Language & Communication 99 (2024) 313–332

Contents lists available at ScienceDirect

Language & Communication

journal homepage: www.elsevier.com/locate/langcom

Instructed perception and action: The mutual accomplishment of manual know-how in using VR games

Arja Piirainen-Marsh*, Margarethe Olbertz-Siitonen

The University of Jyväskylä, Department of Language and Communication Studies, Finland

A R T I C L E I N F O

Article history:

Keywords:
Manual know-how
Multimodal interaction
Instructions
Instructed action
Virtual reality

A B S T R A C T

This paper investigates how manual know-how is manifested in instructed action in a technologized setting characterised by distributed spaces and bodies. Drawing on data from a temporary game lab where co-present participants try out virtual reality games, it analyses how an experienced VR user instructs novice participants in accomplishing the steps required to set up the game, in interacting with virtual objects and resolving troubles with tasks in the virtual environment. We show how instructions are fitted to and serve to advance the practical activity, and how they are recipient designed to teach novice users the manual and bodily techniques for interacting with the virtual game and its features.

1. Introduction

Immersive VR environments require not only adjustment to a fractured ecology characterised by distributed spaces and bodies (Luff et al., 2003; Kohonen-Aho and Haddington, 2023), but also training in manual skills that enable players to bring about changes in the virtual environment. Games are operated through hand-held controllers and players need to learn the correct controller techniques for the game to become playable (Hodges 2017). For novice users, the relationship between physical, manual actions and the events they generate in the virtual environment in real-time interaction is a continual challenge. The haptic techniques through which controllers are operated are not taught by the game itself but often involve informal instruction and guidance from more experienced players or using other resources such as instructional videos. This paper examines how manual know-how is manifested and transmitted in the practices through which an experienced VR user instructs and guides novice players in the task of setting up the VR system, using the tactile interface (controllers) and interacting with features of the game. More specifically, we analyse the practices through which the novice users are instructed on the manual and bodily techniques for accomplishing the tasks through which the game's playability is established and maintained.

Starting and engaging with a VR game requires knowledge not just of the technology in use, but also expertise that is manifested, built, and developed through coordination of embodied courses of action: handling, testing and using the controller(s), calibrating the body for the requirements of the game, attending to game mechanics and interacting with virtual objects. Such activities become intelligible in relation to structures of the environment (e.g. Goodwin 2000, 2018), which are complex and fragmented. The virtual environment projected by the head mounted display is only available to the user, whose agency is contingent on *dual embodiment* (Kohonen-Aho and Haddington 2023; Cleland 2010), i.e. extension of their body in

* Corresponding author. Department of Language and Communication Studies, P.O. Box 35, FIN-40014 University of Jyväskylä, Finland.
E-mail address: arja.piirainen-marsh@jyu.fi (A. Piirainen-Marsh).

https://doi.org/10.1016/j.langcom.2024.10.011

the virtual world and the ability to use the resources afforded by the physical and virtual environment to make sense of activities and produce actions. The manual and bodily techniques to be learnt and used are contingent on material and technological features of the VR system in use. For example, material features of the controllers and the quality of the immersive technology impact affordances for action.

VR provides a simulated experience of a virtual environment and its scenes, objects and spaces. The technical aim is to replace the user's real sense perceptions (vision, sound, touch) with new ones while enabling them to move and use their body, for instance look around, reach and manipulate objects, in a natural way (Slater and Sanchez-Vives 2016). Standard VR systems, such as the one used in our data, still rely heavily on vision and to some extent sound. For users who are not accustomed to the technology, these are enough to disrupt the user's ability to recognise and make sense of phenomenal details in the environment, and thereby their ability to follow instructions (Garfinkel 2002: 210). Novice users' interactions with VR technology therefore provide occasions to explore how participants deal with the lack of coherence between their embodied know-how and the environment. VR and digital games can be seen as one type of technology of instruction (Lynch and Lindwall, 2024; cf. e.g. Brown and Laurier 2012 for GPS navigation) that provides directions on how to proceed and feedback in visual and auditory form. Novice participants often need guidance in noticing and interpreting such system instructions and understanding how to follow them. Our interest is in the way such system instructions are made understandable and followable in situated interaction whilst negotiating the practical steps of setting up and interacting with the game.

This study is situated in the ethnomethodological and conversation analytic traditions of research into instructions and instructed action. Aligning with recent research on instructions on bodily action (e.g. Lindwall and Ekström 2012, Mondada 2014a, 2014b; de Stefani 2018; Rauniomaa et al., 2018, Harjunpää et al., 2021), it focuses on the embodied practices of issuing and following instructions in the context of using VR technology. We investigate how these practices show orientation to manual know-how that is required to operate the hand-held device and interface and engage with objects and features in the virtual environment. Our focus is on two activity environments: (1) pre-game activities where the participants interactively manage the steps required by the VR system to start the game and make it playable and (2) moments of trouble during accomplishing in-game tasks. The pre-game phase involves instructive actions by the more experienced VR user that guide novice participants through the steps of establishing a connection with the system and using the hand-held controllers. During in-game activities players often display difficulties in understanding how to proceed and how to accomplish a game task. These moments make relevant instructions and instructed action through which the participants work to resolve the trouble. We aim to show how manual know-how is manifested in instructional practices and how the participants draw on fragmented resources to develop an understanding of how to proceed.

2. Instructions and instructed action

Following instructions in different formats (including written directions, maps, etc.), i.e. instructed action, is a central topic in Garfinkel's (2002) program for ethnomethodology, which emphasizes the here-and-now, achieved nature of social order. In EMCA research the term instruction has been used in varied ways to refer to instructional or pedagogical practices in different learning environments, instructions in written form or to social actions akin to directives, i.e. sequence-initial actions that make relevant (complying) future actions either immediately in the here-and-now or later (Lindwall et al., 2015; Mondada 2014a, 2014b). In this study we apply a holistic multimodal approach to instructions as embedded in embodied courses of action through which novice participants are guided through the tasks associated with setting up and using immersive VR.

Recent studies of instructions on bodily actions have brought new insights into verbal and bodily practices of giving and following instructions in settings where manual and bodily skills are taught and learned (see e.g. Lindwall and Ekström 2012 for craft instruction; Hindmarsh et al., 2014, Lindwall and Lymer 2014, 2023 for dental instruction; Mondada 2014a for surgery, Mondada 2014b for cooking classes; De Stefani 2018; Deppermann 2018a, 2018b for driving lessons). This research has shown how actions are formulated as instructions through verbal, gestural and bodily resources assembled into multimodal formats, how they are recipient designed in ways that are sensitive to displayed competencies of the participants and how they become intelligible and acquire their sense through the (complying) actions that they make relevant in specific social and material environments. Linguistically instructions can take a variety of verbal formats depending on the activity type and sequential environment in which they occur (see e.g. Deppermann 2018a, 2018b). Studies of practical actions have shown, however, that a linguistic, sequential approach to instructions does not capture the intricate practices through which participants design their utterances as instructive of what to do next and how these are oriented to by recipients (Lynch and Lindwall 2024).

Following instructions (instructed action, Garfinkel 2002; Lynch and Lindwall 2024) exhibits an embodied and situated understanding of the instructions in the courses of action in which they are produced. In embodied courses of action sequences can involve multiple temporal orders related to the practical, physical activities in progress. Instruction can involve parsing, in other words instructive actions can be timed and designed to address specific sub-actions when carrying out a task step-by-step (Rauniomaa et al., 2018) and they can unfold simultaneously with the bodily activities that they guide in a way that shows close monitoring and analysis of these activities. Studies of instructed manual and physical activities such as sports training, for example, show how instructions attend to techniques of the body (Mauss 1973) and involve bodily enactments and demonstrations that are used to correct an ongoing performance (Evans 2017; Råman and Haddington, 2018; Evans and

Lindwall 2020). Similarly, studies of work activities including technology-mediated tasks (see e.g. Goodwin 2000; Heath and Luff 2000; Lindwall and Lymer 2024) show that instructions for practical action are designed to be seen by the recipients (e.g. through gestures or bodily demonstrations). They can also involve manual or bodily interventions that guide or direct the less skilled participant's performance (Kääntä and Piirainen-Marsh 2013; Mondada 2014b). The practices deployed make visible how instructions involve assessment of the techniques and outcomes of the actions performed in following instructions (Lynch and Lindwall 2024).

