Learning, Culture and Social Interaction 56 (2026) 100956

Contents lists available at ScienceDirect

Learning, Culture and Social Interaction

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

Full length article

The accountability of movement in synchronous hybrid classroom interaction: managing mobility troubles in robot-mediated remote participation[☆]

Teppo Jakonen [a,*], Fredrik Rusk [b]

[a] School of Languages and Translation Studies, University of Turku, Finland
[b] Department of Education, Åbo Akademi University, Finland

A R T I C L E I N F O

Keywords:
Video-mediated interaction
Telepresence robot
Synchronous hybrid teaching
Interactional asymmetries
Embodied learning

A B S T R A C T

Telepresence robots enhance videoconferencing by enabling the user to 'move around' remotely, offering more embodied resources for interaction than speaker gallery-based videoconferencing tools. This study explores interactional troubles related to the affordance of mobility during robot-mediated remote participation in synchronous hybrid classroom interaction. Using multimodal conversation analysis, we analyse video-recorded higher education language teaching classes from Finland in which one student used a Double 3 telepresence robot for remote participation, focusing on how the participants managed activity transitions requiring mediated mobility during one L2 Swedish class. The analysis traces a micro-longitudinal interactional trajectory whereby initial troubles in operating the robot led participants to stop using the remote-controlled mobility feature and resort to carrying the robot from one place to another. We discuss how such a practice of human-assisted mobility emerges interactionally through geographically distributed coordination of the classroom space and what kinds of orientations to the accountability of remote student's action, autonomy and engagement are at stake. The findings illustrate challenges that emerging communication tools such as the telepresence robot can introduce for remote embodied participation and learning, and how teachers and students sometimes come up with unexpected ways of building action with the help of, or despite, technological affordances.

1. Introduction

Video-mediated interaction (VMI) has been recently identified as a potential tool for making education more accessible and inclusive by offering opportunities to expand learning and instruction beyond the physical classroom (European Commission, 2023). At the same time, interactional research on video-mediated teaching and learning activities has increased dramatically across a broad range of learning settings, educational levels and countries (e.g., Çimenli et al., 2022; Jakonen & Jauni, 2024; Oittinen, 2022; Pekarek Doehler & Balaman, 2021; Rusk & Pörn, 2019; Uskokovic & Taleghani-Nikazm, 2022). As VMI solutions and platforms are becoming increasingly capable of mediating a broader range of multimodal human action, they are also expanding the scope of potential multimodal troubles that participants may find a need to coordinate interactionally. The purpose of this article is to explore

☆ This article is part of a Special issue entitled: 'Engagement-in-interaction' published in Learning, Culture and Social Interaction.

* Corresponding author at: Arcanum building, Arcanuminkuja 1, FI-20014, Turku, Finland.
E-mail addresses: teppo.jakonen@utu.fi (T. Jakonen), fredrik.rusk@abo.fi (F. Rusk).

https://doi.org/10.1016/j.lcsi.2025.100956
Received 12 January 2024; Received in revised form 21 August 2025; Accepted 13 October 2025
Available online 1 December 2025

participants' practices for addressing emergent mobility troubles in synchronous hybrid classrooms, which, at least in principle, aim to provide an equivalent instructional experience to on-site students in the classroom and video-mediated remote students (e.g., Beatty, 2019; Raes et al., 2020). Through empirical multimodal analyses of a hybrid setting in which the remote student has a technology-mediated physical presence and the ability to move in the classroom, we aim to better understand the situated and technology-mediated nature of action in novel learning environments that cater for remote students through technology-mediated means.

The study analyses a hybrid classroom where remote student participation takes place by way of a telepresence robot. Telepresence robots are mobility-enhanced and remote-controlled individualised VMI tools equipped with a camera, screen, speakers, microphones and wheels. In educational settings, they have been used to enable homebound or hospitalised students to participate in classroom-based teaching (Page et al., 2021) but also in a more casual and exploratory manner in higher education (Leoste et al., 2022). When the robot is taken to a classroom, a remote student can 'drive around' in the classroom space and control what they see via the robot camera and with whom they interact. For the classroom participants, the robot constitutes a material representation of the remote student, which is quite different from, for instance, being projected on a classroom screen using a gallery-based VMI tool such as Zoom. At the same time as the robot's mobility feature aims to reduce inherent interactional asymmetries of VMI by expanding the remote participant's possibilities for embodied action in a material environment, in our data it also introduces troubles which geographically distributed participants routinely address by providing descriptions of how they see and experience each other, their actions and the classroom material space. In this article, we investigate such practices of spatial coordination, in particular the classroom participants' orientations to the remote student's mediated mobility – its recognisability, rationality, competence and accountability – as interpretative work to maintain a shared understanding of the hybrid class "for all practical purposes" (Garfinkel, 1967).

The troubles of spatial coordination we analyse in this article emerge during classroom transitions. Classroom transitions can be interactionally challenging moments for teachers because they require making projections for the next activity, coordinating the students' shifting attention and transforming participation frameworks such as when changing groups or moving physically from one place to another in the material space (e.g., Jacknick, 2011; Jakonen & Jauni, 2024; Råman, 2017). Transitions in hybrid classrooms remain largely uncharted research territory (but see Jakonen & Jauni, 2024), but their complexity is potentially heightened by the interactional asymmetries between classroom-based and video-mediated remote students. Against this backdrop, telepresence robot-mediated remote participation constitutes a perspicuous setting to explore how participants deal with such interactional challenges and make sense of and use the affordances of novel technologies for educational purposes. Our investigation is guided by the following research questions:

RQ1: How do participants solve emergent mobility troubles during activity transitions in synchronous hybrid teaching in which telepresence robots are used for remote participation?

RQ2: What do participants' ways of coordinating the material space during mobility troubles suggest about the accountability of embodied action in hybrid classroom interaction?

Before the analysis, we frame our study by briefly outlining key multimodal asymmetries in (educational) video-mediated interaction. We will then describe our dataset and the analytical approach before presenting a micro-longitudinal case study that illustrates the maintenance of intersubjectivity and the facilitation of remote participation during technological troubles in activity transitions of a single synchronous hybrid Swedish language lesson.

2. Technology-induced asymmetries in video-mediated interaction (VMI)

Technological issues (big and small) are an inherent part of video-mediated interaction (VMI), and the mediated nature of VMI is evident in several technology-related constraints and interactional distortions. VMI takes place over an internet connection, which means that one basic source for interactional asymmetries is latency-related delay. Delay can create situations in which the fine-grained social organisation and its timings that participants manage with precision in face-to-face interaction require participants' attention (Olbertz-Siitonen, 2015; Rusk & Pörn, 2019; Seuren et al., 2021). Latency issues may also distort and impact interaction from the standpoint of affect by creating frustration that may spill over to how participants view each other's personalities (Olbertz-Siitonen, 2015).

Beyond the transmission of talk, VMI is constrained by technical limitations and design choices of the platforms. One clear visual limitation is that participants' video cameras cannot produce an entire image of the physical space in which participants interact. This means that geographically distributed participants need to deal with the "deictic complexities of nonshared spaces" (Hjulstad, 2016, p. 1). This kind of 'fracturedness' (Luff et al., 2003) of the ecology of action becomes consequential, for example, for pointing gestures. They can be complicated to accomplish in VMI because the recipient of the gesture does not necessarily see the pointed-at object and the performer of the gesture does not see how the gesture looks like to the recipient (Hjulstad, 2016; Luff et al., 2016). Users' VMI platforms are also individualised, as cameras may be mirrored and the layout of buttons, windows and arrangement of other participants on participants' screens may differ (e.g., Valasmo et al., 2023). These kinds of design choices in platforms lead to individualisation of the user experience in VMI (e.g., Gillespie, 2010). The material ecology of VMI is thus at the same time shared (digitally, but in a fractured sense), non-shared (participants are not physically co-located) and different to users by design (platformization). This requires that participants make the best out of what relevant contextual, situational and interactional information is available to them as they jointly negotiate and uphold intersubjectivity (e.g., Hjulstad, 2016).

Nevertheless, it can also be argued that there is no such thing as non-mediated perception and interaction. That is, even in the 'primordial' (Schegloff, 1996) setting of co-present social interaction, a perception of the world, talk, embodied conduct and interaction are mediated through participants' senses and embodied practices of sensing (Mondada, 2019). Possibilities to use senses are not always symmetrically available to co-present participants, whose sensory worlds may be quite different, as is the case for example in the interaction between a person with visual impairment and a sighted individual (e.g., Due, 2023). For the analyst, this introduces a challenge in avoiding unwarranted ableist assumptions about participants' sensory capacities and the ways in which the world they inhabit is available and 'experience-able' to them. From an EMCA (Ethnomethodology and Conversation Analysis) perspective on VMI, there is a need to empirically and analytically show when, and how, members themselves orient to technologies used to mediate interaction and the possible technology-induced asymmetries between geographically dispersed participants (e.g., Arminen et al., 2016; Giles et al., 2015, 2017).

