

Embodied Interaction During Mental Health Consultations: Some Observations on Grounding and Repair

Jing Hui Law, Patrick G.T. Healey and Rosella P. Galindo Esparza

Queen Mary, University of London
Wolfson Institute of Population Health and
School of Electronic Engineering and Computer Science
London E1 4NS
jing.law@qmul.ac.uk

Abstract

Shared physical space is an important resource for face-to-face interaction. People use the position and orientation of their bodies—relative to each other and relative to the physical environment—to determine who is part of a conversation, to manage conversational roles (e.g. speaker, addressee, side-participant) and to help co-ordinate turn-taking. These embodied uses of shared space also extend to more fine-grained aspects of interaction, such as gestures and body movements, to support topic management, orchestration of turns and grounding. This paper explores the role of embodied resources in (mis)communication in a corpus of mental health consultations. We illustrate some of the specific ways in which clinicians and patients can exploit embodiment and the position of objects in shared space to diagnose and manage moments of misunderstanding.

1 Background

Non-verbal signals are integral to natural human interaction. The best known are the facial expressions of emotion (e.g. anger, fear, sadness, surprise, happiness) (Ekman, 1979; Chovil, 1991b). However, there are also a range of non-verbal signals that are specific to conversation (Bavelas et al., 1995; Chovil, 1991a; Kaulard et al., 2012). These include large-scale configurations of body position and orientation that can tell us e.g., who is participating in a conversation, what their role is (e.g. speaker, addressee, listener or bystander) and their relative levels of interest and engagement (Schefflen, 1973; Kendon, 2010; Bull, 2016). There are also a range of small-scale conversational gestures. For example, the use of hand gestures to hold or yield a turn, to enlist help with finding a word—or expression and the use of facial gestures such as raised eyebrows to emphasise particular words or to display “thinking” (Bavelas et al., 1995; Ekman, 1979; Chovil, 1991a).

Embodied signals can be produced in parallel with verbal contributions by both the speaker and by other participants (Bavelas, 2007; Bavelas et al., 1995; Deppermann et al., 2021). This facilitates real-time, incremental checking and feedback. Speakers can produce gestures that complement or augment their speech and listeners can simultaneously display their reactions through concurrent backchannel signals (Chovil, 1991b,a). These concurrent signals shape a speaker’s turn in real-time and—if problems are apparent—can cause a speaker to rephrase, change direction or cut-off their turn (Goodwin, 1979).

Some non-verbal signals are associated with potential problems with shared understanding. A frown and briefly averted gaze before a turn can suggest that a speaker is about to say something potentially problematic (Kaukomaa et al., 2014) and gaze aversion by an addressee following a turn can prompt the speaker to rephrase what they said (Kendrick and Holler, 2017). Some facial gestures, such as raised eyebrows and widening of the eyes, can act as stand-alone clarification requests (Kendrick, 2015; Seo and Koshik, 2010). Similarly, temporary suspension of hand movements described as non-verbal ‘holds’ or ‘freezes’ can provide signals of ongoing repairs (Seo and Koshik, 2010; Floyd et al., 2016; Bavelas et al., 1995). Quantitative data from motion captured conversations shows that the velocity and height of head and hand movements changes during both self-repairs/disfluencies and other-repairs—and that these changes are different for speakers and listeners (Healey et al., 2013, 2015; Özkan et al., 2021).

2 Communication in Healthcare

Communication in healthcare settings is critical to the quality of patient-clinician relationships, affecting outcomes such as patient satisfaction and treat-

ment adherence. It is arguably even more important in mental healthcare settings, since talk is the primary means of diagnosis and treatment (McCabe and Healey, 2018; Mahmoodi et al., 2020; McCabe et al., 2013; McVittie et al., 2020; Wu, 2020). However, difficulties in balancing good communication practices with work pressures have often been reported in the NHS (NHS Improvement, 2018).

This has prompted an interest in developing protocols and tools to help healthcare professionals organise their thoughts and structure consultations. However, overly rigid protocols can have a problematic impact on the naturalness of healthcare professional’s interaction, limiting their ability to adapt to individual patient’s needs, concerns or understanding (NHS Improvement, 2018). At its worst, the artificial structure imposed by the protocol disrupts the flow of conversation and can become counterproductive for clinical interactions (NHS Improvement, 2018).

Here we examine these issues in the context of a tablet application (DIALOG+) designed to promote communication in face-to-face mental health consultations. We focus on a detailed qualitative analysis of the moments where misunderstandings arise—and the combination of verbal and non-verbal resources that are used to address such problems. Our analysis shows how the position of the physical device and people’s orientation toward it plays a role in both causing and mitigating misunderstandings.