Instructive and instructed actions are closely intertwined with the tasks and activities that they are part of and the competencies that participants display when they work out ways to proceed with these tasks (Garfinkel 1996, 2002; Nishizaka 2006, 2014; Rauniomaa et al., 2018). A characteristic feature of instructions is that they involve diverse resources to segment the task or activity into steps and unpack the skills to be practiced or learnt (Lindwall et al., 2015). The practices deployed in this are designed to enable the novice participant to carry out the required actions in a coordinated, stepwise manner (Lindwall and Lymer 2024) and to resolve troubles that arise in the process. Our analysis contributes to understanding how instructions are recipient designed and made intelligible in conditions of distributed resources and restricted access to visual fields. The analysis reveals how manual know-how is manifest in practices of instruction through which the novice users learn to use game controllers and the body to connect with the VR system and interact with the game.

3. Data and method

The data comes from informal, co-located VR-gaming sessions where young adults were invited to try out virtual reality games. The main purpose of these meetings was to observe and gather video-material of social interaction involving VR to examine the practical accomplishment of getting acquainted with unfamiliar technology and immersive games. Although organised for research purposes, the set-up of the game lab did not include any predetermined tasks or structure. Rather, the interaction unfolded freely enabling natural engagement with the technology. Nevertheless, the recordings are clearly embedded in a 'purposive', or 'premediated' ecology of action, which is reflected in the situated organization of turns at play (i.e., one-at-a-time), and in the way that the encounters are organised as instructional environments. Our data, therefore, can be described as semi-natural or "quasi-experimental" (Kohonen-Aho and Haddington, 2023), placing it somewhere at the center of the continuum from naturalistic to experimental data collection practices in interaction research (Kendrick 2017). Aligning with previous studies on technology-in-interaction that have used data from laboratory-like settings (e.g. Heath and Luff, 2018), we believe that the data can provide new insight into instructed action in a perspicuous setting that remains understudied.

The game sessions lasted between 135 and 155 min, amounting to a total of approximately 500 min of recorded material. The participants were university students and members of the research group. The data are multilingual in that the language of the games was English, while the languages of interaction were Finnish and English, including situated alternation between them. All participants gave written consent for the use of the material for research and publications. The available equipment (using a PS4 console, as well as one PlayStation VR headset, a wireless controller, and one set of PlayStation Move motion controllers) enabled only one person to engage with VR. However, others had certain audio-visual access to the player's perspective via an external screen and loudspeakers. The participants switched between playing and spectating, i.e., employed a one-at-a-time turn-taking system for trying out a variety of games. Despite some having gaming experience, most participants had little or no prior experience of using VR technology. One participant, an involved colleague (Max), was a proficient VR user and was positioned as the party with sufficient knowledge to instruct others. The games were selected by the participants in situ and ranged from manual puzzles, platformers, and simulations, through to shooting games.

To get an encompassing view of participants' physical and virtual activities, we recorded the situations from several perspectives, rendering synchronized video material that displays the player in the physical environment from the front, a wider view from the back of the lab, and the screen with the game feed. This enabled analysis of the player's physical movements and actions as well as how these actions were realised in the virtual environment. The data was transcribed using the conventions developed by Mondada (2018) for multimodal conversation analysis. In addition, Finnish examples include a free translation into English placed directly below the first line in italics. We included the VR system in the transcripts when relevant, as was the case when participants oriented to textual, graphic, or voiced instructions (see Appendix A). In some examples we refer to actions carried with virtual hands (VH; left: LVH, right: RVH) as distinct from manual actions in the physical environment. The names of the participants are pseudonyms.

The analysis is grounded in multimodal conversation analysis (Mondada, 2019) as a method for investigating situated accomplishment and organisation of action through multiple modalities including resources of the body, materiality and space. The excerpts analysed in this paper were chosen as representative cases of ways in which participants orient to manual know-how in the pre-game phase and in the game when interacting with virtual objects and attempting to carry out bodily action in the virtual environment.

4. Analysis

In the sections to follow we examine how instructive and instructed actions are assembled in the practical activities of using immersive VR technology and interacting with a game. Our aim is to elucidate how instructive and instructed actions are situated in the tasks that they are part of and how the practices used show orientation to manual know-how as well as to distributed spaces and asymmetries of knowledge and expertise. The analysis is organised into three parts to illustrate the manual and other techniques of the body that are targeted by instructions and to unpack the skills that are involved. Section 4.1 focuses on the pre-game phase of activity and illustrates how the participants interactively manage the practical and technical steps required to set up and calibrate the game. For novice users, establishing connection with the game and accomplishing game tasks requires acquiring not only new manual skills but also guided perception. We show what kinds of bodily and interactional work is required to adjust to the fragmented resources and how novices are instructed in understanding features and techniques of the controllers. The following sections focus on moments of trouble experienced during in-game activities, specifically when interacting with virtual objects (4.2) and trying to achieve virtual mobility and whole-body action (e.g. climbing) (4.3). The examples illustrate how the participants orient to lack of congruence between bodily actions in the real world and the way these are actualized in the virtual environment. The expert player's instructions guide and correct novice players' use of the tactile interface and use of the physical body to accomplish tasks in the virtual environment.

4.1. Guided perception and manual techniques

This section examines how the participants orient to asymmetries of manual and bodily expertise in managing the steps required to set up the game and make it playable. The practical tasks this involves include adjusting the head mounted display, taking hold of and manipulating the controller(s) and positioning the body to enable tracking by the VR technology. Instructions by the more experienced participant orient to advancing the practical tasks and are designed to guide or assist the VR user in carrying out the required steps appropriately and at the right time for the activity to proceed.

In the examples presented below, the player is wearing the VR headset and has restricted visual access to their own and others' physical bodies and the physical space. In these initial moments the player recurrently shows orientation to the lack of coherence between their physical body and the virtual body, especially its most relevant parts, the hands and arms. The first example illustrates how this is observable in the way that the player orients to the virtual hands as extensions of her body and how this becomes a resource in negotiating the transfer of the game controllers from the more experienced player to the novice. Tessa, the user wearing the headset, stands facing the VR equipment and Max, the experienced player, stands next to her holding two controllers. The excerpt begins with Tessa's noticing of the virtual hands, which are also faintly visible on the screen (lines 1–2, Fig. 1).

Excerpt 1

Figure 1

Figure 2

```
01   Tes    •♦*#oh I ♦*#see my: (.)
     tes    •both hands on glasses
     max     ♦*gaze screen, holds two controllers
     max           ♦*gaze & body shift closer to T-->
     fig      #fig1   #fig2
```

Figure 3 Figure 4 Figure 5 Figure 6

```
02   Tes       *ha•#nd[s
     tes          •lowers BH -->
     max    ->*
     fig           #fig3

03   Max             *[yea bec•#ause *#they are •here# [hh hh hh
     max             *holds contr. in front of T, close to T's hands
     tes                   -->•RH touches RH controller, then grasps it
     tes                                          •takes LH controller
     max                        *....releases controllers-->
     fig                     #fig4  #fig5           #fig6

04   Tes                                                 [hh *hh hh
     max                                                     *withdraws

05   Max    *♦your hands are here.
     max    *walks further from T,
     max     ♦turns gaze to screen-->
```

Tessa has been occupied with adjusting the headset and still has both her hands on the glasses while the system is starting (l. 1., Fig. 1). Lines 1–2 show how Tessa shifts her attention from the headset as a material object to the virtual pair of hands now visible to her through the head-mounted display. With the 'oh'-prefaced noticing, she attends to the change in her visual field and with the rest of the turn orients to the virtual hands as an extension of her body in the virtual environment. An embodied shift occurs during the final turn component when she lets go of the headset and lowers both her hands. Through this turn, Max also shifts his position and gets ready for the next step, that is handing the two controllers he is holding to Tessa. This is visible in the way that Max shifts his gaze to her (Fig. 2), moves closer and holds the controllers so close to Tessa that they are within touching distance from her hands as she lowers them (Figs. 3–4). Max's turn in l. 3 attends to Tessa's noticing and verbally announces the hand-over of the controllers. The transfer takes place concurrently and smoothly when Tessa first touches and then takes each controller one by one as the turn unfolds.