While there are perhaps always some asymmetries in interaction (regarding senses, knowledge, power, etc.), in the data analysed in this article, the most central embodied asymmetries relate to vision, audio, mobility and tactility *with respect to the action that unfolds in the classroom*. The perspective of a remote student using a telepresence robot on the 'local' classroom space reminds how players in first-person perspective games (Rusk et al., 2023) or immersive virtual reality (Kohonen-Aho & Haddington, 2023) have an individualised but mediated perspective to unfolding action. Such technology-mediated settings highlight the material, spatial and embodied 'layers' of intersubjectivity (shared understanding) in geographically distributed collaboration. Making sense of embodied actions and acting together in distributed environments involves making one's own (emic) perspective, perceptions and experiences available to others, particularly when technological issues become interactionally consequential. This kind of intersubjective work may also involve parsing participants' individual perspectives together to co-construct what the group perceive as a multi-body entity, which some previous studies have referred to as interperceptivity (Rusk et al., 2023; von Wedelstaedt, 2020).

As the remotely controlled telepresence robot enables movement, it also expands the domain of accountability for the remote student. Unlike when a remote student is visible in the class on a fixed screen in the classroom (such as on a video gallery projected on the wall), the possibility of mobility means that the remote student can be expected to *use movement to produce action in a competent, rational and recognisable manner*. This also means that classroom participants need to orient to the robot's movements to make sense of what the remote student is up to as well as manage troubles stemming from technology glitches or platform design issues that may at times affect the remote student's control over robot movements. Therefore, to better understand possible educational affordances and challenges of using telepresence robots as a tool for synchronous hybrid teaching, this study analyses how participants in a hybrid context collaboratively manage mobility troubles during activity transitions and considers what their management might tell us more generally about the accountability of action in hybrid classroom interaction, in which some experience the classroom environment, participants and events in it 'directly' and others in a technology-mediated manner. Using a case study approach, we focus the analysis specifically on how classroom participants make sense of a remote student's (RS) talk, the robot's position and movements to maintain a shared understanding of classroom transitions during one lesson, and how they support the RS's agency and participation through various corrective actions targeting the robot's position and movements.

3. Data and methodological orientation

Our empirical data comes from a database of 15 h of video-recorded L2 teaching (English, Swedish, German and Finnish) in the Finnish higher education context in which 1–2 students participated remotely in classroom-based teaching with the help of a Double 2 or Double 3 telepresence robot. The Double robot was in the classroom during the lessons, and either an individual student or a pair of students connected to the device via the internet and operated it at a time. We were interested in how participants would use new technology in a real-life educational context. The robot was made available to interested teachers and students, and we let the participants decide how to use it in the teaching. One practical reason for this was also our wish to see if such a robot could be used as a tool for a relatively casual form of synchronous hybrid instruction that would not require much preparation in advance from the teacher. The dataset contains lessons in which students participated remotely because of a temporary injury, travel or a study abroad period, but also lessons during which the teachers and students took turns to try out robot-mediated participation from another room on campus.

We draw on multimodal conversation analysis (CA). A growing number of CA studies have begun to analyse the social organisation of mediated interactions (e.g., special issues by Arminen et al., 2016; Giles et al., 2017). Aligning with those studies, we aim to understand and discover possible affordances and challenges of using telepresence robots in classrooms and the ways in which participants maintain shared understanding of the unfolding situation as well as relevant next actions and events in the classroom (i.e., intersubjectivity) using the robot as a shared channel. In our setting, intersubjectivity is complicated by the fact that the remote student has a different audio-visual perspective on the classroom (see Luff et al., 2003; Rusk & Pörn, 2019, regarding similar issues in other video-mediated interaction data). The remote student perceives and experiences the classroom from a first-person perspective through what the telepresence robot mediates to their device (laptop or mobile phone). Even more importantly, the visual perspective of the remote student is partly opaque to the in-classroom students in the Double telepresence robot because they do not see the video frame that is streamed to the remote student, which makes interpreting the remote student's perception and perceptual focus a sort of 'black box' for them. In this study, the multimodal CA orientation enables us to analyse how the participants jointly 'do' interperceptivity through the mediated world of the remote-controlled robot by explicating how they perceive the unfolding situation and by being responsive to similar descriptions by the other party on the other side of the screen.

Analysing the dataset, we became interested in how the participants cooperated in navigating the robot and oriented to its movements in the classroom despite, or perhaps because of, being geographically distributed. Our focus was subsequently directed to navigation problems, and we noticed that in some cases a classroom participant would resort to human-assisted mobility (Cekaite et al., 2021) and carry the robot from one place to another in the classroom instead of relying on navigation by the remote student. We identified all such cases in the entire data corpus, forming a collection of 10 instances. Given that the very point of using a telepresence robot is to enhance remote participation by adding an opportunity for autonomous remote-controlled movement, these situations represent a type of technological breakdown in which participants (re-)configure the mode of remote participation and interactional resources relevant to it. As the rarity of the phenomenon suggests, these cases are somewhat deviant in the corpus. From a CA perspective, their analysis can thus shed considerable light on structures and practices of social action by showing what happens when some routinely occurring phenomenon, practice or resource does not occur or is not available. We further noticed that half of the cases of carrying the robot took place during the activity transitions of one Swedish lesson ('Akademisk svenska', *Academic Swedish*).

In this article, we adopt a case study orientation and present our findings by way of analysing these transitions in close detail to establish how and why carrying the robot emerged micro-longitudinally as a locally meaningful practical solution during the lesson for these participants, who were relatively novice users of this particular technology. In the analysed extracts, the remote student was unable to join the Swedish classroom physically because she was out of town on the day of the class. She was using the robot for the first time, but she had earlier seen another student use it once in the same classroom. The remote student had no long-term need for remote participation, nor did she express or display on the recording any significant challenges related to using VMI technology or perceiving the classroom space. As a material space, this classroom was very familiar to her and other students, as the previous classes had been held there during the course.

The data have been recorded with two classroom cameras and screen capture software on the remote participant's computer. The analysis is based on a careful examination of all these video sources; however, we have made the methodological choice to transcribe our extracts using the screen capture recordings to show the remote participant's visual perspective of the classroom space, events and participants. In the transcripts we present, we have paid particularly close attention to multimodally annotating the remote students' screen-based actions such as cursor movements and clicks and the robot's movements that result from these actions. With these methodological choices, we aim to make the remote participant's (emic) perspective analytically available, including aspects of remote engagement that manifest through her screen-based actions of trying to navigate the Double 3 robot, even if the robot does not always act accordingly or make these actions visible to in-classroom participants. Where relevant, we also illustrate key spatial concerns, movements and classroom participants' embodied actions (e.g., grabbing the robot) by showing screenshots from the classroom cameras as part of the extracts. We have used the conventions of multimodal CA to transcribe the events (Mondada, 2019), protecting participants' identities through pseudonymisation of names and by blurring their faces and other possible personal identifiers in screenshots.

4. Carrying the robot: from a repair solution addressing technological trouble to a local norm

In this section, we show how over four successive classroom transitions within c. 30 min, carrying the robot becomes established as a locally negotiated practice for managing situations in which there is a need for remote mobility. Fig. 1 shows the classroom layout (not to scale), illustrating how the six classroom students (CS) who attended the class that day were organised around two desks, and where the robot was positioned at the beginning of Extract 1a. The room had a whiteboard at the front and a TV screen at the back, on both of which the teacher projected teaching materials. As the continuous line in Fig. 1 indicates, the remote student initially began to move the robot herself during activity transitions (Extracts 1a–1b) but later transitions in the class were completed so that the teacher would carry the robot to the relevant place (dashed line). Altogether, these transitions constitute a micro-longitudinal trajectory that demonstrates how initial troubles in using a novel technology turn out to frame the remote student's engagement in subsequent classroom activities and the kinds of members' competencies that other participants can accountably expect from her.

