

Social Interaction. Video-Based Studies of Human Sociality.
2020 Vol. 3, Issue 3
ISBN: 2446-3620
DOI: 10.7146/si.v3i3.122250

Social Interaction ***Video-Based Studies of Human Sociality***

**Body Part Highlighting:
Exploring two types of embodied practices in two sub-types of showing
sequences in video-mediated consultations**

Brian L. Due and Simon B. Lange

University of Copenhagen

Abstract

Consultations in healthcare settings involve an initial phase of “history-taking”, during which the healthcare professional examines the client for symptoms by asking questions, making the client show symptoms on his or her own body, and performing bodily examinations. But how can bodily symptoms be identified when the interaction is video-mediated and sensory access is limited? One key resource here is “body showings”. However, research suggests that video-mediated teleconsultations reduce body showings due to both technical difficulties and sensory obstruction. In this paper, we provide a contrary case that shows two types of practices employed for successful history-taking through body-part showings. Based on an analysis of an “evocative showing sequence” (Licoppe, 2017), we present two types of gestural highlighting practices, via two types of showing sub-sequences: 1) “mimicable body part highlighting”, which occurs in a sequence of “adapting-body-to-frame”; and 2) “direct body part highlighting”, which occurs in a sequence of “adapting-frame-to-body”. The paper uses a single case to discuss how gestures work in a video-mediated context and how sensory judgements are not just a property of the healthcare professional, but are distributed to clients who are able to creatively adapt to situated contingencies in order to accomplish common understanding about the symptoms. The data consist of video-recorded, video-mediated physiotherapy consultations in Denmark, analysed using ethnomethodological conversation analysis (EMCA). The paper contributes to EMCA research on mediated interaction and embodied, gestural and sensorial practices.

Keywords: multimodality; embodied interaction; video-mediated interaction; highlighting; showing sequences

1. Introductionⁱ

During the “history-taking” phase (Byrne & Long, 1976; Robinson, 2003), a medical professional gathers information about a patient’s pain and illness by asking questions and by looking at and touching the client’s body, in order to identify symptoms and formulate a diagnosis (Heath, 1992). During this process, the client also typically shows body parts or otherwise highlights bodily areas of concern (Maslen, 2016). While healthcare is increasingly performed as telemedicine, which involves practices of “remote care” (Goodridge & Marciniuk, 2016), there is also growing concern about the loss of hands-on examination: research suggests that clients appreciate embodied engagements that represent medical attention and care, especially in the form of touch (Hunt, 2014). Touch is seen as demonstrating empathy, politeness, warmth and that the doctor is genuine in their approach – it is “a human thing to do” (Cocksedge et al., 2013). Research has illustrated the important role played by intimacy of professionals’ bodies and hands when “servicing the body” of the client (Due et al., 2020). So what happens in a video-mediated context where touch and other inter-bodily – or intercorporeal (Meyer et al., 2017) – resources are constrained? In such cases, do bodily showings transform into something else?

In a review article, Lupton and Maslen (2017) conclude: “Greater insights into the sensory transformations inherent in telemedicine are vital in a socio-political climate in which many governments are seeking to expand their telemedicine offerings to their citizens” (Lupton & Maslen, 2017, p. 1558). The current paper seeks to contribute with this kind of knowledge by showing how participants employ two alternative practices and types of sequences (Schegloff & Sacks, 1973) that differ slightly from face-to-face consultations. This gives rise to a more specific research question: how do healthcare professionals in situated video-mediated consultations co-operate with clients in order to achieve a common understanding about the body part or pain spot concerned?

Research suggests that, in most cases involving video-mediated situations, visual access and showings of body parts, where these could be anticipated and relevant, *do not* occur – or occur with repair initiations (Stommel et al., 2020). This lack of visual access and resulting lack of sensory information makes reaching a diagnosis more difficult for the healthcare professional (see Stommel et al. in this issue). However, in some cases, participants display creative adaption to the mediated affordances, and are seen to engage in practices that transform the limited sensorial access via different semiotic means. Based on ethnomethodological multimodal conversation analysis of a single case, this paper aims to show how a successfully assessed activity of history-taking is accomplished via two types of practices and two sub-types of showing sequences, which are unlikely to be produced in a face-to-face consultation.

The *activity* we deal with is therefore history-taking in a mediated healthcare setting. The *practices* we are looking at within this activity are body-part showings that involve environmentally coupled gestures (Goodwin, 2017, p. 197), albeit in a mediated environment. We call these 1) “mimicable body-part highlighting” and 2) “direct body-part highlighting”. We will explore these practices as embedded within two different sub-types of showing *sequences*. The first is when a body part is adapted to fit the camera frame (“adapting-body-to-frame”), the second is when the camera frame is adapted to show a body part (“adapting-frame-to-body”). In the single case provided in this article, the participants produce these two practices of body-part highlighting in the two different types of sequences, respectively, as a progression aimed at dealing with difficulties of understanding. We show how there is an orderliness in the specific production of these two types of sequences, i.e. first adapting-body-to-frame, and if that does not work, then adapting-frame-to-body. This will be discussed in the conclusion. We do not, however, imply that the two practices of body-part highlighting may only occur in these specific sequences. Nonetheless, we suspect that mimicking is rarely found in sequences where the frame is adapted to the body (we have zero occurrences in our corpus), as a direct showing of the body part is possible.

2. History-taking and multisensoriality

The showing sequence under scrutiny in this article occurs during a phase of the physiotherapy consultation that is similar to *history-taking* in medical interaction (Byrne & Long, 1976; Robinson, 2003). While the history-taking activity has been described at length in relation to doctor-patient interactions, the EMCA literature on physiotherapy consultations is more scarce (see, however, Parry, 2004, 2009, 2010, 2013). In our data corpus, history-taking is a distinct interactional activity in which the patient, via a situated identity of history-giver, is co-constructed as having the primary right of the floor, as visible in multi-unit TCU constructions, the elaboration of symptoms, and the providing of information to which he or she has epistemic rights and obligations. The role of the physiotherapist as history-taker is constituted by taking notes, asking information-seeking questions and producing receipts, continuers and acknowledgement tokens. No previous studies, to our knowledge, have investigated history-taking activities in video-mediated situations.

