Instructed and Instructive Actions

The Situated Production, Reproduction, and Subversion of Social Order

Edited by Michael Lynch and Oskar Lindwall

Chapter 2

Detail, Granularity, and Laic Analysis in Instructional Demonstrations

2 Detail, Granularity, and Laic Analysis in Instructional Demonstrations

Oskar Lindwall and Gustav Lymer

Introduction

The overarching interest of the present chapter is in the *description of embodied courses of action.* More specifically, we focus on the instructional descriptions of what sometimes is referred to as manual or instrumental actions; that is, actions done by the hands and for other purposes than communication. The three examples that we use are taken from two different settings: an introductory course in endodontics and a YouTube tutorial on how to crochet. The instructional demonstrations found in these settings make perspicuous several themes central to this volume: how demonstrations rely on what "any member would know"; how they are contingent on competences that are yet to be instructed; how they constitute members' analyses of skills and practices; how they are hopelessly incomplete; how they provide "mock-ups" of the activities they set out to demonstrate; and how they, therefore, are specifically useful for instruction. Although the chapter touches on each of these themes, the cases that we focus on are chosen because they show distinct relationships between descriptions and embodied courses of action. In all the examples, instructional descriptions are occasioned by manual actions, but they vary in the extent to which the sense of a description relies on the details of the displayed actions, and while instrumental actions in some demonstrations are produced independently of their description, there are other situations where descriptions and embodied courses of action mutually elaborate each other.

   In addition to this interest in descriptions as part of instructional demonstrations, we also turn to our own practices of description. The next section begins by situating our interest in the description of actions as a standing concern within ethnomethodology and conversation analysis. Throughout this chapter, we discuss how professional sociological analysis trades on and differs from the analysis produced by members themselves in the course of demonstrations: how our analyses of action are shaped by the fact that "ordinary cultural members are the first analysts on the scene" (Macbeth 2007: 200), and how the visual and embodied details of "say-shown demonstrations" (Burns 2012: 184) unavoidably are different from those that fit the printed page.

Descriptions and the recognizable sense of actions and activities

The relation between descriptions and actions has been central to ethnomethodology from the outset. In *Studies in Ethnomethodology*, Garfinkel (1967) discusses a task in which students were to provide descriptions of the common understandings implicit in the actions making up everyday conversations. For each submitted attempt, Garfinkel urged the students to develop their descriptions further, to make them more detailed, and to iron out any ambiguities. The students soon realized that the task was endless, and they complained about the fact that each addition to the descriptions yielded further material for explication: "the writing itself developed the conversation as a branching texture of relevant matters" (Garfinkel 1967: 26). For Garfinkel, the exercise demonstrated, among other things, the essential incompleteness of accounts, and how the "recognized sense" of an utterance in conversation cannot lie in an explication of meaning abstracted from the action, but in an appreciation of the particular way the action was executed, the evident "method of speaking." Consequently, such descriptions should not be treated as approximations of a substantive content – "what the parties talked about" as underlying "what they said." Instead, Garfinkel treated the students' descriptions as *instructions*:

> [T]heir written explanations consisted of their attempts to *instruct me in how to use what the parties said as a method for seeing what the conversationalists said*. I suggest that I had asked the students to furnish me with *instructions for recognizing* what the parties were actually and certainly saying.
>
> (Garfinkel 1967: 29, emphasis added)

The notion of instruction introduced by Garfinkel in this passage provides a useful entry point to the substantive topic of this chapter: instructional descriptions of actions. His students engaged in attempts to explicate "what any member knows."[1] These explications were instructions for recognizing something that was already plainly visible to the competent member, for whom the recognized sense of an action does not depend upon, or require, descriptive elaboration. In instructional demonstrations addressed to novices, however, descriptions of actions serve more substantive ends, such as guiding the novice to see what is "actually and certainly" done. As an object abstracted from its circumstances, any description or account is hopelessly incomplete (see Garfinkel & Sacks 1970). As illustrated by Garfinkel's exercise, each added element to a description provides material for further elaboration. In almost any other circumstance, however, the task is not to provide endless elaboration, but to furnish descriptions specifically designed for that occasion. With "person reference" as an initial example, Schegloff (2000a) points to the relevance of studies that examine how members refer to or formulate elements of their immediate environment or past experience. A speaker refers to a mutual friend, for instance, not by providing an exhaustive description of who that person is, but by using a first name or some other category

or description that shows that the two parties know whom they are talking about. In a footnote, Schegloff offers the following quotation taken from a story told by a man to two friends:

"He said, 'well I drove it down to this car show, uh someplace in Ohio.' And uh, he got down in it, and the engine heated up and blew on the way back. Took it up, tore the damn thing apart, and found a rag stuffed in the radiator hose."

(Schegloff 2000a: 718, note 8)

Schegloff notes how the story changes during the telling. Before the point where the engine blew, the whole trip is glossed as "I drove it down to this car show," whereas the latter part, "took it up, tore the damn thing apart, and found a rag stuffed in the radiator hose," represents a finer granularity, which coincides with the point and climax of the story. Schegloff argues that examining how members formulate elements of their environment, including actions and activities, and the varying granularity with which this is done, gives access to some of the terms and orders of relevance that shape the experience of this environment. Following the recommendation to examine how members formulate their immediate environments, but in ways adapted to our interest in embodied courses of action, we now turn to a set of examples of members' descriptions of manual activities.