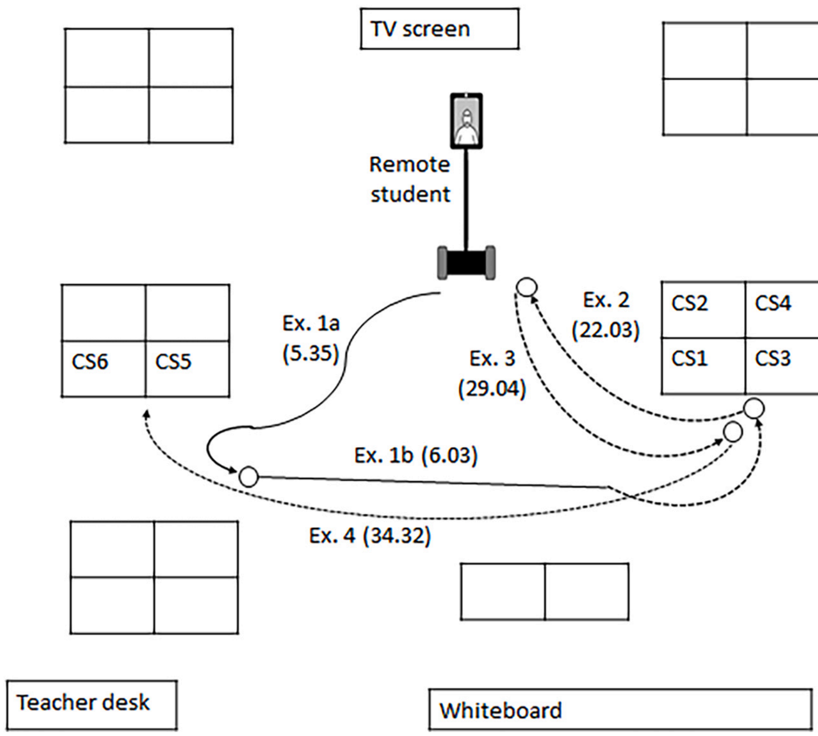
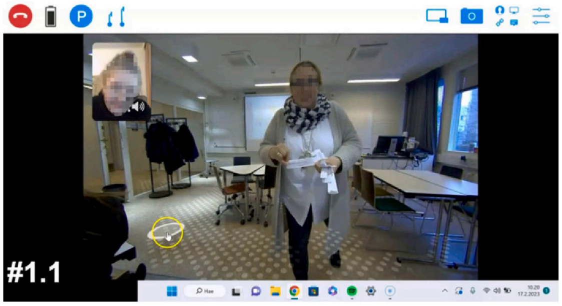


Fig. 1. Approximate movement trajectories of the robot in Extracts 1a–4. The continuous line indicates navigation by the remote student, and the dashed line shows transitions in which the teacher carries the robot. The number in brackets indicates the beginning time of each extract counting from the beginning of the lesson recording. CS = classroom student.

Extract 1a comes from the beginning of the Swedish lesson and shows a transition from the teacher's whole-class instruction to a pair/group discussion task. The teacher has just told the RS that she can drive the robot to any of the classroom groups, and she shows her the paper slips that are used in the task (l. 1–2). As she begins to hand out the slips to classroom participants and continues to talk about the task, the RS makes a series of unsuccessful navigation attempts to get to her group. She tries to use the robot's click-to-drive function, which, in principle, enables the robot to be navigated anywhere in the classroom by clicking on one of the mixed reality dots visible on the remote user's computer screen (as visible in transcript image 1.1). Selecting a destination this way should direct the robot to that place, but for some technical reason, this does not happen in Extract 1a.

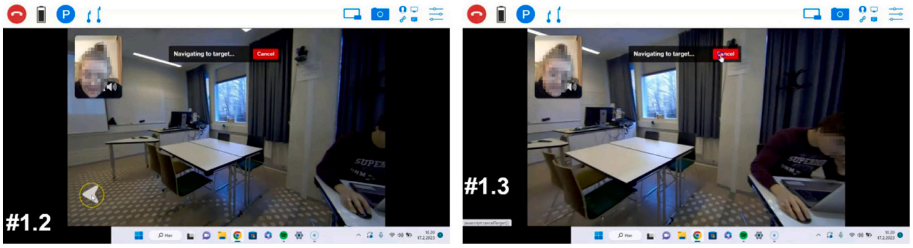
Extract 1a. The initial trouble

```
01 T     =nu får du bara (.) rulla, (.) till nån av *grupperna.
         'Now you can just roll/wheel to one of the groups'
   rs                                             *moves cursor->
02       (0.3)
03 T    ^så öö, (0.8) *ger jag ö- ää (.)
         'So erm I will give erm'
   t     ^walks to rs->
   rs              ->*
04       det är såna här lappar *så nån får ^lä↓sa åt dig sen.#=
         'It's these kinds of paper slips so someone can read to you'
   t                                       ->^shows slip->
   rs                            *moves cursor around->
   fig                                                         #fig1.1
```




```
05 RS    =mhyhy?
06       (0.6)^
   t       ->^walks away to hand out slips>>
07 T     (hela ti)den så-
         'All the time so-'
08       (0.4)
09 T     där har ni *påståenden?*¤
         'Here you have the claims'
   rs              ->*selects dot*
   robot                         ¤turns right->
10       (1.9)
11 T     är ni (.) ett- *(0.9)
         'Are you one-'
   rs                   *moves cursor on top of left arrow->
12       en (0.5) grupp¤ och så är det den <°andra¤ &gruppen# där.°>
         'One group so then there is another group there'
   robot           ->¤turns left----------------¤turns right->
   robot                                        &arrow emerges->
   fig                                                   #fig1.2
```

```
13          ¤(0.4)
   robot ->¤sways left and right in its position->
14 T      °så där var så *&goda.°
          'So there you go please'

   rs                 ->*moves cursor to the middle of the screen->
   robot               ->&
15        (1.9)
16 T      och *↑idéen# är nu *alltså att-¤ (1.2)&(0.2)
          'And the idea is now so that-'

   rs       ->*clicks cancel *moves cursor to left->
   robot                              ->¤
   robot                                       &shows arrow>>
   fig                #fig1.3
17          *en i ¤paret (0.4) tar en lapp? (0.7)
            'one in a pair takes a slip'

   rs     ->*clicks left arrow 15 times>>
   robot           ¤turns left>>
```



. (*continued*).

Although not visible to her co-participants in the classroom, the screen-based actions of RS show that she begins to prepare for the transition to the group activity already while the teacher is instructing her individually for the upcoming task (lines 1–4). She moves her cursor above one of the mixed reality dots on the left at the same time as the teacher is showing her the task materials (Fig. 1.1). The teacher disengages from the dyadic interaction with RS and goes to hand out the paper slips to the classroom groups (line 6), thereby treating RS as someone who can transition to her group on her own.

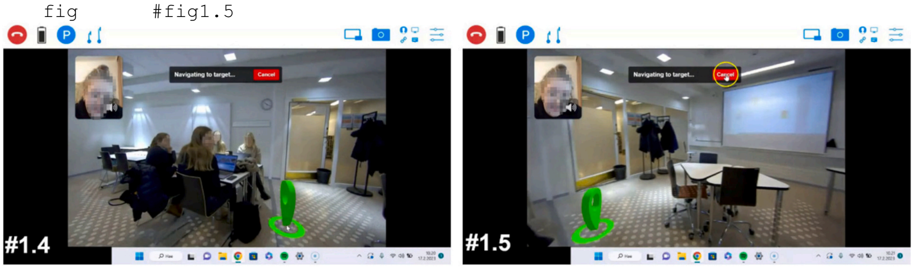
RS selects her destination by clicking on the dot, which prompts the robot to move there (line 9). In addition to RS's verbal turn in line 5, these screen-based actions thus show her situationally appropriate participation and engagement in the form of selecting a group for the upcoming task. The alignment of the teacher's and RS's conduct also displays their joint awareness of the unfolding transition.

The robot begins to move in a swaying manner at the end of line 9, but instead of going towards the selected dot, it begins to turn right as is visible in Fig. 1.2 (and Fig. 1 earlier). RS orients to this by moving her mouse cursor to the left side of the screen (line 11) where the robot can be manually turned by clicking on an arrow that emerges at the edge of the screen when the cursor hovers above it (Fig. 1.2). RS keeps the mouse in that position but does not click on the arrow for a while as the robot is swaying left and right (line 13), making little movement ahead. She then moves the cursor to the middle area of the screen and, after a short wait, terminates the automated navigation (line 16) by clicking on the red Cancel icon (Fig. 1.3). RS corrects the navigation trouble by moving the cursor back to the left side of the screen and, once the arrow becomes visible (line 16), by clicking on it several times to turn the robot towards her group (line 17 onwards).

All in all, RS's screen-based actions in Extract 1a show how she manages her technology-mediated participation during a classroom transition. This includes addressing emerging technological troubles when the robot does not go where it should be going. However, these screen-based actions are not available to the classroom participants, and they can only make sense of RS's participation and engagement in the situation by inspecting the robot's movement trajectories, the RS's talk and her facial expressions mediated by the camera. In other words, the classroom participants may recognise the swaying movement trajectory of the robot as possibly a signal of a technical problem, a user problem or both. In any case, they do not have access to the fact that, in principle, RS handled everything 'correctly' on the level of the user interface. Indeed, the continuation of the situation in Extract 1b shows how the navigation trouble gets noticed by the teacher (line 24). She treats the robot's movement trajectory as a signal of a problem in the remote student's engagement and lifts the robot to the remote student's classroom group members.