The first example illustrates several features of instructions embedded within practical courses of action. Max's turn (l. 3) is part of a multimodal configuration that is instructive of and accomplishes the next step, i.e. passing of the controllers to Tessa to move the activity forward. It is occasioned by and built on Tessa's noticing, using it as a resource for verbally announcing the next step in the activity. The comment both accounts for the change in Tessa's visual field and aligns with her claim of ownership of the virtual hands, while also orienting to the imminent transfer of the controllers. Both participants treat the shift into the virtual space and passing of objects as an occasion for playful commentary, as manifested by shared laughter.

One of the most important tasks for the player is to get a haptic sense of the controller(s) and the features that are relevant for carrying out actions in the virtual world. When instructing the user about controller functionality, the more experienced participant deploys verbal directives and descriptions coupled with tactile interventions that (re)position the player's hands appropriately for effective manipulation of the controllers. These interventions involve referential touching of the controllers and sometimes the player's hands to help them locate relevant features (see also Olbertz-Siitonen and Piirainen-Marsh, 2023). The next example illustrates how tactile engagement with a specific feature of the controller (the trigger button) followed by manual guiding instruct the player about techniques of the object in preparation for interacting with the game. Max hands over both controllers to Sun one by one, but instead of releasing the second one, he holds on to it and begins to verbally instruct her on the relevant features (l. 2).

Excerpt 2


Figure 7


Figure 8

```
01  MAX   +ja anna #vasen käsi?# (*0.5) tuossa.
          and give the left hand? (0.5) there.
    sun   +...L hand towards controller, open palm
    max                             *places contr. in Sun's LH,
                                    holds controller w/RH--->
    fig             #fig7       #fig8
```


Figure 9


Figure 10

```
02  MAX   *eli #täällä (0.5) on se *lii#pasin. (.) the trigger.*
           so here (0.5) is the trigger
    max   *touches controller trigger with LH,
    max                             *moves Sun's finger on it  -->*
    fig         #fig9                    #fig10

03  SUN    [mhm,]

04  MAX   *[öö  ] (0.3) *#se on (0.3) ♦*#eiks s- se ol-
                        it is (0.3)  wasn't i- it wa-
    max   *releases BH  *points left  ♦*gaze to Lea, body torque
                                      points with R thumb tw Lea
                         #fig11        #fig12
```


Figure 11


Figure 12

```
05        *se oli #millä kaikki (tehtiin.)
           it was the one with which (one did) all.
    max   *circling gestures tw Sun, index fingers extended
                  #fig13
```


Figure 13

```
06  LEA   joo.
          yea.

07  MAX   ♦*eli kaikki tapahtuu °sillä° (0.6)
           so everything happens with that
    max   ♦turns gaze tw screen
    max    *lowers hands
```

Max references the trigger (l. 2) while touching the relevant button with his left hand (Fig. 9) as Sun holds it in her hand. Having located the trigger by touch, he then gently guides Sun's finger onto the trigger (Fig. 10) thereby allowing her to identify and get a haptic sense of it. Sun acknowledges this vocally in line 3. Lines 4–6 show a side sequence where Max seeks another player's (Lea) confirmation for his understanding of how the trigger is used in this game. After receiving confirmation, Max resumes full bodily orientation to Sun and shifts his gaze towards the screen. Concurrently with this he provides an upshot referring to the trigger as the relevant controller feature through which "everything happens" (l. 7). The coupling of indexical referencing and touching of the controller, and manual repositioning of the player's finger serve as instructive actions in that they orient to the player's lack of expertise on controller features and the manual techniques required for play. In the circumstances where the player has restricted visual access to the hand-held controller, tactile practices enable the player to acquire a haptic sense and understanding of the interface (Hodges, 2017) and learn controller techniques through their fingers (Parisi, 2011). The understanding of these techniques is demonstrated later in the interaction when the player independently uses the relevant feature (presses the trigger) in a competent and timely way.

To make the game playable, the participants also need to ensure that the player's physical body is correctly positioned to enable the VR technology to track its movements. The next example illustrates how Max guides Sun through checking her body position so that it matches the technical requirements. Max has handed two motion controllers to Sun and instructed her to "look down" to check that she is "standing in a circle" to enable tracking by the system. When the system indicates completion of this task, the participants move to the next technical step: measuring the distance between the user's extended arms when holding and using the controllers. To make this possible, the player needs to extend their arms fully and press the trigger on the motion controller to complete the task. The system instructs the user by displaying written and graphic instructions, but novice participants frequently have trouble noticing and understanding when and how to do the required actions. Here Max's verbal turns orient to such trouble by mediating and explaining the system instructions to make them followable. Max shows expertise in the pre-game set up procedures by verbalizing an instruction already before it is visually displayed (l. 1).

Excerpt 3

Figure 14


Figure 15

```
01   Max   ♦then it Ø#asks *you: (0.3) next to (0.4) reach out
     max   ♦gaze: screen -->
     max                   *....moves backward in squatting pos.
     VR             Øtask complete
     fig             #fig14

02         (0.4) with yourØ #↑a+rms?♦
     max                          ♦gaze: Sun
     sun                      +raises BH, adjusts VR glasses-->
     VR                 Øfully extend arms-->
     fig                  #fig15
```


Figure 16


Figure 17

```
03         it (0.6) #tri:es to measure how (.) how um:
    VR       -->pull both triggers
                    #fig 16

04         ♦(0.6+) +>how far you can< reach (0.8)
    max    ♦glances at screen, then gaze: SUN--->
    sun     --->+  +releases headset

05  Max    +so reach out #with your ↑arms, (.)
    sun    +...extends arms--->
    fig                  #fig17

06         and +then press the trigger.
    sun    --->+arms fully extended, hold--->
```

Figure 18

Figure 19

```
07         (1.3)+ +#(1.0)
    sun     --->+ +presses trigger, retracts arms
    fig             #fig18

08  Max    #good.+ (0.8)
    sun       -->+
    fig    #fig19
```

Max's turn in lines 1–2 is initiated during a moment of transition when the previous task has just been completed (indicated both by a sound and visual display 'task complete', Fig 14.) but before the next step begins. The turn is hearable as a preparatory announcement of the new step. It projects a system instruction which begins to unfold on the final component of the first turn unit (l. 2, Fig. 15). The instruction becomes visible gradually, so that a graphic representation becomes visible first, followed by two written directives. Details of Max's turn show orientation to the temporality of the game system and close monitoring of Sun's bodily conduct. At first Max follows the screen with his gaze while verbally formulating his turn. The sound stretch (you) and the two pauses attend to the time lapse that occurs in the transition from the previous task to the projected next technical step. When the VR instruction begins to unfold, Max turns his gaze to Sun, who does not attend to the visual display but instead starts to adjust the headset (l. 2). As the system instruction continues to unfold, Max continues his turn by explaining the purpose of the system instruction while Sun manipulates the headset (l. 3–4, Fig. 16). During Max's explanation Sun begins to show embodied orientation to the instruction by letting go of the headset, lowering her arms and not moving. Concurrently with Max's directive that reformulates the system instruction (l. 5), Sun begins to extend both her arms (Fig. 17) so that they reach the fully extended position at the end of the turn unit. When Max reformulates the second part of the instruction ("and then press the trigger", l. 6), Sun holds her arms fully extended but does not comply until the turn reaches completion (l. 7, Fig. 18). In this way she visibly relies on the more experienced participant's guidance, which, in turn, draws on monitoring the screen as well as the player's bodily conduct to make the system instruction understandable and followable. Concurrently with the system reaction marking the action successful (Fig. 19), Max's positive evaluation (l. 8) accepts her actions and closes the sequence.

Excerpt 3 highlights the way that instructions involve continual assessment of the environment and the novice user's conduct and parsing of bodily tasks into sub-actions that need to be accomplished in a timely way and in a specific order (Rauniomaa et al., 2018). The experienced player's turn in l. 1–2 orients to the VR system as the primary source of instruction and invites the player to pay attention to the directions provided visually. A delay in the player's response occasions an explanation that provides more general information of the technical step to be accomplished. The instruction to carry out the task required by the system (l. 5–6) is not initiated until the player visibly begins to orient to the visual display and it is temporally aligned with the bodily action of the player. The practice of parsing the system instruction into two actions supports the orderly accomplishment of the relevant actions.

4.2. Grasping and manipulating virtual objects

In using digital or immersive VR games the technology and material qualities of the controllers determine what kind of correspondence manual actions have with on-screen action (Hodges 2017). Interacting with objects in the virtual environment therefore requires manual and bodily skills that enable efficient use of the controller to interact with features of the game and to maintain its playability. This section examines how instructions by the more experienced player become relevant and are oriented to during moments of trouble that arise when the player attempts to interact with objects in the virtual environment.

