

# Non players' embodied practices of engagement in videogaming

Heike Baldauf-Quilliatre & Isabel Colón de Carvajal

Viittausohje:

Baldauf-Quilliatre, H., & Colón de Carvajal, I. (2024). Non players' embodied practices of engagement in videogaming [Ei-pelaajien keholliset videopeleihin osallistumisen käytännöt]. *Prologi – Viestinnän ja vuorovaikutuksen tieteellinen aikakauslehti*, 20(1), 65–88. <https://doi.org/10.33352/prlg.120564>

To cite this article:

Baldauf-Quilliatre, H., & Colón de Carvajal, I. (2024). Non players' embodied practices of engagement in videogaming. *Prologi – Journal of Communication and Social Interaction*, 20(1), 65–88. <https://doi.org/10.33352/prlg.120564>

# Prologi

– Viestinnän ja vuorovaikutuksen  
tieteellinen aikakauslehti

[journal.fi/prologi/](https://journal.fi/prologi/)

ruotsiksi: Prologi – Tidskrift för Kommunikation och Social Interaktion  
englanniksi: Prologi – Journal of Communication and Social Interaction

Julkaisija: Prologos ry.



Avoin julkaisu / Open Access  
ISSN 2342-3684 / verkko

---

---

## Article

---

---

Prologi, 20(1)

65–88

<https://doi.org/10.33352/prlg.120564>

# Non players' embodied practices of engagement in videogaming

Heike Baldauf-Quilliatre

Professor

University Lyon 2, France

[heike.baldaufquilliatre@ens-lyon.fr](mailto:heike.baldaufquilliatre@ens-lyon.fr)

Isabel Colón de Carvajal

Associate professor

Laboratoire ICAR (UMR 5191), France

[isabelle.colondecarvajal@ens-lyon.fr](mailto:isabelle.colondecarvajal@ens-lyon.fr)

received 13.9.2022 / accepted 28.5.2024 / published 3.9.2024

---

### Abstract

Using technology in interaction means actively manipulating a technological device while interacting with others, but it can also refer to a situation where one person is employing a technological device in the presence of others and therefore in a potentially interactive situation. This is, for instance, the case when one acts on a screen in the presence of other physically co-present participants. Our paper deals with a particular situation of co-presence and technology use: a couple is sitting side by side on the sofa in the living room, one is playing a single-player adventure game on a large TV screen, the other is playing another game on a tablet. For about 40 years, the concept of “active spectators” developed in media sciences has highlighted the fact that spectators “do” something. But, few studies have dealt with the interactional practices used to display this “activeness”. This paper explores the construction of spectatorship with a special focus on spectating videogaming as a particular way of using technology in the sense of spectating at what is being done with a technological device. We propose a micro-analysis of the above-mentioned gaming-situation in order to show how the non-player’s engagement in the player’s gaming is co-constructed by both participants. A particular focus was placed on gaze and gaze shifts as a resource to display togetherness and potential availability.

**KEYWORDS:** Embodied practices, engagement, monitoring, non-player, multimodal conversation analysis, videogaming

## Introduction

Videogaming is a very popular activity and numerous technological devices from the touch screen on the smartphone to a complex virtual reality apparatus are widely used for gaming in everyday life. Different studies have shown how players organize their gaming activities in physical presence or at a distance and how they interact to accomplish actions inside or out of the game (e.g., Aarsand & Aronsson, 2009; Mondada, 2012; Piirainen-Marsh, 2010; Reeves et al., 2017; Tekin, 2021). The importance of non-players and their practices came into the focus with regard to the social role of videogaming (Voids & Greenberg, 2009). Non-players might not use a technological device, but players and non-players together co-construct the way the technological device is used. More precisely, as active participants, non-players contribute to the gaming as co-players (Olbertz-Siitonen et al., 2021) or spectators by somehow commenting on the players' gaming actions. Thus, recent studies ask how non-players participate in gaming and which role(s) they take (Baldauf-Quilliatre & Colon de Carvajal, 2021; Tekin & Reeves, 2017). In this study, we want to pursue these first analyses and investigate a particular situation where the non-player's participation is not self-evident: a couple is sitting on the sofa, one is playing a single-player videogame on a large television screen, the other one is using a tablet and doing something else. We will investigate how the single non-player displays engagement in the videogaming without abandoning her activity on the tablet, which resources she uses to manifest her availability for interaction and how both participants co-construct togetherness despite their individual activities. Our study contributes to a better understanding of the construction of spectatorship from a conversation analytic point of view and therefore broadens understanding of gam-

ing as social activity, even if the game itself is a single-player game. Additionally, we want to draw attention to "togetherness"<sup>1</sup> in a technology-related setting, namely as achieved through a multitude of tiny sequences and the use of embodied resources.