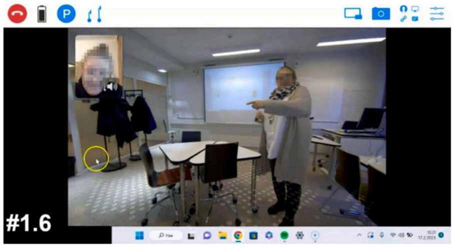
Extract 1b. Where are you headed?

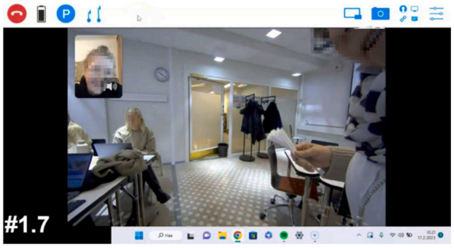
```
17 T     *en i ¤paret (0.4) tar en lapp? (0.7)
          'One in a pair takes a slip'

   rs     *clicks left arrow 15 times->
   robot       ¤turns left->

18       och så står det till exempel kunden har alltid rätt.
          'And so it says for example that the customer is always right'

19       (å då >säger<) (0.5) *öö¤, (0.3)
          'and then say erm'

   rs                       ->*moves cursor->
   robot                       ->¤

20       ↑jag* an↓ser* att# ¤↑kunden har alltid rätt ¤där*för att? (1.0)
          'I'm of the opinion that the customer is always right because'

   rs       ->*selects*                               *moves cursor->
   robot                    ¤moves forward-----------¤
   fig                    #fig1.4

21       och sen (.) *ska kompisen* komma med (ja) så säger du <det> men,
          'And then the partner will join like yes you say that but'

   rs              ->*selects dot *

22       (0.6) komma med ¤motargument.*
          'Join with a counterargument'

   robot                ¤turns right->
   rs                                *moves cursor and clicks cancel->

23       (0.7)

24 T     åt #vilket håll* är du på väg Minna¤*
          'In which direction are you heading Minna'

   rs                ->*                   *moves cursor to left->
   robot                                 ->¤
   fig       #fig1.5
```



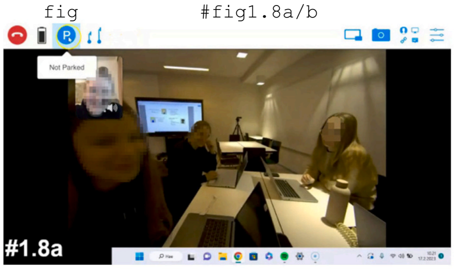

```
25       ska du <hit> eller,#
          'Are you going here or'

   fig                      #fig1.6

26       (0.6)

27 RS      *det är lite svårt att gå där?
            'It is a bit difficult to go there'

   rs     ->*clicks left arrow seven times->
```

```
28          (1.7)
29  T       ¤ja? m- öö (0.9) ska jag lyfta dig *dit.¤
             'Yes? Shall I lift you there'
    robot ¤turns left----------------------------¤
    rs                                       ->*
30          (0.5)
31  RS      ↑tack hhh hehehe
            'Thanks'
32  T&Ss    ((laughter))
33  T       öö, (1.3) nu (.) tar jag dig runt halsen# ^hhh hehehehehe
            'Now I take you by the throat'
    t                                              ^carries robot->
    fig                                          #fig1.7
```






```
34  RS      ((laughter))
35  T       så där^#
            'There'
    t            ->^
    fig             #fig1.8a/b
```




```
36  T       nu fick- (.) bli ni ett par med tre då så att var så goda.
            'Now (you) got- you will be a pair of three then so there you go'
37          (3.5) ((classroom students smile))
```

. (*continued*).

When her group members are visible on the screen, RS makes two more attempts at using the automated navigation. She first selects a dot in the space at the end of the group of desks (line 20, Fig. 1.4). The robot moves forward a bit but soon stops movement. The automatic navigation is still in progress when RS clicks on yet another dot further ahead (line 21). After some delay, the robot begins to turn right, away from the intended destination, and RS immediately cancels the automatic navigation during line 22 (as visible in Fig. 1.5).

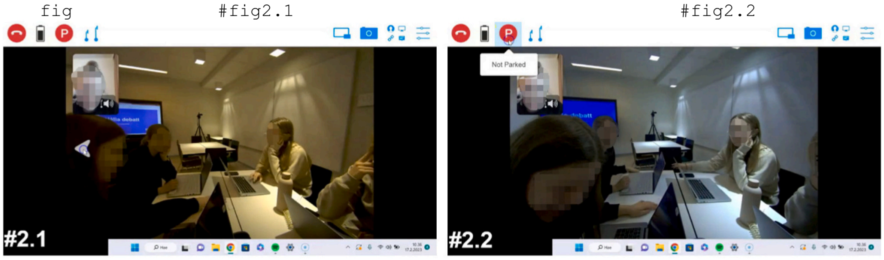
At this point, the teacher treats RS's transition to the group as problematic, an inference that is based on the movement trajectory of the robot, as she does not have access to RS's screen-based actions. Moving in front of the robot's camera, she makes herself visible to RS and asks where RS is going, pointing towards a candidate destination (lines 24–25, Fig. 1.6). Notice that at the beginning of Extract 1a, the teacher let RS choose which group she would like to join for the upcoming activity. In institutional interaction, asking where the recipient is going can be a way to treat their conduct as problematic (see Lin, 2023). The teacher's question can thus be seen to orient to clarifying two things: whether the somewhat aimless robot movement is deliberate and controlled by the remote student, and which of the two in-class groups RS has chosen for herself. In her response, RS confirms that "going (in) there" (line 27) is difficult. At the same time, she keeps trying to manually navigate the robot towards her group by repeatedly clicking on the left arrow on the screen.

Over lines 24–35, the teacher and the remote student negotiate an ad hoc practice of resolving the mobility trouble by carrying the robot to the group. They display in several ways an orientation to carrying the robot as an extraordinary and accountable course of action. Firstly, this workaround solution that bypasses the robot's mobility features is introduced through an offer-acceptance sequence to "lift" RS to her group (lines 29, 31). Secondly, the teacher, RS and classroom students mark the offer and carrying of the robot as a laughable matter. Laughter is a commonly used resource for managing interactional trouble and breaches of social norms (e.g., Glenn & Holt, 2013), and here the laughter seems to make the unusual and socially delicate nature of carrying "a student" visible. Thirdly, the teacher verbally treats the robot as an anthropomorphic participant. As she is beginning to carry the robot (Fig. 1.7), she laughingly describes her embodied action as a "grab by the throat" (line 33), a formulation that projects human-like qualities to the robot. Notice also how the design of the teacher's earlier offer refers to the participant-to-be-carried with the pronoun "you" (line 29). All in all, these two turns exemplify practices for constructing an "indeterminate referential identity" (Ochs et al., 1996, p. 328) and, here, treat the addressed participant as a socio-material assemblage that consists of the student and the robot (see Due, 2021). Carrying the robot to the place at the end of the desks where RS initially tried to navigate (Fig. 1.8a/b), the teacher assigns her as part of a "pair" of three students (line 36). As the transition from whole-class talk to the group has been completed, task work can now begin.

Extract 2, some minutes later, shows the end of the group-based warm-up task that begins in Extract 1b. In line 1, the teacher begins a transition back to whole-class talk to instruct the main activity of the lesson, a debate. As she initiates talk, she is behind the remote student and not visible on her screen. The remote student orients to the upcoming transition by trying to move the robot (l. 5), but the movement results in an unwanted collision that is ultimately resolved again by carrying the robot. Here, carrying is no longer treated as a socially problematic practice.