Example 4 shows how the participants deal with an extended display of trouble in understanding what kind of action with the physical body is required to grasp objects in the virtual environment. The example comes from the first moments of using *National Geographic Explore VR*. When the game has loaded, the player enters *Base Camp* on *Mount Everest* and is instructed to start exploring the environment and objects that are visible around her (l. 1–4). Our focus is on the way that resources of the physical body and the controllers as focal objects are used to deal with the ensuing trouble.

Excerpt 4

```
01  Max   okay. *n I think now you can: (.) •stand ↑up
    max         *approaches Tes from the back
    tes                                     •...stands up

02  Tes    ok•[ay?
    tes   -->•

03  Max       [(k s so,) yea, yea. and look around what,
```

Figure 20

```
04        what •you #have (there.)
    tes        •bends knees, reaches both VH tw objects,
               tries to pick up a bowl on table, fails
    fig             #fig20
```

Figure 21

Figure 22

```
05  Tes   u::hm •(1.0) just (give me some #time to )•
    tes         •bends knees slightly, reaches both VH tw objects,
                    tries to pick up the bowl, fails•
    fig                                   #fig21

06         (.) •[try it.
    tes        •bends knees, reaches both VH tw objects, tries to
              pick up the bowl, fails-->

07  Max        [I think# (1.0)•
    tes                    -->•
    fig                #fig22
```

Figure 23

```
08   Max   I think uh you have to,*• (.) #did-
     max                          *bends left, gaze tw T's hands
     tes                           •turns tw Max-->
     fig                                 #fig23

09   Max   *okay.
     max   *straightens

10   Tes   uh? •[hh
     tes       •turns back to face the screen

11   Max        [try again. try again.

12         try to (.) [touch something.]
```

Figure 24

Figure 25

```
13   Tes              •[eh with: (.)    ]•#-this [•or#]
     tes              •turns tw Max -->
                      holds controllers tw M,
     tes                                 •touches RH top button



                                        w/index finger
     tes                                         •touches LH bottom
                                                 button w/thumb twice
     fig                                  #fig24     #fig25

14   Max                                          [with] those.

15         [I    ] think with those with the- (.) [yes.     ]

16   Tes   [hm.::]                                  [(x those)] hh hh

17         •(this.    )
     tes   •touches RH trigger w/index finger
```

Figure 26

```
18   Max   yea •(.) like #that
     tes       •lifts both hands up, palms facing her
     fig                 #fig26
```

Tessa follows Max's instructions by standing up and attending to the objects in the virtual environment. She tries to pick up a bowl that is placed on a table by using her physical body to reach for the object, bending her knees slightly and extending her arms (l. 4, Fig. 20). This is represented in the virtual space as two virtual hands (VH) wearing mittens reaching towards the object. However, her attempt to grab and pick up the object fails, which generates a vocal response from Tessa (l. 5). She then makes two more attempts to pick up the object while producing a verbal turn that indicates the trouble she is experiencing (l. 5–7, Fig. 21 and 22). When these attempts also fail, Max intervenes with a turn that is framed as an instruction but is not completed ("I think you have to (.) did-," l. 7). Instead of indicating what Tessa should do Max initiates a new action: already before the verbal initiation (the auxiliary 'did' which projects a question), he bends his torso sideways and directs his gaze to Tessa's hands, seemingly to assess how she is holding and using the controllers (Fig. 23). Tessa attends to Max's turn by turning towards him (l. 8–9) and initiating repair, but quickly turns back towards the screen when the turn is not continued, and Max closes the embodied checking activity with an 'okay' (l. 9) and encourages her to try again.

Tessa's trouble resurfaces when she turns to face Max again and holds both controllers up towards him, while performing an embodied understanding check (l. 13, Fig. 24 and 25). She formulates a declarative turn coupled with referential touch to check which button to press to be able to interact with the virtual objects. Although the verbal turn only refers to one button ("this button", l. 13), with touch she references two different buttons. On the demonstrative pronoun 'this' she uses her index finger to touch the top button of the controller in her right hand (the trigger) and immediately after this touches a button at the bottom of the other controller with the thumb of her left hand. The use of touch not only demonstrates that Tessa treats the problem as arising from her lack of understanding of correct controller techniques, but also shapes the way that the utterance is recognisable as an action. Max treats the action as a request for confirmation for an understanding that both buttons need to be used and promptly gives an affirmative response (l. 14). Still facing Max, Tessa then touches the button on the right-hand controller again (l. 17. The verbal utterance coupled with this action is difficult to hear but includes a partial repetition of Max's response. This occasions a further confirmation from Max (l. 18). Tessa attends to the virtual hands (Fig. 26), after which she turns towards the screen, holds the controllers up and presses both buttons as she attempts to pick up the closest object. However, moments later it turns out that the problem is a technical one and is not resolved until Max recalibrates the game. Nonetheless the example illustrates the persistent interactional efforts the participants make to deal with the lack of congruence between manual actions with the controller and the activity of picking up objects in the virtual environment. Failure to grasp, pick up and move virtual objects is treated as a display of lack of competence in the techniques of the controller. Max's actions show how he continually attends to and assesses the novice's attempts and provides verbal guidance to support them. The player's use of manual actions with the controller to check understanding show orientation to the displayed expertise.

The next example (Ex. 5) further illustrates how a player draws on the more experienced player's guidance in working out how to manipulate game-specific virtual objects. The game, *Fantastic Contraption* (created by Radial Games), presents the player with a first-person view of a space with colourful materials (e. g. wheels, beams, sticks) that they can use to build things. The player uses two motion controls to operate virtual tools that enable them to pick up, move and manipulate the objects and fit them together to construct a working machine. The game instructs the player with written directives that indicate what the player is expected to do and graphic features that visibly point to relevant objects. The excerpt below (5a) shows how Henna and Max negotiate a problem with controlling and manipulating the virtual tools. Following this, we focus on the interactional work required to interpret and follow a system instruction that directs the player to the next in-game task (Ex. 5b).

Excerpt (5a) comes from the beginning moments of the game. The visual display shows a written directive which instructs the player to *Grab this stick and move it*. In addition, a visually less prominent text informs the player that *The T-button grabs* (Fig. 27). Henna is using the controllers and attempts to grab a stick using the virtual tool operated by the controller in her right hand. Max initiates an instructional sequence with a turn linguistically designed as an inclusive imperative ("let's try," l. 1).

Excerpt 5a

```
01  Max   ~no kokeillaan. ~(0.5)
           let's try
    hen   ~tries to grab stick w/tool in RH
    hen                    ~moves tool up and down the stick -->

02  Max   eli kun sä otat siitä (.) valkosen ympyrän kohalta~
          (part) when you grab it where the white circle is
    hen                                                  -->~
```

Figure 27

```
03        ~kiini, (.) ni saatko sä #tartuttua siihe.
                      (part) can you grasp it
    hen   ~touches white circle, pulls-->
    fig                           #fig27
```

Figure 28

Figure 29

Figure 30

```
04        ~#(0.5) ~#(1.0) ~#(1.0)
    hen   ~attempts to grab stick
    hen           ~moves tool in LH closer to stick
    hen                   ~pulls RH tool to the right and
                          brings it back
    fig    #fig28  #fig29  #fig30

((lines 5-7 omitted, side sequence involving another participant))

08        ~(2.0)
    hen   ~pulls at stick with one tool -->

09  Max   käännäpä sitä~ ~ympäri sitä ohjainta.~
          turn it round the controller
    hen            -->~ ~holds RH tool still-->~
```

Figure 31

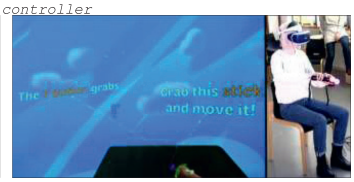

Figure 32

```
10        ~Ø#(0.6)
    hen   ~turns RH controller round
    VR     Ø t-button becomes visible -->
    fig     #fig31

11  Max   #joo. ~(.) eli se haluu et [sä käytät sitä~
          yea         so it wants you to use
    hen         ~turns RH controller back        -->~
    fig   #fig32

12  Hen                              [↑a::
```