Research from face-to-face healthcare interactions has repeatedly illustrated numerous aspects of the multimodal work that goes into the different phases of the interaction, from consultations to operation theatres (e.g. Heath, 1986; Ruusuvuori, 2001; Heritage & Clayman, 2010; Koschmann et al., 2011; Deppermann, 2014; Nielsen, 2016; Heath et al., 2018). Multisensorial practices have been shown to be vital in diagnostic work in general (Maslen, 2015, 2016), while new developments in EMCA research have specifically dealt with multisensorial work in interaction, arguing that sensorial experiences do not just

consist of private feelings, but may also be publicly available (cf. Liberman, 2013; Nishizaka, 2017; Mondada, 2019, 2020; Gibson & Vom Lehn, 2020). However, it has been argued that this sensorial work undergoes a transformation when the consultation is video-mediated (Maslen, 2017). Obviously, there can be no smelling or tactile sensations when the parties are physically separated. The only sensory experiences available to both participants are visual and auditory. Research suggests that this constrains healthcare work, but no research has so far explored the creative ways in which participants emergently adapt to the situation and use the affordances of the available senses, body parts, and material and technological structures to successfully achieve the activity. As this article will show, the highlighting of embodied practices is vital to this process.

3. Showings in video-mediated contexts

Of special relevance to this article is the work on showings (Licoppe, 2017; Licoppe et al., 2017; Licoppe & Morel, 2014). This research demonstrates that showings are a ubiquitous practice in video-mediated interaction. In terms of numerical frequency, showings are reported to happen on average once during every mundane Skype conversation between friends and family (Licoppe et al. 2017, p. 5297). Research has also shown how moving the camera to show something is a mutual accomplishment between interactants, and that it can index familiarity for participants (Licoppe & Morel, 2014).

Showing practices can be divided into at least two distinct categories: they can be either “gestural showings”, where the showing display is a contribution to ongoing talk (“*talking-and-showing*”), or “showing sequences”, where the showing display is the focus of the interaction (“*showing-and-talking*”) (Licoppe, 2017, p. 64). Licoppe introduces a further important distinction between different types of showing sequences, as he divides them into “informative showings”, which position the recipient as not having relevant knowledge about the object shown, and “evocative showings”, which cast the recipient as knowledgeable (Licoppe 2017). However, all the showings analysed so far in the literature are related to various objects as being the showable, rather than the body or its parts per se. Nonetheless, Licoppe’s concepts are still very useful, and in this article, we will analyze two sub-types of “evocative showing sequences”.

The showing sequence we deal with resembles, to some extent, what Nishizaka (2014, 2017) has described as “differentiating sequences” – that is, when participants recognize something related to the body through the specific pairing of an invitation to perceive and a claim of having perceived. The type of showing sequence we will deal with may, generally and somewhat loosely, be termed differentiating, in the sense that, the participants are oriented towards differentiating a specific body part to enable the perception of just this particular part. However, while there are claims of recognition in the data, the sequence

type is not exactly the same with regards to its specific structure and the specific type of embedded actions, as described by Nishizaka. Instead, we suggest that the main showing (and the type of “differentiating”) sequence can be understood as comprising two sub-types of showing sequences, which entail two types of embodied highlighting practices.

4. Direct highlighting and mimicable highlighting

In our data, the body, through its gestures, is a vital resource for accomplishing showings. Interactional research has shown that gestures are not just isolated and static semiotic types, but that “representation by gesture must be approached from the signifier’s side, as constructions or fabrications, not as expressions of independently performed mental abstractions from perceptual realities” (Streeck, 2009, p. 149). The gesturing actions do not just receive their meanings from the hand alone, but must be seen in close connection to a semiotic ecology of other actions and material surroundings (Goodwin, 2000), and as part of multimodal gestalts (Mondada, 2014).

Building on Mandel (1977) Kendon (1988) and Streeck (2008), gestures must be understood as indexical and iconic signs, where “the referent of a sign is the meaning that it carries in a given context of use. The base of the sign, on the other hand, is the object or action that the production of the sign is derived from” (Kendon, 1988: 83). For the purpose of this paper, it is important to bear this distinction in mind when unpacking the two mediated highlighting practices used for showing body parts, because each has the same base (a particular pain spot on the foot), but different referents.

The hands can be used to highlight areas of concern. Highlighting practices consist of a gestural pointing, plus a tracing and drawing of an area. In his seminal paper on professional vision, Goodwin (1994) described highlighting as a practice of “making specific phenomena in a complex perceptual field salient by marking them in some fashion” (p. 606). This involves “methods used to divide a domain of scrutiny into a figure and a ground, so that events relevant to the activity of the moment stand out” (p. 610). In Goodwin’s analysis of the videotape of the beating of Rodney King, he emphasizes how the attorneys manipulated the perceptual field through their method of video analysis. He shows how the work of the viewer is radically changed when the format changes. We will show, based on a completely different kind of setting, how the exploitation of the technical affordances similarly has bearings on the highlighting practices, as it changes the viewer’s perception-ability. In other words, the body’s position relative to the video camera structures the perception of the body parts (cf. Nishizaka, 2017).

However, unlike in Goodwin's analysis of Mr. King, we do not analyze this highlighting as being manipulative in any negative way, but as a form of creative adaptation (for other similar uses, see Due & Toft, forth.). In addition, we will show how the practices are accomplished by using a gesturing hand with a pointing index finger, along with a tracing (and drawing) of lines used for "feature-extracting" (Streeck, 2009, p. 66). One straightforward type consists of highlighting the areas of the object (body part) that are under scrutiny. We call this "direct highlighting". Another type consists of highlighting a simile, i.e. an object (body part) that represents something else. We call this "mimicable highlighting".

Mimicable highlighting is at stake when a body part is used as a model in a depicting practice. Streeck presents a variety of depicting practices, and describes the mimic versions as "practices in which the hand – or the entire body – simply enacts what some body did or could do on another occasion" (Streeck, 2009, p. 149). Mimic depictions have an "original-copy" relationship between signified and signifier and are enactments of something else. They are a kind of "make-believe" (Goffman, 1974). However, Streeck does not describe a version of these practices as specifically related to the body, or as "body modeling", as Kendon (1988) talks about. In "body modeling", the body – or, more commonly, a particular body part – is used as a model to mimically refer to the sign's base. A body part may be used to illustrate, or mimic, something else, and enactment in some form may be combined with body-modeling gestures. Nishizaka (2017) describes one type of body modeling, citing an example from an obstetric clinic in which a pregnant woman is instructed how to position her body by the professional, who uses her own "body as a model".