Consider first Figure 2.1 and Extract 2.1, taken from a recording of a demonstration seminar in a course in endodontics – or root canal procedures – for students at a dental education program. In Figure 2.1, there are two frame grabs where a seminar leader uses a stylus to point at a live stream of the early phases of an endodontic procedure that is taking place in an adjacent room. Extract 2.1 is an English translation of the Swedish original and provides explicative descriptions of the ongoing actions visible in the live video.
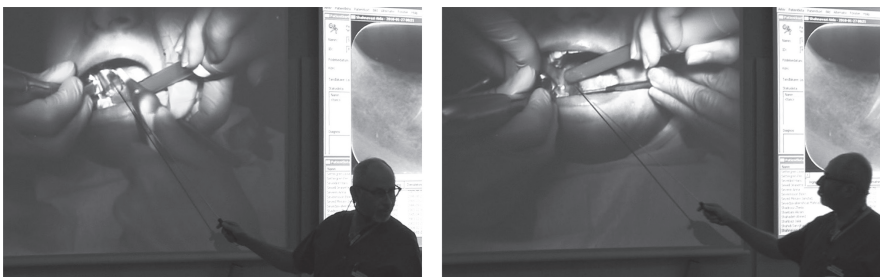
Figure 2.1 On the left, the seminar leader points at the mirror while saying "blasting all the time on the mirror." On the right, he points to the mirror while saying "use it for other things."

Extract 2.1   [END100311-00:09:06]

INS: You see the dental assistant, she's blasting all the time on the mirror. Because when I now drill with this speed-drill a lot of water is spurting. And then I must have the mirror because otherwise I can't see. And then if the assistant blasts on the mirror she is blasting away the film of water so then one can see well in the mirror. ((8 second pause)) And the mirror, here you can see that you can use it for other things: keep the tongue away […] All of us who have worked as a dentist for a couple of years have drilled at least one patient in the tongue. […] The patient swallows, then the tongue goes up, swish. It heals, nothing much happens.

While Schegloff's example illustrates how members describe past experiences, and how the granularity of the description is correlated with the production of the story climax, this instructor's description is occasioned by the ongoing performance: in the details of the video, the seminar leader continuously searches for and finds topics that can be turned into instruction. One could note that the instructor initially describes the visible actions of the dental assistant in plain terms: "blasting all the time on the mirror." "Blasting" (Sw. *blästra*) is an action categorization, and thus the description indicates in a preliminary fashion *how to see* a visible element in the video: that is, how the dental assistant is using the air water syringe to continuously "blast" air on the mirror. Although not going very far in furnishing the audience with "instructions for recognizing" what the assistant is doing as a transparently motivated and functional action within the procedure, this scenic account tells the students what to look at and sets up the relevance of listening to what is said next, as further instructions on how to see and recognize what is done. In line with this, something further, similar to what Garfinkel's students produced in "explicating common knowledge," immediately ensues. The instructor moves on to expand the reasons for and motivations behind the action, as embedded within the relevancies of the procedure: that the blasting is done to clear the view of the operating dentist and that the problem addressed through "blasting" arises because of the use of a particular type of drill.

After the first description, and following a longer pause during which the seminar leader inspects the video, a second descriptive segment follows: "And the mirror, here you can see that you can use it for other things. Keep the tongue away." Although brief, this description characterizes the action in terms of its effect: the tongue is kept away. Already in this brief gloss, then, we find elements of an analysis in terms of recognizably accountable motives. As in the prior segment, the instructor expands shortly after, explicating and contextualizing, among other things, what the tongue is kept away from, and why this is important. In both these segments, the narrowly descriptive part of the account is relatively brief, in comparison to remarks that situate the described action in terms of its motivation, typicality, and various contextual ramifications. Apparently, members in this instructional setting treat explications of motivations and the rest as necessary for

their descriptions to operate, paraphrasing Garfinkel (1967), as instructions for recognizing what the dentist is actually and certainly doing.

Returning to the issue of granularity raised by Schegloff (2000a), the level of detail in these members' descriptions becomes finer when explicating what is *not immediately visible* in the jointly observable action. Categorization of visible actions is done relatively briefly; once pointed out, anyone can see the "blasting" as plainly and transparently what the nurse is doing. It is also plainly visible (again, once pointed out) that the mirror is used to reposition the tongue. With the aid of the instructor's description, this repositioning can be understood as done to "keep the tongue away." Granularity increases, however, in the ensuing expansions. Here, we are reminded of Garfinkel's exercise discussed above and the essential incompleteness of the accounts that his students produced; just where to stop in explicating endodontic context and competence seems to be an open-ended question for the members of this scene, moderated perhaps by the practical concerns of "keeping up" with the ongoing operation.

To further illustrate the availability to the expert observer of "the invisible," we would like to draw attention to the recurrent use of a psychological vocabulary in these narratives. The instructor repeatedly (in the larger data set from which the example above is drawn) describes the dentist's actions in terms of what they "want," "know," "intend," and so on. Of course, it is not *individual* motivations that are explicated, neither in the use of a psychological lexicon nor in the categorization of instrumental actions. The dentist need not be interrogated to enable this kind of narration. Rather, it is the recognizable sense and instrumentality of actions that are unpacked; that is, those aspects of actions that adhere to the normatively expected actionable order of the professional procedure. The seeming transparency of "mind," also evident implicitly in any use of an action category, is really the public observability of accountable action. There are, however, limits to professionally shared vision (Goodwin 1994), as illustrated in the following section.

