

Co-Operation is a Feature of Sociality, not an Attribute of People

“We inhabit each other’s actions.” (Goodwin, cover)

Jutta Wiesemann, Klaus Amann

Rather than philosophizing on classic sociological questions like how society is structured, what holds it together, or how it can function at all, Charles Goodwin repeatedly posed the concrete, empirical question: How is social order continuously manifested? He sought answers to this question by analyzing the minutae of observable human actions. His observations led him to describe the relationships of actors to their actions as a (reciprocal) habitation. Inhabiting means a living experience. As a concept, inhabiting acknowledges the interrelationship between the action taking place at any given moment and its cultural context, which has already been produced, and at the same time is continually being created anew. Goodwin conceptualizes his inhabitants as actors living in the observable present, who constantly draw on the former achievements of their predecessors and make use of available material and cultural resources in order to accomplish their own actions. Furthermore, inhabitants’ current actions impact upon others present, so that actors “[...] actively participate in the detailed organization of each other’s action as it unfolds through time.”(7) Being able to use, adapt, and transform resources from the past and the present enables what Goodwin calls co-operative action. “By building our own actions by using resources provided by others, we live in a world where we inhabit each other’s actions.” (78) This inhabiting is always conceived of as shared and reciprocal. The mutually constitutive nature of inhabiting our own and each other’s actions necessarily implies that actors are al-

ways able to recognise “[...] how another is analyzing and understanding the world that is the focus of action.” (319)

Goodwin’s starting point and the core of his empirical approach is linguistic communication: “talk”. This can be found everywhere in the inhabited social world. As far as he is concerned, it makes no difference who produces this talk, who exchanges it with whom, and in which institutional or other contexts the specific talk arises. For his research, the co-operatively emerging speech is of interest, not the speakers. Nor does he focus on language as a linguistically reconstructed and ordered system. Rather, Goodwin regards language as a resource that is constantly being transformed by those who are obliged to use it together for communication, producing “public” talk as they do so.

While audible utterances – which co-produce mutual comprehension – are thus at the core of his approach to sociology, they are nonetheless always embedded within and dependent upon other actions in a variety of ways. No linguistic utterance is self-explanatory or comprehensible in itself. In other words: I can neither say what I mean, nor can another person understand what I am saying. For something that is said to be understood, certain prerequisites must be met – and these are unique to any particular situation. Each situation has its own configuration of historical, social, and, most importantly, interactively embodied specificities. In particular, the physical-gestural phenomena that accompany every instance of talk must be included in its microanalysis, if that talk is to be understood. “Here, I want to explore the possibility that all of these phenomena are different manifestations of the distinctive ways in which human beings build co-operative, accumulative action in concert with each other.” (2) His unifying model of everything and everyone rests upon the notion of accumulation: “co-operative, accumulative action.” Social life is a kind of workshop, in which accumulative action is inevitable: it has always been carried out, and always will be. Here, Goodwin’s methodologically valuable contention is that this accumulative action can be observed. The same assumption forms

the basis of the classic ethnomethodological premise: that the public nature of social interaction is what makes it verifiable.

As Goodwin elaborates upon the complexity of co-operative action, he draws an analogy to developments in biological cellular research over decades. While DNA was conceptualized for a long time as the language or code of life in which evolution had inscribed itself, the current understanding is that an actual interpretation of that code can only be realized within actual living cells and organisms.¹ What is encoded can only be expressed – i.e., converted into cell activity – in accordance with a real, specific situation. And analogous to what a cellular biologist might describe, a Goodwinian observer does not describe real life as enacted by human actors as cooperation in the sense of deliberate or intentional collaboration. Instead, actions are described as a coordinated series of interdependent actions that occur in time, and bring forth sociality.

Hopscotch and Public Visibility

“[C]ontextual configurations provide a systematic framework for investigating the public visibility of the body as a dynamically unfolding, interactively organized locus for the production and display of meaning and action.” (170)

We know that hopscotch does not normally attract the interest of the public (or of scientists), except when it is done by First Ladies. This photo published by the White House expresses a multi-dimensional political message. There is something ‘behind’ the medium/image that we can interpret as a message from the White House.

