

Ulrike Tikvah Kissmann
Die Sozialität des Visuellen

Fundierung der hermeneutischen Videoanalyse und
materiale Untersuchungen

226 S. · br. · ISBN 978-3-942393-83-6 · 29,90 €

© Velbrück Wissenschaft 2014

(aus der Einleitung)

In der Soziologie ist Intention an Bewusstsein und Sprache geknüpft (vgl. z. B. Weber 1972). Andere Formen von Intentionalität wie diejenigen der leiblichen Habitualität werden unberücksichtigt gelassen. Während z. B. Neurowissenschaftler wie António Damásio (2000) Emotionen als vorsprachlich anerkennen, hat es sich in der Soziologie kaum durchgesetzt, dass es objektives Verhalten jenseits von Sprache gibt. Ausnahmen sind beispielsweise kultursoziologische Veröffentlichungen zum Essen und Schmecken, die Bezug zu Bruno Latour nehmen (vgl. z. B. Hennion 2007; Latour 1999). Latour hat einen neuen Handlungsbegriff entwickelt, der Sozialität nicht an das Bewusstsein eines einzelnen Subjekts knüpft, sondern eine Sozialität zwischen Objekten und Körpern ermöglicht. In der vorliegenden Arbeit wird über die »Zwischenleiblichkeit« von Maurice Merleau-Ponty ein vergleichbarer Weg eingeschlagen. Seine bipolare Phänomenologie legt die Grundlagen für objektives Verhalten jenseits von Sprache. Sie macht subjektübergreifendes Sinnschließen möglich, weil aufgrund der »Zwischenleiblichkeit« der »Pol« oder Schwerpunkt nicht mehr nur beim Subjekt liegt. Auf diese Weise können bestimmte Gesten oder eine bestimmte Mimik auf einer vorsprachlichen Ebene objektiv als Zorn oder Trauer gedeutet werden. Sie zeichnen sich also durch einen objektiven Sinngehalt aus, der für eine spezifische Gemeinschaft oder Kultur charakteristisch ist. Dieser gemeinsam geteilte Sinn ist auf einer vorbewussten Ebene intersubjektiv zugänglich, ohne dass in der Situation Sprache verwendet werden muss. Vor diesem Hintergrund versteht sich die vorliegende Schrift als ein Plädoyer, das soziohistorische Apriori von Thomas Luckmann (1980) auf diejenigen Wissensbestände zu erweitern, die für Visualität und Leiblichkeit charakteristisch sind. Es sollte demnach auch jene Regeln umfassen, die vorreflexiv und spontan sind und nicht nur diejenigen, die – wie z. B. das Inzestverbot oder die Dreifaltigkeit Gottes – dem theoretischen Wissen angehören. Die hier gemeinten Regeln beruhen auf praktischem Wissen, das im leiblichen Verhalten entsteht und in den Leib eingeschrieben wird als gewohnheitsmäßige Motorik oder Wahrnehmung.

Ziel der vorliegenden Arbeit ist es deshalb, visuelle Verhaltensäußerungen aus der Abhängigkeit von der bewussten, intendierten Handlung zu lösen und sie als eigenständige Handlungen anzuerkennen. Es geht um die Emanzipation der Gebärden vom Logozentrismus der Sprache. Zu diesem Zweck werden die Bedingungen aufgezeigt, die vorliegen müssen, damit visuellen Verhaltensäußerungen der Status von Handlungen zuerkannt werden kann und sie als eigenständige Handlungen Gegenstand von Sozialität sein können. Die vorliegende Arbeit liefert einen Beitrag zur Praxistheorie, weil der Handlungsbegriff auf vorbewusst angeeignete, visuell-leibliche Verhaltensweisen erweitert wird. Intersubjektiv geteilte Deutungsmuster werden nicht primär in der direkten Face-to-

© Velbrück Wissenschaft 2014

Face Kommunikation erworben und aufrechterhalten, sondern über den leiblichen Funktionszusammenhang in die gewohnheitsmäßige Motorik und Wahrnehmung eingeschrieben. Die Aktintentionalität von Alfred Schütz wird zu einer fungierenden Intentionalität weiterentwickelt. Letztere entfaltet ihre Wirkung nicht über den Bewusstseinsakt eines aktiven Ichs, sondern über die Passivität des Leibes. Es wird auf diese Weise eine phänomenologische Praxistheorie entworfen, in der Gesten und Mimik als eigenständige Handlungen intersubjektiv zugänglich sind. Vor diesem Hintergrund werden Praxistheoretiker wie Theodore Schatzki dafür kritisiert, dass sie visuelle Verhaltensäußerungen nur als Bestandteil der »Tätigkeit im Vollzug« konzipieren (vgl. Schatzki 1996). Der vordergründige Bezug zu »Materialität« täuscht über den Umstand hinweg, dass der eigenständige propositionale Gehalt von Gebärden nicht unabhängig von der bewussten, intendierten Handlung rekonstruiert werden kann. Schatzki hat sich mit dieser Argumentation auf Habermas bezogen, für den visuelle Verhaltensäußerungen wie das Kopfnicken-um-zuzustimmen keine eigenständigen Handlungen darstellen. Das Kopfnicken-um-zuzustimmen würde für Jürgen Habermas nur dann eine Handlung bezeichnen, wenn sie einer Intention folgt wie das Kopfnicken in der heilgymnastischen Übung (vgl. Habermas 1984).