Access and constitutive detail

In the dental seminar, it is not only the seminar leader who provides descriptions of actions, but often the dentist performing the operation as well. These two perspectives produce two different layers of verbal explication. As Extract 2.2 illustrates, the parties' access to details of the procedure differs, as do the kinds of descriptive accounts they offer. For a competent onlooker, it is possible to see that the dentist in Extract 2.2 (and Figure 2.2) is examining how far down he can get into the root canal. The details toward which the dentist orients while doing the procedure, however, are not all available from the perspective of an observer. Although some degree of shared competence is a prerequisite for seeing what the members are doing, the question is also one of perspective: of access to temporally evolving phenomenal fields and constitutive details. As Macbeth (2012: 200) observes in his notes on the play of basketball in its circumstantial detail, there are phenomena that "cannot be found from anywhere off the court,

Extract 2.2   [END101104-27:13:23]

```
01 DEN:   då känner ja  *efter #lite grann, om man kommer, hur långt-
          then I feel a bit, if one gets, how far-
   deR:                             #moves file to tooth-->
   deL:   >>moves mirror*holds mirror at root canal--->
02        (1.2)#(0.4)
   deR:   -->#positions file in root canal—>
03        ner man *skulle kunna  *komma #sådär spontant.
          down one should get like spontaneously.
   deR:                                 -->#watch-winding
                                            movements-->
   deL:       -->*removes mirror*
04        ja, där tar de ju emot,
          yeah, there is some resistance,
05        (1.4)
06        direkt känner man.
          immediately one feels.
07        (7.0)
08 INS:   ser ni försiktiga såna här watch-winding-rörelser va.
          you see, these careful watch-winding movements y'know
09        (0.4)
10        [väldigt försiktigt.]
          [very careful.      ]
11 DEN:   [lirkar lite grann  ] #så.
          [twiddling a bit    ] there.
   deR:                         -->#removes file
```

no matter how closely you sit to the sidelines" (see also Garfinkel 2022, Part 1, Appendices 2 and 3).

In addition to showing that issues of access and constitutive detail are not only critical to the production of descriptions in instructional demonstrations, the extract

Figure 2.2 On the left, a magnified view of the thumb and forefinger of an endodontic specialist who is working the file down through the length of the root canal using a rotating movement of the instrument. On the right, the seminar leader is commenting on the procedure.

also serves to illustrate some of the interplay between our analytic accounts (e.g., descriptions of actions in transcripts) and the analyses produced by members. As noted by Mondada (2016: 361), "multimodal transcripts" raise questions about relevant description, "both for the co-participants and for the overhearing/seeing/sensing observer." In Extract 2.2, we find transcripts of the verbal contributions of the operating dentists (DEN) and the seminar leader (INS) based on the conventions developed by Jefferson (1984). Besides the transcription of the talk, and English translations of the talk, two lines are added for the manual actions of the dentist (see Mondada 2016): one for the left hand (annotated as "*deL*" in the transcript) and one for the right hand ("*deR*").

Throughout Extract 2.2, the operating dentist provides an online commentary on his own unfolding actions that reflexively build the procedure. This work does not involve the execution of a formal plan but is an incremental and tentative exploration of an endodontic scene. The dentist's talk is highly indexical in its relation to the context of the ongoing, incremental, and embodied exploration of the tooth. It refers to the minute details of how it "feels" to move the file in the tooth (lines 1–6). Laminated over the dentist's verbalized exploration of the tooth, the seminar leader's contribution provides a categorization of visible professional conduct, which enables the dentist's actions at that point to be recognizable as the endodontic technique known as "watch-winding" (lines 8–10). The seminar leader makes evident that she has performed the technique innumerable times herself and can appreciate what it means that there is some resistance. Nevertheless, it is not possible for her to produce a moment-to-moment commentary on the ongoing procedure in the same way as the operating dentist would do. Extract 2.2 thus provides two different accounts of the embodied actions of the dentist: first, we have the operating dentist's online commentaries, and second, we have the seminar leader's categorization and characterization of the visible actions as "careful watch-winding movements." The dentist has access to a first-person perspective on the cavity, which strongly relies on tactile experiences, whereas the seminar leader comments on a gestalt consisting of recognizable endodontic actions in context.

In relation to Extract 2.1, we noted that the instructor's descriptions emphasize what is not immediately visible in the observed actions. In the dentist's descriptions in Extract 2.2, we can see the articulation of the embodied movements in terms of their instrumental, tactile, and explorative character; while the file is inserted into the canal and rotated, the dentist says that he is *feeling* "how far down one should get." He also articulates the relevant *sensation* in relation to this project: that he encounters "resistance" at a certain point. While the first segment would be visible and accessible to the competent observer, the precise sensation of resistance is tied to the first-hand perspective of the operating dentist. Still, the sensation is central to understanding the development of the action sequence; most notably indicated by the fact that the file is removed shortly after it is mentioned.

This raises the issue of how we as analysts show and describe manual actions in the extract. What we have at our disposal is, first, conversation analytic transcription conventions and, second, still images or other graphic means of showing "nonverbal" aspects of the interaction. Conversation analytic transcription techniques

