliert ihre Legitimation, wenn man beide Betrachtungsstile als mögliche wissenschaftliche Reflexe auf den generellen Symbolgebrauch des Menschen in der Herstellung seiner Welt versteht. Von dieser Einheit her nach Differenzen der Symbolsysteme und ihren spezifischen Leistungsfähigkeiten zu fragen, läuft stärker auf Kooperation hinaus, die dann auch die auf den jeweiligen Symboltyp bezogenen, nicht ins andere ‚Medium‘ übersetzbaren Eigenleistungen herausarbeiten kann.

3. Das nicht nur bewegte Bild - zur Ikonik und der Frage der Repräsentation

3.1 Ikonik

Die scheinbare Natürlichkeit einer Audio-Videoaufnahme, die – so die einfache

Unterstellung – eine in Wirklichkeit abgelaufene Interaktionssituation ‚abbildet‘, ist hier auf ihre Künstlichkeit, die Differenz von interaktiv erlebter zu nachträglich gesehener und gehörter Wirklichkeit zu untersuchen. Nur so ist abzuschätzen, was wie getan wird, wenn ein Video aufgezeichnet und was wie verstanden wird, wenn ein Video betrachtet wird. Was also repräsentiert ein Video mit welchen Mitteln, und was geschieht, wenn man es sich ansieht oder wenn man es gemeinsam betrachtet und darüber redet?

Die bisherigen Ausführungen lassen ohne weiteres die Annahme zu, dass der Videogebrauch ein besonderer Symbolgebrauch ist, bei dem sich Sprechen, stehendes und bewegtes Bild und sinnvolles Handeln für die Betrachter zu Bedeutungen verdichten. Was die Überlagerung von verschiedenen Symbolsystemen leistet – also auch welche Grenzen sie hat –, lässt sich möglicherweise einfacher erkennen, wenn Reduktionen auf einzelne Symbolklassen vorgenommen werden. Das ist allerdings insofern problematisch, als tatsächliche Wahrnehmung, wie bereits dargelegt, sowohl im Vollzug in der Lebenswelt wie auch im Sonderfall der Betrachtung einer AV-Aufzeichnung als eine ganzheitliche Wahrnehmung, deren Komponenten im Vollzug nicht zu entwirren sind, konstituiert wird. Der Betrachter selber weiß nie, was gerade seine bestimmte Wahrnehmung steuert: Sind es die auslösenden Bildimpressionen, die gesehenen Handlungsabläufe, seine sozialen Erinnerungs- und Erwartungskonventionen, seine Aufschlüsselung sozialer Positionen der Akteure durch Anzeichen jeder Art, die Sprach- und Kommunikationswahrnehmungen, basale ästhetische Appräsentationen oder kulturelle Bedeutungskonventionen optischer Zeichen und ihre Verteilung in einem Bildfeld? Dennoch besteht die Hoffnung, den Prozess des Sehens wissenschaftlich besser zu verstehen, wenn man einzelne dieser Fragenkreise für sich behandelt, um sie dann später wieder zusammenzufügen.

Geht man, um die Dinge zu vereinfachen, in die lange Geschichte des Bildermachens und -sehens, steht am Anfang und für die meiste Zeit bis in das 20. Jahrhundert das *stehende Bild*. Schon der Begriff Bild impliziert, dass es sich um eine nicht bewegte Darstellung handelt, während es für die jüngste Möglichkeit bewegter Bildproduktion zwar einige spezifische Ausdrücke wie *Film* oder *Video* (im Englischen auch *movies*, *pictures*, *cinema*), aber keinen guten allgemeinen Begriff gibt.

Die *Zweidimensionalität* der Bildpräsentation ist bis heute die Regel. Das unterscheidet sie erheblich vom binokularen und damit räumlichen natürlichen Sehen in der Lebenswelt. Ebenso gravierend ist die Differenz der Zeitlichkeit des natürlichen Sehens zur

Bildwahrnehmung. Ist es dort die leibliche Ausführung und Wahrnehmung einer Bewegung in der Welt, bietet das stehende Bild im Unterschied zum alltagsweltlichen Sehen quasi zeitfrei einen Jetzt-Moment. Dieses ist, wie oben bereits ausgeführt, trotz seiner realitätsstiftenden Präsenz eine Abstraktion, die punktuelle Präsenz ist bedeutungslos ohne vorher und nachher. So steht die eingefrorene Bewegung doch für einen Zeitablauf, der mit gesehen sein will. Etwas Ähnliches gilt für die prinzipielle Stummheit des Bildes. Sie wird deshalb prinzipiell nicht als Problem empfunden, weil der Betrachter im Bildsujet das Bild zum Sprechen bringt. Das Bild kann so in die Nähe zu anderen Bildern und ihrer Symbolik, aber auch zu Texten, die das Bild weiter erschließen, gerückt werden. Sprechen *aus* dem Bild und *über* das Bild ist als Überwecheln der Symbolebene des Bildlichen ins Sprachliche möglich. Ebenso lässt sich ein Überwecheln in das Symbolsystem des Rituals erkennen, wenn etwa eine religiöse Ikone als Adoration der Gottheit fungiert oder – mündlicher – wenn ein Kunstliebhaber durch ein Kunstwerk ergriffen und angerührt wird und so eine Verbindung zu der höhersymbolischen Welt der Kunst entsteht.

Für all dieses ist der visuell identifizierbare Abbildcharakter keineswegs die Voraussetzung zur Konstitution eines Bildsujets. Die Grenzfälle eines getreu abbildenden Schnappschusses, bei dem ein Gesicht mit halb geöffneten Lidern und momentan verdrehten Augen als misslungene Darstellung der Person verworfen wird oder eine mit wenigen Bleistiftstrichen erzeugte Karikatur, die als geniale Repräsentation der Persönlichkeit wahrgenommen wird, belegen dies.

Wer ein Bild erzeugt, etwa malt, muss sich Gedanken machen, wie er welche Wirkung bei Betrachtern – die oft auch die Bilder bestellen und bezahlen sollen - erzeugen will. Die Sujets der Darstellung und ihre Funktion scheinen dabei die Formen zu steuern. So bringen etwa die religiösen Bilder des Mittelalters ein in christlichen Texten bereits vorliegendes Thema oder eine Narration zur Darstellung. Sie tun dies zugunsten einer *memoria*, die nicht nur etwas Vergangenes vergegenwärtigt, sondern eine aktuelle religiöse Erbauung anzielt und Trost oder Hoffnungs spenden soll. Die Betrachter werden so in das reale kirchliche Leben und die transzendente ‚Gemeinde der Gläubigen‘ integriert. Weiterhin stellt die Ausführung der Bilder jenseits der Sujets im engeren Sinne auch das Selbstverständnis, die Befürchtungen und Hoffnungen der mittelalterlichen Gesellschaft mit dar. Für die Bilder der Renaissance, die Sujets aus der griechisch-römischen Antike aufgreifen, gilt analog: Hier wird in den Sujets und der Art der Ausführung in der Anknüpfung an ein weit

zurückliegendes kulturelles Erbe eine mögliche aktuelle Selbstkonstruktion der Bildungselite mit ihren gesellschaftlichen Vorstellungen vergegenwärtigt. Die Kunstinhalte sollen sie nachhaltig prägen, den Reichtum des eigenen kulturellen Lebens darstellen und gleichzeitig das alltägliche Ambiente erhöhen und schmücken.

Wenn es das ist, was diese Bilder ‚wollen‘ (Mitchell 2005), ist die Frage, *wie* können sie es als Bilder, also mit bildlichen Mitteln, erreichen. Die Betrachtung hochmittelalterlicher Fresken oder von Bildern aus der Renaissance ist dabei besonders aufschlussreich, weil sie uns Gegenwartsmenschen vom Sujet und Stil meist völlig fremd sind, also etwas vergegenwärtigen, das wir nicht verstehen. Sie ziehen uns gleichwohl, wenn wir uns ihnen aussetzen, beim Betrachten in einen Sog, den aufgeworfenen Rätseln nachzugehen, also Hermeneutik zu betreiben. Es gibt in diesem Feld seit Anfang des zwanzigsten Jahrhunderts eine differenzierte kunstgeschichtliche, ikonologische Diskussion, die sich intensiv über dieses Problem erstreckt (Boehm 2006, 2008; Imdahl 1980, 1994; Panofsky 1990, 1992, 1997, 2006; Raphael 1989b, 1989a); hier kann davon nur so viel dargestellt werden, dass eine vertiefte Beschäftigung mit diesen Traditionen auch in sozialwissenschaftlichen Kontexten der Bild- und Videoanalyse angeregt wird.

Die Frage lautet also zum einen, mit welchen *formalen Mitteln* wird im zweidimensionalen ‚Raum‘ ein statisches Bildobjekt so strukturiert, dass es sich als Bildsujet von etwas anderem im Bildobjekt unterscheidet und Bedeutung annehmen kann. Zum ist zu fragen, *wie diese Bedeutung als Re-Präsentation oder Kreation kultureller Inhalte* entsteht.

Die erste Teilfrage ist von besonderem Interesse, weil es dabei um die formale Bauweise von Bildern geht, die gerade als Eigenschaft des Bildobjekts nicht explizit wahrgenommen wird, jedoch gleichwohl für die Impression des Bildes verantwortlich zu sein scheint. Es ist, überspitzt gesagt, die manipulative Ebene des Bildes, die dem Betrachter nicht verfügbar ist, sondern optisch einfach zu wirken scheint. Sie formiert das Sehen durch schiefe Transparenz und kann nur analytisch herausgearbeitet werden. Sie entzieht sich weitgehend in der unbefragten Haltung des alltäglichen Blicks einem Wissen um seine Entstehung und einer reflexiven Bildbetrachtungskompetenz, bei der auch das Bildobjekt Thema werden könnte. Es ist fraglich, ob diese formalen Bauprinzipien den Experten der Bildproduktion wirklich voll kalkuliert zur Verfügung stehen, auch wenn sie sie nutzen oder - wie im Falle der modernen Kunst - damit spielen. Es spricht viel dafür, dass sie weitgehend einfach im experimentierenden Produzieren gefunden werden, indem gegebene

Traditionen variiert werden. Sie zeigen sich also im Werk, bewähren sich für die Betrachter und können erst in der nachträglichen Analyse oder einer fast minutiös den Entstehungsprozess nachschaffenden ‚Bildbeschreibung‘ (Raphael 1989a) des Bildobjekts nachhaltig als formale Regeln fixiert werden.

Auf eine bildbezogene Darstellung dieser Formalia, die nur in demonstrierenden Einzelfallrekonstruktionen am Bild selber möglich ist, muss hier verzichtet werden. Es sollen nur die wichtigsten Hinsichten genannt werden. Abstrakt gesprochen geht es um die Darstellung einer Reihe von verschachtelten und aufeinander bezogenen bildlichen Unterscheidungen und Kontraste. Es handelt sich um ein Zeigen aus dem, was es nicht zeigt, um eine *Deixis der Negation*, ein Zeigen, das ein Verbergen voraussetzt (Boehm 2008: 67ff.).

Die erste wichtige Reduktion, ein Weglassen und Pointieren zugleich, liegt in der Festlegung des Bildausschnittes, des ‚Fensters‘ oder der Projektionsfläche des Blicks. Was außerhalb liegt, der weitere Horizont oder die anderen Geschichten, wird nicht (mehr) gezeigt, auch wenn es durchaus als wichtig und vorhanden unterstellt wird. Anders als beim schweifenden Blick auf eine Szene in der Welt, in die sich der Betrachter auch leiblich hineinbegeben kann, wird hier eine willkürliche Blickbegrenzung eingezogen, die einerseits eine erste erzwungene Fokussierung des Sehens in Richtung auf ‚etwas‘ ist. Andererseits ist sie Anlass einer mitlaufenden Appräsentation dessen, was auch da sein könnte. Grenzt sich der Bildausschnitt aus einer weiten Landschaft, einem gemauerten Innenraum, einem aus Holz geschnitzten gotischen Bogen oder einer ebenmäßigen Farb- oder Goldfläche aus – stets ist dem Betrachter etwas mitgegeben, was die Bildimpression mit bedingt. Unterstützt wird diese erste Entscheidung durch die Materialität und Vorbereitung des Trägerobjekts, das ebenfalls einer Vorentscheidung oder Konventionen der physischen Bildauffassung folgt. Male ich auf einen Stein, eine Wand, ein Brett, ein Leinen, ein Papier – immer wird damit auch die erste Auswahl des Bildausschnitts gefestigt.

Diese erste Unterscheidung des sichtbaren und unsichtbaren Bildanteils setzt sich im Prinzip in allen identifizierbaren Bildelementen fort und begründet, „dass jedes Bild seine Bestimmungskraft aus der Liaison mit dem Unbestimmten zieht“. (a.a.O.: 49)

Im ‚Durchgang‘ durch das Bildobjekt fällt als nächstes die Differenz auf zwischen figürlichen Bildelementen auf, die sich vor dem - vielleicht schon durch den Bildausschnitt angelegten - Hintergrund erheben. Es gibt ein Spannungsmoment zwischen den Formen und Figuren und dem Hintergrund. Die Figuren und Konfigurationen werden gesehen, sie ver-

decken den vielleicht nur angedeuteten Hintergrund, schneiden ihn aber nicht aus oder durchlöchern ihn, sondern beziehen Kraft aus ihm. Der oft diffuse oder unbestimmte Hintergrund sorgt als wahrgenommenes Raumkontinuum oder offener Horizont dafür, dass sich die Konfigurationen herausheben und einen ‚Ort‘ haben. Der Hintergrund wird in der Fläche des Bildes im Laufe der Kunstgeschichte des späteren Mittelalters allmählich gestaltet als ein scheinbar dreidimensionaler Raum, in dem sich die Formen, die das Sujet herstellen, ‚bewegen‘.