Figure 33

```
13  Max   liipasi~Ønta.# [sori.] ~(se olikin se T-button)
          that trigger.   sorry. (it was the T-button after all)
    VR            Øbling
    hen          ~grabs stick
    hen                           ~extends it to the left-->
    fig                 #fig33
```

Max's turn (l. 1) frames the activity as a joint enterprise. Max continues to guide Henna by drawing attention to a white circle visible on the screen and asking if she can grasp it (l. 2–3). Concurrently with Max's turn, Henna has been trying to use the tool, moving it up and down the targeted stick. On hearing the reference to the white circle, she begins to comply with Max's interrogatively formatted instruction. She moves the tool close to the white circle, touches it and tries to grab and pull the stick using the controller in her right hand (l. 3–4, Figs. 27 and 28). She then uses both controllers and tries to grab it with both the tools (l. 4, Figs. 29 and 30), but fails. After a short side sequence, during which she continues these efforts, Max intervenes with an instruction targeting the use of the controller ("turn it round", l. 9) and Henna complies by turning the controller in her right hand around (l. 10). This has the effect of showing the "underside" of the virtual tool where the T-button is prominent in bright yellow colour (Fig. 31 and 32). Turning the controller also makes the text referring to the T-button visually more salient and shows a dotted line from the text to the focal button. Max displays understanding of the visual display by providing a verbal instruction to use the trigger button (l. 11, 13). Max's turn is linguistically tied to the visual display with the preface 'eli', which is recurrently used in turns that reformulate a prior speaker's turn (Sorjonen, 2018) and a reference to the system ("it wants you to use"). Concurrently with Max's turn, Henna turns the controller back to its normal position and produces a vocal response displaying new understanding. Having located the T-button, she is now able to use it to grab the stick and start manipulating it (Fig. 33). Max's brief apology in l. 13 refers to his own prior misunderstanding of which button to use, which may have contributed to the player's trouble.

Having resolved this problem, the participants are presented with a new task: connecting one stick to another. Henna's turn in l. 16 expresses uncertainty about how to continue when the system shows two written directives (*Connect me to the other stick* and *Connect to me*) while she is still manipulating the same stick.

Excerpt 5b

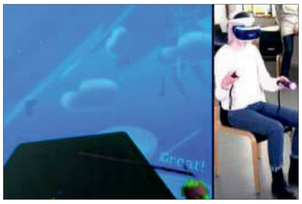

Figure 34

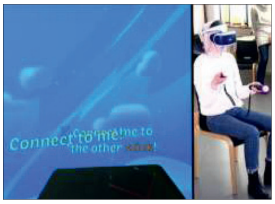

Figure 35

```
15   Hen                   [o↑kei]

16   Hen    Øo↑kei (.) ~#mitäØs,:::
                         what,
     hen                ~extends stick tw the right lower corner
     VR     Ø'Great'           Ø'Connect to me'
                               'Connect me to the other stick'
     fig                 #fig34

17   Max       ~~°aa° (.) #joo.
                         yea.
     hen    -->~~releases stick
            small head movements-->
     fig                 #fig35

18          (2.5)
     hen    -->~sits still-->
```

Figure 36

```
19   Hen    onks täs joku tehtävä~~[vai onks tää #vaa (ru-)
            is there a task here     or is this just
                              -->~~tilts head
                                  bends torso to the right
     fig                                        #fig36
```

```
20   Max                            [on. on. connect to me.
                                     there is. there is. connect to me.
```

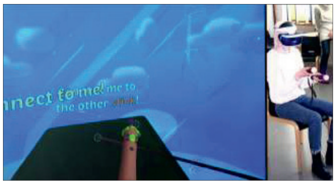

Figure 37

Figure 38

```
21          ↑connect me to the other ~#stick.
     hen                            ~moves tool to short stick
     fig                             #fig37

22   Max    ~eli nyt sun täytyy ~#ottaa se (.)
             so now you have to take it
     hen    ~touches stick      ~starts to extend it
     fig                         #fig38
```

Figure 39

```
23   Max    ~°liit#tää Øse.°                ~
              connect it.
     hen    ~connects stick to long stick->~
     VR               Øbling, 'Nice!'
     fig          #fig39
```

Henna's question on l. 16 indicates that it is not clear to her what she is expected to do next. The level prosody and stretching of the final sound on the question word seem to orient to the new system instructions that appear in her visual field (Fig. 35). Both Henna and Max react to the instructions. Henna releases the stick that she has been manipulating, while Max produces a vocal response ("aa") and followed by the response particle "joo" (l. 17).[1] Henna's lack of activity during the 2.5 s silence (l. 18) suggests that she is having trouble understanding how to proceed. This becomes evident when she formulates a question about a game assignment while examining the virtual environment by tilting her head and slightly bending her torso to the right (l. 19, Fig. 36). Max's answer builds on the two-part system instruction, which he reads aloud (l. 20–21) and then reformulates by translating it into Finnish (l. 22–23), thereby drawing attention to the written directives as relevant for the next task. Already on the final component of Max's answer turn, where he reproduces the system instruction (l. 21, Fig. 37), Henna starts to orient to the instruction by moving the right-hand tool to a short stick.

The latter part of Max's turn temporally aligns with and closely attends to Henna's in-game action. Concurrently with the reformulation of the instruction in Finnish (l. 22–23), Henna extends the stick and connects it with the long stick. Max's utterance is carefully coordinated with these actions so that the verbal components ("take it", l. 22, and "connect it", l. 23) are produced simultaneously with and indexically refer to these actions (Fig. 38 and 39). Rather than an instruction for a next action, it is hearable as running commentary whereby he observes and continually assesses Henna's in-game actions. With her actions on the virtual objects Henna displays close attention to the visible details of the system instruction in interpreting and following them. In this case the sequence closes with an evaluation generated by the game: an audible bling sound accompanied with a written assessment ('Nice', Fig. 39).

The examples in this section illustrate the interactional work involved in following instructions while attempting to interact with virtual objects. Interacting with objects in the game involves skilled perception and use of features of the controllers to manipulate virtual hands or tools. The participants' orientation to manual know-how is particularly clear in the tactile and manual practices deployed by the experienced player in instructing controller use. Instructive interventions through touch and manual guiding provide online guidance on how to locate relevant features and use them to interact with the virtual environment. They guide perception by enabling the player to use touch to perceive the controller features that are crucial for performing in-game actions. The examples also highlight how verbal instructions by the experienced player attend to displayed difficulties in seeing, understanding and following instructions delivered by the system. Verbal formulations targeting on-screen instructions draw attention to system instructions in an anticipatory way or they can orient to the player's observable difficulties as displayed in delay or lack of activity in the virtual space or bodily displays. The timing of the players' complying actions shows how they variously orient to the expert's guidance but may also display independent understanding of the system instruction by performing the complying action simultaneously with the more experienced player's turns.

[1] Due to a restricted camera view it is difficult to interpret to what extent Max's reactions respond to Henna's displayed trouble. The first particle 'aa' is produced quietly and may not be related to Henna's turn. However, the particle 'joo' more clearly acknowledges trouble and projects further action, although the source of the trouble may not be clear yet. It is worth noting that Henna's movement make the visual text difficult to read from the screen.

4.3. Bodily action and movement in virtual space

The following excerpt illustrates how instructions by the experienced participant are designed to help resolve the lack of coherence between the physical body and the actions that are relevant for moving and performing "physical activity" virtually. Virtual movement and bodily actions require correct techniques of the controller and use of the hands and arms to accomplish tasks that in the physical world would require other parts of the body.

Tessa has started climbing Mount Everest in *National Geographic Explore VR*. The game instructs the player through voiced instructions and feedback produced by virtual characters (VC), portrayed as guides. Controller features are used to interact with objects. The feature of teleporting allows the player to relocate themselves in the virtual space. The passage starts after a virtual character (an alpinist facing Tessa) praises Tessa for doing a good job with a prior task involving climbing up a ladder. After this the VC instructs her to continue, pointing upward to the left. In line 1, Tessa acknowledges this in-game directive by turning her head up and left, thereby tracing the character's pointing gesture, which brings into view a new ladder leading further up the mountain. Next, she tries to touch the ladder with both virtual hands (Figs. 40 and 41), holding them out and moving the fingers to grab (repeating the same pulling movements that were required in the previous task), but fails to get a hold of the side rails. This prompts a series of instructions from Max. The following examples thus illustrate an instructional intervention that is occasioned by recruitment, i.e. methods of assistance that to respond to manifest trouble (Drew and Kendrick, 2018; Kendrick, 2021). Excerpts 6a and 6b show how Max uses parsing to break up the task into bodily actions that are adjusted to the requirements of the game. In Ex. 6a he first addresses the location of the virtual body relative to the ladder (l. 2, 4–5), after which he issues corrective instructions to redirect Tessa's bodily actions and prepares to teach the correct controller techniques for achieving movement in the virtual environment (l. 7–9).