Die Dominanz der Sprache wird auch bei Judith Butler deutlich. Durch ihren Bezug zu Derrida und seiner Rezeption der Sprechakttheorie entwickelt sie Performativität als eine Praxis des Zitierens, also der sprachlichen Wiederholungen. Ihr Konzept der Performanz lässt unberücksichtigt, dass sowohl Verletzungen als auch widerständige Handlungen in Form von visuellen Verhaltensäußerungen auftreten (vgl. Butler 2011, 1998). Goffmans Begriff der »performance« unterscheidet sich dagegen dadurch, dass der Gebrauch der Rede analysiert wird und deshalb Satz und Kontext der Rede, also Sprache und Visualität, gleichermaßen in die Konzeption der »performance« aufgenommen werden (vgl. z. B. Goffman 2005b). Sein Performanzbegriff nimmt Bezug auf das »Sprechereignis« der Anthropologen Hymes und Gumperz, die sich dezidiert gegen die Sprechakttheorie abgegrenzt haben.

Goffman wird der Forderung nach einer Emanzipation der visuellen Verhaltensäußerungen aus der Abhängigkeit der Sprache durchaus gerecht. Gleichwohl lässt sich mit seinen Schriften nicht begründen, warum Gebärden als eigenständige Handlungen Gegenstand von Sozialität sein können. Für die Beantwortung dieser Frage müssen visuelle Verhaltensäußerungen unabhängig von ihrem Zusammenspiel mit Sprache betrachtet werden. Empirisch ist dies durch die zunehmende Verfügbarkeit von Videogeräten möglich. Sie erlauben es, Interaktionen zu filmen und die Videodaten bezüglich ihrer multimodalen Elemente zu analysieren.

Mithilfe der hermeneutischen Videoanalyse wird in der hier vorliegenden Arbeit der Frage nachgegangen, ob visuelle Verhaltensäußerungen auf einer vorreflexiven Ebene intersubjektiv zugänglich sind. Es wird empirisch gezeigt, dass es objektive Gesten und Mimik gibt, denen innerhalb einer Gemeinschaft oder Kultur ein und derselbe Sinn zugeschrieben wird. Die Emanzipation der visuellen Verhaltensäußerungen von der Sprache wird theoretisch durch die Leibphänomenologie von Maurice Merleau-Ponty fundiert (vgl. z. B. Merleau-Ponty 1966). Seine Schriften sind im Rahmen der phänomenologischen Bewegung zu verstehen, die es sich zur Aufgabe gemacht hat, die objektiv-wissenschaftlichen Idealisierungen zu hinterfragen und sie auf die ursprüngliche Erfahrung in der Lebenswelt zurückzuführen. Visuelle Verhaltensäußerungen werden als »ursprünglicher« und als der Sprache vorgelagert betrachtet, so dass sie einen handlungspraktischen Zugang zu Realität eröffnen im Gegensatz zum theoretischen

Zugang durch Sprache. Auch die Ethnomethodologie hat eine Haltung zu Realität eingenommen, die sie nicht einfach als objektiv gegeben betrachtet, sondern die mittels der »Haltung der offiziellen Neutralität« (vgl. Garfinkel 1967) ihre Herstellungsmechanismen untersucht. Sowohl mit der Phänomenologie als auch mit der Ethnomethodologie lässt sich der gleichberechtigte Status von visuellen Verhaltensaüßerungen gegenüber den bewussten, intendierten Handlungen theoretisch begründen.