can be said to aim for a transparent rendering of conversationalists' "methods of speaking," to recall Garfinkel's point discussed above. Furthermore, conversation analysis also adheres to an analytic language that avoids explicating "content" in Garfinkel's sense, that is, intentions, psychological traits, background knowledge, motivations, and so on. Instead, transcriptions and analytic accounts explicate and provide access to the sequential embeddedness of turns and their structural organization so that the "recognized sense" of conversational actions is hopefully discernible to the reader.
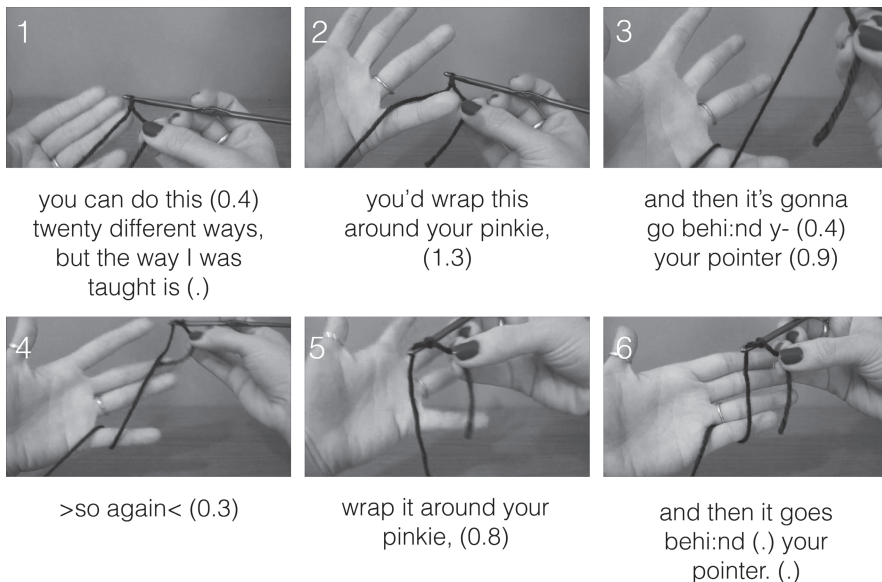
Embodied aspects of interaction have been subjected to the same treatment in the development of multimodal transcripts, most notably the system introduced by Mondada (2016). This system has been highly generative for those who take an interest in the ways in which gesture and gaze are finely synchronized with verbal action. With embodied courses of action, however, the question arises as to how these actions are to be described. While talk can be transcribed verbatim, without directly imposing a particular *categorization* of the actions at the transcription stage (e.g., whether the turn-at-talk is a question, request, or something else), other actions will have to be *described* in some way, picked from a range of possible alternatives.[2] One problem that confronts us is thus, simply put, what to include in the "multimodal" lines of the transcript. After the file has been moved to the tooth and positioned in the root canal, the dentist starts to manipulate the instrument in a way we categorize as "watch-winding movements," beginning on line 3 and continuing to line 11. The transcript marks the onset and offset of the activity in relation to verbal actions. We could start to elaborate on these descriptions, for instance, by attempting to produce a transcript that captures how the thumb and index finger move clockwise and counterclockwise, the degree of the rotation, and even, to some extent, the force that is applied. This would provide lengthy, but in Ryle's (1971) sense *thin*, descriptions of bodies, limbs, and tools that move in a three-dimensional space. Alternatively, we could aim for *thicker* descriptions that would explicate the sense of the actions, how watch-winding movements are different from other reaming and filing actions, how they are used to reach the working length of the root canal, what "working length" means in this particular context, *et cetera*.

In principle, and as we have noted above, our analytic explications of actions could go on indefinitely. Apart from the essential incompleteness of any such account, which by no means detracts from their possible usefulness in ethnographic work,[3] we want to point to a recurrent problem that we run into when producing our analytic descriptions and representations of these instructional activities. This is the chronic sense of absence, on the written page, of the "indexical ground" which provides members' descriptions with their "gestalt coherence"[4]: the actual activities being described, as they play out in real time. The members' categorizations and descriptions in the endodontic seminars provide observers with resources for recognizing what is being done in front of them. As analysts, we can offer additional resources to the reader, thereby engaging in an activity somewhat parallel to the seminar leader. Still, there is a sense in which the "say-shown details" (Burns 2012: 184) of the demonstration are impossible to recover through descriptions or

static images. In Extract 2.2, we saw how the instructor categorized the actions as "watch-winding" and characterized them as "careful." The students should know about watch-winding as a technique from lectures and textbooks, but this is the first time that they have seen it done live. The classification ("watch-winding") and the assessment ("careful") thereby gain their relevant sense when here applied to the details of the visual field to which students and the instructor have shared access. This central point of the demonstration – that the "carefulness" is to be seen in the dentists' performance as it unfolds on the video screen – can be *said* but not *shown* regardless of how extensively we expand our analytic accounts, and regardless of whether we do so "thinly" or "thickly."

Segmentation, redundancy, and followability

We now shift focus from seminars in dental education to an online video tutorial on handicraft techniques. Figure 2.3 is taken from a video called "How to crochet for beginners," and it shows how the yarn is held as a preparatory step before demonstrating how a single crochet is made. As stated in the title of the video, the demonstration is designed for novices. Anyone with some competence in crocheting would already know how to hold the yarn and how to make these stitches. The orientation to a novice audience is reflected in the granularity and pace of the instructions, in the use of repetitions, and in the exaggerated movement of hands, tools, and material. Besides the obvious fact that this demonstration is about handicraft

you can do this (0.4) twenty different ways, but the way I was taught is (.)	you'd wrap this around your pinkie, (1.3)	and then it's gonna go behi:nd y- (0.4) your pointer (0.9)
>so again< (0.3)	wrap it around your pinkie, (0.8)	and then it goes behi:nd (.) your pointer. (.)

Figure 2.3 A video with the title "How to crochet for beginners: single crochet stitch." The excerpt is taken 43 seconds into the video: https://youtu.be/BCDA44Sijx4?t=43.

techniques and not endodontic procedures, some further organizational differences between the next case and the prior ones can be noted: the video has been recorded and uploaded to YouTube; it is not a demonstration done for a live audience; the activity is produced as a "mock up" (Garfinkel & Sacks 1970: 363) designed for instructional purposes and not as a display of the "real" procedure as it naturally unfolds; and it represents a step-by-step analysis of a certain component procedure.