2.1 DIALOG+

The DIALOG+ protocol is designed to support and structure conversations in routine community mental healthcare consultations for patients suffering from psychosis (Priebe et al., 2015, 2017). This intervention applies principles of solution-focused therapy to promote assessment of all relevant aspects of patients’ lives. It also uses this information to help patients initiate change and improve their situations. The overall aim is to improve the therapeutic benefits of the consultation process (Priebe et al., 2015, 2017).

The DIALOG+ application is a tablet-based system, built around a central screen with a sequence of eleven quality of life questions (items) that cover various aspects of patients’ mental and physical health, job situation, relationships, medication and practical help received. Clients and clinicians work through the list together—and as the users select each item, a slider appears, so that clients can pro-



Figure 1: Screenshot of the DIALOG+ Application User Interface

vide their current satisfaction rating for each item. After performing the ratings, a structured, iterative 4-step process is followed of (1) choosing items that the patients would like to discuss, (2) understanding what determines the ratings, (3) considering options for what can be done to improve their satisfaction with these items and (4) agreeing on some next steps and action plans that can be adopted by the clients, or with assistance from the clinician, family members and/or support workers, to get to potential solutions.

Clinical trials suggest the intervention is effective, with patients reporting fewer general psychopathological symptoms, fewer unaddressed needs, higher levels of treatment satisfaction and better objectively measured social outcomes—such as in terms of housing situations and employment status—at one-year follow ups (Priebe et al., 2015, 2017). The DIALOG+ app has been translated into over 15 languages and has been implemented and tested in various studies in more than 10 countries. Nevertheless, little is understood about how the intervention is actually incorporated into interactional practices within consultations. This is important for refining clinician communication skills and also because new versions of the application are being developed, including for remote use.

Since the COVID-19 pandemic, many consultation services in healthcare have moved online, at least temporarily (Liberati et al., 2021; Khan et al., 2021). In mental healthcare settings, this has led to reported difficulties in the establishment or maintenance of meaningful trust and rapport between clients and clinicians (Liberati et al., 2021; Olwill et al., 2021; Khan et al., 2021). The use of standardised tools in mediated communication

is a potential concern, given the complexities of balancing flexibility with system-driven structures of service delivery (Drew et al., 2021). In this context, it is important to understand the interactional features of DIALOG+ and how this may impact on its delivery online. To gain a more granular perspective, we explore the details of the DIALOG+ consultations to observe what exactly happens in these interactions.

3 The current study

The data for this study comes from video recordings of mental health consultations. Extracts from these interactions are described below using conversation analysis (CA) techniques. CA involves detailed qualitative analysis of conversations in naturally occurring circumstances—and the material for such analyses is recorded and/or transcribed talk (Sacks et al., 1978; Sacks, 1992; Silverman, 1998). The focus of the analysis is on the organisation of fine-grained features of interaction such as timing, pauses, repetitions, restarts and the details of concurrent non-verbal signals such as gaze and gesture.

The examples considered below are selected to illustrate the types and trajectories of (mis)communication observed in this corpus, how they are affected by the presence of DIALOG+ application, and how this affects the management of shared understanding. The two overarching research questions in this study are:

1. Where and how does shared understanding break down in the DIALOG+ interactions?
2. What role does embodiment play in detecting and dealing with these breakdowns?

4 Methods

Design. The DIALOG+ trials are described elsewhere (Priebe et al., 2017). Our dataset consists of 40 video recordings of 32 clinical consultations; 16 are from a control group receiving treatment-as-usual and 17 from an intervention group using DIALOG+. The average length of the recordings is 30.12 minutes in the control group and 39.68 minutes in the intervention group.

Participants and Ethics. Participants who consented to join the original DIALOG+ trials provided informed consent to be video-taped or audio-taped on at least one session for use in future studies on potential improvements to the DIALOG+

procedure and technology. The study protocol, including data collection and storage procedures, was audited by the National Research Ethics Service (NRES) London, Stanmore (12/LO/1145). All personal data has been removed from the extracts and faces have been blurred.

Procedure and data analysis. The authors familiarised themselves with the data by watching and listening to the recordings repeatedly, making brief notes about the general form of the conversations and any features that appeared striking. A second pass then focused on selecting particular episodes with overt evidence of misunderstanding. The clearest of these were transcribed in detail using Jefferson's orthography which includes features like pauses, overlap and intonation (Sacks et al., 1978; Ekberg, 2021). Embodied conduct that is relevant to the interaction—such as nods, gestures and postures shifts—were also included in the transcriptions.