Goodwin draws on a specific instance of “playing hopscotch on the street” to explain his sociological concept of co-operation. He is not interested in what lies ‘behind’ the publicly visible performance.³ What he finds interesting sociologically is the visible and audible, the co-op-



Fig. 1: Michelle Obama playing hopscotch²

erative how of the performance. As he sees it, this “how” is generated in different ways by actions as they are enacted over time: as talk, with bodies, and with the materiality of the hopscotch grid. The result of his analysis is a microscopically crafted reconstruction of directed or oriented attention. “I will call some particular subset of possible fields that is being oriented to at a particular moment as relevant to the organization of a particular action a *contextual configuration*.” (180)

This supposedly trivial example makes it clear that the analyzed visible how cannot be deduced or explained in terms of an inferred intentionality on behalf of the participants, but only from what is mutually “shown” (made visible). “[It] is an ongoing contingent accomplishment, something not under the control of a single party.” (ibid.)

Goodwin’s analytical goal is not to discover intentionality ‘behind’ the observable, or the choreographing of individuals/persons. Rather, it is to reveal and analyze the techno-logics that shape the situated

co-operating of sociality. According to Goodwin, interaction ultimately means that visible co-operative action is in progress. “[...] [P]ublic visibility is also crucial to analysis of how the body is used to perform action within interaction.” (187) Significantly, this means that the practical actions of a single isolated individual are not enough to produce social praxis that can be understood as such. There must always be a “reflexive awareness of each other” (185) that is itself accessible for observation and analysis.

The “public visibility” of actions is a key methodological assumption for Goodwin, rooted in his own work in social theory. “Public” refers to an interactively necessary, technical characteristic of human action. Whether we like it or not, we make our activities accessible to others when we act in situations in which others can observe us. Goodwin’s description of this as a ‘technical’ characteristic reflects his take on observation, especially with regard to the embodied dimensions of action. At the same time, it is influenced by the way he uses audiovisual recordings, which play a central role in his research.

Goodwin proceeds to develop a profoundly technical conceptualization of human action. Co-operating is the generalized, underlying characteristic of a fundamental, “technical” productive process. Wherever he looks – through the lens of the linguistic empiricist – he sees complex, technical productive processes; processes that only make sense if they are seen as co-operation (both parts of the term are stressed equally: co- and operation). Inherent to this technicist conceptualization of action is a generalized concept of materiality. The latter goes beyond an everyday understanding of materiality as ‘matter’ or ‘substance’, as opposed to non-material phenomena such as intentions, meanings, or speech. Goodwin uses the notion of talk in particular to argue that words, tones of voice, gestures, looks and so on must be understood as culturally specific (raw) materials from which social situations can be co-operatively produced. “This process of building new action by performing systematic operations on something created by someone else is what is being

investigated here as co-operative action.” (431) “Performing systematic operations” (2) is a technical *modus operandi* that brings forth sociality.

In his writings on technology and philosophy, Ernst Cassirer had already advocated an approach along such lines in 1930: “If, instead of beginning from the existence of technological objects, we were to begin from technological efficacy and shift our gaze from the mere product to the mode and type of production and to the lawfulness revealed in it, then technology would lose the narrow, limited and fragmentary character that otherwise seems to adhere to it.” (20) He continues: “The ‘form’ of the world, whether in thought or action, whether in language or in effective activity, is not simply received and accepted by the human being; rather, it must be ‘built’ by him. (24)

Following Cassirer, we can identify the unity in the thematic diversity of Goodwin’s work: whatever the context, for Goodwin it is always about the elaborate technical analysis of co-operative ‘forms’ and ‘formations’ of (inter-)action. The (paradigmatic) starting point for these analyses is always a visible situation that is witnessed by a co-present observer. Ultimately, sociality is not an abstract notion, but something concrete that unfolds in time, is actively formed, and as such is intelligible.

Thus, Goodwin sees no need to speculate on the hidden origins of human action in the workings of the internal mind. Focusing instead on the diversity of observable forms of action, he is concerned with “arrangements of mutual elaboration”. (429)

Further Methodological Consequences and Camera Ethnography

Goodwin’s empirical approach demonstrates the efficiency of a radically microsociological (linguistic) strategy of discovery. This strategy strongly emphasizes the situatedness of action. To make situated action accessible for observation and analysis beyond the moment it takes place, Goodwin relies on the audiovisual recording and documen-

tation of co-operative actions. The recording situation itself is a shared 'situated doing' that brings together observers, recording devices, and observed actors. However, these different participants do not usually share the same agenda. The technically-equipped observer, unlike the observed, participates in order to produce and preserve some kind of "substrate" of the jointly experienced activities in a technical form that can be worked with further at a later date.⁴