The instructor introduces this part of the demonstration by saying that this is just one of several ways of holding the yarn, whereafter she repeats the demonstration of the technique two times. In Figure 2.3, the first demonstration is represented on the first line and the second on the second line. Also, note how the achievement of the holding of the yarn is segmented into two parts. The first part is the "wrapping around the pinkie" and the other part is the "yarn going behind the pointer." In both cases, the two parts are separated by "and then." If one looks at a skilled craftsperson who weaves a strand of yarn between his or her fingers before starting to crochet, this does not come across as an action with clearly defined subparts. In the production of a step-by-step demonstration like this one, however, embodied actions are formulated as distinct steps in a sequence, and these formulations, in turn, shape how the actions are done and, as a result, seen. There is here a reflexive relation between words and manual actions: not only do the manual actions constitute a basis for what is described, but the instructive descriptions also shape what is done and shown with the hands. In the short sequence represented in Figure 2.3, the hands are first found in a waiting position; the fingers of the left hand are then spread widely before the yarn is wrapped around the little finger; the movements are segmented into parts with a recognizable beginning and an end; there is a slight pointing gesture with the left index finger when "your pointer" is mentioned the first time; and so on. In the case of the dental specialist, the seminar leader commented on the "careful watch-winding movements." This can be contrasted with the exaggerated character of the movements and gestures in Figure 2.3, which rather than displaying expert performance in its natural context are produced as "mock-ups" of the technique, so as to allow the viewer to visually discern the formulated steps as distinct phases.

Although the images in Figure 2.3 are unable to capture the actions as they unfold moment-to-moment in the video, they give some idea of what the video is showing and how words and displayed actions are organized as a coherent whole. As noted in relation to Extract 2.2, the instructive potential of a demonstration can be found in the coherence of embodied displays and verbal characterizations – the meaning of "careful watch-winding movements" derives from watching what the dentist is doing, and, interchangeably, the dentist's embodied actions are understood with reference to verbal descriptions. The video tutorial in Figure 2.3, like innumerable other demonstrations found on YouTube and elsewhere, is characterized by a tight connection between speech and action reminiscent of what, in Swedish schools of film production and TV journalism, is referred to as "Orange-TV" (Swe. "*Apelsin-TV*"). The maxim that comes with this concept is that holding an orange while saying "I'm holding an orange" is bad TV, the idea being that gratuitous verbal explicitness makes the production less watchable. While this might be true for

films and TV shows, instructional videos are not primarily produced or consumed for their watchability, but for their followability. In these cases, the talk and video have complementary roles: the verbal commentaries describe what is shown while the visual details specify what is said, which partly explains why "*Apelsin-TV*" is an omnipresent feature in instructional videos. The seeming redundancy of the verbal description parses and highlights relevant parts of the visual field. In that sense, the description articulates facets of what is already visible, the relevance of which might otherwise be missed. Differently put, the issue of redundancy arises only when the audience is watching from the sidelines, without any intent to follow what the demonstration instructs, or when they are already competent to see without prompting what was "there all along." The sense of redundancy might also emerge in a transcript of the demonstration; for example, if the instructor's verbal formulation "you'd wrap this around your pinky" were tied to a description of the embodied performance, such as "wraps yarn with right hand around little finger of left hand."

The fact that demonstrations are produced for their followability rather than their watchability raises issues tied to transcription and our own practices of looking, listening, and describing. In a response to a post on an electronic newsgroup discussing transcription, Schegloff (2000b) argues that the greatest value of transcription does not attach to the resulting transcript, which he thinks should be treated as a "mnemonic device for what is best engaged on tape," but to the "practice and process of transcribing itself." According to this argument, the value of the transcript is mainly to be found in the hearing that transcription makes possible and the intimate observational access to the data which it fosters. As a conversation analyst, Schegloff argues that there is "no better way of coming to hear what was actually said and done than listening closely and repeatedly under the discipline of committing to paper what you hear, at the level of detail that has become characteristic of good work in CA." When we are investigating instructional demonstrations, the practice of closely watching and listening to the demonstrations is central for us to understand what is happening. In this specific regard, our position might not be that different from the dental students who are following a demonstrated procedure through the instructive comments of the dentists. Like the users of video tutorials, and unlike the students in the dental class, we also have the possibility to replay the video repeatedly in order to figure out what is said and done. That these tutorial videos are produced for their followability and not their watchability, however, points to the relevance of other ways of engaging with them. Instead of merely listening and watching, an alternative way of coming to understand what is said and done would be to also follow the instructions.

This suggestion has a close affinity with prior work in ethnomethodology which involves a "tutorial" orientation to phenomena (Bjelić 1995; Bjelic & Lynch 1992; Garfinkel 2002), in which one "discovers the rule of the practice through one's own work with relevant material fields" (Bjelić 1995: 191). Applying this perspective to video analysis, Sormani (2016) contrasts traditional transcript-based approaches with what he calls a practice-based approach. In his characterization, transcript-based video analysis tends to focus on the coordination of multiple resources and

how these resources are deployed in different circumstances, but they do not necessarily address how these resources are used in the achievement of practical tasks. The alternative suggested by Sormani is for the researcher to reenact what the video shows and to treat the various difficulties encountered in the reenactment as "tutorial problems" (Garfinkel 2002, Ch. 4) with the potential of informing the researcher about the practices and achievements involved.[5] According to Sormani, this practice-based approach could be applied to any video recording of social activities, but we would argue that it acquires particular significance in cases like ours, when the videos, as such, are produced to be followed. In real-world settings, the actual following of a particular tutorial is unavoidably situated in a larger context of practical problems that set up the relevance of online resources. In these cases, there is a close relationship between the watching and the doing. As demonstrated by prior studies (e.g., Garfinkel 2002; Lindwall, Lymer & Greiffenhagen 2015; Livingston 2008), instruction-following involves turning a set of instructions into embodied actions and, in doing so, figuring out what corresponds to what, what to do next, and how to overcome difficulties that arise. In other words, the instruction follower needs to work out an "embodied correspondence" between the instructions and the local context of the activity (Livingston, 2008). This introduces concerns and issues that simply cannot be found in the video itself.[6]