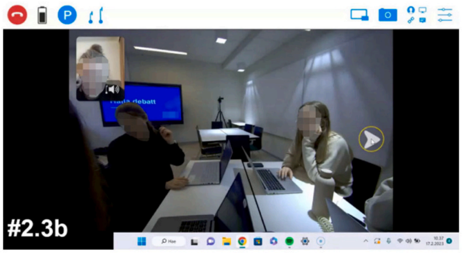
Extract 2. Now you moved our table

```
01 T      börjar ni bli värma i kläderna? (1.0) och argumentera?
          'Are you beginning to warm up for the debate?'
02 CS2    joo
          'Yeah'
03 T      joo? (.) [bra.]
          'Yeah good'
04 RS              [mmhm]
05 T      *då:: ää ((clears throat)) *ska vi (.) ta- titta på en power (.) point
          'Then we will look at a power point'
   rs     *moves cursor to left arrow*
06        (0.3) *hur# *vi ¤ska ↓göra idag *för att [hål↓la-]¤*#
          'How we will do today to hold-'
07 CS1                                             [.hh (apua) hhh he he he he
                                                        'Help'
08 CS2                                             [hhh he he he he
   rs           *clicks arrow*             *clicks "park"-----*
   robot                   ¤turns and pushes table forward----¤
   fig               #fig2.1                                     #fig2.2
```

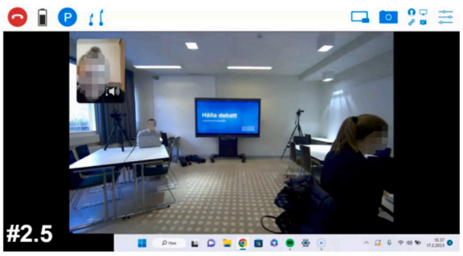



```
09        (8.0) ((classroom students laugh; RS clicks "park" three times and the
          left arrow once))
10 CS1    £sä siir(h)sit meiä pöy(h)tää vähä eteenpäi£
          'You moved our table forward a bit'
11 RS     *okei [hh (.) £okei, (0.3) ikävää? (.) sori?£* hh hehehehe]*
          'Okay okay. Too bad. Sorry.'
12 CSs          [((laughter))                                     ]
   rs     *moves arrow to mid-screen and back to left--*clicks arrow *
13 CS1    £ei se [(haittaa£ hhh) he he he he he he he he he he he he ]
          'It doesn't matter'
14 T            *[allt okej. *öö Minna om du vill (.) komma åt nåt håll=]
                 'Everything okay? Minna if you want to come someplace'
   rs           *moves cursor*
15        =så säg till så flyttar vi på dig.
          'tell it and we will move you'
```

16 *(0.3)*
rs *clicks arrow*
17 T säg vart du ↓vill.
'Say where you want to go'
18 (0.3)
19 T *sano minne *haluat ni &me siirretään.&
'Say where you want to go and we will move you'
rs *cursor to mid-screen*
cs1 &moves and withdraws hand twd robot&
20 (1.0)
21 RS ääm, *(1.0) jostai mist näkee ton, *(0.5) <taulun.>*
'Somewhere where you can see that board'
rs *cursor on classroom TV screen*cursor on left arrow*
22 &(0.8)*(1.0)*(0.6)
cs1 &puts hand around the robot->
rs *cursor on right arrow*
23 T mhm?#& ää ^[(vänta nu)]&
'Umm wait'
24 CS1? [( )]
cs1 ->&withdraws hand&
t ^stands up and walks to the robot->
fig #fig2.3a/b



25 (0.5)
26 T ↑jag tror du ser den bäst- (0.6) TV *skärmen är bättre tror jag.*
'I think you see it best- the TV screen is better I think'
rs *cursor on TV screen--------*
27 ^(0.3) eller^ hur,
'Isn't it?'
t ->^grabs robot^carries robot->
28 CS1 hh hehehehe#
fig #fig2.4
29 (1.2)
30 T här. (.) är det bra (här bort)^#
'Here. Is it good here away?'
t ->^
fig #fig2.5



31 (0.3)
32 RS ja tack (0.3) det är bra.
'Yes thanks. It's good.'

. (continued).

Following the teacher's talk that projects a transition (line 1), RS first moves her mouse cursor to the left-hand side of her screen during line 5. As in Extract 1a, this area of the screen displays an arrow with which the robot can be turned. The "PowerPoint slide" that the teacher is referring to in line 5 is visible both in front of the classroom and on the TV screen at the back, which RS can see on her screen but is unlikely to be able to decipher the text as it is far away (see Fig. 2.1). Hovering over the left arrow, RS clicks on it once (line 6). In this position at the end of the group desks, turning the robot would enable RS to either turn it 180 degrees to face the teacher and the whiteboard at the front of the classroom or allow her to navigate to a place where she would see the TV at the back of the room better. The attempt to turn the robot therefore displays RS's alignment with the emerging classroom transition and shows her moment-by-moment engagement with the instruction.

However, the attempt turns out to be troublesome. Instead of turning around, the robot bumps into the desks of CS1 and CS2 with a force that pushes the desks forward.[1] The classroom students react to this surprising turn of events by bursting out in extended and very loud laughter (line 7 onwards), using it as a resource for managing an unexpected transgression of the social classroom order (Jakonen, 2016; Looney & Kim, 2018). CS3 also orients to the collision by placing her hand at the end of the moved desks (visible in Fig. 2.2). RS's facial expression changes from somewhat neutral to one in which her eyes are wide open and her mouth makes a tense grin as the robot is moving the tables forward. RS furthermore reacts to the collision by trying to correct the movement: she clicks on the parking button several times and the left arrow again during line 8 and the ensuing long silence.

The classroom participants orient to the remote student's participation as incompetent. Besides the very long joint laughter, this is visible in how CS1 verbally informs RS that she has moved the desks (line 10). This treats the impact of the robot movement as unknown to RS. The turn also attributes the agency and blame for the problematic action to the remote student ("*you moved*") and not the robot. RS receives the informing turn as news ("*okay*") and apologises for it. She does not verbally claim an awareness of the robot's collision. An orientation to trouble is also visible in the teacher's check whether everything is "okay" (line 14).

The collision is treated as RS's attempt to move in the classroom space. The teacher makes relevant the workaround method of carrying the device that was negotiated in Extracts 1a–1b to bypass the robot's mobility feature. She first offers to help move the RS (lines 14–15) and then, in a more obligating manner, requests the RS to name a place where she wants to move, first in Swedish (line 17) and then in Finnish (line 19). Similar to Extract 1b, the teacher thus treats RS's intended destination as unclear to her, but, in contrast, here she no longer asks whether RS wants to be moved and therefore treats human-assisted mobility as the preferred course of action. By recognising RS's conduct as an attempt to move in the classroom space, the teacher orients to it as task-appropriate engagement, even if she orients to its cause as unclear to her.

It should be noted that the shift of attention from group talk to examining slides on the screen calls for a very different procedure from the classroom students and the robot-mediated remote student. For classroom students, the transition does not require movement from one place to another because they can reconfigure the participation framework from peer to whole-class interaction by simply shifting their gaze to the teacher and the screen. When RS names a place (line 21), CS1 makes preparations to move the robot by placing her hand around the robot's metal pole, as if scrutinising where to lift it (line 22, visible in Fig. 2.3a). She withdraws her hand as the teacher begins her response to RS (line 23) and walks towards the group. The teacher approaches the robot from behind, announces a place to RS (line 26) and lifts the robot (Fig. 2.4) without verbalising the lifting or making herself visible to the remote student as in Extract 1b. The teacher places the robot in a position in which the remote student can see the TV screen at the back of the room (Fig. 2.5) and checks whether the place is suitable for RS (line 30).

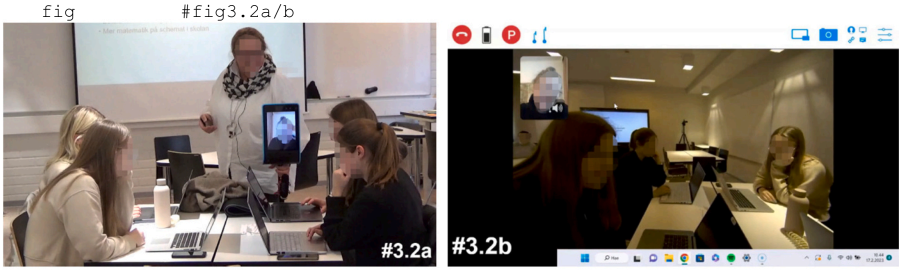
The two different types of mobility trouble shown in Extracts 1a–2 were ultimately resolved by carrying the robot to a relevant place concerning the ongoing activity. Through these troubles, the participants established an intersubjective routine for managing transitions by overriding the robot's affordance for remote-controlled mobility. In the rest of the lesson, the practice became more routinised such as in Extract 3, which shows the transition back to group work after the teacher's whole-class task instruction.

[1] Technically, this may result from the fact that the teacher has in Extract 1b placed the robot so close to the desks that there is no room for it to manoeuvre. Moreover, the robot's kickstands are deployed when RS clicks on the left arrow, which means that the robot will first need to release them and balance itself on its wheels before it can turn around. It may be that the combination of the robot trying to balance itself so close to a material object and RS's command to turn it makes it push against the desks, although its object detection sensors should prevent a collision.