5 Results

5.1 Overview

The recorded conversations were mostly follow-up sessions to previous consultations. Although DIALOG+ is designed to encourage input by both parties to the interaction in practice, the clinicians typically took the initiative and controlled all input to the device. Patients only rarely touched the tablet or laptop—and the devices were often positioned in ways that obscured the patients' view of the screen (although see Figure 2).

The full DIALOG+ protocol is not strictly followed in the consultations. All participating clinicians were trained in the use of the protocol—and the application in front of them is also structured according to the protocol and contains prompts. However, steps are sometimes skipped, merged or discussed in varying orders. These deviations typically arise as adaptations to the immediate conversational context, particularly where clinicians' and patients' expectations or interpretations are misaligned. In general, this merging and skipping was designed to prioritise the flow of the conversation over the protocol.

A specific structural problem arises where patients mention a concern early in a consultation and then this same concern is reintroduced later in the conversation by the DIALOG+ protocol. This creates a sequential problem in the conversation—and the pragmatic effect of these repetitions is to

imply that either the patient's original description was somehow inadequate, or that the clinician had not been listening (see also below).

Another recurring issue is understanding the relevance of patients' responses to questions about particular items. There are often natural interconnections between, for example, people's accommodation situation and their family situation (see e.g., Excerpt 1 and Excerpt 2). The protocol questions imply a level of conceptual independence between different quality of life domains, but this may be misaligned with the concerns and practical circumstances of individual patients.

A third common source of trouble is clinicians' concurrent typing or note-taking, which was interspersed throughout discussions. These activities disrupt the flow of conversation and sometimes cause a misalignment on what should come next in the conversation.

In the control group, clinicians usually make notes throughout the consultations, but spend less time on this than clinicians in the intervention group (partly because typing on the tablet keyboard is a slow process). Similar issues in the patients' lives are explored—but more rapidly and lightly in conversation. This sometimes led to sessions with no action plans agreed upon. In other sessions, time is spent actually taking action on and resolving a particular issue the patient is facing (e.g. making calls; filling out forms). There are cases of clinicians finding out that pressing issues in the patient's life have been missed from previous consultations, simply because a particular topic never came up. Additionally, there are fewer opportunities to review or make conclusions about what has been discussed during the sessions, contributing to consultations that appear even less structured.

5.2 Understanding and Misunderstanding

As noted, the standardised format of the protocol questions can be a source of trouble. One problem arises from differences in the interpretation of apparently simple phrases. For example, the interpretation of "practical help" was a source of difficulty in more than one consultation. Each patient's personal circumstances are different and can involve problems that are partly or wholly outside the practical, ethical and institutional competence of the clinicians. Patients typically recognise that these limits exist but they do not have the context to be able to decide what "practical help" they can rea-

sonably seek. In addition, different clinicians place different boundaries around what they consider to be practically possible. These multiple sources of indeterminacy lead to clinicians to paraphrase and extend these key phrases (see e.g. below) or list possible examples of what they consider might constitute "practical help".

In CA terms, many of the problems created by the protocol relate to sequential appropriateness. For example, a clinician and patient are discussing the latter's physical health, specifically the pain in his leg that had been troubling him for a while. The patient mentions that he had been taking ibuprofen to help relieve the pain, but the medication has not been very effective. He also mentions that he is going to see his General Practitioner. The clinician then asks the patient what he thinks the "best case scenario" would be. The patient has already talked about the steps he is taking to improve the situation and finds it difficult to interpret the question in context. He pauses and eventually explicitly says he does not understand the question. In response, the clinician tries to reformulate "best case scenario" as something that the patient would "look forward to" or see as being "more satisfying" for his physical health. In response, the patient tuts, sits back and repeats that the pain in his leg has been "holding [him] back" from his daily activities; he wants to get rid of it. The timing of the responses, the repetition and the posture changes convey his sense of frustration at the apparently irrelevant question.

5.3 Typing, Distraction and Repetition

Points of potential miscommunication are often associated with clinicians dividing their attention between typing on the tablet and engaging with the patients' conversations. The notes are often only partially visible, in the sense that the patient can see that the clinician is typing, but often not *what* they are typing. Although the patients pay attention to the concurrent activity and usually suspend speaking, they do still sometimes make follow-up comments—but this may or may not tie into what the clinician is typing. Clinicians sometimes try to compensate, e.g. using concurrent *outlouds* of what is being written (Heath and Luff, 1992) to mitigate this. The concurrent activities place significant cognitive demands on the clinician and make it more difficult for them to track what is said. In the example immediately above and in Extracts 1 and 2 below, the clinician hears the patient but has

trouble integrating what they are saying with the current step in the protocol, one that is often not visible to the patient.