The situatedness of co-operative actions, however, is about more than just the bodily co-presence of humans in a physical-spatial sense. In our view, participants' ongoing activities must also be understood as that which creates and maintains a shared social place by utilizing various resources. In order to create such a place, participants must jointly make interactional, material, and occasionally technically- or media-related preparations. This is particularly relevant when the situation is not just a face-to-face encounter as would be paradigmatic according to Goffman, but one in which mediatised forms of communication play a central role in restructuring interactional orders.⁵

In our ongoing project "Early Childhood and Smartphone. Family Interaction Order, Learning Processes and Cooperation", we observe media appropriation and use in family settings with children aged between 0 and 6 years old. As a research strategy, we concentrate on the smartphone as the mediator and central focus of interaction processes involving adults and children. As a ubiquitously available, globally distributed and used medium, the smartphone contributes to and fosters the "cooperation" of family members (Schüttpelz 2016). Our conceptualization of cooperation, rooted in media theory, underlies our endeavours to describe the "mutual making of joint goals, means, and processes" (ibid. 6) enabled by media, which we trace in the interactions we observe. Our aim is to observe and identify the conditions within which families use various (not only digital) media, and the practices that they develop. We take a radically empirical stance by observing actual action rather than speculating on intentions or evaluating pre-

scriptively. Following Goodwin, we are less concerned with the individual motives of family actors or the pedagogical effects of what we observe than with gaining an in-depth, structural understanding of actual real situations of appropriation and use. We conceptualize the smartphone as a medium of production – not least in opposition to the idea that it offers limited pre-defined functions, the operation of which adults simply teach to small children as part of a learning process understood as a unilateral imparting of knowledge.

As we observe the use of media, we trace – following Goodwin – co-operative actions as they unfold. These actions become recognizable as necessarily teachable and learnable modes of use. The participants are dependent on one another as they orient, develop, and coordinate their interactive practices. The resources utilized in the actions that make up this “mutually constitutive” co-operation are – as our observations show – manifold ways of using the body: facial expressions, gestures, and entire choreographies of bodies oriented towards the respective medium.

Goodwin dedicates an entire chapter of his book to an exploration – drawing on research from the past few decades in the sociology of scientific knowledge – of how **professional vision** is established among scientific communities, such as geologists, geochemists, and archaeologists. (Goodwin 2017, Part V: 325-428) “However, to function in the social life of a profession the ability to see relevant structure in a complex environment must be organized, not as an idiosyncratic individual ability, but instead as systematic public practice.” (ibid. 349) In our own research, **camera ethnography**, a specific visual anthropological approach developed by Bina Mohn, is a key method that enables us to establish professional vision as a “systematic public practice”. With camera ethnography, we aim to see and to show “relevant structure in a complex environment” in sequences of family interaction involving the media we are interested in. The following three stills from one of our research films illustrate this specific way of seeing and showing (Fig. 2).



Fig. 2: We can observe what is visible, including the way that mother and child show each other their awareness of each other (video stills by Astrid Vogelpohl 2017: a, b, c)

These stills are taken from a camera observation made in one of the families we have been working with since the beginning of our research project. As recipients of any text, or indeed any (still or moving) image, readers are dependent on various operations undertaken by its author(s), including their selective decisions. For their part, our camera ethnographers certainly have to ensure that what they create is visible and legible. And that further images / videos of actions observed in this research context will also be visible and legible. This process of making visible and legible encompasses not only the usual practical, methodological, and technical procedures required for audiovisual research (such as establishing contact/obtaining consent, making appointments, preparing equipment, and so on), but also – and this is what sets camera ethnography apart from other ways of recording video – it requires actively observational camerawork.⁶ In any observational situation, the camera ethnographer is dependent on the “public visibility” of the actors’ actions. But that does not require an extra effort on behalf of the actors, rather, it is the fundamental, technical *modus operandi* of sociality to make one’s actions visible to those present. Actively observational camerawork requires (in addition to technical proficiency) a researcher to sustain a concentrated, constant ethnographic orientation towards the visibility of actors’ actions as they unfold over time. The framing of the first still, **a** (Fig.2), reveals this ethnographic orientation: the camera is directed towards the toddler, woman, and smartphone from a position slightly below the eye level of the mother, who is kneeling in the middle of the room. Mother, child, and smartphone are in focus, and remain so in the two subsequent stills (**b** and **c**). Yet the frame includes more than just these three focused subjects; it also shows the contents of the living room including a range of toys on the right, which the child has already turned away from in still **a**.