Es geht also einerseits um die Perspektivität des Bildes und die Anordnung der Figuren gegeneinander. Während letzteres - vor allem im Falle ‚gegenständlicher‘ Malerei - seine Ordnung weitgehend aus den Ordnungen der Lebenswelt, den leiblichen Interaktionen und jeweils gesellschaftlichen Konventionen bezieht, ist die Entwicklung der räumlichen Perspektivität im Bild ein voraussetzungsreiches Produkt der Malerei der italienischen Renaissance und der niederländischen *ars nova* im 15. Jahrhundert. Die Übersetzung der aus dem binokularen Sehen geläufigen dreidimensionalen Raumwahrnehmung in eine zweidimensionale Reduktion erforderte eine genaue Auseinandersetzung mit optischen Projektionen, Blickrichtungen, festen und variablen Standpunkten des Sehens. War erst einmal die Umsetzung der vom aufrecht stehenden Menschen als ‚normale‘ Sichtperspektive empfundene Räumlichkeit bildlich über einen in der unteren Bildhälfte liegenden Fluchtpunkt, also ein perspektivischer Realismus, gelungen, konnten anschließend Variationen von Übereckperspektiven, Vogel-, Auf-, und Untersichtsperspektiven oder gar von mehrfachen Perspektiven in einem Bild entwickelt werden. Sie erlauben dem Betrachter eine je spezifische Raumwahrnehmung, die er vielleicht in seiner lebensweltlichen Wirklichkeit mangels Zugänglichkeit nicht einmal erleben kann. Die moderne Kunst (wie auch der Cartoon oder Zeichentrickfilm) hat schließlich die perspektivische Rationalisierung von Bildern wieder aufgelöst, auch wenn sie über Kameras und deren massenhaften Gebrauch im Alltag einen zweifelhaften Siegeszug mit dem impliziten Versprechen scheinbar getreuer Abbildung angetreten hat.

Analysiert man die *Bildfläche als Ebene*, kommen schließlich die Konfiguration von Flächen und Linien ins Spiel. Es ist diese planimetrische Betrachtung, die scheinbar schnell zur Dekomposition und damit dem besseren Verständnis des formalen Funktionierens eines Bildes beiträgt. Ein auf der Basis stehendes Dreieck im Bildmittelfeld signalisierende ruhende Wichtigkeit, von links nach rechts (also in europäischer Leserichtung)

aufsteigende oder absteigende Linien werden mit Agitiertheit oder Beruhigung verbunden, verschränkte oder gegeneinander abgegrenzte Flächen, Proportionen (goldener Schnitt), geschwungene oder zerrissene Linien vermitteln offenbar für alle (eingeübten) Betrachter vergleichbare emotionale Bildwerte wie etwa Harmonie, Konflikt, Bewegtheit. Diese planimetrischen Merkmale tragen insgesamt zur Konstitution des Sujets bei. Schließlich wäre noch über die Farbgebung, die strukturelle Gestaltung von Strich und Fläche (mit den jeweiligen Werkzeug- und Materialspuren) wie die Modulation des Lichts als Gegebenheiten des Bildobjekts zu sprechen, die jeweils Gefühlswerte vermitteln und zum Gesamteindruck beitragen.

Alle diese Merkmale des Bildobjekts kann man als ‚Form‘ begreifen, die zunächst ein Blickzentrum schaffen soll. Das Auge soll auf ihm ruhen, der Blick um es kreisen und verdichtet wieder zum Blickzentrum zurückkehren. Die Zentrierung des Blicks im formalen Bildaufbau löst Appräsentationen aus, ruft Themen auf und birgt somit die eigentlichen Bildinhalte. Die formalen Strukturen lassen sich jeweils getrennt beschreiben, führen aber nicht direkt und zwangsweise zur Bedeutung bzw. dem Verstehen des Bildes, obgleich sie im Zusammenwirken eine einheitliche Bildimpression leisten. Sie ist jedoch erst die Voraussetzung für die Re-Präsentation oder auch Schaffung eines kulturellen Bildinhalts. In der dreistufigen Arbeitsweise Panofskys, des Begründers einer modernen Ikonologie, bilden die formalen Charakteristika die erste Ebene der *vorikonographischen Beschreibung* (Panofsky 2006: 29, 56).

Dem folgen als weitere notwendige und noch voraussetzungsreichere Operationen des Verstehens die *ikonographische Analyse* und die *ikonologische Interpretation*. Selbst wenn es sich beim alltäglichen Verstehen eines Bildes, das nicht als Kunstwerk einzustufen ist, um einfachere Verstehensprozesse handeln sollte, erscheint die hermeneutische Stufenfolge ebenfalls geeignet, zur erklären, wie man Bilder versteht.

Nach der formalen Beschreibung des Bildobjekts geht es nun bei der nächsten Erkenntnisoperation der *ikonographischen Analyse* darum, sich dem Bildsujet dadurch zu nähern, dass die Figuren und ihre Konstellationen als kulturell ‚bekannt‘ identifiziert werden. Dies bezieht sich auf den Kulturkreis der ursprünglichen Bildproduktion und deren Gebrauch. Die Bedeutung des Bildes, also die Erkenntnis des Bildsujets, wird hier etwa auf der Ebene erreicht, auf der sie sich einem Zeitgenossen und Angehörigen des sozialen Milieus, in dem das Bild ursprünglich seinen Ort hatte, unmittelbar hergestellt haben mag. Es geht ein-

mal um das Erkennen von Themen oder Narrationen aus den Diskursen der Gesellschaft, auf die sich das Bildsujet bezieht. Ist die Kultur historisch weit entfernt, stellt sich von selbst keine Bildbedeutung her, und das dazu nötige Wissen muss erarbeitet werden. Dazu sind für einen heutigen Betrachter umfangreiche historische Kenntnisse von Ereignissen und Texten der betreffenden Kultur und ihrer Visualisierungen erforderlich. Zum anderen sind hier auch stil- und typengeschichtliche Kenntnisse notwendig, sonst kann gar nichts beschrieben werden (a.a.O.: 13), weil der Betrachter die Darstellungskonventionen einer bestimmten Kunstgattung oder –epoche nicht auf ein Sujet hin entschlüsseln kann. In der *ikonologischen Interpretation*, der letzten Erkenntnisebene, die die Bedeutung in den allgemeinsten Rahmen symbolischer Werte stellt, verlangt Panofsky neben den beiden ersten Operationen eine ‚synthetische Intuition‘, die ihrerseits durch kunst- und kulturgeschichtliches Wissen anzureichern ist. Ziel ist es, die allgemeinste Bedeutung des Bildes als Symbol und Ausdruck menschlichen Geistes zu fassen. Panofsky legt Wert darauf, die drei Analyseaspekte nicht als getrennte Operationen nach einander, zusammensetzen, sondern er versteht den kundigen Analyseprozess als ein untrennbares Ganzes, das sich in der Einheit des Bildes selbst zentriert (a.a.O.: 56).

Obgleich die ikonologische Analyse im Objektbereich der Kunst entwickelt wurde, ist sie für jede Art von Bildern und ihre sozialen Kontexte adaptierbar. Man kann allerdings fragen, ob diese Methoden auch „auf die Analyse von Bewegtbildern bzw. von Bilderströmen, die keine Einzelbildbasis mehr besitzen“ (Wiedemann 2005: 365) anwendbar sind. Trotz dieses ernst zu nehmenden Fragezeichens können für die Analyse des *bewegten Bildes*, also einer AV-Aufzeichnung, die genannten Kriterien genutzt werden, soweit simultane, koprärente Bildphänomene untersucht werden. Es geht technisch gesprochen um die formale Analyse und Bedeutungsfindung von ‚Standbildern‘, einzelnen ‚frames‘ oder ‚Fotogrammen‘ (Bohnsack 2009: 151ff.; Raab 2008: 165ff.). Es kommt allerdings für die Videoanalyse zusätzlich zur Bildanalyse eine zentrale Ebene der Bedeutungskonstitution hinzu. Es ist allgemein die Darstellung von Zeitlichkeit oder besser *Sequentialität* (Raab 2008: 109f.). Sie lässt sich weiter differenzieren in die Ereignis- und Interaktionsdarstellung eines Ablaufs und in die Sequentialität der Kommunikation, also der Sprech-Hörereignisse. In beiden Teilaspekten gleicht die AV-Aufzeichnung der Sequentialität der Handlungen, Interaktionen und Kommunikationen in der Lebenswelt. Diese sind deshalb auch filmisch schwer abkürzbar und auf keinen Fall umkehrbar; eine Handlung muss in

ihrem Ablauf gezeigt werden, wenn sie vom Betrachter verstanden werden soll, eine Kommunikation muss als Kommunikation in der zeitlichen Abfolge gezeigt und zu Gehör gebracht werden, sonst ist keine Bedeutung erkennbar.

Es ist gerade dieser Aspekt der Sequentialität im visuellen Medium des Filmischen, der es für sozialwissenschaftliche Erforschung von Handlungs- und Interaktionsvorgängen so interessant macht. Das bewegte Bild von Abläufen und in der Bewegungsperspektive, die durch die Eigenbewegung der Kamera, durch Fahrten, Schwenks und Brennweitenwechsel entsteht, weckt die Erwartung, dass es medial vermittelt etwas zeigen kann, was der Sequentialität von konstitutiven Interaktionen ziemlich genau gerecht wird. Das bewegte Bild kann offenbar etwas als visuell verfolgbaren Prozess vergegenwärtigen, was auch in der sozialen Realität nur als Prozess vorkommt, bei dem sich die Akteure leiblich im Raum bewegen und sich ebenfalls visueller und sprachlicher Mittel der Interaktion bedienen. Vollends gesteigert ist diese Erwartung bei der elektronischen Übermittlung von Bildern in Echtzeit, wo Betrachtungszeit und Ereigniszeit in eins fallen. Diese Life-Sendungen oder Web-Cams vermitteln dem Betrachter den ‚Kitzel‘ des Dabeiseins bei einer noch nicht abgeschlossenen Interaktion, einem Gespräch oder einer Handlung.

Diese Bemerkung macht andererseits deutlich, dass es bei der Darstellung von Abläufen generell auch ein *Kondensierungsproblem* gibt, denn wie lange kann ein Film von etwas angeschaut werden, ohne dass der Betrachter ermüdet oder aussteigt? Es eignen sich offenbar nur kürzere Prozesse oder face-to-face Interaktionen, die sich über Minuten, maximal eins bis zwei Stunden erstrecken für die direkte zeitsynchrone ‚Abbildung‘. Es musste also ein spezifisch filmisches Konstruktionsmittel für dieses Kondensierungsproblem, das bereits in der sprachlichen Darstellung von Ereignissen in der Form der Narration im Symbolsystem Sprache gelöst ist, gefunden werden. Analog zur sprachlichen Narration und komplexeren literarischen Formen wie der Novelle oder des Romans wird eine filmische *Erzählweise* geschaffen, bei der aus dem gefilmten Rohmaterial (footage) kurze Handlungssegmente ausgeschnitten (oder gleich nur solche als ‚Einstellungen‘ produziert werden) und dann so zusammengesetzt werden, dass solche Ablaufstrukturen inkl. gewünschter ‚Vor- oder Rückblenden‘ entstehen, die der gewünschten Bedeutungsvermittlung entsprechen. Es ist also nicht nur die Kondensierung sondern auch die Möglichkeit, mehrere Handlungsstränge oder Reflexionen zu verweben, die die Leistungsfähigkeit eines geschnittenen Films, der sichtbar mehrere Erlebnisebenen und Perspektiven

zusammenführen kann, ausmacht. „Der Schnitt und die Montage bestimmen die Organisation von Raum und Zeit und sind die zentralen Elemente des Zeigens und der Narration, damit visuell-kognitive Interpretationsanleitungen, denn durch sie werden der Blick, das Verstehen und das ‚Wissen‘ des Betrachters gerahmt und gelenkt.“ (Raab 2008: 109) Filmschnitt und Montage entwickeln sich zu so etwas wie einer eigenen „Filmsprache“ (Soeffner/ Raab 1998; Wiedemann 2005: 368), die dem Zeigen eine visuelle Syntax zu geben versucht. Die wichtige Rolle der Filmmusik für die emotionale Bedeutungserzeugung kann hier nicht behandelt werden, sie belegt jedoch einmal mehr die Einheit der Sinne bei jedem Wahrnehmungsakt.

War es schon beim einzelnen Bild schwierig, die Frage zu beantworten, wie die Bedeutung in der Betrachtung erreicht wird, muss die Antwort beim bewegten Bild somit noch komplexer ausfallen. Zwar ist durch den dargestellten Ablauf von etwas selber eine prozessinterne Organisation (Sequenz, zeitlicher und logischer Bezug von Teilelementen der Bewegung und der Handlung) fokussierend und sinnstiftend, doch wofür die Handlung steht und was sie bedeutet, hängt von weiteren Rahmungen ab, die der Fortgang der visuellen Darstellung – wenn überhaupt – nur teilweise liefern kann. *Wofür* steht also das bewegte Bild?