As previously pointed out, a central feature of instructions, including video tutorials, manuals, and recipes, is the decomposition of skills, actions, and activities into steps. In Figure 2.3, the decomposition of the crafting technique into steps and the repetition of the same steps twice are produced for the recognition and followability of the demonstrated procedure. In turning the demonstration into actions, however, another segmentation is introduced when the watching of the demonstration and the doing of what the instructions instruct is coordinated in a stepwise manner (see Tuncer, Lindwall & Brown 2021).[7] Some tasks require that the watching and the following are done in an alternating manner: first, you watch a segment of the demonstration; then, you pause the video and attempt to do what the segment instructed. In other circumstances, it is possible to watch the video and do the task at the same time. In these cases, the video is paused as a way of keeping up with the demonstration and securing the ability to follow what comes next. In both cases, however, the segmentation of the video into parts involves an analysis of what is demonstrated, which relies on, but is not determined by, the decomposition found in the production of the demonstration. Through the decomposition that is endogenous to the demonstration, it is possible to project when one step begins and the other ends, and therefore where it might be relevant to pause. As the purpose of the watching is to follow what is shown, however, the relevance of this segmentation does not emerge from the demonstration itself, but from a gradually emerging understanding of the task, just what is needed to move on, and the contingencies that arise.

What we want to highlight here is a difference caught by two related senses of the word *follow*: following as *seeing-and-understanding* versus following as *reproducing in action*. The contingencies that emerge in the work of following step-by-step instructions, like the need to parse a tutorial video into segments in order

to keep up, are not necessarily recoverable through repeated viewing or detailed transcription. The distinct ways in which the two forms of instructional demonstration relate to followability and instruction-following place the analyst in different positions *vis-à-vis* the description of action. In the endodontic data, through "members' action category analysis" (MACA; see Lindwall & Lynch 2021) we get access to professional classifications ("watch winding"), assessments ("careful"), and rationales ("to keep the tongue away"), which makes it possible for us to "follow" what the dentist is doing.

In this respect, we as analysts are in a similar position as the students watching the seminar. Indeed, the recommendation to watch and listen to what one sees and hears, and to do so "closely and repeatedly under the discipline of committing to paper," would be equally valid for a student of dentistry as it is for the conversation analyst. For the analyst, moreover, the recommendation provides an entry point into the work of root canal procedures as well as to studies of "professional vision" (Goodwin 1994) and similar topics. Regardless of how closely the demonstrations are observed, however, a different set of issues will emerge when students later attempt to do what the demonstrations have shown. Our analysis of the crocheting tutorial points to some of these issues. Taken together, the two settings suggest that instructed observation – as one kind of orientation in instructional demonstrations – can be contrasted with instruction-following, where the demonstrated technique is to be reproduced in its constitutive details. These are not hard-and-fast categories, of course, but they point towards differences in orientation, relevant for the organization of instructional activities, and for the work of producing descriptions of embodied action.

Concluding comments: Instructed observation and following instructions

Sacks' (1963) metaphor of the "commentator machine" introduces distinctions between different types of encounter with social objects. The metaphor casts social life as being akin to the kinds of machines found "at industrial and scientific exhibitions" (p. 5) – with a doing part and a saying part, where the latter describes what the former is doing. As pointed out in the introduction to this chapter, ethnomethodology starts with the observation that the subject matters of social science – social objects of various kinds – are always already described and interpreted by the parties to their production. Sacks explicates this notion in terms of differences and parallelisms between commonsense and sociological descriptions.

What Sacks terms the "common sense" perspective is exemplified by his brief description of the machine: the saying part describes what the doing part is doing. From such a perspective, "[a] successful encounter consists […] in using one's background to learn what the object is doing and how it proceeds (its means and its ends). The payoff of such success consists in a perceived adaptation to its activities" (1963: 8). Regarding the adequacy of the descriptions offered, they "need only be 'good enough' to permit the encounter to proceed. Possible misunderstandings may be left until they actually raise difficulties, and when they raise

difficulties, they need only be solved as far as is necessary for the encounter to proceed" (1963: 9).

A quite different kind of encounter is produced when the observer is someone who knows the language *and* knows what the machine is doing. Such a person has for the most part no need for the "saying part" as a resource for learning about the object. Instead, the possibility is opened of assessing the correspondence between the saying and the doing. The observer can see errors in the description and errors in the doing, note verbal ambiguities and actions in need of clarification, and so on. Sacks ties this perspective to the typical sociological orientation to "remedy" noted problems in order to reconcile the observed doing with the flawed understandings embodied in the endogenous descriptions provided by the saying part. For sociology, adequacy becomes a principled problem to be resolved through theory, and not a practical problem to be resolved over the course of the encounter.

In our endodontic seminars, we might say that the operating dentist is a commentator machine, of sorts, performing the operation while describing their own actions. There is clearly a category of observers corresponding to Sacks' commonsense perspective, that is, the students who use the offered descriptions as resources for learning "what the object is doing." There is also, however, an observer clearly competent in both the activities and in the natural language produced by the object, namely, the seminar leader. The latter notes possible ambiguities and things that may need to be clarified and also, at times, irregularities in the performed activities (e.g., the omission of taking an X-ray image when it is normally required, see Lindwall & Lymer 2014). The descriptions provided by the seminar leader, however, are still oriented to criteria for adequacy that are practical; they need only be "enough" for the students to follow the procedure.