We first focus on the dynamics of engagement and disengagement in the gaming activity – at certain moments of the game and throughout the game as a whole – in order to show how the engagement in another person's activity on a technological device is co-constructed in a larger time frame. We are particularly interested in gaze and gesture as resources to display sensorial engagement (Mondada, 2022) in the materiality of the gaming, by touching and manipulating the controller, as well as potential availability for interaction. In a subsequent section we will illustrate different practices of constructing engagement, more precisely of mutually displaying engagement, as for instance, responding to a recruitment for assistance or searching for understanding and offering assistance. We therefore (1) provide an overview of the "gaze organization" in a 45-minutes gaming session and (2) do a qualitative analysis of a segment of mostly silent gaming as well as (3) examine the sequential and multimodal interaction of two larger extracts from this gaming session that show frequently employed practices of co-constructing engagement.

### Non-players engagement in gaming activities: Monitoring and showing availability for interaction

Non-players are players who are waiting for their turn, players who have lost their previous game and therefore cannot play anymore, individuals who are co-present watching the game or individuals who simply share the same phys-

---

<sup>1</sup> The concept goes back to Merleau-Ponty's (1960/1964) "compréence" and has been used in conversation analysis to draw on the construction of "being together" in terms of spending time together (Kremer-Sadlik et al., 2008), sharing activities (vom Lehn, 2013) and/or sensations (Cekaite & Kvis Holm, 2017).

ical space and therefore have access to the gaming. In this study, we only focus on the last type: physically co-present non-players with access to the gaming activities. Since they share the same physical space, both non-players and players are aware of each other and perceive each other as player or non-player respectively. This means that they have to co-orient their sensory perception, co-ordinate their bodily movements and co-operate in interaction (Hausendorf, 2013; Meyer & Jucker, 2022). To do so, they need to mutually display engagement in the co-participant's activities (Goodwin, 1981).

“Engagement/disengagement” goes back to Goffman's (1957) description of involvement and is closely related to the concept of participation framework (Goodwin & Goodwin, 2004). “Participation” is described as participation framework to investigate “how multiple parties build action together while both attending to, and helping to construct, relevant action and context (Goodwin & Goodwin 2004, p. 240); “being engaged/disengaged” means that one displays that he or she takes part in the interaction in a particular way. Recently, Peräkylä and colleagues (2022) investigated engagement in psychiatric interviews through two temporal trajectories: the engagement in next actions and the engagement in the process during the entire series of interviews. In our setting, engagement in the gaming activity also needs to be considered on different trajectories: locally (e.g., current gaming actions), and more globally, with regard to gaming strategies. Peräkylä and colleagues (2022, pp. 259–260) unpack engagement into three closely related and analysable components, such as collaboration in a joint action, bodily orientation to a co-participant and a shared moral order. Thus, engagement can be understood as a multimodal Gestalt (Peräkylä et al. 2022, p. 292; see also Mondada, 2018). We draw on these components to show

how non-players orient to the gaming activity even when they do not comment on the gaming verbally.

Especially when players and non-players stay together in the same place for a while, they also participate in the construction of a particular encounter, where individuals might be engaged in separate activities but mutually display their presence, in other words they display “togetherness”. Vom Lehn and colleagues (2013) show how museum visitors construct togetherness by transforming side-by-side arrangements into a joint activity as a concerted onward movement by mutually displaying a shift in orientation and readiness to move on. Similar to museum visitors, the participants in a gaming encounter can momentarily transit from an unfocused gathering to focused interaction within a social encounter by shifts in orientation, for example by gazing at the gaming screen. But in contrast to museum visitors, non-players and players, in the data we investigate, are not engaged in a similar activity, such as visiting an exposition, but in two different activities, that is gaming for the player, acting on the tablet for the non-player. Constructing togetherness is thereby closely related to the display of “availability”. For Heath (1982, p. 154), the display of availability “is an action that creates, for its recipient, a range of undifferentiated opportunities in which to initiate action” when and where they wish. The multimodal and joint construction of availability has since then been widely described for different situations: for interaction with technological devices (see for instance Gehle et al., 2017; Kohonen-Aho & Vatanen, 2021; Weilemann, 2003), but also with regard to the transition from a gathering to an encounter (Mondada, 2009) or from one encounter to another (De Stefani & Mondada, 2018). In contrast to these studies, we investigate an unfocused interaction where participants constantly transit to

small bits of focused interaction and back to unfocused interaction, a situation described by Goffman (1967, p. 145) as “open state of talk”.