3.2 Repräsentation – von was?

Der Alltagsverstand geht davon aus, dass ein Bild etwas abbildet. Es scheint also zunächst banal, dass das (bewegte) Bild das zeigt, für das es steht, also etwas repräsentiert. Es kann eine Sache, eine Person, eine Handlung, eine Geschichte etc. sein, deren Darstellung *und* Betrachtung etwas vergegenwärtigen und damit eine bestimmte Bedeutung vermitteln sollen. Wo liegt das Problem? Ein Teil des Problems liegt darin, dass schon die Frage, inwieweit Wirklichkeit repräsentiert werde *oder* konstruiert werde, ein Problem ist. Sagt das Bild mehr über die *Wirklichkeit* oder über den *Bildproduzenten* oder über den *Bildbetrachter*, der ein Bild auslegt indem er es mit seinen Augen sieht und seinen Worten beschreibt, aus? Man wagt sich angesichts des selbstverständlichen persönlichen und massenmedialen Bildergebrauchs nicht an die Reflexion oder Unterminierung der Vorstellung einer schlichten Repräsentation oder Abbildung, weil dies der Annahme, man könne Wirklichkeit einfach durch Bilder zeigen und sehen, entgegenläuft. Gleichzeitig weiß jeder und nicht nur der professionelle Produzent von Bildern und Videofilmen, dass

allein der notwendige Ausschnitt oder verfügbare Szenen nicht immer das zeigen (können), worum es geht. Sie verfälschen, nehmen Partei oder können je nach Standpunkt des Betrachters sogar als Unwahrheit wahrgenommen werden. Die vorangegangene Analyse zur Wahrnehmung (s. 2.1) hat bereits verdeutlicht, dass die Entstehung und Perzeption von Bildinhalten, also die Konstitution von Sujets alles andere als ein Abbildungsvorgang ist. Das Sujet einer Bildwahrnehmung, so konnte man von Husserl lernen, wird nicht abbildhaft wie auf einer Projektionsleinwand durch die Optik des Auges im Inneren des Betrachters erzeugt, sondern es entsteht im komplexen Zusammenspiel von Bewusstseinsleistungen, Erinnerungen, Erwartungen, in die auch die gesellschaftlichen Diskurse eingehen, und den visuellen Bildangeboten, die ihrerseits Konstruktionsregeln folgen, die nicht einfach Wirklichkeit abbilden wollen, sondern bestimmte soziale Zwecke (z.B. Werbung, politische Information, Unterhaltung, Schaffung angenehmer Erinnerungen, Diagnostik in therapeutischer Absicht, Schaffung sozialer Netzwerke, das Beilegen oder Schüren von Konflikten etc.) verfolgen.

Auch ohne einen Rückgriff auf diese wahrnehmungstheoretischen Traditionen hat sich durch die Praxis der Bildproduktionen und des Bildgebrauchs die naive Selbstverständlichkeit von Abbildannahmen aufgeweicht. Vor allem auf der Seite der professionellen Bildproduzenten ist das Bewusstsein, dass Bilder *eher wirksame Konstruktionen als einfache Abbildungen* sind, voll entwickelt. „Bilder funktionieren nicht wie starre Spiegel, die eine stets vorauszusetzende Realität wiederholen, sie sind keine Doubles.“ (Boehm 2006: 332) Auch der Gebrauch der Bilder, seine Einbettung in bestimmte Argumentationsvorgänge und mediale Produkttypen (von der wissenschaftlichen Aufzeichnung in bildgebenden Verfahren bis zur Werbung) hat gezeigt, dass eine Abbildvorstellung nicht ausreicht, oft falsche Erwartungen weckt (oder bedient) (Mitchell 1994: 418-425).

Unter dem Titel ‚*Krise der Repräsentation*‘ wirkt die Thematik vor allem aus der wissenschaftstheoretischen Selbstreflexion der Ethnologie in die Sozialwissenschaften hinein (Berg/ Fuchs 1993; Hirschauer/ Amman 1997). Die Problematisierung der eigenen Beschreibungsstandpunkte hat in der Ethnologie zur Erkenntnis geführt, dass jede wissenschaftliche Beschreibung (inkl. der Erzeugung von Bildern) als Konstruktion eines fremdkulturellen Beobachters die beschriebene Welt nicht einfach erklärt oder hermeneutisch versteht, sondern sie sich gerade dadurch aus der Sichtweise (westlicher) Wissenschaftsstandards aneignet, sie mithin aus der Sicht der beobachteten Kultur verfremdet, ‚ver-

ändert'. Das traditionelle, aufgeklärte Wissenschaftsverständnis als Erforschung der Wahrheit durch einen objektiven und interessefreien Beobachter wurde in diesem Methodendiskurs nachhaltig erschüttert. Auch das hermeneutische Verfahren, das sich eigentlich als Gegenmethodologie zu dem dualistischen Modell positivistischer Fremdbeobachtung versteht, erweist sich als ambig, weil die Frage aufkommt, „wie weit damit nur die Praxis Anderer objektiviert und damit auf Distanz gehalten oder wie weit die eigene Praxis und Forschungsbeziehung einbezogen wird“. (Fuchs/ Berg 1993: 20)

Die verzweigte Diskussion, die nicht nur auf der Ebene wissenschaftstheoretischer Selbstreflexion stattfindet, sondern auch dialogische Forschungspraktiken, Modelle der indigenen Selbstbeobachtung und -darstellung, sowie dialogische, poetische und imaginative Textgenres in der Forschung erprobt, kann hier nicht verfolgt werden. Allgemein ist für unseren Zusammenhang festzuhalten, dass sich sozialwissenschaftliche Forschung auch gerade in der Nutzung von bildgebenden Verfahren jeder Art von einer impliziten Abbildtheorie und der Annahme realistischer Repräsentation durch Bilddokumente methodologisch und forschungspraktisch verabschieden muss. Selbst die scheinbar harmlosesten, laienhaften und ungeschnittenen dokumentarischen Videofilme von sozialen Praxen, widerspiegeln nicht einfach die Wirklichkeit sozialen Handelns, sondern sie sind selber Konstruktionen der Bildproduktion durch die Forschung. Ihre Tragfähigkeit als Rekonstruktionen sozialer Praxis muss erst im methodisch kontrollierten selbstkritischen Verfahren der Bild- und Sprachauslegung hergestellt werden. Um dies – etwa in hermeneutischen Interpretationen (s. Abschnitt 4) – zu erreichen, ist nicht nur zu fragen, welcher sozialer Konstruktionen sich die *Akteure* in der seh- und hörbaren Interaktion bedienen, sondern es ist auch zu fragen, wie der wissenschaftliche *Beobachter* mit den Akteuren interagiert und wie er vor allem in seiner Forschungsinteraktion im ‚Alltag der Forschung‘ den Sinn der Forschung konstruiert. Wie gehen die visuellen Beobachtungsinstrumente und ihre Auslegung in die Konstruktion der Beobachtung ein? Es ist mithin eine Beobachtungsebene zweiter Ordnung einzunehmen, die nicht die Sozialwelt, sondern die Beobachtung der Sozialwelt durch Wissenschaftler beobachtet.

Elisabeth Mohn hat in beispielhafter Weise am Thema des Dokumentarfilms die Probleme und Möglichkeiten des „starken“ (intuitive oder automatisierte Aufnahmepraxis) und des „schwachen“ (starke Autorschaft und Interpretation) Dokumentierens ausgelotet (Mohn 2002). Sie hat nachdrücklich erneut belegt, dass jede Repräsentation auch Fiktion ist und

somit stets die Frage mitlaufen muss, welchen Standpunkt und welches Interesse die Dokumentaristen mit einbringen. Einfaches Zeigen, das angeblich interpretationsfrei sei, ist ebenso problematisch, wie stark interpretierte Darstellung, die im Prozess der Aneignung des beobachteten Weltausschnitts diesen verändert, verfremdet und kommentiert. Mohn plädiert in der Herstellung des Filmdokuments für „dichtes Zeigen“ oder „paradoxes Dokumentieren“ (a.a.O.: 125ff.) , realistische und imaginative Filmsequenzen und Kommentare werden montiert und aus ihrem Oszillieren empfängt der Betrachter Anregungen, selbst weiter zu interpretieren, um welche Wirklichkeit es geht. In ihrem eigenen wissenschaftlichen *Text* präsentiert sie imaginierte Dialoge zur Reflexion und Darstellung paradoxer Aussagen. Bei der Beobachtung der Beobachtung, also der Analyse wissenschaftlicher Interpretationspraxis (dem „Alltag der Wissenschaft“ a.a.O.: 159ff.) greift sie auf Mannheims und Garfinkels Konzept der *dokumentarischen Methode der Interpretation* zurück. Dieses allgemeine Verstehenskonzept folgt der Annahme, dass sich sinnstiftende Interaktion immer im Wechselbezug von partiellem Zeigen (dokumentieren von etwas als etwas) und Interpretieren realisiert. Der wahrgenommene Sinn entsteht durch das Zusammenspiel von Eindruck und Deutung als Bewusstseinsleistungen, die wiederum durch nächste Rede- und Interaktionszüge für die Interaktionspartner ausgedrückt und wahrnehmbar gemacht werden und so fort. So bilden sich Interaktionssequenzen, bei denen im Wechselspiel von Eindruck und Ausdruck Sinn produziert wird. Diese Formalstruktur des Verstehens und der Sinnbildung, die sich in die oben dargelegte Husserlsche Wahrnehmungsanalyse einfügt, ist die alltägliche ‚Methode‘ der Bedeutungsvermittlung und –schaffung. Sie ist auch im ‚Alltag der Wissenschaft‘ in Geltung, da auch hier die Bedeutung erzeugenden Interaktionen Interaktionen sind. Damit kann trotz aller Bemühungen auch die wissenschaftliche Wissensbildung keinen prinzipiellen Sonderstatus einer interpretations- oder interaktionsneutralen Erzeugung beanspruchen, wie Mohn zu Recht gegenüber einer wissenschaftlichen Methodisierung und höherstufiger Hierarchie des wissenschaftlichen Verstehens über dem Alltagswissen vertritt (a.a.O.: 162ff.) – ein Anspruch den etwa Bohnsack auch gegenüber wissenssoziologisch hermeneutischen Ansätzen, die er „Common-Sense Theorien“ nennt, geltend macht (Bohnsack 2009: 134f.). Mohn schlägt ein „instrumentelles dokumentarisches Interpretieren“ vor und belegt das Verfahren an einer reflexiven Interpretationspraxis einer Bielefelder Arbeitsgruppe zur Visuellen Soziologie Ende der 1990er Jahre. Eine Selbstanweisung lautet: „Betrachte Auf-

zeichnungen nicht allein als Dokumente *von* Situationen, sondern als Dokumente *für* soziologische Entdeckungen, [...] überführe deine Materialien in evokative Verfahren, die Imaginationen hervorrufen.“ (a.a.O.: 196f.) Als forschungspraktisches Fazit ihrer Kritik an realistischen und interpretativen Programmatiken empfiehlt Mohn einen häufigen Verfahrenswechsel zwischen starkem Dokumentieren und Anti-Dokumentieren in Abhängigkeit von Forschungsphasen. Auch die zeitweilige (keineswegs permanente) Selbstreflexion als Beschreibung der Praxis gelingender Verständigung im Forschungsprozess drängt die einseitigenden Kosten orthodoxer Praktiken zurück. Das Spielen mit verschiedenen dokumentierenden und interpretativen Verfahren, ein häufiger Registerwechsel, erlaubt nach Mohn sowohl *probeweisen Realismus* als auch das *Ausprobieren von Interpretationen* im Rückbezug auf Daten aus dem Forschungsfeld und anhand von (visuellen) Daten aus dem Interaktionsfeld der sich verständigenden Wissenschaftler.

4. Videoanalyse als Interaktionsanalyse – forschungspraktische Verfahren

Die letzten Überlegungen haben uns bereits dicht an Fragen der Forschungspraxis geführt. Wenn jede Forschung immer auch als Rekonstruktion sozialer Wirklichkeit eine *Konstruktion* ist, besteht die Aufgabe nicht einfach darin, die Dokumentationen und Abbildungsinstrumente wie AV-Aufzeichnungen zu schärfen, sinnvoll einzusetzen und interpretierend in wissenschaftliche Texte umzuschreiben. Vielmehr sind die Forschungshandlungen nach zwei Richtungen zu gestalten. Einmal zielen sie auf die Praxen der dokumentierten Alltagswelt, zum anderen gehen sie selbstreflexiv auf die Praxen der wissenschaftlichen Interpretation ein. Es sind mithin zwei verschränkte Interaktionsszenarien und deren Praktiken zu berücksichtigen: Das Feld des Objektbereichs, also *alltagsweltliche* (oder professionelle) Interaktionen, die als Untersuchungsbereich ausgewählt werden *und* das Szenario der eigenen Beobachtungsaktivität. Als Objektbereich kommt alles infrage, was aus theoretischen oder praktischen Gründen erforscht werden soll. Im Bereich der sozialen Arbeit wären dies etwa die Interaktion in einer Gruppe von Jugendlichen, eine Familieninteraktion, ein Beratungsgespräch, eine Arzt-Patienten-Interaktion, eine sozial-therapeutische Session, etc. Hauptfrage ist immer, wie in diesen Interaktionen gesprächslokal der jeweilige Handlungstyp gefüllt oder auch problematisch gestaltet wird. Das Feld der *wissenschaftlichen* Interaktion, die sich mit diesen Materialien in einem Forschungsdiskurs auseinandersetzt leistet im Vergleich zu den autochthonen Sinnproduktionen im Objektfeld

zugespitzte, erweiterte oder alternative Sinnproduktionen, deren Erzeugungsregeln ebenfalls zu untersuchen sind. Es muss also auf der Suche nach Forschungsverfahren ein Ansatz gefunden werden, der diesen Sinnstiftungs- und Beobachtungsprozessen in *beiden* Feldern gerecht wird.