These insights into gestures with indexical and often also iconic structures have been researched only in the context of face-to-face interaction. One exception to this is the version of mediated mimicable actions described in Due et al. (2019), who refer to aspects of visual and gestural recipient design as a "mimicable embodied demonstration". These kinds of actions are designed by a participant in the pursuit of making the recipient mimic the same kind of action, using a body part as a model. However, the practices found and reported in the present paper are different, as they are not recipient-designed to be copied – rather, they have an illustrative function in the pursuit of achieving understanding about the symptoms.

In the case presented in this paper, each of the two described practices consist of members' choices, produced within particular types of showing sequences. The client chooses first to engage in a sequence of adapting-body-to-frame, using the practice of mimicable highlighting. When that fails, it is remediated by a different kind of sequence, of adapting-frame-to-body, within which the client picks a practice of direct highlighting. The order of this kind of sequential choice

will be discussed after the analysis, in relation to the norm of a talking head configuration.

5. Data and representation

The data in this article are based on a corpus of video recordings collected by researchers at the Centre for Interaction Research and Communication Design, UCPH. The corpus consists of over 100 hours of recordings of naturally occurring video-mediated interaction between professionals and clients in different settings. The particular fragments in this article are from a subset of the larger corpus. This smaller corpus consists of five hours of video recordings from online physiotherapy consultations. The initial analysis of the video material revealed how the client's production of gesturing actions adapted to the camera frame as showings is a pervasive phenomenon. This is unsurprising, as the purpose of the consultations is to identify symptoms, provide diagnoses and suggest physiotherapy-based exercises. While most of the interactional work is usually conducted using descriptive talk in combination with simple pointing practices, other, more creative practices also occur. Two such practices will be described in this article, based on a single case.

The interaction was recorded from the practitioner's side. The technical setup comprised four GoPro cameras located in the practitioner's office, and a recording of the screen through which the video is mediated at the practitioner's location. In order to give the reader a sense of the recorded ecology, Table 1 presents a screenshot of three cameras and a cropped version of the screen recording, which adds extra information.

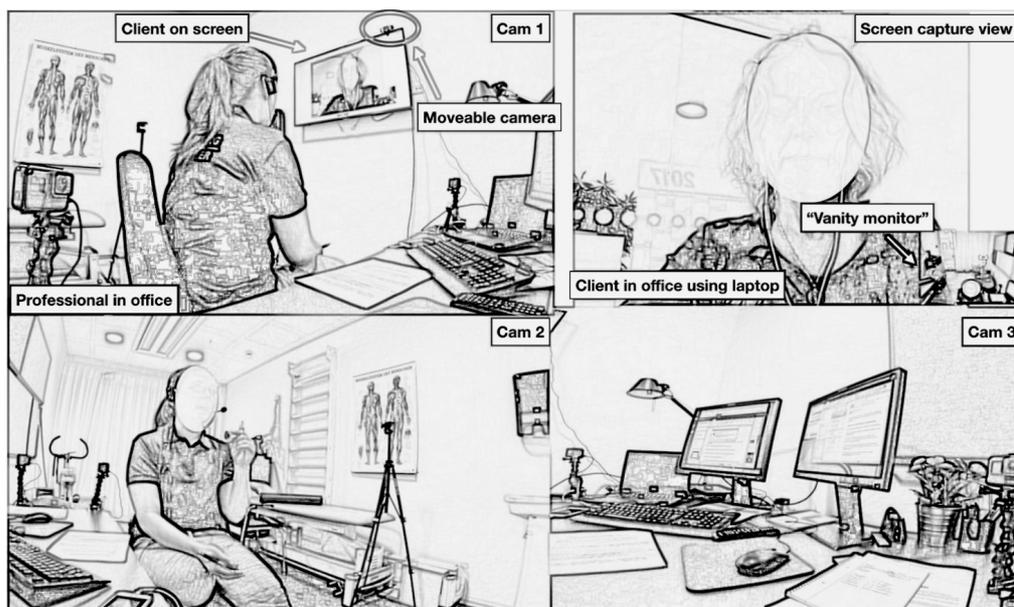


Figure 1. Ecologies in the interaction.

The different camera angles and the screen capture footage were synchronised on the basis of audio waveforms in Final Cut Pro. The transcription of speech largely follows the conventions developed by Gail Jefferson (Jefferson, 1985, 2004), while the multimodal conduct follows the Mondadian framework (Mondada, 2018a). In order to comply with the General Data Protection Regulation (GDPR), the video stills in the transcripts have been sketched out in SketchBook Pro. This graphical representation (Laurier, 2014) secures the anonymity of the participants while preserving the gaze direction and other embodied conduct relevant to the analysis. All recordings were collected with the informed consent of all participants, and other relevant parties and have been anonymised accordingly.

6. Analysis 1: *Mimicable body-part highlighting*, using a body part as a representational model in *adapting-body-to-frame*

The client, C, has been explaining her symptoms to the physiotherapist, P, for a while and is specifying the location of pain in her foot. To do this, she uses her left arm as a model to illustrate where it hurts.

Excerpt 1.1 “My left foot”

8 ((C has described her problem for a while))
 9 C: ∠%Min venstre #fod å #så det #benet ing,#%
 My left foot and then it is the leg right,
 %-----right index finger pointing-----%
 *holds left arm still---->1.22
 Fig #1 #2 #3 #4



Fig 1



Fig 2



Fig 3



Fig 4

10 P: Ja
 Yes
 11 % (0.3)
 P: %points to specific point on arm ---->

The excerpt begins with C providing P with information about her medical problem: she has been experiencing pain in her foot. While the pain is located in a specific part of the foot, it is not an anatomical area for which she seems to have a discrete lexical name. She explains the location of her pain by a gestural showing, designed as a body-part mimicking, using her left arm and hand as a model that mimics her leg and foot (l. 9). The model is constructed by holding her right arm up in vertical manner, so that her hand is positioned with fingers pointing horizontally and palm downwards, mimicking a foot and leg (see figures 1–4, l. 9). The way in which she holds her arm is an example of the careful recipient design of video-in-interaction: she positions her arm so that it fits within the visual field requirements of the video camera, thereby adapting her body to the format. This showing sequence is thus recognizable as a sequence of *adapting-body-to-frame*. It is also noticeable how she simultaneously keeps her head and upper torso in sight, thereby orienting to the norm of talking head configurations in video-mediated interactions (Licoppe & Morel, 2012).