Excerpt 6a

```
01          •(4.0)
     tes    •looks up and L, torque L tw ladder, reaches both VHs tw
            bottom of the ladder, #tries to take hold w both VHs,
            shifts head and torso to the L
     max    -- squatting position, looks at screen-->
     fig.                          #fig40
```


Figure 40


Figure 41

```
02   MAX    I think •y- you have to# •uhm: *(tch)
     tes            •reaches left VH tw object next to ladder,
            tries to grab
     tes                             •retracts left VH
     max                                   *closes eyes, opens
     fig                           #fig41

03   TES    •hmhm?
     tes    •stops, holds position-->

04   MAX    move •closer to the l-ladder,
     tes         •slowly turns R tw ladder, leans forward slightly-->

05          so look at the ladder, (.)
```


Figure 42

```
06   TES    (I'm #locking?•*)
     tes              -->•
     max               -->*
     fig         #fig42

07   MAX    *•uhm but *no you can't
     max    *gets up  *moves next to Tes, extends RH
     tes     •stands still-->

08          you can't *move yourself in a way.*
     max              *--- RH gesture tw T--->*
     tes          -->
```


Figure 43

```
09          *you have to move• •#virtually. (.)* ♦*so.
     max    *extends RH tw Tes, palm down   -->* ♦*...
     tes                  -->• •extends both VHs tw ladder
     fig                         #fig43
```

In the beginning of Ex. 6a Tessa notably scans the area for potential solutions to the problem, indicating difficulties with moving forward (note the shift of gaze and torso, line 1). While Tessa unsuccessfully tries to grab an object to the left of the ladder, Max, who is squatting on the left behind her and looking at the screen, begins to offer verbal guidance. He projects an instruction ("you have to") of what action is needed to proceed in the game (l. 2). This is followed by an instruction directing Tessa to move closer to the ladder and a specification requesting her to look at the ladder (l. 4–5). By holding her position and then, in fine coordination with Max's instruction, turning right and leaning forward slightly (l. 4–6, Fig. 42), Tessa not only demonstrates close listening but also a practical understanding of 'moving closer'. With this movement she draws on the physical body to move the virtual body closer to the ladder. However, Max's corrective instructions treat Tessa's embodied response as inappropriate for the technical task and his own instructions in need of further explanation (l. 7–9).

The categorical design of Max's turn in lines 7 to 9 is noteworthy as he not only refutes Tessa's attempt to physically move towards the ladder, marking it as impossible ("you can't", lines 7 and 8), but also refers to the necessity of using resources afforded by the system to proceed ("you have to move virtually", line 9). He thus orients to the affordances of the game, acknowledging and verbalizing a mismatch between physical and virtual embodiment ("move yourself" vs. "move virtually"). At the same time, he stands up, moves forward and reaches out towards Tessa's left hand (Fig. 43), projecting tactile intervention on the left controller. Tessa, on the other hand, holds still until – in concord with Max's "virtually" – she resumes movement by extending both arms in the direction of the ladder, thereby indicating close monitoring of and alignment with Max's verbal activities. She visibly anticipates involvement of the arms and hands and thus displays an initial understanding of what the directive of 'moving virtually' possibly means in the physical environment and in terms of using the physical body.

Max's "so" with falling intonation (line 9) further announces a shift in his orientation, which materializes next in the instructive form of parsing the task of traversing virtually into several consecutive moves (Ex. 6b, lines 10–13). More precisely, these steps lead Tessa to establish eye-hand coordination through gazing at a certain spot at the bottom of the ladder, modifying the direction of the controller, and pressing a button to interact with the ladder subsequently.

Excerpt 6b

```
10  MAX   •look at the #ladder •look at the uh:* (.) #floor,
    max   -->glances at VR glasses, points tw screen-->
    max                                           *downward head&
                                                  hand gesture
    tes   •gaze tw ladder, pedal movement w/both hands
                             •both extended VHs slightly tw L
    fig               #fig44                          #fig45
```


Figure 44


Figure 45

```
11        •#(1.0)
    tes   •head down, looks down
    fig    #fig46
```


Figure 46

```
12  MAX   ♦*and then, *•we #press this: Ø#middle button,
    max   ♦gaze: screen, glances down, then screen
    max    *places RH on top of T's LH;
    max               *lowers T's LH, touches button
    tes                •lowers RH
    VR                                 Øgreen circle at left handle
    fig                     #fig47       #fig48
```

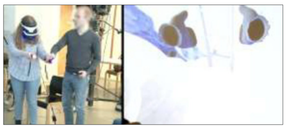
Figure 47

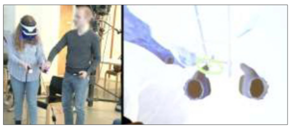
Figure 48

```
13        I *Øthink there,# to [move you clo*se,]
    max     *presses button w/index fg      *releases T's hand
    VR       Øscreen shows ladder very close, VHs behind ladder
    fig                    #fig49

14  TES                          [uh •huh huh     ] huh [huh huh    ]
    tes                              •body movement

15  MAX                                                 [heh heh the]

16        £I hh sorry.£
```

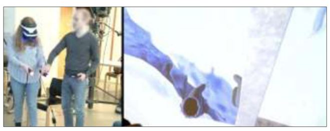
Figure 49

As before, Tessa's instructed actions (in the form of successive embodied responses) are finely tuned with Max's talk. In line with his directives in line 10, she looks first at the ladder in front of her (Fig. 44 and 45) and then adjusts her gaze downwards, which brings into view the snowy ground and the bottom of the ladder (Fig. 46). However, simultaneously to Max's "look at the ladder" (l. 10), Tessa again displays an embodied (yet unsuccessful) candidate understanding of how to proceed by exerting pedalling movements with both her arms, thereby demonstrating some agency in figuring out in-game functionalities herself and – again (see line 9) – visibly anticipating a solution that requires manual action in the physical world to relocate the virtual body.

When Tessa gazes at the bottom of the ladder, Max continues by introducing the next step, which includes lowering the left controller to point at the bottom of the ladder as well as locating and pressing a button in the middle of the hand-held device (l. 12–13). In doing so, Max draws on manual guiding and provides Tessa a hands on experience of the required actions. He places his right hand, palm down, on Tessa's left hand ("and then", l. 12) and lightly pushes her left arm down ("we press this:", Fig. 47). Tessa immediately aligns with this by lowering the other arm. These actions activate a green target (Fig. 48), located at the bottom of the left rail – a way of the system to mark object interactability. Now Max presses the button on the controller with his index finger (line 13). Interestingly, he uses the first-person plural pronoun "we", when projecting this action in line 12. Indeed, as they are both holding the controller now, manipulating the device momentarily becomes an intercorporeal activity to jointly mobilize the virtual body. When pushing the button, Tessa's virtual body instantly moves close to the ladder such that her hands in VR appear behind the ladder (Fig. 49). Tessa reacts with a sudden upward turn of her physical body accompanied with laugher indicating surprise and possibly an experience of disalignment in terms of speed and dual embodiment (l. 14–16). The sequence closes with Tessa moving up the ladder in VR and commentary by Max who at the same time withdraws and thus treats the trouble as resolved.

5. Discussion

As part of his explorations of instructions and instructed action Garfinkel (2002) used an experiment where his students wore a mask with inverted lenses while trying to do ordinary everyday tasks. By disrupting the coherence of "phenomenal details" of the surround and the embodied experience, this experiment revealed the achieved nature of our phenomenal field and mundane actions (Garfinkel 2002: 209–210). Garfinkel shows how communicative actions such as instructions become intelligible and followable in the embodied work of practical activities and how this relies on the achieved coherence of our perceived environment. The accomplishment of tasks involves interactional work through which the participants (re) construct the perceptual structuring of the environment (e.g. Goodwin 1994; Luff et al., 2003; Nishizaka 2006, 2014; Lindwall and Lymer 2024). This paper has analysed how the practical work of instructed action gets accomplished in the technologized setting of using Virtual Reality. We hope to have shown how instructional practices manifest manual and bodily know-how that is specific to the tasks of setting up a VR game and making it playable as well as interacting with objects and the virtual environment.