In beiden Bereichen handelt es sich um *Interaktionen und Bedeutungskonstitution*, vereinfacht gesagt Sprechen in Interaktionen. Das Zusammenspiel von Handeln und Wissensbildung verläuft in beiden Bereichen nach den gleichen Strukturierungsprinzipien. Daher sind auch die gleichen Beobachtungsverfahren und Rekonstruktionstechniken sowohl für den Objektbereich wie für die Beobachtung der Forschungshandlungen selbst einsetzbar. Will man es schematisch in einer Beobachtungsterminologie (etwa Luhmann 1990: 68ff.; 1995: 92-164; 1997: 36-59) ausdrücken, nutzen die *Akteure* in der Sozialwelt das gleiche Strukturierungsprinzip, um Beobachtungen (einschließlich Selbstbeobachtungen), Handlungen und Sinnerzeugung zu verweben, wie die *Forscher*, die solche Prozesse sozialer Strukturierung untersuchen (Beobachtungsebene erster Ordnung), wie auch die Forscher, die das Forschungshandeln auf der Beobachtungsebene zweiter Ordnung untersuchen. Den gleichen Gedanken verfolgen bei allen sonstigen gravierenden Unterschieden die ethnomethodologischen und wissenssoziologischen Forschungsvarianten, die davon ausgehen, dass konkrete sinnkonstituierende Praktiken (,situated practices') Interaktionen und Sprechen miteinander verweben, und somit Handlungsziele und die Herstellung einer gemeinsamen Welt (sicht) verwirklicht werden (Bergmann 1994; Drew/ Heritage 1992; Goodwin 1986; Goodwin/ Goodwin 1996; Goodwin 1995; Schegloff 1998, 2006). Sozialwissenschaftliche Forschung begibt sich daher sozusagen an den ,Quellpunkt' gesellschaftlicher Konstitution, der Ausbildung von Sozialität bei den Akteuren und von gesellschaftlichen Einrichtungen, wenn sie die Kommunikation in konkreten Interaktionssituationen betrachtet und nach den Konstruktionsvorgängen von sinnvollen Handlungen fragt. Was auch immer einer solchen Situation vorausgegangen sein mag oder ihr folgt, es muss möglich sein, in der Betrachtung der ausgewählten Interaktion die nachhaltigen Bauprinzipien der Kommunikation zu erkennen, die erwartbar auch künftig gelten und daher einen gewissen Allgemeingrad haben. Es geht darum zu erkennen und zu sagen, worum es den Akteuren geht, und im Bezug auf die beobachteten Handlungen Unterscheidungen zu treffen wie: Sinn/Unsinn; gelingende/nicht gelingende Kooperation; Typus der Handlung, etc.

Welche Forschungsmittel stehen zur Verfügung? Wenn diese Aufgabenbeschreibung für

alle drei genannten Handlungsbereiche der Alltagswelt (inklusive der Berufs- und Arbeitswelt), der wissenschaftlichen Beobachtung und der Beobachtung der Beobachtung zutrifft, müssen die Mittel *prinzipiell ähnlich* und auch jedem interaktions- und kommunikationsfähigen Mitglied der Gesellschaft verfügbar sein. Dies betrifft vor allem die Formalstruktur der Bedeutungsbildung entlang von Handlungszügen in zeitlich nicht umkehrbaren Sequenzen.

Die Mittel können andererseits ‚graduell‘ verschieden sein, was die jeweiligen Handlungserfordernisse und damit verbundene Entschleunigung und Verlangsamung der Sinnbildung angeht. Handeln und Problemlösen in Alltagssituationen unterliegt einem anderen Erwartungsdruck und einem kleineren Zeitfenster als die wissenschaftliche Beobachtung desselben Handelns; wissenschaftliche Interaktion in einer Arbeitsgruppe unterliegt wieder einem höheren Entscheidungs- und Produktionsdruck als deren Reflexion und Beobachtung in einer wissenschaftstheoretischen Betrachtung der zweiten Beobachtungsebene. Die jeweils beim Überwechseln von der (Inter-) Aktionsebene in die Beobachtungsebene erzielten Verlangsamungen werden als Zeitgewinn zur Entwicklung alternativer Interpretationen, das heißt alternativer Sinnmöglichkeiten und Handlungsoptionen, genutzt. Der Beobachter hat Zeit, sich zu entscheiden, ob er das oder jenes sieht, ob er das oder jenes in Anschluss erwarten kann. Die Beobachtungsfähigkeit kann im häufigen Vollzug ohne direkte Handlungskonsequenzen und –kosten trainiert und weiter methodisiert werden. Die Beobachter müssen sich vor allem nicht wie die beobachteten Akteure auf *eine* Lesart und den daraus folgenden Handlungszug (und sei es nur probeweise) festlegen, sondern es können mehrere zu den Beobachtungen passende Lesarten aufrechterhalten werden. Im Fortgang der Beobachtung der Handlungskette kann so entschieden werden, welche belegbaren Interpretationen und Erwartungen an Handlungen und Äußerungen tatsächlich eintreten und welche Schlüsse sich daraus für die Handlungsstrukturen und Relevanzen der Akteure ergeben, die ja als solches niemals direkt sichtbar sind, weil sie sozusagen dem Bewusstsein zuzurechnen sind. So kommt es dazu, dass ein Beobachter etwas sieht und erkennt, was die Akteure selbst nicht sehen.

Dieses Wissen kann kumuliert werden als wissenschaftliches Wissen, es kann Bestandteil von professionellem Wissen helfender Berufe werden oder auch ‚lokal‘ im direkten Kontakt zu den beobachteten Akteuren – etwa in sozialpädagogischen Situationen, einer Psychotherapie oder in Ausbildungssettings (vgl. auch oben Abschnitt 1) - zurück

kommuniziert werden. Immer entstehen in den jeweiligen Verwendungs- und Kommunikationszusammenhängen des aus Beobachtung gewonnenen Wissens neue Hinsichten, Irritationen, Veränderungen, Destabilisierungen oder Erweiterungen der bisherigen Handlungsschemata und Deutungen von Wirklichkeit.

Soweit diese Darstellung plausibel erscheint, verbindet dieses Konzept stimmig eine Konstitutionstheorie sozialer Strukturen mit einem mehrbenenfähigen hermeneutischen Ansatz, der in allen Beobachtungsbereichen von Sinnkonstitution sowohl ‚naturwüchsig‘ qua Interaktion und Kommunikation stillschweigend wirkt als auch in einer Forschungspraxis methodisch gesteigert angewendet werden kann.

Ein solcher interaktionsbezogener hermeneutischer Grundansatz (Hitzler 1999; Hitzler/Honer 1997) ist derzeit bei allen ausgearbeiteten Verfahren der Videoanalyse, die keine Messtechniken einsetzen, zu beobachten (Friebertshäuser u. a. 2007; Heath 1992, 1997; Heath/Luff 1992, 1996; Knoblauch u. a. 2006; Marotzki/ Stoetzer 2007; Raab 2008; oder auch die Fernsehanalysen von Reichertz 2009), auch wenn unterschiedliche Nähen zu Oevermanns Konzept und Techniken einer objektiven Hermeneutik (Oevermann 1991, 1993, 2000, 2001; technisch einführend Wernet 2006) gewahrt werden und es differierende Einschätzungen der höheren Leistungsfähigkeit wissenschaftlicher gegenüber alltagsweltlichen Praxen sowie gegenseitige Abgrenzungen in Details der Methoden gibt (Bohnsack 2009). Auch der Verfasser dieses Beitrags reiht sich in seinen Lehrforschungsprojekten (Fischer u. a. 2004) und von ihm angeregten und betreuten Forschungs- und Promotionsvorhaben (Bayas-Linke 2009; Müller 2006; Schwedhelm 2008) hier ein.

Welche Materialien sollen nun für eine Videoanalyse – vor allem im Bereich der sozialwissenschaftlichen Erforschung Sozialer Arbeit - erzeugt werden? Wie sind sie für die hermeneutische Analyse aufzubereiten? Wie kann letztere forschungslogisch und auslegungsinteraktiv umgesetzt werden? Welche spezifischen Publikationsprobleme der Ergebnisse können wie bewältigt werden? Die (zum Teil sehr gewichtigen) Fragen werden knapp angerissen, ohne dass die Erwartung geweckt werden soll, die Fülle notwendiger Details der forschungspraktischen Durchführung könne hier - wie in einem Manual - vermittelt werden. Es handelt sich um hoch diskursive Auswertungstechniken, deren praktisches Erlernen eines interaktiven Lernsettings und längerer Einübung bedarf.

Welche Materialien sollen erzeugt werden, also *welche* Videofilme sollen *wie* gedreht werden?

Nach dem bisher Ausgeführten liegt es nahe, für die Untersuchung von Interaktionen eben solche Videoaufnahmen zu wählen, die für die Soziale Arbeit relevante natürliche Interaktionen zum Inhalt haben. Dies wird in der Regel heißen, dass nicht bereits verteilungsmedial (im Fernsehen oder Internet) zugängliche oder innerhalb des interessierenden sozialen Milieus selber ‚angefallene‘ Filmproduktionen analysiert werden, sondern eigenes Material im Forschungsprozess in einem dokumentarisch-ethnographischen Stil erzeugt wird. Auch wenn in medienpädagogischen Produktionen die Sozialweltakteure selbst Filme planen (Treatment, Storyboard, Dialogskripte) und zur Kamera greifen, um ihre Welt zu reflektieren, und solche Videos durchaus in gleicher hermeneutischer Methodik analysierbar sind, ist für die Erforschung von sozialen Strukturierungen die Video-Dokumentation ‚natürlicher‘ Interaktionen (einschließlich einer guten Tonaufzeichnung) vorzuziehen. Dass sich hier viele Fragen nicht automatisch lösen, wenn man sich in einer natürlichen Situation filmend hinter die Kamera stellt, ist bereits oben bei der Diskussion um die Krise der Repräsentation und anlässlich der Studie zum dokumentarischen Filmen von Elisabeth Mohn skizziert worden. Es ist zunächst wichtig, dass Interaktionen möglichst umfänglich gefilmt werden, damit die materialinternen Verweis- und Prüfmöglichkeiten in der folgenden Rekonstruktionsphase reichhaltig sind. Weiterhin ist eine eher statische Aufnahmetechnik (anhaltend lange, die Interaktion nicht unterbrechende Einstellung, Bevorzugung der Bildtotale, wenige Schwenks, keinen Zoom, keine Kamerafahrten, -bewegungen oder Zooms) zu bevorzugen, da jede ‚interessantere‘ Kameraführung immer auch schon Auswahl und Akzentuierung und damit Bedeutungserteilung und Interpretation ist. Diese soll aber erst in der Analysephase breit und im Diskurs mit anderen Beobachtern aufgefächert werden. Die aufnahmetechnische Schaffung von Blickzentren, die ja gerade in der Kunstproduktion eine riesige Rolle spielt – ich erinnere an die obigen Ausführungen Panofskys (Panofsky 1979b, 1979a, 1992, 1997), Imdahls (Imdahl 1980, 1994, 1996a, 1996b) und Schumms (Beller 2002: 226; Schumm 2004) – wird hier nach Möglichkeit ausgesetzt. (Allerdings ist zu beachten, auch eine Totalperspektive ist eine Interpretation, die Bildwerte vermittelt.) Es wird also eine Aufnahmetechnik verwendet, die nicht wie die Blickbewegungen des Auges hin und her streift (und damit auch Bildsegmente zeitweise ausschließt), sondern der ‚Blick aufs Ganze‘ wird bevorzugt, damit nicht schon Interpretationen beim Aufnehmen suggeriert werden. Damit wird auch der Standpunkt des Forschers (wie seine leibliche Präsenz) als partizipatives und interpretatives Moment der

Datenerzeugung an den Rand oder außerhalb der Interaktion gelegt. Diese Präsenz ist natürlich für die Interaktanten durchaus in der Beobachterposition eines Fremden vorhanden und wird in der Interpretation mit in Rechnung gestellt. Will man auch die Beobachteranwesenheit in die spätere Datenanalyse auch visuell und nicht nur als Kontextinformation einbeziehen, muss auch von filmenden Forscher soviel Material erzeugt werden (z.B. am Beginn der Aufzeichnung), dass die Analysekommunikation darauf zurückgreifen kann. Dennoch bleiben bei diesem Vorgehen eine Reihe von Problemen. Schon der übliche Einsatz von nur einer Kamera, die ja auch nur einen visuellen Standpunkt erzeugt, der bestimmte Interaktanten von vorn, andere von hinten darstellt und sie unter Umständen damit bevor- oder benachteiligt, ist eine Einschränkung. Sie erstreckt sich auch auf die dadurch bedingte einseitige Raumperspektive, die keineswegs neutral wahrgenommen wird. In Spielfilmen wird das Problem vor allem bei Dialogen durch einen Wechsel der Kameraperspektive analog der wechselnden Sprecher-Hörer-Perspektiven, durch Einstellung und Gegeneinstellung gelöst; die ethnographische Kamera könnte das nur durch Verdoppelung der Aufnahmetechnik erreichen und müsste vor allem eine kaum hinnehmbare Beeinträchtigung der natürlichen Situation in Kauf nehmen. Auch der Schnitt und die Analyse von zwei Kameraaufzeichnungen verkomplizieren den Forschungsprozess, so dass in der Regel trotz der Nachteile mit nur einer Kamera operiert wird.