Extract 3. I will lift the remote student back

```
01  T      å sen bestämmer vi (.) vilket ämne (.)
           'And then we decide which topic'
02         så vi ska börja förbereda för debatt.
           'So we will begin to prepare for debate'
03         (1.0)
04  T      så, (.) öö (.)
05         jag ^lyfter tillbaks Minna hit# till*     (0.6)^ hh  ^bordet.
           'So I lift Minna back here to the table'
    t          ^walks to robot----------------------------^smiles^
    rs                                           *zooms out, smiles faintly->
    fig                                                  #fig3.1a/b
```


```
06          ^(0.9)
    t      ->^lifts robot and carries it to group, smiles->
07  T      *så där.
           'There'
    rs     *grins broadly->
08         &(0.5)
    cs1&2  &smile->
09  CS?    ( )
10  RS?    hhh he he
11         (1.4)^#(0.6)&*
    t         ->^
    rs                  ->*
    cs1&2              ->&
    fig         #fig3.2a/b
```




```
12  T      så en kort diskussion (.) vad vill ni hålla debatt om.
           'So a quick discussion about what you want to debate'
```

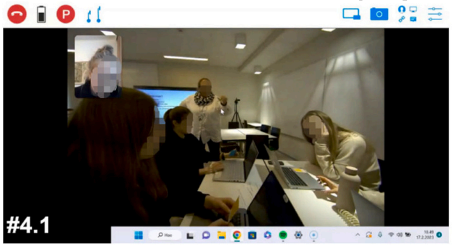
It is noticeable that the remote participant displays neither trouble related to situation-appropriate mobility nor a lack of awareness of what is going on, yet she is carried back to her group. The teacher's turn in lines 1–2 projects a transition to "preparing" for the debate activity but it does not yet provide a go-ahead for beginning the preparation in groups (and the other students in the class have not indeed yet begun such preparation at this moment). During this time, the remote student has been following the teacher's slides from the TV screen at the back of the classroom, using the robot's zoom function (Fig. 3.1b). In line 5, the teacher specifically orients to the remote student: she walks to the robot (Fig. 3.1a) and announces in a louder voice that she will move the remote student. As in Extract 1b, the teacher uses somewhat flexible practices of referential identification (Ochs et al., 1996) in her turn: Although she refers to the remote student with her name - which suggests her turn might be a commentary addressed to the whole class - the loud voice is a design feature that shows that the turn is above all meant to be heard by the remote student. When the teacher announces the unilateral movement, she also briefly smiles and chuckles (line 5). Smiling can be a way to flag one's turn as humorous (Çopur & Brandt, 2024), and here the smile and laughter token can be seen to deal with the socially delicate nature of carrying "a student", much like in Extract 1b.

Unlike in Extracts 1a–2, here the teacher thus does not offer or negotiate assistance in a moment of visible trouble during a transition but instead unilaterally moves the robot before any trouble has emerged. The remote student orients to the somewhat forced transition that reduces her agency over mobility by zooming out during line 5 and by beginning to smile. As the teacher lifts the robot and carries it to the table, she and the other participants mark the action as humorous by smiling (line 8 onwards; Figs. 3.2a/b). This seems to treat teacher-assisted mobility as a normative breach of the classroom order (e.g., Hazel & Mortensen, 2017).

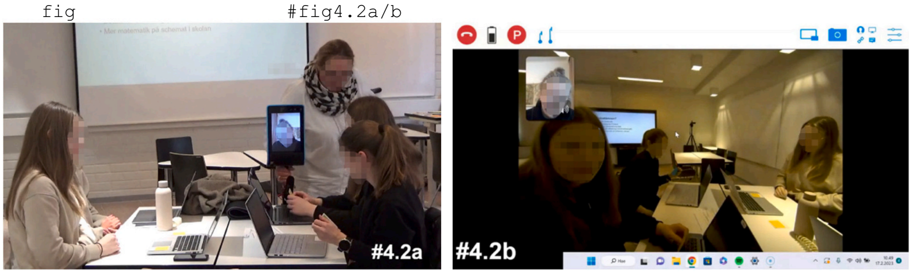
Although in Extract 3 the robot was carried without any indication of trouble in that situation, for the participants, the troubles in Extracts 1a–2 that happened earlier in the class are still a relevant aspect of their micro-level interactional history and a sense-making resource for adjusting themselves to the new form of participation. A similar unilateral carrying takes place in Extract 4, which shows how the teacher is reorganising the students into the eventual debate teams approximately 5 min after Extract 3. The teacher has just provided the in-class students with either a green or a yellow card, which shows whether they belong to the team that will argue for or against a topic at the following debate. She has also already verbally assigned RS to the green team as she in line 2 begins to instruct the students to go sit in their new teams and to carry the RS next to another group of desks (line 6).

Extract 4. I will take you again

```
01        (1.6)
02 T      nå ni ^gula# sitter^ bra
          'You Yellows are okay where you sit'
   t            ^points------^
   fig                 #fig4.1
03        ^men dom gröna får sätta sig nån annan^ stans (.) °nu°
          'But the Greens can seat themselves someplace else (.) now'
   t      ^turns towards other group desk-------^scans with gaze->
```




```
04        (0.8)
05 T      *öö ^(0.3) <Jenna,>& (1.0) ^<Min:na,> ^(0.9) <och> (1.0)^ (   )&
          'Umm Jenna, Minna and (   )'
   t        ->^walks, points at Jenna^gaze-robot^turns, gaze scans^grabs robot->
   rs     *smiles->
   cs1                       &smiles, glances ahead and behind----------&
06 T      .hh jag tar dig# igen nu ^he he
          'I'll take you again now he he'
   t                          ->^carries robot to other group->
   fig                  #fig4.2a/b
```




```
07        (0.6)*
   rs        ->*
08 RS     <ta[ck> (.)] för hjälpet.
          'Thanks for the help'
09 T         [så där.]
             'There'
10        (1.1)
11 T      *nu, ^hh he he he
          'Now'
   t         ->^
   rs     *smiles>>
```

In line 2, the teacher announces that the Yellow team can stay where they are. She simultaneously makes a sweeping pointing gesture towards the in-class students in RS's desk group (Fig. 4.1), in which three students have just been assigned to the Yellow team. This implies that RS's group, the Green team, will sit somewhere else. The teacher instructs this immediately afterwards while scanning the classroom with her gaze, as if identifying a suitable place in the classroom for the team (line 3). Altogether, these teacher actions project a need to move but, unlike in Extracts 1a–2, RS does not attempt to navigate the robot elsewhere from its current location. However, she smiles as the teacher, after a brief silence, starts to name the students in the Green team. While doing so, the teacher walks between the two groups of desks and orients to the students in the Green team in an embodied manner, pointing at Jenna, glancing at the robot and scanning the classroom with her gaze as she pivots just behind the robot to face the back of the classroom. Soon after, the classroom student who is seated nearest to the robot (CS1) also smiles and looks around (line 5). These smiles seem to target the projected need to move physically for the next task phase and the cumbersome nature of mobility to the robot-mediated RS.

As the students in the Green team have been verbally nominated, the teacher moves her hand around the robot's pole and announces in line 6 that she will move the robot (see Figs. 4.2 a and b). The design of the turn ("dig" *you*) and the louder volume address the turn to RS. Similar to the teacher's announcement of moving the robot in Extract 3, the turn conveys a unilateral, non-negotiated decision that is preceded by no indication of RS's mobility trouble nor any attempt to move the robot autonomously. The teacher still treats moving the robot as an accountable matter, but her account is more condensed than in Extract 3. Altogether, the brevity and lack of negotiation in Extract 4 suggest that carrying the robot has become the default method for managing activity transitions for the robot-mediated participant. As the robot is carried and placed next to other students in the Green team for debate preparation (lines 6–11), the teacher and RS nevertheless mark the transition as a laughable and extraordinary course of affairs through light chuckles (teacher) and by smiling (RS).

A similar pattern of non-negotiated yet verbally accounted-for carrying of the robot occurs once more in the final transition of the lesson when the participants gather together for the debate activity roughly 20 min later. In that situation, the teacher uses a similar verbal formulation to announce the movement ("Minna jag lyfter dig nu", *Minna I will lift you now*) simultaneously with grabbing the robot, and RS makes no attempt to move autonomously. Due to space constraints, we will not analyse that extract here.

5. Concluding discussion

This study has investigated how participants resolve emergent mobility troubles during activity transitions in telepresence robot-mediated synchronous hybrid classrooms, and what such troubles suggest about the joint maintenance of intersubjectivity and participation in hybrid classroom interaction. By analysing four successive activity transitions during one university-level Swedish as L2 lesson, we have shown how the participants managed transitions by carrying the robot from one place to another. Such a social practice of human-assisted mobility is highly different from the designed use of the robot. It can be seen as an example of how participants adapt themselves to situations in which technology that is supposed to facilitate communication (such as AAC devices, see e.g., Hengst et al., 2016) fails to meet the needs of situated social interaction. We argue that the local routinisation of the practice and the kinds of troubles it addresses illustrate not only quite typical failures of communication technology but also the creative ways in which participants make sense of novel technological affordances and create locally meaningful cultures-of-use around them (Thorne, 2016). Even though our data showcases a very specific kind of technology in use, the human-assisted mobility featured in it can be seen as a locally meaningful solution to the more general and inherent interactional asymmetry related to embodied participation in geographically distributed synchronous hybrid learning.