Repetition of questions can be especially problematic. In one case, a patient had complained, several times, that her medication is causing her a lot of physical symptoms—including vomiting—and that she needs to change her medication. However, the clinician is following the protocol and only noting responses relevant to the current item. When the clinician then asks about the vomiting problem and what she could do about it, the patient becomes visibly frustrated; she tosses her hands up in the air, says “I don’t know” and sighs.

The preceding examples illustrate some common sources of trouble in the consultations and also the integration of verbal and non-verbal resources used in response to them. An important feature of the way misunderstandings are addressed in these interactions is the way they make use of the shared space.

5.4 The Tablet as a Resource for Coordination

The layout of chairs and tables in the consultation rooms has a direct influence on how the participants orient to each other (Kendon, 2010). Direct face-to-face positioning is rare. The typical arrangement is an l-shape with two chairs arranged at a 45-60 degree angle and a table in-between (see Figure 2). The tablet is typically placed on the table and angled towards the clinician. The importance of its position in-between the participants and the influence of this on the management of the conversations is significant in all the intervention videos.

Changes in posture, reflecting shifts in orientation between the tablet and the other person present, help to mark important changes in participants’ focus and level of engagement (Bull, 2016). One example is the way clinicians start a new question by simultaneously displaying a shift of attention from the tablet to the patient by sitting back and turning their head toward the patient. Similarly, gestures to the tablet and gestures placed between the tablet and the other participant are used to propose or reintroduce items on the tablet screen as relevant to the ongoing discussion, e.g. as a prompt to shift from complaints to the possible solutions that need to be entered in a dialogue box (see Figure 2 and 3).

The significance of gestures is also illustrated by

examples of how misunderstandings occur when the tablet is *not* used effectively for reference coordination. At discussion stages, some clinicians review the action plans agreed on with their patients for each item, before moving on to discuss the next item. They usually wrap things up by asking patients if there is anything else they want to add, while gesturing towards that item on the tablet. In the absence of such gestures, patients can interpret the question more widely. For example, a clinician and patient were discussing action plans for the latter’s job situation when the clinician asks, while typing, if the patient thought there was anything she could do for him. The patient starts describing how she could possibly get in touch with the housing association in charge of his case; something that the patient had mentioned earlier during the consultation because he wanted to speed up his relocation (with his wife and child) to permanent housing. At this point, however, the clinician points to the item on the tablet that is specific to job situation to help reinterpret their original question—in CA terms a form of third-position repair.

In another example, the wrap-up discussion is on physical health, when the clinician asks—again while typing—if there is anything else the patient wants to tell her. The clinician glances up briefly after asking this question and notices the patient looking down at the tablet. She then adds a more explicit statement about the item she was referring to. This coincides with the patient’s request for clarification. As such, these examples illustrate how the tablet display not only has a role in managing ambiguity, but also what people are engaged with, as a means of diagnosing potential trouble sources.

5.5 Embodied Resources for Managing Misunderstanding

The complex configuration of bodies and artefacts in shared space during misunderstandings can be illustrated by considering two extracts from one consultation.

In Excerpt 1, C reorients from the screen to P and asks about P’s family situation, to which she responds with a comment about accommodation. C fails to understand the relevance of the response, partly because they have just talked about a different aspect of accommodation and he initiates a repair in line 3 by quickly pointing to the screen whilst still looking at his notes (2) on the first “its” before verbally repeating the topic. P tries again



Figure 2: C Points to Screen with Left Hand

in overlap on line 4 and when she gets the floor in line 5—and reformulates her point about accommodation preceded by a small gesture for emphasis. C nods twice towards P to acknowledge the importance of P’s reformulation which contains a significant complaint about being denied access to her possessions but then queries who “they” are, possibly thinking this could be a reference to family, which P answers in overlap (“housing”). At this point C is unable to integrate P’s turns into his understanding of what is going on. He switches to a generic clarification question at line 8. In line 9, P repeats that she was locked out and after the truncated question in line 10 C clearly orients to P. P’s gestures then become more emphatic, peaking at “things” and then reducing as C turns back to his notes.

Excerpt 1: Discussion of Relationships and Family (Key: C = Clinician, P = Patient, Underlining = emphasis, ↑ = rising intonation, :: = lengthened sound, // = onset of overlapping speech, [] = non-verbal action, (.) = short pause).