In the following turns (l. 12–16), C continues her multimodal description of the location of her pain, using the established body part model as a substrate (C. Goodwin, 2018, p. 38) for further verbal and gestural elaboration.

Excerpt 1.2 “My knuckle”

12 C: Så sidder min kno *hertude*# (.)%*Δ*cirka samme sted *∟*som den der*∟*%*†**Δ*
 Then my knuckle is out here (.) almost the same place as that one there
 ∟raises arm-∟
 -->%-turns hand-%---circles finger on bone---%
 †-----to screen-----†
 Δ-----screen-nods-----Δ

P:
 Fig #5



Fig 5
 13 †(0.4)
 †to left hand --->
 14 C: %Så: det lige her #under.
 Then it is right under here.
 %circles index around specific area--->

Fig. #6



Fig 6
 15 (0.3)

C has established the body-part mimicking model and received a minimal aligning response from the physiotherapist (l. 10), which encourages her to further multimodally describe the location of the pain. Through the video-mediated format, the physiotherapist can only make a diagnosis based on two senses: vision and hearing. The client can be seen to orient to this by carefully lifting her arm so that it is visible in the camera frame, which she is able to monitor herself using the vanity screen.

The talk and gesture are inextricably linked through deictic expressions that refer to the base of the sign such as *herude* (“out here”, l. 12), *den der* (“that one there”, l. 12), *her under* (“under here”, l. 14) and *der og så rundt om* (“there and then around”, l. 16). These expressions are carefully produced to occur simultaneously with the gesturing of the hands, i.e. pointing with the index finger and tracing lines around the “knuckle” (l. 12), creating a *complex multimodal gestalt* (Mondada, 2014) via which C describes her symptoms to P. Clearly, although participants orient to the body (hand) as a representational model for something else (the foot), this is also a highlighting practice that makes specific part of the hand (foot!) stand out from the ground.

Excerpt 1.3 “yes okay”

16 >#De:t sån lissom om de:t<% der og så #rundt om.
 >It's like it's< there and then around.
 ---%underlines spot--->

Fig.#7



Fig 7

#8



Fig 8

17 †#(0.3)
 -->†to screen--->
 Fig #9



Fig 9

18 P: Ja(.) okay.=
Yes (.) okay.=

Upon producing “om” (“around”, l. 16), there is transition relevance, as the TCU is possibly syntactically complete, with falling intonation. After the turn, C shifts

her gaze from the model to her screen (l. 17, fig. 9), thereby visibly orienting away from describing her model, and adding to the transition relevance. Looking away from the model and to the screen can be seen as C treating her entire description of her symptoms as complete, thereby inviting a response from P that both acknowledges the information received and demonstrates understanding of the description (l. 18).

To sum up: The depiction of the pain spot uses a mimicable highlighting practice based on an “original-copy” relationship between signified and signifier, for the purpose of achieving a common understanding about the body part concerned. This is accomplished in a specific type of showing sequence: *adapting-body-to-frame*. The client is observably oriented towards providing information about a specific pain spot to the physiotherapist, and she does this in and through the accomplishment of a showing sequence in which a body part (the hand) is adapted to fit the frame. However, the body part used in the highlighting is not the real object of concern, but a simile. As the client does not receive the response she anticipated, she restructures both the sequence type and the highlighting practice. This will be evident in the following analysis.

7. Analysis 2: *Direct body-part highlighting, adapting-frame-to-body*

This analysis will show how a lack of achieving common understanding about the pain spot on the foot prompts the client to shift practice and sequence type. From using the body as a mimicking model, thereby separating the base from the referent, she now restructures the material configuration from *adapting-body-to-frame* to *adapting-frame-to-body*, which enables her to go directly to the base and highlight the exact spot on the foot.

Excerpt 2.1 “are you following”

19 C: =er du med?%
=are you following?
-->%

20 P: Ja omkring malleolen.
Yes around the malleolus.
21 †(0.7)%(0.7)
-->†to left hand
%....---->

22 C: Nej he- %her ved den der kno.%∠
No he- here by that knuckle.
-->%-taps spot w index-%
----∠

23 +(0.4)
+lowers arms, turns, brings laptop to floor---->

24 P: [Ja]
[Yes]

25 C: [Nu] prøver jeg [lige (???)
[Now] I'll just [try (???)

26 P: [de:t
[it's

27 (.)

28 P: (De:t godt.)=
(That's good.)=

29 C: Nu er vi jo live ikke,
Now that we're live right,
30 (0.4)

31 P: Ja.
Yes.
32 (1.3)

33 P: Men det [er de:t selve] den der +knogle% (.)ø:h
But it [is it that bone there (.) u:h
---->+

34 C: [*jeg prøver lige°*]
[*° I'll just try°*]
-->+

35 ∠(.)
∠moves hand to foot ---->

While the physiotherapist acknowledges the information provided by the client so far (l. 18), the client still checks that they are on the same page regarding their understandings about the exact spot on the foot where the pain is located. Latching on to P's turn, C initiates repair with an interrogative format that explicitly targets a potential understanding issue (“are you following?”, l. 19). This repair initiation makes it explicitly relevant for P to demonstrate her understanding of C's description so far. P's turn in line 20 (“Yes around the malleolus.”) would seem to qualify as repair proper by referring to the correct anatomical nomenclature for the bony prominence on each of the ankles, “the malleolus”, which C has referred to as a “knuckle” (line 12). She hereby provides both a demonstration and claim of understanding and professional competence, based on readily knowing precisely where it hurts. However, this turn is treated as insufficient, presumably because C is not familiar with the anatomical term.