Similar to spectators or co-players, non-players in gaming interactions more or less simultaneously watch the game and potentially interact with the player.<sup>2</sup> Haddington and colleagues (2014, p. 6) investigate multi-activity as “collective, collaborative, and intersubjective process”. They argue that “the activity or activities in which a person is engaged cannot be separated and isolated from the engagement of others, from multiple participation frameworks [...]” (p. 13). More specifically, participants can orient to different activities simultaneously, they can disengage from one activity to engage in another, they can stop or postpone an activity to orient to another one etc. (p. 24). With regard to gaming interactions, this implies, first, that all participants, non-players and players, co-construct togetherness - despite their different activities - through a display of mutual engagement in the co-participants’ activities. Second, it means that non-players who are engaged in another activity than watching the gaming can use different types of resources to engage simultaneously in the two activities. They can also develop particular practices to pause or postpone their own activity and engage in the gaming. One of the practices for doing so is “monitoring”.

Monitoring in multi-activity interactions was first described for workplace settings (Goodwin & Goodwin, 1996; Heath & Luff, 1996). The concept is introduced by Goodwin (1980) who describes how co-participants mutually monitor their actions through gaze and then readjust their turns. More recently, Luff and Heath (2001, p. 152) explain monitoring in workplace settings “as ‘peripheral awareness’ where individuals appear sensitive to other activities and

events whilst often engaged in other tasks”. In this sense, monitoring seems to be a particular practice that participants use to display their local sensitivity. While in some multi-activity settings and especially in workplace interactions, participants keep a sustained orientation to one activity through monitoring while other activities are constructed as additional (Nishizaka, 2014), in gaming interactions, the non-player keeps, through more or less regular monitoring, an orientation to the gaming activity, but this activity is not necessarily constructed as primary.

With regard to workplace interactions, Kamunen and Haddington (2020) draw attention to monitoring as a resource for a jointly coordinated transition to a follow-up activity. They argue that participants show through gaze shifts and body positioning their orientation to “an observably emerging event” (p. 116) and additionally complete the embodied conduct with verbal or embodied prompts of their co-participants to jointly transition from the ongoing activity to the emerging one. In a similar way, the non-player’s monitoring practices might indicate a transition from individual silent gaming into an interactional exchange. But in the gaming activity, the monitoring non-player does not necessarily prompt the player, and the player does not necessarily collaborate and start the transitioning.

Gaze shifts are also important with regard to the differentiation between “noticing” and “watching” as described by Kidwell and Reynolds (2022) based on the analysis of different types of video-recorded settings: “noticing gaze shifts” are announced by quick head turns while lower body and torso remain in the initial position. They alert that something “attention worthy” is happening and that this happening is displayed as a newly discovered event (p. 6).

---

<sup>2</sup> See also Holly & Baldauf (2002) for the specificities of “talking while watching”.

Watching, in contrast, can be recognized by a “sustained and committed arrangement” of the whole body as well as “the sustained direction of gaze”; it alerts that something already discovered is worthy of sustained attention (Kidwell and Reynolds, 2022, p. 14). This differentiation allows us to distinguish more precisely the non-player’s gaze and its functions within the monitoring practices, as for instance quick gaze shifts to the gaming screen and a longer gazing preceded by verbal turns and/or followed by an interactional exchange.

By using mobile eye tracking glasses and an external camera, Stukenbrock (2020) observes “gaze following” (e.g., the addressee follows the speaker’s gaze) and “gaze monitoring” (e.g., the speaker checks the addressee’s gaze orientation). She describes gaze following as “an interactional practice of tapping into a coparticipant’s gaze and exploiting it as a resource to gather information on *where*, *how*, and for *how long* he or she is looking, and to infer *what* he or she is looking at, and *why*”, whereas gaze monitoring is defined as “an interactional resource [...] to gather real-time evidence on whether joint attention is emerging and to incrementally add material when they anticipate that intersubjectivity is threatened” (p. 20, emphasis in the original). In our analyses we are interested in the way non-players follow or check what happens on the screen or what the co-participant is doing with the controller. Therefore, the distinction between “exploiting [the player’s actions] to gathering information” through the non-players’ gaze to the screen and “gather real-time evidence” on emerging joint attention through the players’ gaze to the non-players becomes crucial.<sup>3</sup>

With regard to previous studies on gaming, we claim that non-players’ monitoring practices are particularly important as display of availability for interaction and therefore as display of engagement – concerning both, gaze direction, and gaze duration. The direction of the non-player’s gaze, for instance, to the screen, to the player(s), to other non-player(s), to consoles, controllers or other technological devices, etc. as well as its duration project different types of actions and sequences.