The story that this series of three stills tells and makes legible is one in which a child is invited to engage in a joint activity with the smartphone.



Fig. 3: Co-operative action (d, e, f)

In the subsequent series (Fig. 3), the ethnographer concentrates exclusively on showing the focused actors (mother, child, smartphone). The mother's hands release the smartphone, but remain ready to catch it should it fall. The joint activity is made publicly visible by the clear positioning of the camera's ethnographic selection. Still **f** (Fig. 3) shows the changed positioning of the mother and the child in relation to the smartphone as compared to still **a** (Fig. 2).

The image sequence as a whole also serves to illustrate the meaning of "mutual copresence" within Goodwin's understanding of sociality: "Building action in a state of mutual copresence is a central site for the ongoing, dynamic constitution of human sociality, and the place where a host of phenomena, including language and other forms of semiosis intertwined with it, emerge within the mundane social world." (248)

Beyond Language and Text

"However, to function in the social life of a profession the ability to see relevant structure in a complex environment must be organized, not as an idiosyncratic individual ability, but instead as systematic public practice." (349) In the sub-chapter "Calibrating Professional Vision", Goodwin highlights the importance, for diverse communities of experts such as scientists and lawyers, of establishing a shared, standardized professional vision. In relation to his own profession, however, such endeavours are conspicuously absent in Goodwin's book. There is no methodological self-reflection, for example, or deconstruction of what led him to create certain graphic visualizations or to draw particular conclusions. Despite relying mainly on audiovisual material for his empirical sources, Goodwin does not address the active use of the camera as an instrument for creating professional visibility. With our use of camera ethnography, by contrast, we conceptualize what happens during a situation of observation as the first, active phase of a situated production of professional vision, which is followed by further analytical phases of working on and with the observational material. This approach opens

up new perspectives on knowledge and knowledge production, which can be communicated in audiovisual forms that can reach far beyond the realm of written texts or graphic signs.

Notes

- 1 James E. Darnell, Harvey Lodish, David Baltimore: *Molekulare Zellbiologie*. de Gruyter, Berlin u. a. 1993, ISBN 3-11-011934-X (4. Auflage. Harvey Lodish: *Molekulare Zellbiologie*. Spektrum Akademischer Verlag, Heidelberg u. a. 2001)
- 2 Source: The White House from Washington, DC - P110610CK-0355, <https://commons.wikimedia.org/w/index.php?curid=4633235>
- 3 In his farewell lecture, systems theory theoretician Niklas Luhmann asserted that sociology needs to abandon the distinction between "What is the case?" and "What lies behind it?" (Luhmann and Fuchs 1994).
- 4 A further observer watching such a recording situation would be able to discern at least two simultaneous situations: actions of the "first order" and of the "second order", whereby the latter is concerned with encapsulating the observability of the former and fixing it as a stable document.
- 5 Justifying his claim that face-to-face interaction is the norm from which other situations deviate, Erving Goffman writes: "Such a phenomenon as talking to oneself, or talking to unrati-fied recipients as in the case of collusive communication, or telephone talk, must first be seen as a departure from the norm, else its structure and significance will be lost." (Goffman 1964: 13).
- 6 A more detailed account is given by Mohn (2013)

References

Cassirer, Ernst (1930): *Form and Technology*. Accessed 19 September 2018. https://monoskop.org/File:Cassirer_Ernst_1930_2012_Form_and_Technology.pdf

Goffman, Erving (1964): "The Neglected Situation", in: *American An-*

thropologist (New Series) 66 (6), Part 2: The Ethnography of Communication, pp. 133-16.

Luhmann, Niclas and Stephen Fuchs (1994): "What is the Case?' and 'What Lies behind It?' The Two Sociologies

and the Theory of Society”, in: *Sociological Theory* 12 (2), pp. 126-139.

Mohn, Bina Elisabeth (2013): “Differenzen zeigender Ethnographie. Blickschneisen und Schnittstellen der Kamera-Ethnographie”, in: Schnettler,

Bernt / Baer, Alejandro (Hg.): “Themenheft Visuelle Soziologie”, *Soziale Welt* 64 (1-2), pp. 171-189.

Schüttpelz, Erhard (2017): “Infrastructural Media and Public Media”, in: *Media in Action* 1(1), pp. 13-61.