Another difference between the seminar leader and Sacks' account of the sociological observer is that there are aspects of the observed procedure that, as noted above, elude even the most competent "professional vision" (Goodwin 1994): visual details and tactile dimensions available only from the first-person perspective of the operating dentist. A further difference between our setting and Sacks' metaphoric one can be noted: we, as professional analysts, are not competent in any strong sense in the endodontics that underlies both the doing and the saying of the "machine," as well as the seminar leader's elaborations. As analysts, we learn what the object is doing through ethnographic work, and on the basis of members' descriptions, in a way that aligns our perspective most of all with that of the students.

Leaving the parallel to Sacks' metaphoric setting, and returning to our distinction between two ways of following instructional demonstrations, it is possible to argue that ethnomethodology has little interest in such seminars or tutorials in themselves: that instructional demonstrations make up just one kind of item in an open-ended list of "docile objects" (Garfinkel 2002), where manuals, recipes, and rules would belong to the same family; that these demonstrations therefore only are relevant as the first part of a pair, where the practical activity of instruction-following would be the second part; and that, in order to investigate instructed action ethnomethodologically, we need to turn to settings where instructional demonstrations are being actively followed in the sense of doing what these demonstrations

instruct. To meet this requirement, we could set out to follow tutorial videos ourselves; we could record people who come to terms with the instructions as part of some practical project; or we could record the instructional interaction between a teacher and a novice. As testified by several chapters in this handbook, there is much to learn from such approaches. For Garfinkel, however, instructed action is a more encompassing formulation about the praxeologies of social worlds than is instruction-following. In line with prior ethnomethodological studies (e.g., Burns 2012; Garfinkel 2002, Ch. 7), we thus believe that there are lessons to learn from these instructional demonstrations that do not require us to attempt endodontic procedures ourselves, or to study video records of novices being instructed in doing them; that the demonstrations are not mere docile objects since they consist of the *in vivo* work of say-shown demonstrations; that such settings are perspicuous in that they provide us with laic and professional analyses of embodied courses of actions; and that members' instructional descriptions, including the use of action categories and the varying granularity employed, give access to some of the terms and orders of relevance that shape not only experience but also the production of everyday and professional worlds in common.

Acknowledgment

Notes

1 The notion of "what any member knows" (Garfinkel 1967: 24) or the "socially-sanctioned-facts-of-life-in-society-that-any-bona-fide-member-of-the-society-knows" (Garfinkel 1967: 76) is recurrent in the writings of both Garfinkel and Sacks. Garfinkel expands on the issue in the following way:

> With respect to the problematic character of practical actions and to the practical adequacy of their inquiries, members take for granted that a member must at the outset 'know' the settings in which he is to operate if his practices are to serve as measures to bring particular, located features of those settings to recognizable account. They treat as the most passing matters of fact that members' accounts of every sort, in all their logical modes, with all of their uses, and for every method for their assembly are constituent features of the settings they make observable. Members know, require, count on, and make use of this reflexivity to produce, accomplish recognize, or demonstrate rational–adequacy–for–all–practical–purposes of their procedures and findings.
>
> (Garfinkel 1967: 8)

2 Clearly, the transcription of verbal action must also be done *in some way*, picked from a range of possible alternatives (e.g., Ochs 1979). As this chapter hopefully makes clear, however, the description of action involves other issues and considerations than those that applies to the transcription of talk-in-interaction.

3 Ryle's (1971) distinction between "thin" and "thick" descriptions brings us further to possible points of contact between members' descriptions and the kinds of analytic accounts we produce as professional analysts. For instance, Crabtree et al. (2012), in a discussion of the kinds of description that are necessary to produce praxeological accounts in design ethnography, argue for Ryle's notion of thick description, meaning a description which provides a "*recognisable* account of what a person or persons are doing" (p. 193, italics in original). For Crabtree et al., the goal is to produce ethnographic descriptions which unpack and convey the accountable recognizability of work activities. Such a description would include a very high level of detail, going beyond brief glosses. The authors' example relates to "searching the internet," where "typing in words" would be a "thin description." A thick description, by contrast, explicates the various practices that go into searching and the projects which motivate particular acts in a larger contexture of relevant activities. Ryle's version of thick description is then contrasted with Geertz' use of the same term, which emphasized abstraction, "scholarly artifice" (Geertz 1973, cited in Crabtree et al. 2012: 193), and "the constructions we imagine [members] to place upon what they live through" (ibid.).

4 The idea of a "gestalt coherence" as it is used here draws on Garfinkel's ethnomethodological "misreading" of Gurwitsch (see Eisenmann and Lynch 2021; Garfinkel 2021).

5 Garfinkel (2002, 248) introduces "tutorial problems" as "the empirical grounds for Ethnomethodology's central claims," which "are used to look for the local, endogenously produced, and accountable appearances of social facts."

6 In fact, these concerns begin even before an actual video is consulted. The first step is to find a relevant video for the task at hand. In many cases, the selection of the video is unproblematic, but in the attempt of fixing a broken dishwasher, for instance, just writing "dishwasher repair" does not suffice. It is necessary to assess whether the model shown in the video is the same as, or sufficiently similar to, the one that is to be repaired. In cases where the problem is unclear, it might be relevant to watch a video where the problem first is diagnosed; for a dishwasher that does not drain, it could be the drain pump and motor, the belt, the piston and nut assembly, the drain hose, *et cetera*. Some of these issues will require specific tools or replacement parts, and before starting the process, it is therefore relevant to skim through the video just to see whether it is at all possible to follow the tutorial, or if extra tools or parts need to be ordered in advance. In larger repair jobs, there is typically an expectancy that these issues will emerge and that substantial time will be spent looking for tools and makeshift objects. In watching a video on how to crack an egg with one hand, in contrast, there would clearly be an expectancy that eggs are required, but that two ping pong balls and a coin would be required for practicing the technique would probably come as a surprise.