Consequently, C does not acknowledge the turn as demonstrating understanding, and extends the sequence by recycling what she has already explained (“he- here by that knuckle”, l. 22). Just after producing the word “knuckle”, she puts her arms down, disengaging from the body-part mimicking model. In the ensuing silence between spoken turns, she lowers her arms and

starts bringing the laptop to the floor. Eventually, she places the laptop on the floor, adjusts the camera angle and leaves the computer in this position (l. 33), creating a new visual frame for her upcoming highlighting (l. 36–44). The camera action, i.e. moving the computer to the ground, is a remedial action that seems aimed at restoring intersubjectivity in a context in which the co-participants orient to potential problems of understanding. The new practice of direct highlighting also relies on deictic expressions such as *den her* (“this here”, l. 36), *her* (“here”, l.39), *her foran* (“here in front”, l. 41) and *her rundt* (“around here”, l.44), as well as accompanying gestures (that seem to be recycled from earlier), creating a complex multimodal gestalt that cannot be understood without reference to all of its parts.

Excerpt 2.2 “this here right”

36 C: **Ja: den her ik/kå.#**
Yes, this here right.

-->∠holds index on bone --->

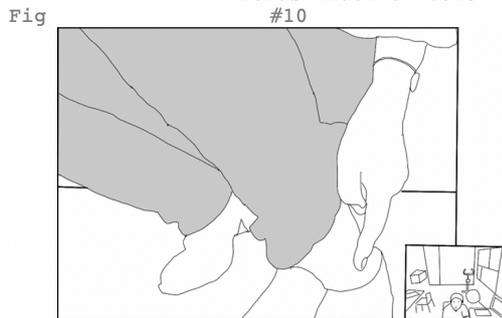


Fig 10

37 (.)

38 P: **JA,=**
YES,=

39 C: **Knogle der sidder her.**

Bone that is here.

40 P: **ΔJa.Δ=**

Yes.=

Δ-1-Δ 1:head tilt

41 C: **∠=å så gør det #ondt her foran,**

=n' then it hurts here in front,

-->∠taps spot on instep--->

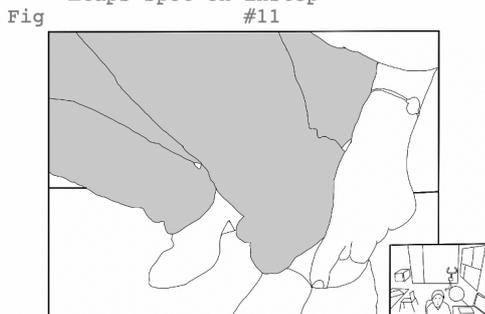


Fig 11

42 P: **Ja.**

Yes.

43 (0.6)

The gestures are environmentally coupled (C. Goodwin, 2007), and thus the visual ecology to which P has access, i.e. the whole visual frame, is necessarily an integral part of how C performs the highlighting. In contrast to the former practices of body-part mimicking, in which the base was separated from the signified referent (the foot), the moving of the computer to the ground and the tipping of the screen to secure visual access through the camera are embodied actions that utilize the technological affordances of mobility (moving the camera to the showable (Licoppe et al., 2017)) to accomplish a new gestural practice. This sequential organization, in which the frame is adapted to fit the specific body part of concern, enables a direct connection between the base (foot) and the signified referent (foot). It is notable that the work of the fingers is carefully crafted, in order to highlight the specific area on the foot where it hurts, at the same time as it is verbally and indexically described. While saying “this here right” (l. 36), she points with her index finger directly to the specific body part, the bone. While saying “it hurts here” (l. 41), she taps on a spot on her instep, and while

saying “around here” (l. 44), she traces her finger around the bone, thereby highlighting it to make it stand out from the ground.

Excerpt 2.3 “around here”

44 C: ∫å #så her rundt,#
 n' then around here,
 --><underlines ankle-->
 Fig #12 #12



45 P: ∆Ja.∆
 Yes.
 ∆-2-∆ 2: upward head nod
 ((turns omitted from transcript))

56 *P: =Men det var egentlig rart du os lige viste det
 =But it was actually nice that you also showed me
 (.)

57

58 *P: Ø:h fordi så [ku jeg]se hvorhen omkring på foden det var.
 U:h because then I could see where on the foot it was.

59 *C: [okay]
 [okay]

It is notable how this direct highlighting in a sequence of adapting-frame-to-body is a co-operative achievement that produces claims of recognition from the recipient (“YES”, l. 38 + “yes” l. 40 + “yes” l. 42 + “yes” l. 45). The whole showing sequence thus has a recipient design to which the recipient responds systematically by displaying alignment and understanding. However, whereas the first practice of mimicable highlighting of a body part presumably *did* accomplish the work of providing exact information to the physiotherapist, as she acknowledged through her use of the term “malleolus”, the client did not orient to this understanding as being achieved. Therefore, remedial actions of adapting-frame-to-body by moving it to the ground were produced in order to make possible a more direct body-part highlighting. Now, the whole showing sequence is finally assessed, with an emphasis on the importance of the restructuring of the showable from being the simile (hand) to being the exact object (foot): (“=But it was actually nice that you also showed me U:h because then I could see where on the foot it was”, l. 56 and 58). This leaves us with some open questions: Why were there two different sequences and practices? What do they do differently? Why in that order?

8. Discussion: Showing sequences subjected to a preference for a talking head configuration

The analysis of this single case involves several interconnected features that we must both separate and understand as complex gestalts: a) an overall activity

concerned with history-taking, b) a sociomaterial setting of video mediation with a fractured sensorial ecology, c) two different sub-types of showing sequences that we labelled adapting-body-to-frame versus adapting-frame-to-body, and d) two different types of embodied highlighting practices.

The sequential organisation and the embodied practices are the participants' resources for accomplishing the activity. Although we did not find in our data corpus any other precisely similar cases – in which both the different sub-types of adaptation sequences and the embodied highlighting practices occurred in just this way – we do not consider this a deviant case, but a compact one. We found both the types of sequences, including their specific order, and also the types of embodied mimicable and direct highlighting practices elsewhere in the corpus, just not all together in one single case like the one presented. Thus, although we cannot conclude that these sub-types of showing sequences entail exactly these types of embodied highlighting practices in this precise order, this single case still provides us with a basis for investigating all of these as co-occurring (see Schegloff's discussions on single cases (1968, 1987)), while knowledge about their individual designs can be found in isolation elsewhere.