The analysis elucidates the sequential development and assemblages of resources that are deployed in calibrating the game and the body during pre-game tasks (4.1) and the methods through which participants resolve troubles related to in-game activities, specifically interacting with virtual objects (4.2) and performing virtual bodily actions such as climbing (4.3). These activities make relevant techniques of the body (Mauss 1973) and expertise (i.e. knowledge of how things are done, Arminen and Simonen 2021) that are not taught by the game but are transmitted in interaction. Knowing how to operate the game controllers and how to follow instructions delivered by the system is manifested in the social and interactional organisation of the activity as an instructional environment. We have shown how the more experienced participant uses configurations of talk, touch and manual actions to instruct and assist novice players in adjusting to the perceptual environment and how he guides their use of the game controllers, thereby enabling them to locate features and recognise their relevance for play. The players unproblematically orient to Max's expertise by both requesting assistance and complying with his instructions. Details of interaction show, however, how expertise can vary and fluctuate in interaction. Our data show, for example, that Max may initiate a side sequence with a spectating participant to solicit confirmation for his understanding of a relevant detail (Ex. 2). On other occasions, uses inclusive verbal formats (e.g. let's try, Ex. 5a) that frame the activity as a joint problem-solving activity. In addition, mutual orientation to a player's emerging expertise in VR use is visible in the temporal organisation of actions, as is the case when an instruction for next action and the projected practical action are performed (almost) concurrently.

In the VR settings in focus tactile experience of the interface is crucial for participation. Controllers serve as technological prostheses (Cleland 2010, citing McLuhan, 1967) that enable extension of abilities into the virtual space and actions of the body in the virtual environment. Developing a tactile perception (Arnold 2012) of the controllers is a precondition for interacting with the system and performing in-game actions. The bodily techniques through which games are made playable and operated are contingent on the material and technological affordances of the system and may be game specific. The range of variation between such techniques is visible in the array of manuals, websites, videos and other resources that provide information and instructions on how to operate the controllers or tools for playing a specific game. Our analysis elucidates how the participants orient to manual and bodily expertise and practice the skills involved in engaging with VR in different phases of the activity. In the circumstances where the head-mounted device restricts players' access to their own physical body, touching the object and directing the player's hand are integral components of instructions. Coupled with indexical expressions and verbal formulations of projected actions, they make the instructions transparent and demonstrate techniques for efficient use of the controllers. However, manual actions targeting controller use do not just demonstrate how to perform actions but also direct their accomplishment. In this way the completion of an instruction and the instructed action merge into one intercorporeal achievement (Evans 2017). While the pre-game phase is characterised by instructive actions that involve touching and manual repositioning of the hand to demonstrate controller techniques and move the activity forward, controller use can become a focal activity also during moments of trouble in in-game activities. The participants orient to problems with interacting virtual objects as troubles associated with controller techniques. Touch and manual interventions provide effective resources for initiating remedial action and resolving troubles (Ex. 4 and 6).

Our analysis also sheds light on the complex sequential and simultaneous organisation of instructed action as the participants make efforts to establish and maintain connection with the system and achieve coherence across the physical and virtual spaces. We have shown how the participants orient to the VR system and the real-time directions it provides as they work out how to proceed and what actions are required to move forward and interact with the game. System instructions are delivered through different modalities (written text, graphics, sound) and can be idiosyncratic to specific games. The inherent inadequacy of such instructions is visible in the troubles that novice players display in noticing and interpreting them and the interactional work it takes to make them followable. Players orient to the experienced participant's verbal guidance in learning how to see and understand the local relevance of system instructions (Ex. 3, 5, and 6). The analysis elucidates the work that is required to form a situated understanding of how to proceed when guided by a technology of instruction (Lynch and Lindwall 2024; Brown and Laurier 2012). This also involves instructed perception (Goodwin 1994; Nishizaka 2021): the more experienced participant guides perception by projecting system instructions before they appear, by reading them aloud and reformulating them to make them salient and highlight their relevance in the moment. As the players have independent access to the system, they may respond early and demonstrate new understanding by performing the task simultaneously or even before verbal instructions are completed. The analysis highlights how the sequential organization of instructed action is tied to the accomplishment of tasks and the different temporalities of talk, embodied and instrumental action as well as the affordances of the technology.

Our data has revealed how manual know-how is actualized in instructed action and the specific ways that novice participants are guided in developing the skills required to use a tactile interface and their body to make things happen in a virtual environment. Our findings align with recent calls for research insights into the relevance of different qualities of knowing for accomplishing action (Arminen and Simonen 2021) and forms of knowledge that resides in the body (Ehmer and Brône 2021). We have shown how manual know-how involves coordination of different resources including skilled perception, tactile expertise and bodily practices that depart from the natural use of the body in non-technological settings. The examples illustrate how such know-how is transmitted in instructed action through which the participants learn to use the controllers and their extended body to touch, grasp, and manipulate objects, relocate their virtual body and climb virtually. The analysis highlights how instructive and instructed actions are assembled using resources of the physical and virtual environment in locally relevant ways. The practices used deal with the asymmetrical access to resources and lack of embodied fluency (Lynch and Lindwall 2024) resulting from the VR set up. Our findings are limited to interactions with the standard VR technology used in this study. As new, more advanced VR technologies become more widely available, there is a need for more research to investigate how the practices and techniques of the body are manifested and learnt in such environments.

Declaration of competing interest

None.

CRediT authorship contribution statement

Arja Piirainen-Marsh: Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Margarethe Olbertz-Siitonen:** Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Appendix A

Transcription conventions

Talk is transcribed with the conventions developed by Gail Jefferson. Embodied actions are transcribed based on the following conventions developed by Lorenza Mondada (2018).

Gestures and descriptions of embodied actions are delimited between two identical symbols and are synchronized with corresponding stretches of talk/lapses of time:

* * • • + + ~ ~
* embodied actions by Max.
◆ Max's gaze.
• embodied actions by Tessa (Tes).
+ embodied actions by Sunny (Sun).

∼ embodied actions by Henna (Hen).

Ø written directions, feedback sounds by the VR system.

Additional notations:

- > The action described continues across subsequent lines until the same symbol is -> reached.

Participants doing the embodied action when they are not the speaker are identified by name written in lower case (tes, sun, hen).

Manual actions and movements by human participants are identified with the symbols designated to the participants and reference to the physical body. For example: RH (right hand), LH (left hand), BH (both hands).

On-screen actions that show parts of the virtual body, especially hands, are indicated with VH (virtual hand), left VH (left virtual hand), right VH (right virtual hand), and both VHs (both virtual hands).

A hash # indicates the moment at which a screen image has been taken to show how it is positioned within talk.

References

Arminen, I., Simonen, M., 2021. Expertise as a domain in interaction. Discourse Stud. 23 (5), 577–596. https://doi.org/10.1177/1461445621101679.

Arnold, L., 2012. Dialogic embodied action: using gesture to organize sequence and participation in instructional interaction. Res. Lang. Soc. Interact. 45 (3), 269–296. https://doi.org/10.1080/08351813.2012.699256.

Brown, B., Laurier, E., 2012. The normal-natural troubles of driving with GPS. In: CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM Association for Computing Machinery, Austin, pp. 1621–1630. https://doi.org/10.1145/2207676.2208285.

Cleland, K., 2010. Prosthetic bodies and virtual cyborgs. Second Nature: International Journal of Creative Media 2, 74–101.

De Stefani, E., 2018. Formulating direction: navigational instructions in driving lessons. Int. J. Appl. Ling. 28 (2), 283–303. https://doi.org/10.1111/ijal.12197.

Deppermann, A., 2018a. Changes in turn-design over interactional histories: the case of instructions in driving school lessons. In: Deppermann, A., Streeck, J. (Eds.), Time in Embodied Interaction: Synchronicity and Sequentiality of Multimodal Resources. John Benjamins, pp. 293–324. https://doi.org/10.1075/pbns.293.09dep.

Deppermann, A., 2018b. Instruction practices in German driving lessons: differential use of declaratives and imperatives. Int. J. Appl. Ling. 28 (2), 265–282. https://doi.org/10.1111/ijal.12198.