1 C: [selects item] looking at [sits back, looks at P] what makes you dissatisfied with your family situation [moves forward to pick up pen]
 2 P: its it was originally to do with my accommodation (pause)
 3 C: well no it’s to [points at screen] // it’s to do with your family yeah ↑
 4 P: // but when when it
 5 P: when they um:: [raises hands slightly] when they didn’t let me get my things and the power of attorney was lost [C

nods twice] (.)
 6 C: when your say they they wouldn’t let you get // your things
 7 P: // the housing
 8 C: [head down] explain to me what you mean
 9 P: // was locked out of my hostel
 10 C: // is this to do with your quest-
 11 P: I was locked out of my hostel [hands raised] and all my things [hands wide apart] were in there clothes [C’s head goes back to notes, P’s hands come back together] I lost everything [extra gesture then hands drop to lap]

Excerpt 2 comes from the same session approximately four minutes later. In the intervening interaction, C has focused on how P might retrieve her belongings. However, C is still having trouble understanding why P is raising the issue about personal belongings and power of attorney, given the protocol item (*Friends and Family*) currently in front of them. C tries to reintroduce this item using a flat hand gesture towards the screen on “family”. P sits forward and directly reformulates this, using a particularly direct form of other-initiated other-repair, accompanied with an emphatic hand gesture. C quickly acknowledges this with a nod and verbal acceptance, but is still showing signs of trouble. He asks a question about an address shared previously, which P answers, partly by taking hold of her bag where she has kept it, but this still doesn’t resolve the issue.

At line 7, C gestures to the screen and then produces a finger pointing up ‘hold’ gesture positioned between them 3. He then produces some filled pauses, makes a small flick of the pointing gesture and explains he needs to remind himself while looking at the screen. At this point, P sits back simultaneously with C dropping his finger point to gesture back to the screen. C’s Line 8 is formatted as reasoning out loud, but the truncated um: and long pause invites a possible response from P. In line 9, C then directly asks why P wants power of attorney. At this point the connection to *Family and Friends* finally becomes clear. P needs the missing power of attorney so she can attend to her (living)



Figure 3: C Points Up ‘Hold’ Gesture, Right Hand

grandmother’s financial affairs.

Excerpt 2: Continued Discussion of Actions Around Relationships and Family.

1 C: now this is don’t forget
this is all to do with family
[gesture at screen, P sits
forward] // umm

2 P: //it’s to do with power of
attorney [hands apart gesture]
that I was gettin to =

3 C: = [nods] yeah and its all
to do with power of attorney
[hand moves to screen] umm
because do you remember I give
an address didn’ I [short eye
contact]

4 P: yeah [P reaches to bag, C
looks down touches box] I’ve got
that here.

5 C: and did you ever do
anything with ↑ that

6 P: no that’s what we were
gonna discuss today [pause]

7 C: so:: [3.0s pause while
gestures to screen than puts
finger up in the air] cus right
oka:y let me just remind myself
now [C point back to screen, P
sits back] um:: agreeing on
actions [hand hovers over screen
pause] um:: [pause]

8 C: cos if you had power of
attorney then you’d be able to:
um:: [4.0s pause]

9 C: what is it that you need
power of attorney [P sits
forward] to achieve now =

10 P: = to sort our my
grandmother’s affairs

6 Discussion

The examples presented above raise a number of basic points about the organisation of face-to-face interaction and the use of embodiment in shared space as a resource for communication.

The data presented highlight the tension between the use of the protocol as a standardised assessment instrument and its function as a tool to promote effective conversation. Ideally, quantitative assessments of quality of life should be consistent across different participants and different contexts. In practice, the meaning of the different assessment items and even the meaning of the numbers on the assessment scale varies across consultations. Patients and clinicians routinely engage in active, collaborative re-interpretation of the protocol in order to complete the assessment. Standard phrases such as *best case scenario* and *practical help* take on specific meanings depending on individual circumstances. What is ostensibly the same question means different things to different people and can also mean different things to the same people in different sequential contexts.

The observations show the process of detecting and dealing with differences in interpretation is fundamental to effective communication (Healey et al., 2018). Although this recurrent interpretive work means the application of the DIALOG+ protocol is not strictly standardised, it is part of normal interaction and arguably central to the therapeutic effectiveness of the intervention. The assessment items provide prompts that encourage a wider ranging conversation and greater continuity across sessions than observed in the control groups. The more strictly the protocol is applied, the more friction it would cause to the conversation (Drew et al., 2021; NHS Improvement, 2018). The work people do to bridge the gap between the protocol and the details of individual’s lives can play an important role in uncovering the different combinations of practical circumstances and constraints that influence long term outcomes.