Both evocative showing sub-sequences of adapting-body-to-frame (another example being moving the face/body in front of camera) and adapting-frame-to-body (another example being turning or moving the screen) can occur as single sequences, without the other. They are not mutually dependent. Likewise, "mimicable body-part highlighting" and "direct body-part highlighting" can occur independently as mediated highlighting practices in showings of body parts. However, because in this particular case they occur in a specific order, we need to explore the *haecceity* (Garfinkel, 1996), or the *why that now* (Schegloff & Sacks, 1973) of this organization in order to learn more about its preference organization.

Concerning the embodied highlighting practices, we described how they are both accomplished as highlighting, but with different objects of reference. In the first case, the sign base (foot) was separated from its referent (hand), thereby constituting the practices as mimicable by using one body part as a model for another. In the second case, there was a direct index between the base and the referent. Although the two practices described were different, they involved similar kind of multimodal actions, most notably the use of the index finger for pointing and tracing, in combination with deictic expressions. This reveals overall a strong orientation to the audible and visual senses, as the gestures' visibility is anticipated. It is thus evident that vision is a primordial resource for intersubjectivity in video mediation, which participants use to build and interpret action in showing sequences.

Gaze behaviour can also be monitored in video-mediated interaction, but the direction of the gaze is lost in translocation (Luff et al., 2003; Luff, Heath, Yamashita, Kuzuoka, & Jirotko, 2016). Interestingly, however, the co-participants in a video-mediated encounter often have access to the entire visual perspective (the visual frame) of their interlocutor, via a vanity monitor. This frame is continuously monitored and curated by the co-participants, and obligates each party to create mutually intelligible, meaningful shots (Licoppe & Morel, 2009). The visual frame thus becomes part of the recipient design of actions in video-mediated interaction. This was especially observable in the first practice, when the client monitors her own body-part mimicking practice in order to adapt it to the camera frame.

While the two analysed highlighting practices are distinct, it is possible that they could also be observed occurring in face-to-face interactions. However, in this mediated setting, they have particular traits that make them stand out. Firstly, the fact that they are sequentially linked means that the second practice is performed as remedial, establishing a more direct relation between the base and the referent, i.e. decomposing (Svennevig, 2018; Due et al., 2019) the first complex semiosis into a more direct, indexical, highlighting practice. Secondly, they are dependent on the situation's spatial, mobile, technical and material affordances, which means that participants can be seen to orient to the camera angles and limitations in the showing sequence, leading to a restructuring of the sociomaterial configuration.

The analysis of these two practices in this particular sequential organization thus provides a nuanced answer to the interactional economy of actions in preference organization, in which the first step is to try the “least complicated and costly remedy” (Pomerantz, 1984). We might presume that a mimicking highlighting gesture using a body part as model is a more semiotically complex practice than a direct pointing and highlighting practice – as we know from the differences between iconic gestures that are abstractions from the concrete, indexical matter (Mandel, 1977). So, *why that now?* Why does the client begin with the mimicking highlighting practice, and not the less costly remedy? This might be explained by the “talking heads norm” (Licoppe & Morel, 2012).

According to this norm, any deviation from the talking head is an accountable action. We also see that C accounts for the change of the visual frame when adapting-frame-to-body (“[Now] I’ll just try”, l.25 and “Now that were live right”, l. 29). Therefore, with regards to interactional economy, there seems to be a reverse relation and hierarchical order. The directness of the gesturing practice does not precede the sequential organization – rather, it is the other way around: the sequence organization, through its material configurations, constrains the gesturing practices.

Orienting to a talking-head configuration and *adapting-body-to-frame* is in this case observably preferred over a non-talking head configuration and *adapting-frame-to-body*, even though the mimicking highlighting gesture has a more complex base-referent organization. Presumably, the exact opposite would be the case in a face-to-face encounter, in which a client would presumably point directly to a pain spot as a first and preferred action, without needing to detour into mimicking.

Interestingly, when problems of understanding arise, the client restructures the specific feature of the video-mediated format, i.e. the talking head norm, and adjusts to the rule of interactional economy by producing a more straightforward, indexical gesture. The remedial action of moving the camera frame to a non-talking head configuration might thus be a last resort for securing intersubjectivity and restoring a simpler or more direct semiosis in which the sign base is also the signified referent. Orientations to the camera (Licoppe, 2015) by participants in video-mediated interaction may therefore seem to be a perspicuous site for investigating creative adaptations employed to secure understanding, and as such may spark discussion of the interactional possibilities and the empowerment of the client to which the format gives rise.

The sensory transformations inherent in telemedicine, as discussed at the beginning (Lupton & Maslen, 2017), prompt participants to find alternative ways to describe and focus on symptoms. While research has illustrated the asymmetry of sensory access (Seuren et al., 2020; Stommel, Van Goor, & Stommel, 2020) and the problems of fractured ecology, this article has, on the contrary, shown how creative adaptations entangled in gesturing practices enable a successful symptom description. The restructuring of the camera position can be seen as a compensation strategy for the fact that the participants are not co-present – they lack the reciprocity of perspectives (Schutz, 1967, p. 14) that is a basic assumption in co-present, face-to-face interaction. However, the mobility of the camera enables the participants to re-calibrate the evidential boundaries at a given moment in order to show visual phenomena that are otherwise inaccessible to their co-participants. Where a linguistic reformulation of a turn is a method of controlling *what the other can hear*, moving the camera in certain sequential environments is thus a method for controlling *what the other can see*. Both can be seen as part of an intersubjective framework upon which the participants draw in order to produce intersubjectivity in interaction.

The politics of telemedicine are not new – indeed, they are quite old (Weissert & Silberman, 1996) – but recent discussions of the importance of the loss of sensory access are sparking a new critical approach to the limits of the format (Scott Kruse et al., 2018). Conversely, we have presented the practices and creative adaptations involved in a successful symptom description in which a sensory judgement that could have solely been the healthcare professionals’

responsibility is successfully provided by the client, thereby establishing a setting based on a “sensing of sensors” (Lupton & Maslen, 2017). In this way, video-mediated formats may actually work as a patient-empowering technology.