## Data and methodological approach

This study is based on video recordings of gaming sessions in a non-experimental setting that were collected within the research project “Ludespace”.<sup>4</sup> Altogether, we collected about 20 hours of videogaming in France and in French in eight different situations, involving different numbers of participants. In all cases, the participants decided what to play, how long, with whom, etc.

For this study, we focused on one particular gaming setting of about 90 minutes: Lucie and Greg, a couple with two children not being present during the data collection, were sitting side by side on the sofa at home. Greg played *Tomb Raider: Underworld* (Eidos Interactive, 2008), on Wii (Nintendo, 2006), a single player action-adventure game where the player takes control of Lara Croft. Lucie was occupied with a tablet during the first half of the session. They stayed in the family living room and Greg played on the large TV screen. We were particularly interested in this setting because in the first 45 minutes of the session, both participants carried out their own activity: Greg was playing a single player game, Lucie was acting on

<sup>3</sup> Stukenbrock’s “gaze following” is surely different from “gazing to the screen”, but we think that the distinctive functions she identified, for instance exploiting the co-participant’s actions to gather information, are highly relevant for the screen-based interaction.

<sup>4</sup> The “Ludespace: the spaces of video gaming in France” project was funded by the ANR program “young researchers” during three years (2012–2014). The participants have given permission to publish their photographs for scientific purposes.

her tablet. We asked if, despite these different “individual” activities, social interaction had occurred and if so, how the participants established joint attention and co-constructed engagement in the co-participant’s activity.

The data were recorded with two cameras: one focusing on the large gaming screen, the other on the two participants on the sofa. The setting therefore didn’t give access to the screen of Lucie’s tablet. We could only see when she was acting on her tablet but not what she actually did.

Our research was based on the principles of conversation analysis (Sidnell & Stivers, 2013). CA is a micro-analytic and qualitative research method that focuses on the co-constructing of interaction and on the participants’ practices to mutually display the understanding of their actions. However, within the last decade, CA researchers have started to use quantitative methods, and different studies have given insights into the way to combine both approaches (e.g., Robinson, 2020; Stivers, 2015). Drawing on the methodological reflections developed in these studies, we have used a coding scheme - gaze shifts as well as gaze duration, preceding and follow-up actions, etc. - to get an overview of the organization of the first 45 minutes of the gaming session where Lucie and Greg were both involved in different activities. We did not conduct any statistical analysis but instead, by quantitatively exploring the corpus, we identified recurrent patterns related to gaze shifts and their distribution through the recorded session. Considering Peräkylä and colleagues’ (2022) different temporal trajectories of engagement/disengagement, the exploration of the quantitative corpus allowed us to focus on the “global scale” and to show when gaze shifts occurred and how they were distributed over the gaming session, while the qualitative analysis allowed

us to investigate how engagement was co-constructed “locally”.

### Gaze and gaze shifts during the gaming session involving multiactivity<sup>5</sup>

When we transcribed our data, we observed that Lucie, the spectator, frequently gazed at the gaming screen whereas Greg hardly ever looked at Lucie’s tablet. We wanted to know if these gaze shifts occurred at particular moments relating to specific actions accomplished by the player or events occurring in the game, how they were organized throughout the interaction according to the moment when they occurred, their duration, the social actions that were accomplished before, during or after these gaze shifts etc.

In a more general way, we wanted to see how Lucie’s engagement in the gaming activity was structured and evolved during the 45 minutes where both participants were engaged in two different activities. Lucie’s gaze shifts from her tablet to the gaming screen or the controllers were extremely frequent at some moments and less frequent at others. She sometimes gazed for a rather long time to the screen or the controllers, sometimes she withdrew and reoriented her gaze very quickly. To investigate this “gaze organisation”, that is to say, the direction of gaze shifts (i.e., gaze shifts between participants themselves and between participants and screen) and the way they were maintained or not through the interaction, including the way they were projected and responded to, we coded Lucie’s gaze shifts in ELAN and annotated five phenomena: (1) gaze shifts from her tablet to the gaming screen, (2) gaze shifts from the tablet to the controllers, the co-participant, the recording device, etc.,<sup>6</sup> (3) the element/action