Drew, P., Kendrick, K., 2018. Searching for trouble: recruiting assistance through embodied action. Social Interaction. Video-based studies of Human Sociality 1 (1). https://doi.org/10.7146/si.v1i1.105496.

Ehmer, O., Brône, G., 2021. Instructing body knowledge. Multimodal approaches to interactive practices for knowledge constitutions. Linguistics Vanguard 7 (s4). https://doi.org/10.1515/lingvan-2021-0012.

Evans, B., 2017. Intercorporeal (re)enaction. Instructional correction in basketball practice. In: Meyer, C., Wedelstedt, U. (Eds.), Moving Bodies in Interaction – Interacting Bodies in Motion: Intercorporeality, Interkinaesthesia and Enaction in Sports. John Benjamins, pp. 267–399. https://doi.org/10.1075/ais.8.11eva.

Evans, B., Lindwall, O., 2020. Show them or involve them? Two organizations of embodied instruction. Res. Lang. Soc. Interact. 53 (2), 223–246. https://doi.org/10.1080/08351813.2020.1741290.

Garfinkel, H., 1996. Ethnomethodology's program. Soc. Psychol. Q. 59 (1), 5–21. https://doi.org/10.2307/2787116.

Garfinkel, H., 2002. In: Rawls, E.W. (Ed.), Ethnomethodology's Program: Working Out Durkheim's Aphorism. Rowman & Littlefield.

Goodwin, C., 1994. Professional vision. Am. Anthropol. 96 (3), 606–633. https://doi.org/10.1525/aa.1994.96.3.02a00100.

Goodwin, C., 2000. Action and embodiment within situated human interaction. J. Pragmat. 32, 1489–1522. https://doi.org/10.1016/S0378-2166(99)00096-X.

Goodwin, C., 2018. Co-operative Action. Cambridge University Press. https://doi.org/10.1017/9781139016735.

Harjunpää, K., Deppermann, A., Sorjonen, M-L., 2021. Constructing the Chekhovian inner body in instructions: An interactional history of factuality and agentivity. J. Pragmat. 171, 158–174. https://doi.org/10.1016/j.pragma.2020.09.034.

Heath, C., Luff, P., 2000. Technology in Action. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/CBO9780511489839.

Heath, C., Luff, P., 2018. The naturalistic experiment: video and organizational interaction. Organ. Res. Methods 21 (2), 466–488. https://doi.org/10.1177/1094428117747688.

Hindmarsh, J., Hyland, L., Banerjee, A., 2014. Work to make simulation work: 'realism', instructional correction and the body in training. Discourse Stud. 16 (2), 247–269. https://doi.org/10.1177/1461445613514670.

Hodges, J.A., 2017. How do I hold this thing? Controlling reconstructed Q*berts. New Media Soc. 19 (10), 1581–1598. https://doi.org/10.1177/1461444817717511.

Kääntä, L., Piirainen-Marsh, A., 2013. Manual guiding in peer group interaction: resource for organizing a practical classroom task. Res. Lang. Soc. Interact. 46 (4), 1–22. https://doi.org/10.1080/08351813.2013.839094.

Kendrick, K., 2017. Using conversation analysis in the lab. Res. Lang. Soc. Interact. 50 (1), 1–11. https://doi.org/10.1080/08351813.2017.1267911.

Kendrick, K., 2021. The 'Other' side of recruitment: methods of assistance in social interaction. J. Pragmat. 178, 68–82.

Kohonen-Aho, L., Haddington, P., 2023. From distributed ecologies to distributed bodies in interaction: capturing and analysing 'dual embodiment' in virtual environments. In: Haddington, P., Eilittä, T., Kamunen, A., Kohonen-Aho, L., Oittinen, T., Rautiainen, I., Vatanen, A. (Eds.), Ethnomethodological Conversation Analysis in Motion: Emerging Methods and Technologies. Routledge, pp. 111–131.

Lindwall, O., Ekström, A., 2012. Instruction-in-interaction: the teaching and learning of a manual skill. Hum. Stud. 35 (1), 27–49. https://doi.org/10.1007/s10746-012-9213-5.

Lindwall, O., Lymer, G., 2014. Inquiries of the body: novice questions and the instructable observability of endodontic scenes. Discourse Stud. 16 (2), 271–294. https://doi.org/10.1177/1461445613514672.

Lindwall, O., Lymer, G., 2024. Detail, granularity and laic analysis in instructional demonstrations. In: Lynch, M., Lindwall, O. (Eds.), Instructed and Instructive Actions: the Situated Production, Reproduction and Subversion of Social Order. Routledge, pp. 37–54.

Lindwall, O., Lymer, G., Greiffenhagen, C., 2015. Sequential analysis of instruction. In: Markee, N. (Ed.), The Handbook of Classroom Discourse and Interaction. John Wiley & Sons, pp. 142–157.

Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., Oyama, S., 2003. Fractured ecologies: creating environments for collaboration. Hum. Comput. Interact. 18 (1–2), 51–84. https://doi.org/10.1207/S15327051HCI1812_3.

Lynch, M., Lindwall, O., 2024. Introduction. Instructed and instructive action. In: Lynch, M., Lindwall, O. (Eds.), Instructed and Instructive Actions: the Situated Production, Reproduction and Subversion of Social Order. Routledge, pp. 2–18.

Mauss, M., 1973. Techniques of the body. Econ. Soc. 2 (1), 70–88.

Mondada, L., 2014a. Cooking instructions and the shaping of things in the kitchen. In: Nevile, M., Haddington, P., Heinemann, T., Rauniomaa, M. (Eds.), Interacting with Objects. Benjamins, pp. 199–226.

Mondada, L., 2014b. Instructions in the operating room: how the surgeon directs their assistant's hands. Discourse Stud. 16 (2), 131–161. https://doi.org/10.1177/1461445613515325.

Mondada, L., 2018. Multiple temporalities of language and body in interaction: challenges for transcribing multimodality. Res. Lang. Soc. Interact. 51 (1), 85–106. https://doi.org/10.1080/08351813.2018.1413878.

Mondada, L., 2019. Contemporary issues in conversation analysis: embodiment and materiality, multimodality and multisensoriality in social interaction. J. Pragmat. 145, 47–62. https://doi.org/10.1016/j.pragma.2019.01.016.
Nishizaka, A., 2006. What to learn: the embodied structure of the environment. Res. Lang. Soc. Interact. 39 (2), 119–154. https://doi.org/10.1207/s15327973rlsi3902_1.
Nishizaka, A., 2014. Instructed perception in prenatal ultrasound examinations. Discourse Stud. 16 (2), 217–246. https://doi.org/10.1177/1461445613515354.
Nishizaka, A., 2021. Seeing and knowing in interaction: two distinct resources for action construction. Discourse Stud. 23 (6), 759–777. https://doi.org/10.1177/14614456211017712.
Olbertz-Siitonen, M., Piirainen-Marsh, A., 2023. Bridging physical and virtual ecologies of action: giving and following instructions in co-located VR-gaming sessions. Prologi – J. Commun. Soc. Interact. 20 (1), 137–166. https://doi.org/10.33352/prlg.121525.
Parisi, D., 2011. Game interfaces as bodily techniques. In: Ferdig, R.E. (Ed.), Handbook of Research on Effective Electronic Gaming in Education. Information Science Reference, Hershey, PA, pp. 111–126.
Råman, J., Haddington, P., 2018. Demonstrations in sports training: communicating a technique through parsing and the return-practice in the budo class. Multimodal Communication 7 (2), 1–19. https://doi.org/10.1515/mc-2018-0001.
Rauniomaa, M., Haddington, P., Melander, H., Gazin, A.-D., Broth, M., Cromdal, J., McIlvenny, P., 2018. Parsing tasks for the mobile novice in real-time: orientation to the learner's actions and to spatial and temporal constraints in instructing-on-the-move. J. Pragmat. 128, 30–52. https://doi.org/10.1016/j.pragma.2018.01.005.
Slater, M., Sanchez-Vives, M.V., 2016. Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI 3 (74). https://doi.org/10.3389/frobt.2016.00074.
Sorjonen, M.-L., 2018. Reformulating prior speaker's turn in Finnish. In: Heritage, J., Sorjonen, M.-L. (Eds.), Between Turn and Sequence. Turn-Initial Particles across Languages. John Benjamins, pp. 251–286.