In all four cases analysed here (and in all the remaining cases from our broader data corpus from other lessons), the participants oriented verbally to the embodied practice of carrying the robot. Verbalisation took place through turns such as offering assistance, providing a commentary on ways of touching and carrying the robot and making jokes about the need to move the robot in the first place. Such verbal turns and the accompanying smiles and laughter marked carrying the robot as a deviant practice and treated it as an accountable matter, a breach of the normative classroom order (see also Hazel & Mortensen, 2017; Jakonen, 2016). The accountability of carrying may on the one hand be related to the hybrid and assemblage-like nature of the telepresence robot, which is oriented to both as a technological tool and a representation of the remote student's body (see also Due, 2021). The fact that the robot is often referred to with flexible pronoun use as "you" in the dataset and that participants orient to its material structures as bearing similarity to human body parts (such as the "throat" in Extract 1b) may partly explain why lifting the robot was treated as a normatively accountable matter. On the other hand, accounts for carrying the robot may also stem from the fact that human-assisted mobility is a practice that momentarily reduces the remote student's agency by taking away their possibility to move the robot as they wish and as the device is designed to be used. From this perspective, carrying is a corrective action that implies a socio-technical failure, a lack of members' competence in autonomous movement that the remote student can be accountably expected to demonstrate. The recurring use of smiles and laughter may soften and mitigate such implications.

The practice of carrying the robot initially emerged as a solution for addressing problems of participation of the remote participant, evident as seemingly aimless (Extracts 1a–1b) or uncontrolled (Extract 2) movements, in situations where activity transitions called for spatial repositioning in the classroom. However, this developed stepwise into an interactional routine for handling transitions during the later lesson that did not showcase any such problems (Extracts 3–4). Given that the focal lesson was the second time the robot was used in the class and the remote student used it here for the first time, the analysed cases show how these novice users made sense of the affordances of a novel communication technology and building a local culture-of-use around it. For them, the interactional task at hand in the analysed extracts was the spatial coordination needed for accomplishing activity transitions, which are vulnerable moments in classroom interaction, particularly in synchronous hybrid teaching (see also Jakonen & Jauni, 2024). The routinisation of transitions is also visible in that, over the course of the analysed lesson, transitions involving the remote student became progressively shorter in duration and less negotiated. Indeed, in Extracts 3–4, the teacher no longer initiates human-assisted mobility through offers of assistance to the remote student but simply announces that she will move the robot to the relevant place.

Overriding the robot's mobility can arguably be seen as detrimental to the remote student's agency for movement as it runs counter to the designed affordances and *raison d'être* of the device. As the participants established the new routine, their orientation to the remote student's agency shifted. At the beginning of the lesson, proposing carrying (Extract 1b), asking where the RS "wants" to go and checking whether the new positioning is acceptable (Extract 2) displayed the teacher's orientation to preserving the RS's agency as much as possible. Conversely, her later unilateral announcements of carrying (Extracts 3–4) did not display such an orientation and instead prioritised an efficient transition to the small group. Of course, the RS's agency continued to be a relevant concern in the small group, one that perhaps took its shape through interactional resources other than mobility. Nevertheless, in accomplishing these transitions, the teacher seemed to be striking a balance between control over lesson content and time used on activities, and students' affordances for engagement, agency and participation (e.g., Emanuelsson & Sahlström, 2008; Höglund & Jusslin, 2023). Given that the situations we have analysed are characterised by technology not working as intended and expected, the teacher's routine can be seen as a way to manage her own and the students' workload by minimising the time used on the routine task of moving from one place to another in a class that is becoming more complicated because of technology as a mediating factor (e.g., Moraiti et al., 2024).

Previous research has observed differences in experienced engagement levels between on-site and remote students in hybrid teaching (e.g., Raes, 2022). Compared to other videoconferencing solutions, the telepresence robot enables a more visible and embodied form of participation in the hybrid class than in hybrid designs where remote participants are present as a 'talking head' (Licoppe & Morel, 2012) displayed on the screen of a laptop on the table or a classroom screen in a fixed location. The extracts illustrate vividly how the robot constitutes a technologically mediated 'body' (Keating, 2017), an extension of the remote participant's self that claims a physical - and changing - space in the classroom. This means that classroom participants can make sense of the remote student's moment-by-moment engagement with the ongoing activity by monitoring the movements and navigation of the robot in the classroom space. At the same time, the enhanced embodiment brings with it new kinds of coordination troubles. It seems to require more engagement from the remote student than, for example, participation via Zoom where a participant can always switch off their camera and where tracking the remote participant's focus of attention is less transparent. Our analyses of the remote student's screen-based actions highlight how she worked hard to project the timing of activity transitions and initially navigated the robot in transition-relevant ways. Indeed, the screen-recording data shows that what the classroom participants treated as mobility troubles in Extracts 1a–2 were, or stemmed from, the remote student's appropriately timed engagement displays. These actions show that instead of 'sitting back', distributed embodied learning via the robot also requires considerable work to operate the technology. Studies of telepresence robots in non-educational contexts have suggested (e.g., Takayama, 2011) that, over time, participants can become so routinised to robotic technology that it becomes invisible in use. However, at least in the kind of early and non-routinised use that our extracts showcase, there may also be times when engagement with the technology takes away attention from task-related engagement in the hybrid classroom.

The telepresence robot is a technology for supporting remote embodied learning by enabling the remote participant to use mobility as a possible interactional resource. A direct implication of the ability to move is a need to coordinate participants' mobility trajectories in the classroom space during the unfolding pedagogical activity. The verbal practices preceding carrying the robot such as checking where the remote student is headed (Extract 1b) or telling RS when the robot has bumped into an object (Extract 2) orient to interactional asymmetries related to this form of VMI. They exemplify how the classroom-based participants display inferences about the causes of mobility troubles in ways that seem to treat RS as unaware of the trouble. In this sense, such verbal orientation to mobility and mobility-related troubles can be seen as interactional work to maintain and repair intersubjectivity and participation, in relation to what is happening now and what will happen next. They are motivated by quite apparent interactional asymmetries in the situation and in turn help the participants in distributed spatial coordination despite those asymmetries.

To conclude, this study adds to EMCA studies of mediated interaction (e.g., mobile telephony, video calls and text-based computer messages) an investigation of how participants solve troubles related to space and mobility in video-mediated activities. Maintaining intersubjectivity is a basic concern in social interaction in all kinds of contexts and particularly challenging in complex embodied activities mediated by technology in remote and hybrid education. Our study foregrounds the challenges of embodied participation and learning in synchronous hybrid teaching and illustrates how digital transformation can make human agency increasingly distributed, manifesting itself in and as emergent assemblages of humans, material artefacts and technology (e.g., Thorne et al., 2021). We have zoomed into a fairly specific context and interactional phenomenon, but we hope that the broader implications of our findings for education may be in helping to understand that participants need to conduct complex interactional work to coordinate action in multiparty hybrid settings, and in illustrating some ways to consider the role the human body and embodiment in technology-rich educational settings.

CRediT authorship contribution statement

Teppo Jakonen: Writing – original draft, Visualization, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Fredrik Rusk:** Writing – original draft, Formal analysis, Conceptualization.

Funding

This work was supported by the Research Council of Finland [grant number 343480].

Data availability

Data will be made available on request.

References

Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying mediated interaction. *Research on Language and Social Interaction, 49*(4), 290–309.

Beatty, B. J. (2019). *Hybrid-flexible course design*. https://doi.org/10.59668/33

Cekaite, A., Keisanen, T., Rauniomaa, M., & Siitonen, P. (2021). Human-assisted mobility as an interactional accomplishment. *Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion, 22*, 469–475.

Çimenli, B., Sert, O., & Jenks, C. (2022). Topic maintenance in video-mediated virtual exchanges: Rolling the ball back in L2 interactions. *System, 108*, Article 102834.

Çopur, N., & Brandt, A. (2024). Teacher's squeezed-mouth smile as a practice for managing affiliation in L2 classrooms. *Classroom Discourse, 15*(3), 256–272.

Due, B. L. (2021). RoboDoc: Semiotic resources for achieving face-to-screenface formation with a telepresence robot. *Semiotica, 2021*(238), 253–278.

Due, B. L. (Ed.). (2023). *The practical accomplishment of everyday activities without sight*. Routledge.