In der Soziologie wird gegenwärtig Video als Methode der qualitativen Datenerhebung und Datenanalyse diskutiert (vgl. Loer 2010; Knoblauch 2011). In dem Aufsatz von Knoblauch wird die Videoanalyse von der Fernseh-, Film- oder Medienanalyse unterschieden. Die Videoanalyse zeichnet sich dadurch aus, dass die Videodaten von der Forscherin selbst erhoben werden, während die Fernseh-, Film- oder Medienanalyse von z. B. Reichertz/Englert (2011) oder Keppler (2006) fertige Videodaten verwendet, die von anderen erstellt wurden. Es gibt aber auch Arbeiten wie diejenigen von Bohnsack und Mitarbeiterinnen (vgl. z. B. Wagner-Willi 2005, Klambeck 2007, Bohnsack 2009), die auf der Grundlage ein und derselben Methodologie sowohl Interaktionsanalysen als auch Filmanalysen durchführen. In der hier präsentierten Videoanalyse werden wie in der klassischen Interaktionsanalyse von Heath u. a. (2010) oder auch von Goodwin (2009) die Videos zu Forschungszwecken selbst hergestellt. Der Videoaufzeichnung geht prinzipiell eine mehrwöchige Feldphase voraus, in der teilnehmende Beobachtung durchgeführt und Interviews erhoben werden. Der Erwerb von Kontextwissen über das Forschungsfeld ist sowohl für die Datenerhebung als auch für die Datenanalyse relevant. Das Kontextwissen wird erstens dazu verwendet, um z. B. Entscheidungen zu treffen, wo die Kamera am sinnvollsten zu positionieren ist. Die Stellung der Kamera trägt mit dazu bei, den Forschungsgegenstand zu konstruieren und wird deshalb über das Kontextwissen abgesichert. Zweitens werden, wie in Kapitel 5 deutlich wird, in der Auswertung der nonverbalen Interaktion auf der Ebene der ikonographischen Deutung und später beim abschließenden Vergleich der Analyseschritte Informationen aus dem Forschungskontext miteinbezogen. Die hier präsentierte Videoanalyse kommt also nicht ohne Kontextwissen aus und ist deshalb immer im Rahmen einer Ethnographie durchzuführen. Dennoch wird bewusst nicht der Terminus »Videographie« von Hubert Knoblauch verwendet, weil hier neben der Videoaufzeichnung und Videoanalyse eine fokussierte Ethnographie durchgeführt wird, die auf eigenen methodischen und methodologischen Überlegungen beruht (vgl. z. B. Knoblauch/Tuma 2011).

In der hier präsentierten hermeneutischen Videoanalyse werden der vorikonographische und der ikonographische Sinngehalt voneinander getrennt. Auf diese Weise lässt sich empirisch überprüfen, ob visuelle Verhaltensaüßerungen auf einer vorreflexiven Ebene intersubjektiv zugänglich sind. Die unterschiedlichen Sinnschichten werden über Erwin Panofskys Dreistufenmodell der Interpretation eingeführt (vgl. z. B. Panofsky 1975). Danach entspricht der vorikonographische Sinngehalt dem praktischen Wissen, während die ikonographischen Deutungen auf narrativ-theoretisches Wissen zurückgreifen. Diese Sichtweise auf soziale Realität, in der alle sozialen Gebilde über zwei Formen von Sozialität zugänglich sind, wird durch die Leibphänomenologie von Merleau-Ponty theoretisch begründet. Er unterscheidet die Zwischenleiblichkeit vom Dialog. Die erstgenannte Form von Sozialität bezeichnet das triadische Verhältnis von Eigenleib, Fremdleib sowie der Welt und ist im Bereich des praktischen Wissens angesiedelt. Merleau-Ponty charakterisiert diese Sinngehalte auch als »gestische Bedeutungen«. Im Gegensatz dazu rekurriert der Dialog auf narrativ-theoretisches Wissen und umfasst »begriffliche Bedeutungen« (vgl. Merleau-Ponty 1966). In der hermeneutischen Videoanalyse wird durch die gegenseitige Validierung von

vorikonographischer und ikonographischer Ebene die dritte Sinnebene nach Panofsky, die so genannte Ikonologie erreicht. Hier wird der ikonologische Sinngehalt einer Interaktion bzw. der Habitus der Interaktionspartner rekonstruiert.

Im interpretativen Paradigma von Alfred Schütz ist die »theoretische Distanz« zwischen Alltagsbeobachterin und wissenschaftlicher Beobachterin die Voraussetzung dafür, den Common Sense für wissenschaftliche Erklärungen zu benutzen (vgl. Luckmann 1981). Unter Bezug auf die Desiderate von Bohnsack (2006) und Raab (2008) wird die »theoretische Distanz« der Hermeneutik einer Feinabstimmung unterzogen. Es wird mit Hilfe des Dreistufenmodells von Panofsky und der Ethnomethodologie Garfinkels einerseits gezeigt, worin sich Alltagsbeobachter und wissenschaftlicher Beobachter genau unterscheiden und wie die Standortgebundenheit und die Deutungsinteressen des wissenschaftlichen Beobachters in der Datenanalyse zu berücksichtigen sind. Andererseits wird begründet, an welchen Stellen in der Datenanalyse der Kontext einzuklammern bzw. hinzuzunehmen ist, um die für die hermeneutische Methode charakteristische Suchbewegung zwischen Fremdheit und Vertrautheit zu gewährleisten.