Wie soll nun das gefilmte Rohmaterial (footage) für die Analyse vorbereitet werden? Das Material, das beim Stand der heutigen Aufnahmetechnik digital vorliegt und unmittelbar durch Schnittprogramme am Computer bearbeitet werden kann, wird als erstes lediglich in eine für die Analyse handhabbare Abspielform gebracht und nicht in veränderter Reihenfolge montiert. Die Schnitte begrenzen sich vor allem auf die Auswahl von kurzen Bearbeitungssequenzen (Einheiten von nur wenigen Minuten haben sich bewährt), die ihrerseits ohne Auslassungen in separat hintereinander abspielbare Clips (ca. 20-30 sec) geschnitten werden. Die Erstauswahl des Materials erfolgt nach thematisch-szenischem Interesse, bei dem nach Möglichkeit keine Interaktions- und Redezüge durchschnitten werden; weitere Segmente können nach ersten Analyseergebnissen in gleicher Weise bereit gestellt werden. Entscheidend ist, dass die Auswertung auf der Grundlage der wiederholt abspielbaren und somit beobachtbaren Ablaufdarstellungen *sequentiell* stattfinden kann, ohne dass in dem bearbeiteten Segment Ausschnitte ausgelassen werden.

Der nächste aufgrund bislang nicht verfügbarer Zeichensysteme am wenigsten präzierte

Schritt der Datenaufbereitung ist die Transkription. Die Verwandlung des Interaktionsverlaufs im Zeitablauf des Videodokuments in ein analoges, den Ablauf irgendwie ‚zeitfrei‘ darstellendes Schriftstück wäre eine notwendige Vorbedingung für die Abkoppelung der Analysezeit von der Zeitlichkeit des Dokuments und der realen Interaktion. Diese Abkoppelung ist zwingend, da ja gerade der Zeitgewinn oder der weit größere Zeitbedarf der Analyse im Vergleich zur Spanne der Interaktionszeit ein Entkommen aus den Abkürzungszwängen der alltäglichen Bedeutungserteilung ermöglicht. Dies ist überhaupt die Voraussetzung für den möglichen wissenschaftlichen Ertrag der Vermehrung von Lesarten und Verstehensvarianten. Für die gesprochene Sprache gibt es in enger Anlehnung an die Schriftsprache bewährte und für verschiedene linguistische oder soziologische Aufgaben leistungsfähige Transkriptionssysteme (vgl. umfassend Dittmar 2000). Etwas Vergleichbares konnte noch nicht für Filmdokumente entwickelt werden. So können Interaktionen und ihr Zusammenspiel von sichtbarem Handlungsraum, von leiblich-körperlichen und sprachlichen Handlungsfäden trotz jahrelanger Bemühungen immer noch nicht befriedigend analog textlich dargestellt werden (illustrativ dazu Luckmann 2006). Es ist fraglich, ob es eine solche ‚Partitur‘ überhaupt geben kann. Die visuellen Anknüpfungsmöglichkeiten von Beobachtern sind immens, dagegen verlaufen Versprachlichungen in ihrer Bedeutungswahl immer hoch selektiv und schmälern Interpretationsmöglichkeiten. Ein weiteres Fundamentalproblem liegt auch im Fehlen einer Systematik der Gesten und körperlichen Haltungen, einer ‚Grammatik‘ und ‚Semiotik‘ des Nonverbalen, auf deren Grundlage man zeichenhafte Repräsentationen entwickeln könnte. (Was bei vergleichbarer Problemalge etwa in der Musik als Notation durch Festlegung des Tonrepertoires in Skalen, Rhythmen und Dynamiken im Laufe der letzten Jahrhunderte relativ gut gelang.) So wird man sich in den meisten Fällen mit für den Forschungsanlass selbst geschneiderten partiturartigen Auflistungen weniger Ereigniskategorien im Zeitablauf ohne (Raab 2008: 110 u. 326ff.) oder mit Standbildern über dem Text (Bohnsack 2009: 179ff. u. 241) begnügen und in jedem Fall zusätzlich eine konventionelle Sprachtranskription für die zu analysierenden Segmente herstellen. Diese texthaften Partituren dienen dann eher zur groben Katalogisierung des Materials und dem Auffinden von Szenen im Verlauf des Analyseprozesses. Die eigentliche Interpretationsarbeit erfolgt bei gemeinsamer wiederholter Betrachtung der kleinteilig angelegten Videosegmente, wobei der für die Rekonstruktion notwendige Zeitbedarf durch das sequentielle ‚Stopp and Go‘ und die häufigen Wiederholungen erreicht wird.

Die sich auf der Grundlage der aufbereiteten Materialien vollziehende Interpretationsarbeit ist nach dem bisher Ausgeführten nicht messtechnischer sondern immer *hermeneutischer Art*. Ihre Qualität wird sich daran bemessen lassen, wie weit sie in der Lage ist, die bislang aufgeführten wahrnehmungs-, kunst-, interaktions- und sprachtheoretischen Kontexte für die konkrete Materialbearbeitung und Interaktionsrekonstruktion zu aktivieren.

In der rekonstruktiven Analyse trennt der Verfasser forschungspraktisch, angeregt durch Überlegungen von Erving Goffman und Adam Kendon zu verschiedenen Wahrnehmungskanälen (Goffman 1977; Kendon 1988b, 1988a), nach visuellen, sprachlichen und Ton-Bild-Rekonstruktionen. Die drei Analysestränge werden zunächst ohne Kontextinformationen und auch voneinander dekontextualisiert im Stile der objektiven Hermeneutik nach Möglichkeit in einer Interpretationsgruppe streng sequentiell ausgearbeitet und zu je eigenen Strukturhypothesen verdichtet. In einem letzten Arbeitsgang werden die Teilergebnisse kontrastiert und abgeglichen. Diese Bearbeitungsebenen sind nicht so zu verstehen, als gäbe es in Interaktionen realiter getrennte Ebenen der Bedeutungskonstitution gestisch-visueller und sprachlicher Art. Vielmehr findet im Zeitablauf ständig ein Wechselspiel leiblich-körperlicher Eindrucks- und Ausdrucksphänomene und sprachlicher Austausch statt, in dem Sinn konstituiert und Handlungsstrukturen realisiert werden. Sie leben aus dem sozialen Kontext und machen diesen aus. Die analytisch-rekonstruktive Trennung von visuellem und sprachlichem Bereich dient alleine der heuristischen Wahrnehmungsschärfung und soll gerade die Sinnkonstitution im Wechselspiel besser zugänglich und verständlich machen.

Die *erste Analysephase, die Rekonstruktion von Strukturen visueller Bedeutungsherstellung* ohne Sprechanteile erfolgt auf der Grundlage tonlos abgespielter Interaktionssequenzen, die aus dem Rohmaterial hergestellt wurden. Es wird dazu eine Software zum Abspielen benutzt, die die Ablaufzeiten anzeigt; eine Beamerprojektion ist notwendig, damit eine Interpretationsgruppe am Material arbeiten kann. In der Auslegungspraxis hat es sich bewährt, zunächst mit einem Standbild (dem ersten Frame) der analysierten Sequenz zu beginnen. Die Analysegruppe bildet spontane Hypothesen zur Frage, was das Bild zeigt, was man sieht und worum es in der Szene geht, also zur Frage nach dem *hypothetischen Bildsujet*. Die Betrachter greifen hier in erster Linie auf ihre visuelle Vertrautheit mit Interaktionen ihrer eigenen Kultur, aber auch in gewissem Umfang auf sonstige ihnen verfügbare visuelle Images zurück. Die Tragfähigkeit der Hypothesen erhöht sich mit der

Kenntnis der Kultur oder dem sozialen Milieu, in dem die Interaktion stattfand. Die Standbildanalyse ist auch der Ort im Forschungsprozess, wo die oben (Abschnitt 3.1) skizzierten formalen ikonologischen Merkmale angewendet und ausgelegt werden können, soweit eine entsprechende Ausbildung der Forscher im Erkennen der perspektischen und planimetrischen Merkmale geleistet wurde. Auch ohne diese Expertise lassen sich von den in der Interpretationsgruppe auftretenden Impressionen her Lesarten entwickeln. Welche „Apellstruktur“ (Waldenfels 2006: 252) des Standbildes lässt sich wahrnehmen; gibt es möglicherweise darauf zu sozialen Situationen passende „Antwortregister“? Lassen sich durch die Betrachter Bildzentren isolieren, die bereits noch vor der sichtbaren Bewegung im Video schon hypothetische Bedeutungszuteilungen zulassen, die dann an anderer Stelle auf dem Standbild und im Fortgang des Bewegungsablaufs überprüft werden können (Breckner 2003, 2007; Müller-Dohm 1993, 1997)? Die von den Beobachtern geäußerten Bildwirkungen werden in ihrer Vielfältigkeit oder auch Widersprüchlichkeit nicht konsensuell nivelliert, sie müssen aber stets an Gegebenheiten auf dem Bild festgemacht werden. So kommt es immer zu rudimentären und heuristisch-deiktischen Teilbeschreibungen des Bildobjekts, auch wenn es – wie auch die Probleme der Transkription belegen – nicht wirklich versprachlicht werden kann. Im strengen Sinne gilt: „Was das Bild als solches ist, widersetzt sich aller sprachlichen Substitution.“ (Imdahl 1994: 310) Die Annahmen zum Bildsujet werden - wie stets die Hypothesen in den folgenden Arbeitsgängen - immer für alle Teilnehmer der Interpretationsgruppe sichtbar notiert (auf Flipchart oder in gebeamten Textdateien), was gegenüber dem diskursiven Austausch in der Gruppe eine Fokussierung, Verdichtung und Fixierung bedeutet.

Nach dem Standbild werden sukzessiv kleine Teilsequenzen (meist im Umfang von weniger als 10 Sekunden) untersucht und weiter Hypothesen darüber gebildet, was die Interaktion bedeutet. Woran kann sich der wissenschaftliche Beobachter dabei orientieren? Auch die Interaktanten in der Situation haben dieses Problem zu lösen. Sie müssen sich dazu außer an ihren internalen Relevanzen über Beobachtungen des Blicks, des Gesichts und der Gesten der anderen Interaktanten orientieren. Sie entwickeln aufgrund dieser nonverbalen Anzeichen Hypothesen, worum es jetzt geht, und verhalten sich aufgrund dieser Annahmen dann entsprechend zustimmend, neutral oder ablehnend dazu. Ob die gebildeten Hypothesen zutreffen, wird sich im Fortgang der Interaktion zeigen. In gleicher Weise können sich nun auch die wissenschaftlichen Beobachter methodisch gesteigert im

Fortgang der Betrachtung der Sequenz an den beobachteten *Blicken, Haltungen, Gesten und Gesichtsausdrücken* orientieren und Anhaltspunkte für sich entwickelnde Bedeutungszuweisungen gewinnen (Goodwin 1986). So werden nach und nach Lesarten entwickelt und eingegrenzt, die sich später auch im dritten Arbeitsgang bei der Analyse des Sprechgeschehens – womöglich zeitversetzt – noch bekräftigen lassen. Am Abschluss der ersten Rekonstruktionsphase steht eine schriftliche Zusammenfassung der geprüften Strukturannahme. Auch wenn man schon oft solche Analysen tonloser Videosequenzen gemacht hat, ist es immer wieder erstaunlich, welche stabilen und tragfähigen Strukturhypothesen zur Interaktion sich nur aufgrund der visuellen Bildabfolge ohne jegliche Kontextinformationen erarbeiten lassen.

Bei der *zweiten Analysephase* geht es um die *Strukturen der Bedeutungsherstellung durch das Gespräch*. Als Arbeitsgrundlage dient hier ausschließlich das Sprachtranskript. Ohne auf die vorgängige Bildanalyse zurückzugreifen, wird hier – wie schon im vorigen Schritt im Stil der objektiven Hermeneutik – Redezug um Redezug oder noch kleinteiliger über eine Fülle von Einzelhypothesen eine zusammenfassende Strukturhypothese darüber erarbeitet, was die Interaktanten hier *tun, indem sie so und nicht anders miteinander sprechen*. Dieser Analysestrang zieht keine visuellen Daten in Betracht, er beschränkt sich nur auf die hörbaren Sprechereignisse, die in Form des Transkripts gegeben sind.

Im *dritten Analyseschritt* wird die *Bedeutungsherstellung der Interaktion ganzheitlich audio-visuell* untersucht. Methodisch in der gleichen Weise wie in den beiden vorigen Arbeitsschritten wird sequentiell kleinteilig das ausgewählte Videosegment bearbeitet. Anhand des Zusammenspiels der Ton- und Bildimpressionen kann jetzt hier vor allem darauf geachtet werden, welche Wechselwirkungen (unterstützend, konterkarierend) sich zwischen den verbalen und nonverbalen Äußerungsformen zeigen. Obwohl durch die beiden vorangegangenen Arbeitsschritte bereits geprüfte Interpretationen vorliegen, die den Ertrag dieser Gesamtinterpretation zu mindern scheinen, empfiehlt es sich, doch nicht auf die ersten beiden Arbeitsphasen zu verzichten und die Rekonstruktionen nicht alleine in diesem letzten Schritt leisten zu wollen. Die nach Ton und Bild vollständigen Daten sind sehr komplex, und die beobachtenden Forscher erkennen mehr, wenn sie sich sukzessiv am spezifischen Material innerhalb verschiedener Symbolklassen schulen und in der vorge-schlagenen Reihenfolge der Schritte *dreiphasig* vorgehen. Sie verleihen nur so auch dem Nonverbal-visuellen ein Eigengewicht gegenüber der sich ansonsten schnell

vordrängenden sprachlichen Bedeutungskonstitution. Den Abschluss der Interpretation bildet die verschriftlichte Zusammenführung aller Ergebnisteile.