Emanuelsson, J., & Sahlström, F. (2008). The Price of Participation: Teacher control versus student participation in classroom interaction. *Scandinavian Journal of Educational Research, 52*(2), 205–223. https://doi.org/10.1080/00313830801915853

European Commission. (2023). *Working group on schools (2021-25) "pathways to school success": Blended learning for inclusion. Exploring challenges and enabling factors: key messages and illustrative examples*. Publications Office of the European Union. https://doi.org/10.2766/14836

Garfinkel, H. (1967). *Studies in ethnomethodology*. New Jersey: Prentice Hall, Inc.

Giles, D., Stommel, W., Paulus, T., Lester, J., & Reed, D. (2015). Microanalysis of online data: The methodological development of "digital CA". *Discourse, Context & Media, 7*, 45–51. https://doi.org/10.1016/j.dcm.2014.12.002

Giles, D., Stommel, W., & Paulus, T. M. (2017). The microanalysis of online data: The next stage. *Journal of Pragmatics, 115*, 37–41. https://doi.org/10.1016/j.pragma.2017.02.007

Gillespie, T. (2010). The politics of 'platforms. *New Media & Society, 12*(3), 347–364. https://doi.org/10.1177/1461444809342738

Glenn, P., & Holt, E. (2013). Introduction. In P. Glenn, & E. Holt (Eds.), *Studies of laughter in interaction* (pp. 1–14). Bloomsbury.

Hazel, S., & Mortensen, K. (2017). The classroom moral compass–Participation, engagement and transgression in classroom interaction. *Classroom Discourse, 8*(3), 214–234.

Hengst, J., McCartin, M., Valentino, H., Devanga, S., & Sherrill, M. (2016). Mapping communicative activity: A chat approach to design of pseudo-intelligent mediators for augmentative and alternative communication (AAC). *Outlines: Critical Practice Studies, 17*(1), 5–38.

Hjulstad, J. (2016). Practices of organizing built space in videoconference-mediated interactions. *Research on Language and Social Interaction, 49*(4), 325–341. https://doi.org/10.1080/08351813.2016.1199087

Höglund, H., & Jusslin, S. (2023). Embracing unpredictability. In S. H. Klausen, & N. Mård (Eds.), *Developing a didactic framework across and beyond school subjects* (1st ed., pp. 119–130). Routledge. https://doi.org/10.4324/9781003367260-12.

Jacknick, C. M. (2011). Breaking in is hard to do: How students negotiate classroom activity shifts. *Classroom Discourse, 2*(1), 20–38.

Jakonen, T. (2016). Managing multiple normativities in classroom interaction: Student responses to teacher reproaches for inappropriate language choice in a bilingual classroom. *Linguistics and Education, 33*, 14–27.

Jakonen, T., & Jauni, H. (2024). Managing activity transitions in robot-mediated hybrid language classrooms. *Computer Assisted Language Learning, 37*(4), 872–895.

Keating, E. (2017). Challenges of conducting interaction with technologically mediated bodies. In C. Meyer, J. Streeck, & J. S. Jordan (Eds.), *Intercorporeality: Emerging socialities in interaction* (pp. 303–322). Oxford University Press.

Kohonen-Aho, L., & Haddington, P. (2023). From distributed ecologies to distributed bodies in interaction: Capturing and analysing "dual embodiment" in virtual environments. In P. Haddington, T. Eilittä, A. Kamunen, L. Kohonen-Aho, T. Oittinen, I. Rautiainen, & A. Vatanen (Eds.), *Ethnomethodological conversation analysis in motion* (pp. 111–131). Routledge.

Leoste, J., Kikkas, K., Tammemäe, K., Rebane, M., Laugasson, E., & Hakk, K. (2022). Telepresence robots in higher education–The current state of research. In W. Lepuschitz, M. Merdan, G. Koppensteiner, R. Balogh, & D. Obdržálek (Eds.), *Robotics in education: RiE 2022* (pp. 124–134). Springer.

Licoppe, C., & Morel, J. (2012). Video-in-interaction:"Talking heads" and the multimodal organization of mobile and Skype video calls. *Research on Language and Social Interaction, 45*(4), 399–429.

Lin, Y.-H. (2023). 'Where are you going?' Managing wandering in a residential home in Taiwan. *Journal of Interactional Research in Communication Disorders, 14*(2), 241–267. https://doi.org/10.1558/jircd.24436

Looney, S. D., & Kim, J. (2018). Humor, uncertainty, and affiliation: Cooperative and co-operative action in the university science lab. *Linguistics and Education, 46*, 56–69.

Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., & Oyama, S. (2003). Fractured ecologies: Creating environments for collaboration. *Human-Computer Interaction, 18*(1–2), 51–84. https://doi.org/10.1207/S15327051HCI1812 3

Luff, P., Heath, C., Yamashita, N., Kuzuoka, H., & Jirotka, M. (2016). Embedded reference: Translocating gestures in video-mediated interaction. *Research on Language and Social Interaction, 49*(4), 342–361.

Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics, 145*, 47–62.

Moraiti, K., Bergviken Rensfeldt, A., & Lundin, M. (2024). Digital platform work reinforcing performativity: Teacher responses to work intensification explored through trace ethnography. *Critical Studies in Education*, 1–18. https://doi.org/10.1080/17508487.2024.2409661

Ochs, E., Gonzales, P., & Jacoby, S. (1996). "When I come down I'm in the domain state": Grammar and graphic representation in the interpretive activity of physicists. In E. Ochs, E. Schegloff, & S. Thompson (Eds.), *Interaction and grammar* (pp. 228–269). Cambridge University Press.

Oittinen, T. (2022). Negotiating collaborative and inclusive practices in university students' group-to-group videoconferencing sessions. *Linguistics and Education, 71*, Article 101107.

Olbertz-Siitonen, M. (2015). Transmission delay in technology-mediated interaction at work. *PsychNology Journal, 13*(2-3), 203–234.

Page, A., Charteris, J., & Berman, J. (2021). Telepresence robot use for children with chronic illness in Australian schools: A scoping review and thematic analysis. *International Journal of Social Robotics, 13*(6), 1281–1293.

Pekarek Doehler, S., & Balaman, U. (2021). The routinization of grammar as a social action format: A longitudinal study of video-mediated interactions. *Research on Language and Social Interaction, 54*(2), 183–202.

Raes, A. (2022). Exploring student and teacher experiences in hybrid learning environments: Does presence matter? *Postdigital Science and Education, 4*(1), 138–159.

Raes, A., Detienne, L., Windey, I., & Depaepe, F. (2020). A systematic literature review on synchronous hybrid learning: Gaps identified. *Learning Environments Research, 23*, 269–290.

Råman, J. (2017). The organization of transitions between observing and teaching in the budo class. *Forum Qualitative Sozialforschung, 19*(1), 28.

Rusk, F., & Pörn, M. (2019). Delay in L2 interaction in video-mediated environments in the context of virtual tandem language learning. *Linguistics and Education, 50*, 56–70. https://doi.org/10.1016/j.linged.2019.02.003

Rusk, F., Taylor, N., & Ståhl, M. (2023). The social accomplishment of seeing together in networked team play. *Simulation & Gaming*. https://doi.org/10.1177/10468781231209484

Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 52–133). Cambridge University Press.

Seuren, L. M., Wherton, J., Greenhalgh, T., & Shaw, S. E. (2021). Whose turn is it anyway? Latency and the organization of turn-taking in video-mediated interaction. *Journal of Pragmatics, 172*, 63–78. https://doi.org/10.1016/j.pragma.2020.11.005

Takayama, L. (2011). Toward making robots invisible-in-use. In K. Dautenhahn, & J. Saunders (Eds.), *New frontiers in human-robot interaction* (pp. 111–132). John Benjamins.

Thorne, S. L. (2016). Cultures-of-use and morphologies of communicative action. *Language Learning & Technology, 20*(2), 185–191 (doi:10125/44473).

Thorne, S. L., Hellermann, J., & Jakonen, T. (2021). Rewilding language education: Emergent assemblages and entangled actions. *The Modern Language Journal, 105* (S1), 106–125.

Uskokovic, B., & Taleghani-Nikazm, C. (2022). Talk and embodied conduct in word searches in video-mediated interactions. *Social Interaction. Video-Based Studies of Human Sociality, 5*(1).

Valasmo, V., Paakkari, A., Sahlström, F., Backholm-Nyberg, Y., & Majors, J. (2023). Invisible participation: A sociomaterial analysis of a synchronous distance lesson during emergency remote teaching. *Digital Culture and Education, 14*(4), 112–132.

von Wedelstaedt, U. (2020). The interactional accomplishment of "shootables": Visualisation and decision making before an Apache helicopter attack. *Ethnographic Studies, 17*, 100–124.