Moving to remote delivery directly alters the configuration of resources available to participants in the interaction. One effect will be to change the visibility and control of actions in the application. As noted, input is currently led by clinicians and often not directly visible to the patient. If the ap-

plication is running in a shared window all updates will be immediately visible to both participants—and control of input by one person will not automatically restrict input by the other. This should mitigate, for example, problems with coordinating when a concurrent action (typing notes) starts and finishes—and also what item is currently under discussion. However, the work of (re)interpreting items and actions would still be required. Other possible advantages of remote interaction are the savings in cost and time, as well as the potential for improved access for some patients.

Nonetheless, some important features will be removed by remote interaction. One is the use of the embodiment of the protocol itself as a shared screen positioned in space between the participants. There are two ways in which this matters. First, changes in people's overall physical orientation, through head movements or posture shifts signal important changes in their focus of attention and engagement. This is especially true of the clinicians, who often use posture changes to and from the device and to and from the patient to mark changes in engagement e.g. to introduce a direct question to the patient (see e.g. the first line of Excerpt 1). Secondly, clinicians (and occasionally patients) have the ability to point at a question and, for example, gesture from the question to a person. These movements, in effect, use the shared space to provide a useful spatial map for the embodied coordination of topics (Deppermann et al., 2021; Guxholli et al., 2021; Kitzinger, 2012).

It is also noticeable from the preceding examples that where a gesture is specifically placed in space is significant. The pointing gestures in Figures 2 and 3 have the same form, but their different speed of execution, orientation and placement give them a different interpretation. In an example from the same session not included in the Extracts, there is a rapid shift by the clinician from a short two finger point at the screen to a two finger point at the patient. The form and speed of the gesture is very similar but its significance is different for the participants because of where it is placed, a phenomenon also noted in other face-to-face contexts (Battersby and Healey, 2009; Özyürek, 2002). It also illustrates the ways in which people can use contrasting head and hand orientation as a means of concurrent triangulation of different reference points (e.g. people and objects) to help coordinate understanding (Battersby and Healey, 2010).

These examples also illustrate how these additional, embodied resources seem to become especially significant at the points where shared understanding is threatened (Healey et al., 2013, 2015; Özkan et al., 2021). The space between participants becomes an important extra resource for detecting and dealing with these misunderstandings.

Remote video-mediated interactions flatten the three-dimensional world of face-to-face interaction in shared space into a two-dimensional window (Mlynář et al., 2018). A shared application and a video window impede the forms of interaction highlighted above. Working out what someone is orienting to requires more effort—and although posture shifts and gestures may be visible on camera, they are attenuated and cannot take advantage of relative position in a shared space. People are able to compensate in these situations and it is an open question what the cost of adapting could be. Shared applications that enable people to see each other's cursors can partially replicate a sense of the current focus of attention. This may help with the redirection of attention, but provides a significantly reduced set of cues.

One way in which remote interaction can give greater access is to allow asynchronous updates. For example, allowing patients to add notes or modify ratings in response to events outside the clinical context. This could give patients the opportunity to understand the protocol and application better and also feel that they are on a more equal footing with clinicians. It could also provide clinicians with potentially useful additional context updates for use in the face-to-face sessions. This shift in the distribution of control over the tool might impact on the dynamics of clinical interaction and careful design would be needed to avoid it becoming overused as a communication channel.

There are some qualifications to the findings. The observations are selective and are based on a specific population of patients with a diagnosis of psychosis. Although we think the general principles should apply in other contexts, different issues will be encountered with other communication tools and/or other client groups e.g. telephone delivery (Drew et al., 2021). Nevertheless, this study has shed light on some of the concerns and challenges related to designing health communication protocols and some specific issues for remote settings. In the wake of the pandemic, while remote consultations are unlikely to fully replace

face-to-face consultations, it will no doubt become a feature increasingly integrated into current health systems and used alongside conventional practice (Khan et al., 2021). Care needs to be taken to ensure that developments surrounding technologies like DIALOG+ are balanced with appropriate flexibility, as every nuance in communication between patients and clinicians can have a role to play in influencing the quality of therapeutic relationships and the effectiveness of clinical encounters.

7 Conclusion

Structured protocols are increasingly used in community mental healthcare consultations. Detailed analysis of the interactions using one of these protocols (DIALOG+) and its associated tablet application shows some of the advantages and pitfalls of the approach. The application provides a useful tool to support engagement in the consultations, but in practice, deviations from the protocol play an important role in the success of the consultations. The interactions are characterised by collaborative work done to (re)interpret the assessment items in the context of each client's and each clinician's practical circumstances. Participants use embodiment in shared space as an important, flexible interactional resource in doing this. With remote consultations increasingly integrated into healthcare settings, our findings provide a starting point for thinking about how software like the DIALOG+ application can be redesigned for these environments.