Finally, this research has some interesting implications for practitioners. As we have shown, there is a preference for first adapting-body-to-frame – and then, if that does not work, switching to adapting-frame-to-body. However, research has otherwise suggested that showings of body parts, where these might be anticipated and relevant, are *not* occurring, and the lack of visual access makes it harder for the healthcare professional to reach a diagnosis (see Stommel et al. in this issue). If the alternative is no showing at all, then both the practices of using a body part as a mimicable model for highlighting within the frame, and the direct highlighting of the pain spot by a restructuring of the frame, could be solutions for practitioners. Professionals could be instructed to pre-empt problems of non-showing by asking the client to highlight (i.e. show, point, trace or draw, to make a figure stand out from the ground). Preferably, a client should go straight to the practice of direct highlighting, using the affordance of camera mobility to secure interactional economy. As technology improves, virtual additional and mobile camera apps (e.g. <https://www.mmhmm.app>) could reduce the fractured nature of the ecology and enable participants to keep a “talking head configuration” in front of the screen while simultaneously moving the mobile camera to the pain spot, thereby removing the “conflict” between adapting-body-to-frame versus adapting-frame-to-body.

Transcription key

Transcript follows Mondada (2014) multimodal transcript conventions

Participants: Professional (P), Client (C)

Δ delimits gestures done by P	% delimits gestures done by C (right hand)
• delimits gaze done by P	∟: delimits gestures done by C (left hand)
§ delimits body movements by P	+: delimits body movements done by C
	† delimits gaze done by C

References

- Byrne, P. S., & Long, B. E. L. (1976). *Doctors Talking to Patients: A Study of the Verbal Behaviours of Doctors in the Consultation*. London: Royal College of General Practitioners.
- Cocksedge, S., George, B., Renwick, S., & Chew-Graham, C. A. (2013). Touch in primary care consultations: Qualitative investigation of doctors' and patients' perceptions. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, 63(609), P. 283-290. <https://doi.org/10.3399/bjgp13X665251>

- Deppermann, A. (2014). Multimodal participation in simultaneous joint projects: Interpersonal and intrapersonal coordination in paramedic emergency drills. In *Multiactivity in Social Interaction Beyond multitasking* (eds.) P. Haddington, T. Keisanen, L. Mondada, M. Nevile. Amsterdam: John Benjamins Publishing Co.
- Due, B. L., Lange, S. B., Nielsen, M. F., & Jarlskov, C. (2019). Mimicable embodied demonstration in a decomposed sequence: Two aspects of recipient design in professionals' video-mediated encounters. *Journal of Pragmatics*, *152*, 13–27. <https://doi.org/10.1016/j.pragma.2019.07.015>
- Due, B. L., Lehn, D. V., Webb, H., Heath, C., & Trærup, J. (2020). Servicing the body: Placing glasses on the client's head at the opticians. *Visual Studies*, *35*(2-3), 109–123. <https://doi.org/10.1080/1472586X.2020.1763196>
- Due, B. L., & Toft, T. (forth.). Phygital Highlighting: Achieving joint visual attention when physically co-editing a digital text. *Journal of Pragmatics*.
- Garfinkel, H. (1996). Ethnomethodology's Program. *Social Psychology Quarterly*, *59*(1), 5–21.
- Gibson, W., & Vom Lehn, D. (2020). Seeing as accountable action: The interactional accomplishment of sensorial work. *Current Sociology*, *68*(1), 77–96. <https://doi.org/10.1177/0011392119857460>
- Goffman, E. (1974). *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper & Row.
- Goodridge, D., & Marciniuk, D. (2016). Rural and remote care: Overcoming the challenges of distance. *Chronic Respiratory Disease*, *13*(2), 192–203. <https://doi.org/10.1177/14799723166633414>
- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, *96*(3), 606–633.
- Goodwin, C. (2000). Action and Embodiment Within Situated Human Interaction. *Journal of Pragmatics*, *32*(10), 1489–1522.
- Goodwin, C. (2017). *Co-Operative Action*. New York: Cambridge University Press.
- Heath, C. (1986). *Body movement and speech in medical interaction*. New York: Cambridge University Press.
- Heath, C. (1992). The delivery and reception of diagnosis in the general-practice consultation. In P. Drew and J. Heritage (eds.) *Talk at work* (pp. 235–268). New York: Cambridge University Press.
- Heath, C., Luff, P., Sanchez-Svensson, M., & Nicholls, M. (2018). Exchanging implements: The micro-materialities of multidisciplinary work in the operating theatre. *Sociology of Health & Illness*, *40*(2), 297–313. <https://doi.org/10.1111/1467-9566.12594>
- Heritage, J., & Clayman, S. (2010). *Talk in action: Interactions, identities, and institutions*. Malden (MA): Wiley-Blackwell.
- Hunt, D. P. (2014). Do patients with advanced cancer value the physical examination? *Cancer*, *120*(14), 2077–2079. <https://doi.org/10.1002/cncr.28678>
- Kendon, A. (1988). *Sign languages of Aboriginal Australia: Cultural, semiotic, and communicative perspectives*. Cambridge: Cambridge University Press.
- Koschmann, T., LeBaron, C., Goodwin, C., & Feltoich, P. (2011). "Can You See the Cystic Artery yet?" a Simple Matter of Trust. *Journal of Pragmatics*, *43*(2), 521–541.
- Lieberman, K. (2013). The Phenomenology of Coffee Tasting: Lessons in Practical Objectivity. In K. Lieberman (Ed.), *More Studies in Ethnomethodology* (pp. 215–262). New York: SUNY Press.