In jeder Forschungsphase kann die selbstreflexive Arbeit als Beobachtung zweiter Ordnung eingeschoben werden. Dabei ist zu entscheiden, durch welche Daten (Protokolle, Fotos, Audio- und Videoaufzeichnungen von Forschungsprotokollen, Ergebnistexte) die Beobachtung der Beobachtung gestützt werden soll. Wenn diese ‚Metadaten‘ nicht bereits im Forschungsprozess angefallen sind, müssen sie eigens hergestellt werden, was einen zusätzlichen Aufwand darstellt. Die Frage an dieses Material lautet prinzipiell, *wie* die Interpretationsgruppe *interaktiv zu ihren Ergebnissen* findet, welche Konzepte sie dabei benutzt und ob bei anderen Konzepten oder Interaktionen die Ergebnisse auch anders ausfallen könnten. Eine solche konstruktivistische Reflexion kann nicht permanent mitlaufen, da sie die primäre Forschungsinteraktion unterbricht und auch möglicherweise unterminiert, indem sie Details der Forschungsinteraktion und funktionierende Diskurs- und Entscheidungsroutrinen problematisiert. Der Ertrag ist dennoch beträchtlich, da ein explizites Wissen um die eigenen Konstruktionsmechanismen die Handlungsmöglichkeiten der Forschung erweitert und die Ergebnisse in ein neues Licht rückt. Selbst wenn dabei Irritationen entstehen, lässt sich erwarten, dass diese langfristig nützlich sind.

Je nach Verwendungszweck sind die Forschungsergebnisse der konkreten Videoanalyse dann abschließend weiter wissenschaftlich zu kontextuieren und zu verbreiten. Die Anbindung an einen relevanten Forschungsdiskurs erfolgt meist in Form wissenschaftlicher Publikationen oder durch Vorträge in der Scientific Community. Auch eine interkativ-kommunikative Rückführung der Ergebnisse in dasselbe oder ein vergleichbares Praxisfeld ist im Rahmen professionellen Handelns sinnvoll. Alle diese Verbreitungsformen der Ergebnisse finden in jeweils eigenen Interaktionskreisen statt und haben ihre eigenen situativen Anforderungen und Regeln ausgebildet, denen zu folgen ist, wenn der Ergebnis- und Wissenstransfer gelingen soll. Das kann nicht mehr Gegenstand des Beitrags sein. Es soll jedoch noch kurz auf einige Probleme der Verbreitung von Ergebnissen dieses Forschungstyps eingegangen werden, die unmittelbar mit dem Videomaterial als Datenbasis und seiner hermeneutischen Auslegung zu tun haben.

Das Hauptproblem der Darstellung ergibt sich aus der Bild- und Zeitlichkeit des Videomaterials, die zwar den untersuchten Interaktionen kongenial ist, sich aber nicht befriedigend *textlich* präsentieren lässt. Das Problem, das forschungsintern als unzureichende

Transkription auftaucht, wiederholt sich auf der Ebene der Ergebnispräsentation, weil Wissenschaft (immer noch) Textform fordert. Da sich ein Film nicht in einem Text abbilden lässt, ist auf verbale Beschreibungen zurückzugreifen, um die Interpretationen an visuell sichtbaren Handlungsvollzügen festzumachen. Diese sich den bewegten Bildern nur annähernden Beschreibungen klingen meist umständlicher, als die Bilder anzuschauen sind. Schlimmstenfalls sind sie ohne vorgängige Visualisierung nicht verstehbar. So ist diese Art der Forschung im Vollzug von Beobachtungen und diskursiven Interpretationen spannend, in der Ergebnispräsentation für nicht am Forschungsprozess Beteiligte ergibt sich eine nicht aufhebbare Distanz, die die Evidenz der Ergebnisse schlecht vermittelt. Es erscheint einfacher, einen Film zu machen, als darüber einen wissenschaftlichen Text zu verfassen. Immerhin sind die medialen Möglichkeiten durch preiswerte digitale Bildträger soweit fortgeschritten, dass sich mit solchen Publikationsformen experimentieren lässt, die Filme auf DVDs einbeziehen.

Ein zweites Problem ergibt sich aus der Bildlichkeit, die für die Analyse notwendig ist. Sie macht Personen visuell identifizierbar, verletzt damit aber möglicherweise Persönlichkeitsrechte oder kann auch bei den Forschern Befangenheiten bei der Ergebnisveröffentlichung erzeugen. Hier wird mittlerweile auf vertretbare Weise in der Endpublikation (nicht bei der Datenanalyse) mit graphischen Verfremdungen von Bildern und Gesichtern gearbeitet.

Ein Präsentationsproblem ganz anderer Art folgt aus dem Analysemodus der objektiven Hermeneutik, der im Kern aus der Entwicklung einer Vielzahl von Lesarten besteht, die sequentiell abgearbeitet werden müssen, um zu Gesamtaussagen zu finden. Weder ist der gesamte Interpretationsprozess in seinem Werkstattcharakter den Lesern zumutbar, noch kann ganz auf eine Repräsentation der Interpretationswege verzichtet werden. Eine reine Darstellung von Endergebnissen wird nicht plausibel und verliert Substanz, wenn man den Weg dorthin nicht kennt. So entstehen seit drei Jahrzehnten im Schreiben und Präsentieren vielfältige Darstellungsformen innerhalb der Forschungstradition der objektiven Hermeneutik, die immer noch die Kreativität der Forscher fordern und für die Publikation von Ergebnissen keine Standardformen hervorgebracht haben.

5. Videoanalyse als Forschungsinstrument in der Sozialen Arbeit

Sucht man nach bereits vorliegenden videogestützten Forschungsbeispielen in der Sozialen Arbeit, fällt die Ausbeute zwiespältig aus. Einerseits finden sich im Umkreis der bereits

genannten erziehungswissenschaftlichen und soziologischen Textsammlungen (Ehrenspeck et al.; Friebertshäuser et al.; Knoblauch et al.; Raab) viele Studien, bei denen sich ein Bezug zur Sozialen Arbeit herstellen lässt. Neben den allgemein methodischen Arbeiten (Herrle u. a. 2009; Marotzki 2006) geht es um Kinder (Nentwig-Gesemann 2007), Jugendkultur (Bohnsack/ Nohl 2001; Fritzsche 2003), Erwachsenenbildung (Kade/ Nolda 2007; Nolda 2007), Altern und Generation (Schäffer 2003, 2009) oder Psychosomatik (Klambeck 2007). Auch die Unterrichts- und Schulforschung, die eine Tradition in der sprachbasierten Interaktionsanalyse hat, greift die videogestützte Dokumentation auf (Wagner-Willi 2004, 2005, 2007). Andererseits gibt es in der engeren Forschung der Sozialen Arbeit trotz einer Vielzahl sprachbasierter rekonstruktiver Untersuchungen wenig videogestützte Interaktionsstudien. Im Fachgebiet des Verfassers sind in jüngster Zeit Forschungsarbeiten zum Zusammenhang von Coachinginteraktion und Biographie (Müller 2006), interkultureller Beratung (Schwedhelm 2008) und Führungsinteraktion in Organisationen (Bayas-Linke 2009) abgeschlossen worden. In einer Fachgebietskooperation mit der Medizinsoziologie und der theoretischen Medizin, bei der neben biographieanalytischen auch videogestützte Verfahren eingesetzt wurden, sind salutogene Prozesse auch im Blick auf ihre Förderung untersucht worden (Bahrs/ Matthiessen 2007). Der Typus solcher Forschungsprojekte, die auch professionelle Interventionen anzielen und gesellschaftliche Praxen mit verändern wollen, eignet sich meines Erachtens besonders für videographische Verfahren. Auf dieser Linie kann weiterhin erwartet werden, dass das jüngere Interesse an ethnographischen Methoden in den empirischen Erziehungswissenschaften und der Sozialpädagogik (Cloos/ Thole 2006; Heinzel u. a. 2009) früher oder später neben dem Mittel schriftlicher Protokolle bei Feldbeobachtungen auch stärker zur Nutzung von Videoprotokollen führt.

So nüchtern die Bestandsaufnahme für videogestützte Interaktionsanalysen in Forschungsfeldern der Sozialen Arbeit auch ausfällt, der Beitrag hat gezeigt, dass hier eine voll entwickelte Forschungsmethodik vorliegt, die hervorragend in der rekonstruktiven Sozialarbeitsforschung einsetzbar ist. So bleibt am Ende die Aufforderung, sich dieses visuelle Instrument der Interaktionsanalyse anzueignen und weiter zu entwickeln. Es kommen dafür meines Erachtens drei Felder infrage.

Zum einen eignet sich die Videoanalyse für *alle Adressatenfelder* der Sozialen Arbeit. Über die ganze Lebensspanne, also bei Kindern, Jugendlichen und bei Erwachsenen im

Erwerbs- und Rentenalter, lässt sich die Frage empirisch verfolgen, welche Handlungs- und Erlebensmuster in Interaktionen wirksam sind, die zu Exklusionen, Problemverhalten und dauerhaftem Hilfebedarf durch Institutionen des sozialen Netzes beitragen. Die videogestützte Interaktionsanalyse ist dabei in der Lage, nicht nur problematische Interaktionsmuster, sondern auch Ressourcen aufzuspüren und kann so zum Anknüpfungspunkt für sozialpädagogische Interventionen und Veränderungen in der Lebenswelt der Klienten werden. Die Ergebnisse dieser qualitativen Forschungen sind in ihren Generalisierungen auch für sozialpolitische Strukturmaßnahmen von Bedeutung, da diese valide empirische Darstellungen von Lebenslagen braucht, um den entsprechenden politischen und gesetzlichen Rahmen für Unterstützungen zu schaffen. Weiterhin sind videogestützte Interaktionsanalysen ein primäres Forschungsmittel, um die professionellen Interventionen der Sozialen Arbeit wissenschaftlich zu begleiten, die Beratungsinteraktionen verschiedenster Art im Detail zu evaluieren und so insgesamt die *berufliche und professionelle Identität durch Selbstaufklärung* zu verbessern.

Nicht zuletzt eignet sich videogestützte Interaktionsanalyse für *Grundlagenforschung*. Wie entsteht Sozialität im Zusammenspiel von ‚Innen und Außen‘? Wie wirkt Sozialwelt auf Individuen und wie drücken diese sich leiblich kommunizierend und interagierend aus? Es geht allgemeiner darum, die Konstitution sozialer Strukturen, die Art und Weisen der Welterzeugung durch Handeln, symbolischen Austausch und Herstellung von Orientierungen zu rekonstruieren. Bei diesem anspruchsvollen Vorhaben ist die videogestützte Interaktionsanalyse angemessen komplex und ihre Leistungsfähigkeit ist noch auszubuchstabieren.

Alles in allem versprechen der Einsatz und die Weiterentwicklung der Videoanalyse in der empirischen Forschung einen Wissenszuwachs, der theoretisch weit reicht und zugleich die soziale Praxis befördert.

Anschrift des Verfassers:
Prof. Dr. Wolfram Fischer
Universität Kassel, FB Sozialwesen
Arnold-Bode-Straße 10
34109 Kassel
wfischer@uni-kassel.de

Literatur

- Aarts, Maria (2002): The Marte Meo programme for autism. Six information sessions on how to stimulate social and emotional development; especially helpful for parents and caregivers of children on the autistic spectrum and children with emotional and social development problems. Harderwijk: Aarts Productions.
- Aarts, Maria u. a. (2002): Marte Meo: ein Handbuch. Harderwijk: Aarts Productions.
- Amann, Klaus (1997): Ethnographie jenseits von Kulturdeutung. Über Geigespielen und Molekularbiologie. In: Hirschauer, Stefan/Amann, Klaus (Hrsg.): Die Befremdung der eigenen Kultur. Zur ethnographischen Herausforderung soziologischer Empirie Frankfurt a.M.: Suhrkamp, S. 298-330.
- Apel, Karl-Otto (1990): Ist Intentionalität fundamentaler als sprachliche Bedeutung? Transzendentalpragmatische Argumente gegen die Rückkehr zum semantischen Intentionalismus der Bewußtseinsphilosophie. In: Homburg, Forum Philosophie Bad (Hrsg.): Intentionalität und Verstehen. Frankfurt a.M.: Suhrkamp, S. 13-54.
- Bahrs, Ottomar u. a. (Hrsg.)(1996): Ärztliche Qualitätszirkel - Leitfaden für den niedergelassenen Arzt. 3 Köln: Deutscher Ärzte-Verlag.
- Bahrs, Ottomar/Matthiessen, Peter, F. (Hrsg.)(2007): Gesundheitsfördernde Praxen. Die Chancen einer solutogenetischen Orientierung in der hausärztlichen Praxis. Bern: Verlag Hans Huber.
- Bahrs, Ottomar u. a. (1998): 'I can now face a patient'. The development of a concept to analyse video recorded encounters from general practice in a multidisciplinary quality circle. In: Bensing, J u. a. (Hrsg.): Doctor-patient communication and the quality of care in general practice. Utrecht: NIVEL-Institut, S. 113-130.
- Bayas-Linke, Dirk (2009): Organisation und Interaktion: Rekonstruktion von Führung. Videointeraktionsanalyse und Systemtheorie, eine empirische Komplementarität. Kassel: kassel university press.
- Beller, Hans (Hrsg.)(2002): Handbuch der Filmmontage. München: TR-Verlagsunion.
- Berg, Eberhard/Fuchs, Martin (Hrsg.)(1993): Kultur, soziale Praxis, Text. Die Krise der ethnographischen Repräsentation. Frankfurt: Suhrkamp.
- Bergmann, Jörg R. (1994): Ethnomethodologische Konversationsanalyse. In: Fritz, G./Hundsnurscher, F. (Hrsg.): Handbuch der Dialoganalyse. Tübingen: Max Niemeyer, S. 3-16.
- Blumer, Herbert (1969): Symbolic Interactionism. Perspective and Method. Englewood Cliffs, N.J.: Prentice-Hall.
- Boehm, Gottfried (Hrsg.)(2006): Was ist ein Bild? 4. Aufl. München: Fink Verlag.
- Boehm, Gottfried (2008): Wie Bilder Sinn erzeugen. Die Macht des Zeigens. 2. Aufl., Berlin: Berlin University Press.
- Bohnsack, Ralf (2003): Die dokumentarische Methode in der Bild- und Fotointerpretation. In: Ehrenspeck, Yvonne/Schäffer, Burkhard (Hrsg.): Film- und Fotoanalyse in der Erziehungswissenschaft. Opladen: Leske + Budrich, S. 87-107.
- Bohnsack, Ralf (2007): Zum Verhältnis von Bild- und Textinterpretation in der qualitativen Sozialforschung. In: Friebertshäuser, Barbara u. a. (Hrsg.): Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen, Farmington Hills: Verlag Barbara Budrich, S. 21-45.
- Bohnsack, Ralf (2009): Qualitative Bild- und Videointerpretation. Die dokumentarische Methode. Opladen & Farmington Hills: Barbara Budrich/UTB.