Acknowledgments

We are grateful to the National Institute for Health Research (NIHR) for funding Healey and Galindo-Esparza's contribution to this paper through the project "Remote delivery of an app-based intervention (DIALOG+) in community mental health care-development" (Reference: NIHR201680) and to the Wellcome Trust for funding Law's contribution through the programme "Health Data in Practice: Human-centred Science" (Reference: 218584/Z/19/Z).

References

Stuart Battersby and Patrick G T Healey. 2010. Head and hand movements in the orchestration of dialogue. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 32.

Stuart A Battersby and Patrick G T Healey. 2009. The interactional geometry of a three-way conversation.

In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 31.

Janet B Bavelas. 2007. Face-to-face dialogue as a micro-social context. S. Duncan, E. Levy, & Cassell (Eds.), *Language in mind, body, and context*, pages 127–146.

Janet Beavin Bavelas, Nicole Chovil, Linda Coates, and Lori Roe. 1995. Gestures specialized for dialogue. *Personality and social psychology bulletin*, 21(4):394–405.

Peter E Bull. 2016. *Posture & gesture*, volume 16. Elsevier.

Nicole Chovil. 1991a. Discourse-oriented facial displays in conversation. *Research on Language & Social Interaction*, 25(1-4):163–194.

Nicole Chovil. 1991b. Social determinants of facial displays. *Journal of Nonverbal Behavior*, 15(3):141–154.

Arnulf Deppermann, Lorenza Mondada, and Simona Pekarek Doehler. 2021. Early responses: An introduction.

Paul Drew, Annie Irvine, Michael Barkham, Cintia Faija, Judith Gellatly, Kerry Ardern, JC Armitage, Helen Brooks, Kelly Rushton, Charlotte Welsh, et al. 2021. Telephone delivery of psychological interventions: Balancing protocol with patient-centred care. *Social Science & Medicine*, 277:113818.

Stuart Ekberg. 2021. Proffering connections: Psychologising experience in psychotherapy and everyday life. *Frontiers in Psychology*, page 3686.

Paul Ekman. 1979. About brows. *Emotional and conversational signals in: von Cranach, M./Foppa, K. ua (Hrsg.)(1979): Human Ethology, Cambridge*.

Simeon Floyd, Elizabeth Manrique, Giovanni Rossi, and Francisco Torreira. 2016. Timing of visual bodily behavior in repair sequences: Evidence from three languages. *Discourse Processes*, 53(3):175–204.

Charles Goodwin. 1979. The interactive construction of a sentence in natural conversation. *Everyday language: Studies in ethnomethodology*, 97:101–121.

Aurora Guxholli, Liisa Voutilainen, and Anssi Peräkylä. 2021. Safeguarding the therapeutic alliance: Managing disaffiliation in the course of work with psychotherapeutic projects. *Frontiers in Psychology*, page 3905.

Patrick G T Healey, Jan P De Ruiter, and Gregory J Mills. 2018. Editors' introduction: Miscommunication. *Topics in Cognitive Science*, 10(2):264–278.

Patrick G T Healey, Mary Lavelle, Christine Howes, Stuart Battersby, and Rosemarie McCabe. 2013. How listeners respond to speaker's troubles. In *Proceedings of the annual meeting of the cognitive science society*, volume 35.