- Licoppe, C. (2017). Showing objects in Skype video-mediated conversations: From showing gestures to showing sequences. *Journal of Pragmatics*, 110, 63–82. <https://doi.org/10.1016/j.pragma.2017.01.007>
- Licoppe, C., Luff, P. K., Heath, C., Kuzuoka, H., Yamashita, N., & Tuncer, S. (2017). Showing Objects: Holding and Manipulating Artefacts in Video-mediated Collaborative Settings. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 5295–5306. <https://doi.org/10.1145/3025453.3025848>
- Licoppe, C., & Morel, J. (2014). Mundane Video Directors in Interaction: Showing One's Environment in Skype and Mobile Video Calls. In M. Broth, E. Laurier, & L. Mondada (Eds.), *Studies of Video Practices Video at Work*. Abingdon, Oxon: Routledge.
- Lupton, D., & Maslen, S. (2017). Telemedicine and the senses: A review. *Sociology of Health & Illness*, 39(8), 1557–1571. <https://doi.org/10.1111/1467-9566.12617>
- Mandel, M. (1977). Iconic devices in American Sign Language. In L. Friedman (Ed.), *On the other hand: New perspectives on American Sign Language research* (pp. 57–107). New York: Academic Press.
- Maslen, S. (2015). Researching the Senses as Knowledge. *The Senses and Society*, 10(1), 52–70. <https://doi.org/10.2752/174589315X14161614601565>
- Maslen, S. (2016). Sensory Work of Diagnosis: A Crisis of Legitimacy. *The Senses and Society*, 11(2), 158–176. <https://doi.org/10.1080/17458927.2016.1190065>
- Maslen, S. (2017). Layers of sense: The sensory work of diagnostic sensemaking in digital health. *Digital Health*, 3. <https://doi.org/10.1177/2055207617709101>
- Meyer, C., Streeck, J., & Jordan, J. S. (2017). *Intercorporeality: Emerging Socialities in Interaction*. New York: Oxford University Press.
- Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137–156. <https://doi.org/10.1016/j.pragma.2014.04.004>
- Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, 145, 47–62. <https://doi.org/10.1016/j.pragma.2019.01.016>
- Mondada, L. (2020). Orchestrating Multi-sensoriality in Tasting Sessions: Sensing Bodies, Normativity, and Language. *Symbolic Interaction*. <https://doi.org/10.1002/symb.472>
- Nielsen, S. B. (2016). How Doctors Manage Consulting Computer Records While Interacting With Patients. *Research on Language and Social Interaction*, 49(1), 58–74. <https://doi.org/10.1080/08351813.2016.1126451>
- Nishizaka, A. (2014). Instructed perception in prenatal ultrasound examinations. *Discourse Studies*, 16(2), 217–246. <https://doi.org/10.1177/1461445613515354>
- Nishizaka, A. (2017). The Perceived Body and Embodied Vision in Interaction. *Mind, Culture, and Activity*, 24(2), 110–128. <https://doi.org/10.1080/10749039.2017.1296465>
- Parry, R. (2004). The interactional management of patients' physical incompetence: A conversation analytic study of physiotherapy interactions. *Sociology of Health and Illness*, 26(7), 976–1007. <https://doi.org/10.1111/j.0141-9889.2004.00425.x>
- Parry, R. (2009). Practitioners' accounts for treatment actions and recommendations in physiotherapy: When do they occur, how are they structured, what do they do? *Sociology of Health & Illness*, 31(6), 835–853. <https://doi.org/10.1111/j.1467-9566.2009.01187.x>
- Parry, R. (2010). Practitioners' Accounts for Treatment Actions and Recommendations in Physiotherapy: When do they Occur, how are they Structured, what do they do? In

- Pilnick, A., Hindmarsh, J., Gill, V.T. (eds.) *Communication in Healthcare Settings: Policy, Participation and New Technologies* (pp. 48–65). John Wiley & Sons, Ltd.. Chichester : Wiley-Blackwel. <https://doi.org/10.1002/9781444324020.ch4>
- Parry, R. (2013). Giving Reasons for Doing Something Now or at Some Other Time. *Research on Language and Social Interaction*, 46(2), 105–124. <https://doi.org/10.1080/08351813.2012.754653>
- Robinson, J. D. (2003). An Interactional Structure of Medical Activities During Acute Visits and Its Implications for Patients' Participation. *Health Communication*, 15(1), 27–59. https://doi.org/10.1207/S15327027HC1501_2
- Ruusuvuori, J. (2001). Looking means listening: Coordinating displays of engagement in doctor–patient interaction. *Social Science & Medicine*, 52(7), 1093–1108. [https://doi.org/10.1016/S0277-9536\(00\)00227-6](https://doi.org/10.1016/S0277-9536(00)00227-6)
- Schegloff, E. A. (1968). Sequencing in Conversational Openings. *American Anthropologist*, 70(6), 1075–1095.
- Schegloff, E. A. (1987). Analyzing Single Episodes of Interaction: An Exercise in Conversation Analysis. *Social Psychology Quarterly*, 50(2), 101–114. <https://doi.org/10.2307/2786745>
- Schegloff, E. A., & Sacks, H. L. (1973). Opening up closings. *Semiotica*, 8(4), 289–327.
- Scott Kruse, C., Karem, P., Shifflett, K., Vegi, L., Ravi, K., & Brooks, M. (2018). Evaluating barriers to adopting telemedicine worldwide: A systematic review. *Journal of Telemedicine and Telecare*, 24(1), 4–12. <https://doi.org/10.1177/1357633X16674087>
- Stommel, W. J. P., Goor, H. van, & Stommel, M. W. J. (2020). The Impact of Video-Mediated Communication on Closed Wound Assessments in Postoperative Consultations: Conversation Analytical Study. *Journal of Medical Internet Research*, 22(5), e17791. <https://doi.org/10.2196/17791>
- Streeck, J. (2008). Depicting by gesture. *Gesture*, 8(3), 285–301.
- Streeck, J. (2009). *Gesturecraft: The manu-facture of meaning*. Philadelphia: John Benjamins
- Svennevig, J. (2018). Decomposing Turns to Enhance Understanding by L2 Speakers. *Research on Language and Social Interaction*, 51(4), 398–416. <https://doi.org/10.1080/08351813.2018.1524575>
- Weissert, W. G., & Silberman, S. (1996). Health care on the information highway: The politics of telemedicine. *Telemedicine Journal: The Official Journal of the American Telemedicine Association*, 2(1), 1–15. <https://doi.org/10.1089/tmj.1.1996.2.1>

ⁱ Thank you to researchers at the Center for Interaction Research and Communication Design, University of Copenhagen, for fruitful comments and discussions during data sessions. This research was supported by a grant from the Velux Foundation, Denmark. Grant number: 00013125.