Nach der hier vorliegenden Einleitung werden in Kapitel 2 die performativen Eigenschaften von Visualität in Interaktionen behandelt. Das Zusammenspiel von Sprache und Visualität steht dann im Zentrum der Darstellung, so wie es in Alltagssituationen auftritt ohne die durch Video ermöglichte Trennung der multimodalen Elemente. Aber nicht nur das Gesehene und Gehörte sind Bestandteil der Darstellung, sondern insgesamt alle fünf Sinne. Es wird dann in Goffmans Interaktionssoziologie eingeführt und sein Begriff der »performance« dargestellt im Vergleich zu Butlers Performanzbegriff. Außerdem werden die geschlechtersoziologischen Ansätze präsentiert, die in der Tradition von Goffman stehen. Dazu gehören die sozialpsychologischen Schriften von Cecilia Ridgeway und Shelley Correll, die in Kapitel 6 für die Auswertung der von mir geleiteten Krankenhausstudie benutzt werden.

In Kapitel 3 werden die theoretischen Vorannahmen behandelt, die zu der Durchführung der hermeneutischen Videoanalyse notwendig sind. Dies ist Panofskys Dreistufenmodell der Interpretation mit den drei Sinnschichten der vorikonographischen, ikonographischen und ikonologischen Ebene. Es wird dargestellt, in welcher unterschiedlichen Weise die drei Auslegungsstufen Bildlichkeit ausdrücken. Außerdem wird auch Mannheims Dreistufenmodell von Deutung präsentiert, weil beide Ansätze in Bezug zueinander entwickelt worden sind und deshalb große Ähnlichkeiten aufweisen. Schließlich wird Bourdieus Rezeption der Ikonologie zum Thema gemacht. Es werden die Gemeinsamkeiten und Unterschiede von Bourdieus und Panofskys Habitus herausgearbeitet.

In Kapitel 4 werden die Bedingungen aufgezeigt, die vorliegen müssen, damit visuellen Verhaltensäußerungen der Status von Handlungen zuerkannt werden kann und sie als eigenständige Handlungen Gegenstand von Sozialität sein können. Über die Ethnomethodologie wird theoretisch und methodologisch die AnalyseEinstellung eingeführt, die gegeben sein muss, um die Motivkonstruktionen von visuellen Verhaltensäußerungen zu rekonstruieren. Alfred Schütz hat die zweckrationale Konstruktion eines Um-zu-Motivs als Merkmal einer Handlung betrachtet. Sein Handlungsbegriff wird auf vorbewusst angeeignete visuell-leibliche Verhaltensäußerungen erweitert. Über die Phänomenologie wird theoretisch und konstitutionsanalytisch begründet, warum Gesten und Mimik der Status von Handlungen zuzuschreiben ist und warum sie deshalb als eigenständige Handlungen Gegenstand von Sozialität sein können.

Kapitel 5 stellt die einzelnen Analyseschritte der hermeneutischen Videoanalyse dar. Die präsentierte Sequenz beinhaltet die Interaktion eines leitenden Oberarztes mit einer leitenden OP-Schwester. Sie ist Bestandteil eines Datencorpus mit Videomaterial von insgesamt 400 Stunden aus dem von mir geleiteten DFG-Krankenhausprojekt. Es wird auf die Datenerhebung und die Auswahl der zu analysierenden Sequenz eingegangen. Im Gegensatz zu Raab (2008) und Bohnsack (2009) wird keine Bild-in-Bild Interpretation durchgeführt, sondern eine Segment-in-Segment Interpretation. Dies ermöglicht, dass Interaktionen in ihrem »flow« analysiert werden können. In dem Kapitel wird anhand der Segment-in-Segment Interpretation vorgeführt, wie durch gedankenexperimentelle Antizipation Handlungsalternativen für das nächste Segment entwickelt werden. Anfangs noch enthaltene, später dann ausgeschlossene Deutungsmöglichkeiten erlauben dann eine allmähliche Schließung der Deutungen. Anhand der Ellenbogengeste wird empirisch gezeigt, dass es objektive Gebärden gibt, die auf der vorikonographischen Ebene intersubjektiv zugänglich sind.

In Kapitel 6 wird die Ellenbogengeste in den breiteren Forschungskontext des von mir geleiteten Forschungsprojekts gestellt. Es wird gezeigt, worauf die habituellen Probleme der leitenden OP-Schwester zurückzuführen sind und wie es ihr letztendlich gelingt, die an sie herangetragenen Statuserwartungen zu umgehen. Es werden die Forschungsergebnisse aus den Operationssälen von insgesamt zwei Krankenhäusern dargestellt.

Die Zusammenfassung in Kapitel 7 beinhaltet nicht nur eine Synopse der wichtigsten Ergebnisse der vorliegenden Arbeit, sondern sie stellt mit dem sozio-historischen Apriori für visuelle Verhaltensäußerungen auch den Schluss dar, der aus den hier präsentierten Ergebnissen gezogen werden muss.