7 There are, of course, several ways of watching online tutorials. Some people watch these videos purely for entertainment, with no intention of ever doing what the tutorial shows, much like the way cooking shows are normally seen on broadcast TV. The videos might be watched for inspiration, or for broadening the understanding of some activity or interest. Even when the videos are watched to be followed, these activities are often separated in time; after one or more videos have been watched, it might take minutes, hours, or days until any attempt at doing what the tutorial instructs takes place. The discussion of "pausing" we describe later in this chapter exhibits one of several ways in which video instructions are parsed in the act of following them.

References

Bjelić, Dušan. 1995. 'An Ethnomethodological Clarification of Husserl's Concepts of "Regressive Inquiry" and "Galilean Physics" by Means of Discovering Praxioms', *Human Studies* 18(2–3): 189–225.

Bjelić, Dušan and Michael Lynch. 1992. 'The Work of a (Scientific) Demonstration: Respecifying Newton's and Goethe's Theories of Prismatic color', in G. Watson and R. M. Seiler (eds.) *Text in Context: Contributions to Ethnomethodology*. London: Sage, pp. 53–78.

Burns, Stacy Lee. 2012. '"Lecturing's Work": A Collaborative Study with Harold Garfinkel', *Human Studies* 35(2): 175–192.

Crabtree, Andy, Mark Rouncefield and Peter Tolmie. 2012. *Doing Design Ethnography*. London: Springer.

Eisenmann, Clemens and Michael Lynch. 2021. 'Introduction to Harold Garfinkel's Ethnomethodological "Misreading" of Aron Gurwitsch on the Phenomenal Field', *Human Studies* 44(1): 1–17.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Garfinkel, Harold. 2002. *Ethnomethodology's Program: Working Out Durkheim's Aphorism*. Edited by E.W. Rawls. Lanham, MD: Rowman & Littlefield.

Garfinkel, Harold. 2021. 'Ethnomethodological Misreading of Aron Gurwitsch on the Phenomenal Field', *Human Studies* 44(1): 19–42.

Garfinkel, Harold. 2022. *Studies of Work in the Sciences*. Edited by M. Lynch. London: Routledge.

Garfinkel, Harold and Harvey Sacks. 1970. 'On Formal Structures of Practical Action', in J. C. McKinney and E. A. Tiryakian (eds.) *Theoretical Sociology: Perspectives and Development*. New York: Appleton-Century-Crofts, pp. 338–366.

Geertz, Clifford. 1973. 'Thick Description: Toward an Interpretive Theory of Culture', in C. Geertz (ed.), *The Interpretation of Cultures: Selected Essays*. New York: Basic Books, pp. 3–30.

Goodwin, Charles. 1994. 'Professional Vision', *American Anthropologist* 96(3): 606–633.

Jefferson, Gail. 1984. 'Transcription Notation'. In J. M. Atkinson and J. Heritage (eds.) *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. ix–xvi.

Lindwall, Oskar and Gustav Lymer. 2014. 'Inquiries of the Body: Novice Questions and the Instructable Observability of Endodontic Scenes', *Discourse Studies* 16(2): 271–294.

Lindwall, Oskar, Gustav Lymer and Christian Greiffenhagen. 2015. 'The Sequential Analysis of Instruction', in N. Markee (ed.) *The Handbook of Classroom Discourse and Interaction*. New York: John Wiley, pp. 142–157.

Lindwall, Oskar and Michael Lynch. 2021. '"Are You Asking Me or Are You Telling Me?": Expertise, Evidence, and Blame Attribution in a Post-Game Onterview', *Discourse Studies* 23(4): 652–669.

Livingston, Eric. 2008. *Ethnographies of Reason*. Aldershot, UK: Routledge.

Macbeth, Douglas. 2007. 'Sequential Analysis in an Ethnomethodological Key: Order Without Theory', in L. N. Berlin (ed.) *Theoretical Approaches to Dialogue Analysis: Selected Papers from the IADA Chicago 2004 Conference*. Tübingen: Max Niemeyer Verlag, pp. 199–213.

Macbeth, Douglas. 2012. 'Some Notes on the Play of Basketball in Its Circumstantial Detail, and an Introduction to Their Occasion', *Human Studies* 35(2): 193–208.

Mondada, Lorenza. 2016. 'Challenges of Multimodality: Language and the Body in Social Interaction', *Journal of Sociolinguistics* 20(3): 336–366.

Ochs, Elinor. 1979. 'Transcription as Theory', in E. Ochs and B. B. Schieffelin (eds.) *Developmental Pragmatics*. New York: Academic Press, pp. 43–72.

Ryle, Gilbert. 1971. 'The Thinking of Thoughts. What Is "Le Penseur" Doing?' in G. Ryle, *Collected Papers*, Vol. 2. London: Hutchinson, pp. 480–496.

Sacks, Harvey. 1963. 'Sociological Description', *Berkeley Journal of Sociology* 8(1): 1–17.

Schegloff, Emanuel A. 2000a. 'On granularity', *Annual Review of Sociology* 26: 715–720.

Schegloff, Emanuel A. 2000b. 'Post from Manny Schegloff Re: CA Orthography, Sacks, etc.', *COMSERVE's Ethno Hotline*, #4040 (9 March).

Sormani, Philippe. 2016. 'Practice-Based Video Analysis: A Position Statement', *Società-MutamentoPolitica* 7(14): 103–120.

Tuncer, Sylvaine, Oskar Lindwall and Barry Brown. 2021. 'Making Time: Pausing to Coordinate Video Instructions and Practical tasks', *Symbolic Interaction* 44(3): 603–631.