- Bohnsack, Ralf/Nohl, Arndt-Michael (2001): Jugendkulturen und Aktionismus. Eine rekonstruktive empirische Analyse am Beispiel des Breakdance. In: Zinnecker, Jürgen/Merkens, Hans (Hrsg.): Jahrbuch Jugendforschung. Folge 1. Opladen: Leske & Budrich.
- Breckner, Roswitha (2003): Körper im Bild. Eine methodische Analyse am Beispiel einer Fotografie von Helmut Newton. In: ZBBS (H. 1): 33-60.
- Breckner, Roswitha (2007). Bildwelten - Soziale Welten. Zur Interpretation von Bildern und Fotografien (22 S.), Konferenz Visuelle Soziologie 23.-24. Nov 2007 Wien: www.univie.ac.at/visuellesoziologie/.
- Brümmer, Marita/ter Horst, Klaus (1997): Die Familie im Focus der Kamera. Einführung in die Methode des Video-Home-Trainings (VHT). In: e.V., Evangelischer Erziehungsverband (Hrsg.): Schriftenreihe 4/97 Familienaktivierenden Konzepte. Hannover: Linden-Druck Verlag, S. 57-69.
- Bünder, Peter u. a. (2009): Lehrbuch der Marte-Meo-Methode: Entwicklungsförderung mit Videounterstützung. Göttingen: Vandenhoeck & Ruprecht.
- Cloos, Peter/Thole, Werner (Hrsg.)(2006): Ethnografische Zugänge. Professions- und adressatInnenbezogene Forschung im Kontext von Pädagogik. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Cordes, Ragna (2000): Videogestütztes Trainingsprogramm für Risikofamilien. Entwicklung und Erprobung eines Video-Interaktionstrainings (VIT). Hamburg: Kovac.
- Dittmar, Norbert (2000): Transkribieren. Ein Leitfaden für Forscher und Laien. Opladen: Leske + Budrich.
- Doering, Hilke/Hirschauer, Stefan (1997): Die Biographie der Ding. Eine Ethnographie musealer Repräsentation. In: Hirschauer, Stefan/Amann, Klaus (Hrsg.): Die Befremdung der eigenen Kultur. Zur ethnographischen Herausforderung soziologischer Empirie. Frankfurt a.M.: Suhrkamp, S. 267-297.
- Drew, Paul/Heritage, John (Hrsg.)(1992): Talk at work. Interaction in institutional settings. Cambridge: Cambridge University Press.
- Ehrenspeck, Yvonne/Schäffer, Burkhard (Hrsg.)(2003): Film- und Fotoanalyse in der Erziehungswissenschaft. Opladen: Leske + Budrich.
- Fischer, Wolfram (2003): Körper und Zwischenleiblichkeit als Quelle und Produkt von Sozialität. In: ZBBS (H.1): 9-31.
- Fischer, Wolfram u. a. (2004). Interaktionsanalyse von Videosequenzen. Multimediales online Modul: www.uni-kassel.de/fb4/labor/projekt_interaktion.htm.
- Fischer-Rosenthal, Wolfram (1996): Medizinische Diagnose als offene praktische Beschreibung. Ein Versuch über das Nicht-Wissen und das ärztliche Handeln. In: Bahrs, Ottomar u. a. (Hrsg.): Vom Ablichten zum Im-Bilde-Sein. Würzburg: Königshausen und Neumann, S. 27-53.
- Friebertshäuser, Barbara u. a. (Hrsg.)(2007): Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen, Farmington Hills: Barbara Budrich.
- Fritzsche, Bettina (2003): Pop-Fans. Studie einer Mädchenkultur. Opladen: Leske & Budrich.
- Fuchs, Martin/Berg, Eberhard (1993): Phänomenologie der Differenz. Reflexionsstufen ethnographischer Repräsentation. In: Dies. (Hrsg.): Kultur, soziale Praxis, Text. Die Krise der ethnographischen Repräsentation. Frankfurt a.M.: Suhrkamp, S. 11-108.

- Gens, Hannelore (2002): Aufbau von Kompetenz im konkreten Praxisbezug – Video-Interaktions-Begleitung. Das "integrierte Modell" des Leitens und die "teaching organization". In: Jugendhilfe (H. 6): 290-303.
- Goffman, Erving (1977): Rahmenanalyse. Ein Versuch über die Organisation von Alltagserfahrungen. Frankfurt a. M.: Suhrkamp.
- Goodwin, Charles (1986): Gesture as a resource for the organization of mutual orientation. In: Semiotica, 62: 29-49.
- Goodwin, Charles/Goodwin, Marjorie Harness (1996): Seeing as situated activity: formulating planes. In: Engeström, Y/Middleton, D (Hrsg.): Cognition and communication at work. Cambridge: Cambridge University Press, S. 61-95.
- Goodwin, Marjorie Harness (1995): Assembling a response: setting and collaboratively constructed work talk. In: Have, Paul ten/Psathas, George (Hrsg.): Situated order: studies in the social organization of talk and embodied activities. Washington, DC: University Press of America, S. 173-186.
- Gurwitsch, Aron (1974): Das Bewußtseinsfeld (1957). Berlin, New York: deGruyter.
- Heath, Christian C. (1992): Gesture's descreet tasks: Multiple relevencies in visual conduct and in the contexutalisation of language. In: Auer, P./di Lzlia, A. (Hrsg.): The Contextualization of Language. Amsterdam: Benjamins, S. 101-127.
- Heath, Christian C. (1997): The analysis of activities in face to face interaction using video. In: Silverman, David (Hrsg.): Qualitative research. London: Sage, S. 183-200.
- Heath, Christian C./Luff, P.K. (1992): Explicating face-to-face interaction. In: Gilbert, N.G (Hrsg.): Researching social life. London: Sage, S. 306-326.
- Heath, Christian C./Luff, P.K. (1996): Convergent activities: line control and passenger information on the London Underground. In: Engeström, Y/Middleton, D (Hrsg.): Cognition and communication at work. Cambridge: Cambridge University Press, S. 96-129.
- Heinzel, Friederike u. a. (Hrsg.)(2009): "Auf unsicherem Terrain": Ethnographische Forschung im Kontext des Bildungs- und Sozialwesens. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Herrle, Matthias u. a. (2009): Erziehungswissenschaftliche Videografie. In: Friebertshäuser, Barbara u. a. (Hrsg.): Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft. Neuauflage. Weinheim/ München: Juventa.
- Hirschauer, Stefan/Amman, Klaus (Hrsg.)(1997): Die Befremdung der eigenen Kultur. Zur ethnographischen Herausforderung soziologischer Empirie. Frankfurt a. M.: Suhrkamp.
- Hitzler, Ronald (1999): Hermeneutische Wissenssoziologie Standpunkte zur Theorie der Interpretation. Konstanz: UVK.
- Hitzler, Ronald/Honer, Anne (Hrsg.)(1997): Sozialwissenschaftliche Hermeneutik. Opladen: Leske & Budrich.
- Husserl, Edmund (1966): Zur Phänomenologie des Inneren Zeitbewußtseins (1893-1917). Husserliana Bd. X. Den Haag: Martinus Nijhoff.
- Husserl, Edmund (1976a): Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie. Husserliana Bd. VI. Den Haag: Martinus Nijhoff.
- Husserl, Edmund (1976b): Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie (1913). Husserliana Bd. III/1. Den Haag: Martinus Nijhoff.

- Husserl, Edmund (1980): Phantasie, Bildbewußtsein, Erinnerung. Zur Phänomenologie der anschaulichen Vergegenwärtigung. Texte aus dem Nachlaß (1898-1925). Husserliana Bd. XXIII. Den Haag, Boston, London: Martinus Nijhoff.
- Husserl, Edmund (2001): Die Bernauer Manuskripte über das Zeitbewußtsein (1917/18). Husserliana Bd. XXXIII. Dordrecht: Kluwer Academic Publishers.
- Husserl, Edmund (2004): Wahrnehmung und Aufmerksamkeit. Texte aus dem Nachlass (1893-1912). Husserliana Bd. XXXVIII. Dordrecht: Springer.
- Imdahl, Max (1980): Giotto. Arenafresken. Ikonographie, Ikonologie, Ikonik. München: Fink Verlag.
- Imdahl, Max (1994): Ikonik. Bilder und ihre Anschauung. In: Boehm, Gottfried (Hrsg.): Was ist ein Bild? München: Fink, S. 300-324.
- Imdahl, Max (1996a): Giotto. Zur Frage der ikonischen Sinnstruktur. In: Imdahl, Max (Hrsg.): Gesammelte Schriften Band 3: Reflexion - Theorie - Methode. Frankfurt a.M.: Suhrkamp, S. 424-463.
- Imdahl, Max (1996b): Kontingenz - Komposition - Providenz. Zur Anschauung eines Bildes von Giotto. In: Imdahl, Max (Hrsg.): Gesammelte Schriften Band 3: Reflexion - Theorie - Methode. Frankfurt a.M.: Suhrkamp, S. 464-500.
- Kade, Jochen/Nolda, Sigrid (2007): Das Bild als Kommentar und Irritation. Zur Analyse von Kursen der Erwachsenenbildung/ Weiterbildung auf der Basis von Videodokumentationen. In: Friebertshäuser, Barbara u. a. (Hrsg.): Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen & Framington Hills: Barbara Budrich, S. 159-177.
- Kendon, Adam (1988a): Goffman's approach to face-to-face interaction. In: Drew, Paul/Wootton, Anthony (Hrsg.): Erving Goffman - Exploring the interaction order. Oxford: Polity Press, S. 14-40.
- Kendon, Adam (1988b): How gestures can become like words. In: Poyatos, F. (Hrsg.): Crosscultural Perspectives in Nonverbal Communication. Toronto: C.J. Hogrefe, S. 131-141.
- 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: Vandenhoeck & Ruprecht Unipress.
- Knoblauch, Hubert u. a. (Hrsg.)(2006): Video Analysis: Methodology and Methods. Qualitative Audivisual Data Analysis in Sociology. Frankfurt a.M., Berlin, Bern, Brüssel, New York, Oxford, Wien: Lang.
- Kreuzer, Max/Räder, Helga (1999): Video-Home-Training : Kommunikation im pädagogischen Alltag ; eine erprobte Methode (nicht nur) in der Familienhilfe. 2., erw. Aufl.. Mönchengladbach: Fachhochsch. Niederrhein, Fachbereich Sozialwesen.
- Leist, M. (1998): Video-Home-Training - Ein ressourcenorientiertes Angebot für verhaltensauffällige Kinder und ihre Familien. In: Verhaltenstherapie & psychosoziale Praxis, 1998 (H. 1): 69-88.
- Luckmann, Thomas (1980): Philosophie, Sozialwissenschaft und Alltagsleben (1973). In: Luckmann, Thomas (Hrsg.): Lebenswelt und Gesellschaft. Paderborn, München: Schöningh, S. 9-55.
- Luckmann, Thomas (2002): Wissen und Gesellschaft. Ausgewählte Aufsätze 1981-2002. Konstanz: UVK Verlagsgesellschaft.