- Patrick G T Healey, Nicola Jane Plant, Christine Howes, and Mary Lavelle. 2015. When words fail: Collaborative gestures during clarification dialogues. In *2015 AAAI Spring Symposium Series*.
- Christian Heath and Paul Luff. 1992. Collaboration and control/crisis management and multimedia technology in london underground line control rooms. *Computer Supported Cooperative Work (CSCW)*, 1(1):69–94.
- Timo Kaukoma, Anssi Peräkylä, and Johanna Ruusuvaara. 2014. Foreshadowing a problem: Turn-opening frowns in conversation. *Journal of Pragmatics*, 71:132–147.
- Kathrin Kaulard, Douglas W Cunningham, Heinrich H Bülthoff, and Christian Wallraven. 2012. The mpi facial expression database—a validated database of emotional and conversational facial expressions. *PLoS one*, 7(3):e32321.
- Adam Kendon. 2010. Spacing and orientation in copresent interaction. In *Development of multimodal interfaces: Active listening and synchrony*, pages 1–15. Springer.
- Kobin H Kendrick. 2015. Other-initiated repair in english. *Open Linguistics*, 1(1).
- Kobin H Kendrick and Judith Holler. 2017. Gaze direction signals response preference in conversation. *Research on Language and Social Interaction*, 50(1):12–32.
- Abdul Waheed Khan, Nisha Kader, Samer Hammoudeh, and Majid Alabdulla. 2021. Combating covid-19 pandemic with technology: perceptions of mental health professionals towards telepsychiatry. *Asian Journal of Psychiatry*, 61:102677.
- Celia Kitzinger. 2012. Repair. *The handbook of conversation analysis*, pages 229–256.
- Elisa Liberati, Natalie Richards, Jennie Parker, Janet Willars, David Scott, Nicola Boydell, Vanessa Pinfold, Graham Martin, Mary Dixon-Woods, and Peter Jones. 2021. Remote care for mental health: qualitative study with service users, carers and staff during the covid-19 pandemic. *BMJ open*, 11(4):e049210.
- Neda Mahmoodi, G Jones, Tom Muskett, and Sally Sargeant. 2020. Exploring shared decision making in breast cancer care: A case-based conversation analytic approach. *Communication and Medicine*, 16(1).
- Rose McCabe and Patrick G.T. Healey. 2018. Miscommunication in doctor–patient communication. *Topics in Cognitive Science*, 10(2):409–424.
- Rosemarie McCabe, Patrick GT Healey, Stefan Priebe, Mary Lavelle, David Dodwell, Richard Laugharne, Amelia Snell, and Stephen Bremner. 2013. Shared understanding in psychiatrist–patient communication: Association with treatment adherence in schizophrenia. *Patient education and counseling*, 93(1):73–79.
- Chris McVittie, Slavka Craig, and Margaret Temple. 2020. A conversation analysis of communicative changes in a time-limited psychotherapy group for mothers with post-natal depression. *Psychotherapy Research*, 30(8):1048–1060.
- Jakub Mlynář, Esther González-Martínez, and Denis Lalanne. 2018. Situated organization of video-mediated interaction: A review of ethnomethodological and conversation analytic studies. *Interacting with Computers*, 30(2):73–84.
- NHS Improvement. 2018. Spoken communication and patient safety in the NHS. Technical report, NHS.
- C Olwill, D Mc Nally, and L Douglas. 2021. Psychiatrist experience of remote consultations by telephone in an outpatient psychiatric department during the covid-19 pandemic. *Irish journal of psychological medicine*, 38(2):132–139.
- Elif Ecem Özkan, Tom Gurion, Julian Hough, Patrick GT Healey, and Lorenzo Jamone. 2021. Specific hand motion patterns correlate to miscommunications during dyadic conversations. In *2021 IEEE International Conference on Development and Learning (ICDL)*, pages 1–6. IEEE.
- Asli Özyürek. 2002. Do speakers design their cospeech gestures for their addressees? the effects of addressee location on representational gestures. *Journal of Memory and Language*, 46(4):688–704.
- Stefan Priebe, Eoin Golden, David Kingdon, Serif Omer, Sophie Walsh, Kleomenis Katevas, Paul McCrone, Sandra Eldridge, and Rose McCabe. 2017. Effective patient–clinician interaction to improve treatment outcomes for patients with psychosis: a mixed-methods design. *Programme Grants for Applied Research*, 5(6).
- Stefan Priebe, Lauren Kelley, Serif Omer, Eoin Golden, Sophie Walsh, Husnara Khanom, David Kingdon, Clare Rutterford, Paul McCrone, and Rosemarie McCabe. 2015. The effectiveness of a patient-centred assessment with a solution-focused approach (dialog+) for patients with psychosis: a pragmatic cluster-randomised controlled trial in community care. *Psychotherapy and psychosomatics*, 84(5):304–313.
- Harvey Sacks. 1992. Lectures on conversation: Volume i. *Malden, Massachusetts: Blackwell*.
- Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. 1978. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*, pages 7–55. Elsevier.
- Albert E Scheflen. 1973. *Communicational structure: Analysis of a psychotherapy transaction*. Indiana U. Press.
- Mi-Suk Seo and Irene Koshik. 2010. A conversation analytic study of gestures that engender repair in esl conversational tutoring. *Journal of pragmatics*, 42(8):2219–2239.

David Silverman. 1998. *Harvey Sacks: Social science and conversation analysis*. Oxford University Press on Demand.

Yijin Wu. 2020. Pain talk in hospice care: a conversation analysis. *BMC Palliative Care*, 19(1):1–8.