- Luckmann, Thomas (2006): Some Remarks on Scores in Multimodal Sequential Analysis. In: Knoblauch, Hubert u. a. (Hrsg.): Video Analysis: Methodology and Methods. Qualitative Aduivisual Data Analysis in Sociology. Frankfurt a.M.: Peter Lang, S. 29-34.
- Luhmann, Niklas (1990): Die Wissenschaft der Gesellschaft. Frankfurt a.M.: Suhrkamp.
- Luhmann, Niklas (1995): Die Kunst der Gesellschaft. Frankfurt: Suhrkamp.
- Luhmann, Niklas (1997): Die Gesellschaft der Gesellschaft. 2 Bde. Frankfurt a. M.: Suhrkamp.
- Marotzki, Winfried (2006): Film- und Videoarbeit. In: Bohnsack, Ralf u. a. (Hrsg.): Hauptbegriffe Qualitativer Sozialforschung. Opladen & Farmington Hills: Barbara Budrich, S. 62-65.
- Marotzki, Winfried/Stoetzer, Katja (2007): Die Geschichten hinter den Bildern. Annäherungen an eine Methode und Methodologie der Bildinterpretation in biografie- und bildungstheoretischer Absicht. In: Friebertshäuser, Barbara u. a. (Hrsg.): Bild und Text. Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen, Farmington Hills: Barbara Budrich, S. 47-60.
- Merleau-Ponty, Maurice (1966): Phänomenologie der Wahrnehmung. Berlin: de Gruyter.
- Merleau-Ponty, Maurice (1986): Das Sichtbare und das Unsichtbare. München: Fink.
- Merleau-Ponty, Maurice (2003): Das Primat der Wahrnehmung. Frankfurt a.M.: Suhrkamp.
- Mitchell, W. J. Thomas (1994): Picture theory. Essays on verbal and visual representation. Chicago, London: University of Chicago Press.
- Mitchell, W. J. Thomas (2005): What do pictures want? The lives and loves of images. Chicago, London: University of Chicago Press.
- Mohn, Elisabeth (2002): Filming Culture. Spielarten des Dokumentierens nach der Repräsentationskrise. Stuttgart: Lucius & Lucius.
- Müller, Jutta (2006): Coaching, Biographie und Interaktion. Eine Studie zum Coach in Ausbildung. Opladen: Barbara Budrich Verlag (Diss. Universität Kassel 2005).
- Müller-Dohm, Stefan (1993): Visuelles Verstehen - Konzepte kultursoziologischer Bildhermeneutik. In: Jung, Thomas/Müller-Dohm, Stefan (Hrsg.): "Wirklichkeit" im Deutungsprozess. Verstehen und Methoden in den Kultur- und Sozialwissenschaften. Frankfurt a.M.: Suhrkamp, S. 458-481.
- Müller-Dohm, Stefan (1997): Bildinterpretation als struktural-hermeneutische Symbolanalyse. In: Hitzler, Ronald/Hohner, Anne (Hrsg.): Sozialwissenschaftliche Hermeneutik. Opladen: Leske + Budrich, S. 81-108.
- Nentwig-Gesemann, Iris (2007): Sprach- und Körperdiskurse von Kindern. Verstehen und Verständigung zwischen Textförmigkeit und Ikonizität. In: Friebertshäuser, Barbara u. a. (Hrsg.): Text und Bild - Methoden und Methodologien visueller Sozialforschung in der Erziehungswissenschaft. Opladen & Farmington Hills: Barbara Budrich, S. 105-120.
- Nolda, Sigrid (2007): Videobasierte Kursforschung. Mögliche Erträge von interpretativen Videoanalysen für die Erforschung des organisierten Lernens Erwachsener. In: Zeitschrift für Erziehungswissenschaft, 10 (H. 4): 478-492.
- Nolle, Reinhard/Bauer, Maria (Hrsg.)(2008): Junge Gefangene - Wir sagen aus. Biographische Geschichten, Gedichte und Rap-Songs von jungen Gewalttätern und Gefangenen Kassel: Kassel University Press.

- Nolle, Reinhard/Hildebrandt, Anke (Hrsg.)(2006): Wir sagen aus. Biographische Geschichten und Interviews jugendlicher Gewalttäter und Gefangener. Kassel: kassel university press.
- Oevermann, Ulrich (1991): Genetischer Strukturalismus und das sozialwissenschaftliche Problem der Erklärung der Entstehung des Neuen. In: Müller-Doohm, S. (Hg) (Hrsg.): Jenseits der Utopie. Frankfurt: Suhrkamp, S. 267-336.
- Oevermann, Ulrich (1993): Die objektive Hermeneutik als unverzichtbare methodologische Grundlage für die Analyse von Subjektivität. Zugleich eine Kritik an der Tiefenhermeneutik. In: Jung, Th./Müller-Doohm, S. (Hrsg.): "Wirklichkeit" im Deutungsprozeß. Frankfurt: Suhrkamp, S. 106-189.
- Oevermann, Ulrich (2000): Die Methode der Fallrekonstruktion in der Grundlagenforschung sowie der klinischen und pädagogischen Praxis. In: Kraimer, Klaus (Hrsg.): Die Fallrekonstruktion. Frankfurt a. M.: Suhrkamp, S. 58-156.
- Oevermann, Ulrich (2001): Strukturprobleme supervisorischer Praxis: eine objektiv hermeneutische Sequenzanalyse zur Überprüfung der Professionalisierungstheorie. Frankfurt am Main: Humanities Online.
- Orth, Ernst Wolfgang (1994-95): Zu Husserls Wahrnehmungsbegriff. In: Husserl Studies, 11 (No. 3): 153-168.
- Panofsky, Erwin (1979a): Ikonographie und Ikonologie (1955). In: Kaemmerling, Ekkehard (Hrsg.): Bildende Kunst als Zeichensystem, Bd. 1: Ikonographie und Ikonologie. Theorien, Entwicklung, Probleme. Köln: DuMont, S. 207-225.
- Panofsky, Erwin (1979b): Zum Problem der Beschreibung und Inhaltsdeutung von Werken der bildenden Kunst (1932). In: Kaemmerling, Ekkehard (Hrsg.): Bildenden Kunst als Zeichensystem, Bd. 1: Ikonographie und Ikonologie. Theorien, Entwicklungen Probleme. Köln: DuMont, S. 185-206.
- Panofsky, Erwin (1990): Die Renaissancen der europäischen Kunst. Frankfurt a.M.: Suhrkamp.
- Panofsky, Erwin (1992): Die Perspektive als 'symbolische Form'. In: Panofsky, Erwin (Hrsg.): Aufsätze zu den Grundfragen der Kunstwissenschaft. Berlin, S. 99-167.
- Panofsky, Erwin (1997): Studien zur Ikonologie der Renaissance (1939). 2. Aufl., Köln: DuMont.
- Panofsky, Erwin (2006): Ikonographie und Ikonologie (1932, 1955). Köln: DuMont.
- Plessner, Helmut (1980): Die Einheit der Sinne. Grundlinien einer Ästhesiologie des Geistes (1923). Gesammelte Schriften Bd. III. Frankfurt a.M.: Suhrkamp.
- Plessner, Helmuth (1981): Die Stufen des Organischen und der Mensch. Einleitung in die philosophische Anthropologie. Gesammelte Schriften Bd. 4. 1. Aufl., Frankfurt am Main: Suhrkamp.
- Raab, Jürgen (2008): Visuelle Wissenssoziologie. Theoretische Konzeptionen und materiale Analysen. Konstanz: UVK.
- Raab, Jürgen/Tänzler, Dirk (2006): Video Hermeneutics. In: Knoblauch, Hubert u. a. (Hrsg.): Video Analysis: Methodologie and Methods. Qualitative Audiovisual Data Analysis in Sociology. Frankfurt a.M.: Peter Lang, S. 85-97.
- Raphael, Max (1989a): Bild-Beschreibung. Natur, Raum und Geschichte in der Kunst. Frankfurt a.M.: Suhrkamp.
- Raphael, Max (1989b): Wie will ein Kunstwerk gesehen sein? Frankfurt a.M.: Suhrkamp.
- Reichertz, Jo (2009): Die Macht der Worte und der Medien 2. Aufl., Wiesbaden: VS Verlag für Sozialwissenschaften.

- Sachs-Hombach, Klaus (Hrsg.)(2005): Bildwissenschaft. Disziplinen, Themen, Methoden. Frankfurt a.M.: Suhrkamp.
- Schäffer, Burkhard (2003): Generation - Medien - Bildung. Medienpraxiskulturen im Generationenvergleich. Opladen: Leske & Budrich.
- Schäffer, Burkhard (2009): Abbild - Denkbild - Erfahrungsbild. Methodisch-methodologische Anmerkungen zur Analyse von Altersbildern. In: Ecarius, Jutta/Schäffer, Burkhard (Hrsg.): Typenbildung und Theoriegenerierung. Perspektiven qualitativer Biographie- und Bildungsforschung. Opladen & Farmington Hills: Barbara Budrich.
- Schegloff, Emanuel A. (1998): On talk and its institutional occasions. In: Drew, Paul/Heritage, John (Hrsg.): Talk at work. Interaction in institutional settings. Cambridge: Cambridge University Press, S. 101-134.
- Schegloff, Emanuel A. (2006): Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In: Enfield, N.J./Levinson, S.C. (Hrsg.): Roots of Human Sociality: Culture, cognition and interaction London: Berg, S. 70-96.
- Schepers, Guy (1996): Orion Video Home Training Reader. Roermond: Orion.
- Schepers, Guy/König, Claudia (2000): Video-Home-Training : eine neue Methode der Familienhilfe. Weinheim: Beltz.
- Schumm, Gerhard (2004): Montage. In: Schättle, Horst/Wiedemann, Dieter (Hrsg.): Bewegte Zeiten - bewegte Bilder. Berlin: Vistas, S. 167-169.
- Schütz, Alfred (2003a): Don Quijote und das Problem der Realität (1953). In: Endreß, Martin/Srubar, Ilja (Hrsg.): Alfred Schütz Werkausgabe Band V.1: Theorie der Lebenswelt 1. Die pragmatische Schichtung der Lebenswelt. Konstanz: UVK, S. 285-324.
- Schütz, Alfred (2003b): Symbol, Wirklichkeit und Gesellschaft (1964). In: Knoblauch, Hubert u. a. (Hrsg.): Theorie der Lebenswelt 2. Die kommunikative Ordnung der Lebenswelt. Alfred Schütz Werkausgabe Band V.2. Konstanz: UVK Verlagsgesellschaft, S. 117-198.
- Schütz, Alfred (2003c): Theiresias oder unser Wissen von zukünftigen Ereignissen (1945). In: Endreß, Martin/Srubar, Ilja (Hrsg.): Alfred Schütz Werkausgabe Band V.1: Theorien der Lebenswelt 1. Die pragmatische Schichtung der Lebenswelt. Konstanz: UVK, S. 249-284.
- Schütz, Alfred (2003d): Theorie der Lebenswelt 1. Die pragmatische Schichtung der Lebenswelt. Hg. von Martin Endreß und Ilja Srubar. 9 Konstanz: UVK.
- Schütz, Alfred (2003e): Theorie der Lebenswelt 2. Die kommunikative Ordnung der Lebenswelt. Hg. von Hubert Knoblauch, Roland Kurt und Hans-Georg Soeffner. 9 Konstanz: UVK.
- Schütz, Alfred (2003f): Über die mannigfaltigen Wirklichkeiten (1945). . In: Endreß, Martin/Srubar, Ilja (Hrsg.): Alfred Schütz Werkausgabe Band V.1: Theorie der Lebenswelt 1. Die pragmatische Schichtung der Lebenswelt. Konstanz: UVK, S. 177-250.
- Schütz, Alfred /Luckmann, Thomas (Hrsg.)(1979/1984): Strukturen der Lebenswelt. 2 Bde. Frankfurt a.M.: Suhrkamp.
- Schwedhelm, Elmar (2008): Rekonstruktive Analysen von interkulturellen Beratungsgesprächen in Uganda. Interaktionsanalysen in der internationalen Entwicklungszusammenarbeit. Kassel: kassel university press.

- Soeffner, Hans-Georg (2004): Protozoziologische Überlegungen zur Soziologie des Symbols und des Rituals. In: Schlögl, Rudolf u. a. (Hrsg.): Die Wirklichkeit der Symbole. Konstanz: UVK, S. 41-72.
- Soeffner, Hans-Georg/Raab, Jürgen (1998): Die Medialisierung des Sehens: Schnitt und Montage als Ästhetisierungsmittel medialer Kommunikation. In: Rammert, Werner (Hrsg.): Technik und Sozialtheorie. Frankfurt a.M./ New York: Campus, S. 121-148.
- SPIN/Deutschland (Hrsg.)(1994): Video-Home-Training, Reader 1: Grundlagen zu Theorie und Praxis. Düsseldorf.
- ter Horst, Klaus (2008): Dokumentation von Veränderungen in Hilfeplanverfahren durch den Einsatz der Videotechnik. In: Schwabe, Mathias (Hrsg.): Methoden der Hilfeplanung, Zielentwicklung, Moderation und Aushandlung. 2. Aufl., Frankfurt a.M.: Internationale Gesellschaft für Erzieherische Hilfen (IGFH).
- Wagner-Willi, Monika (2004): Videointerpretation als mehrdimensionale Mikroanalyse am Beispiel schulischer Alltagsszenen. In: Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung (ZBBS) (H. 1): 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: VS Verlag für Sozialwissenschaften.
- Wagner-Willi, Monika (2007): Videoanalysen des Schulalltags. Die dokumentarische Interpretation schulischer Übergangsrituale. In: Bohnsack, Ralf u. a. (Hrsg.): Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. 2. Aufl., Wiesbaden: VS Verlag für Sozialwissenschaften, S. 121-140.
- Waldenfels, Bernhard (2006): Ordnungen des Sichtbaren. In: Boehm, Gottfried (Hrsg.): Was ist ein Bild? München: Fink, S. 233-252.
- Wernet, Andreas (2006): Einführung in die Interpretationstechnik der Objektiven Hermeneutik. 2. Aufl., Wiesbaden: VS Verlag für Sozialwissenschaften.
- Wiedemann, Dieter (2005): Film und Fernsehen. In: Sachs-Hombach, Klaus (Hrsg.): Bildwissenschaft. Disziplinen, Themen, Methoden. Frankfurt a.M.: Suhrkamp, S. 365-380.
- Wiesing, Lambert (2005): Artificielle Präsenz. Studien zur Philosophie des Bildes. Frankfurt a.M.: Suhrkamp.