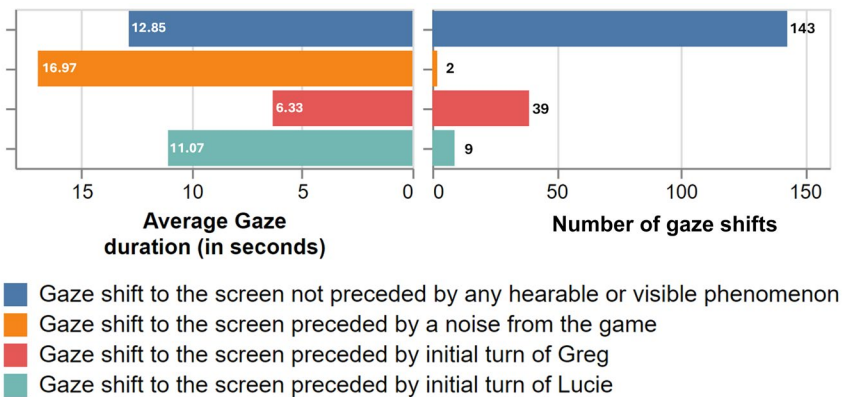
<sup>5</sup> This is the first half of the gaming session, about 45 minutes.

<sup>6</sup> We decided to use only these two codes for gaze since it appeared that the gaze shifts to / withdrawals from the screen are the most interesting.

that preceded the gaze shift and that may have projected it, (4) the type of this preceding element/action and (5) the actions that followed the gaze shift. We thought that the coding allowed us to show the “global trajectory” of engagement and pointed to interesting aspects concerning potential “local practices”.

Our first observation concerned the frequency of Lucie’s gaze shifts and the duration of gazing at the gaming screen: approximately 25 minutes (half of the time of the analysed 45 minutes), Lucie gazes at the gaming screen and the shifts from one direction to the other are rather equally distributed through the entire session: there are no longer phases of (dis)engagement in the gaming activity if we consider gazing at

the gaming screen as sufficient for “being engaged in the gaming”. But this means also that the other 20 minutes, she is looking at and acting on her tablet. During the entire 45 minutes she maintains an engagement in at least two activities, switching from one to another or being engaged in both of them simultaneously. We then looked at what happened before the gaze shifts to Greg’s screen (Figure 1). Altogether, only 20% (n = 39) of the gaze shifts are preceded by a verbal/vocal action accomplished by Greg which means that the player rarely solicits the non-player, neither vocally nor by gesture, gaze or body posture. He nearly always looks at the screen without moving his body. Also, his actions are not always initial ones that project a responsive turn, such as requests or questions.

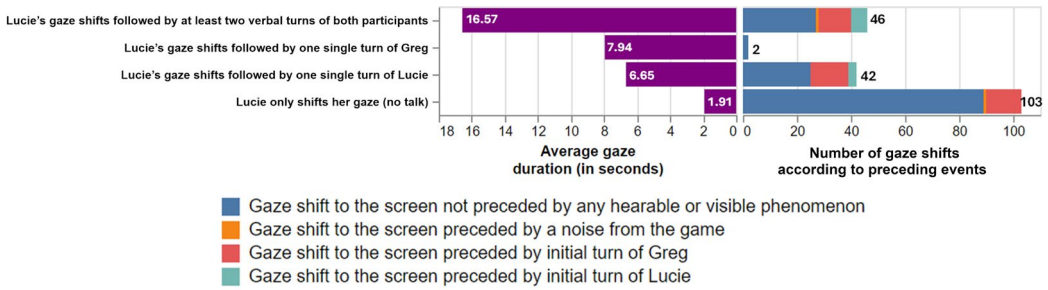


**Figure 1. Non-player’s gaze to the screen during the first 45 minutes**

Most of the time, Lucie shifts her gaze to the gaming screen without being recruited in some way (Kendrick & Drew, 2016): over 70% (n = 143) of Lucie’s gaze shifts do not seem to be projected in any way. This observation raises a number of different questions as, for instance: Do the gaze shifts really occur “out of nothing”? And if so, what does this mean for the co-con-

struction of engagement in the game? Do the non-player’s gaze shifts announce a shift in orientation and project an emerging social action in the gaming frame and display availability for participating in the game? With regard to these questions, we looked at the follow-up actions or events and connected them to the gaze duration (Figure 2).





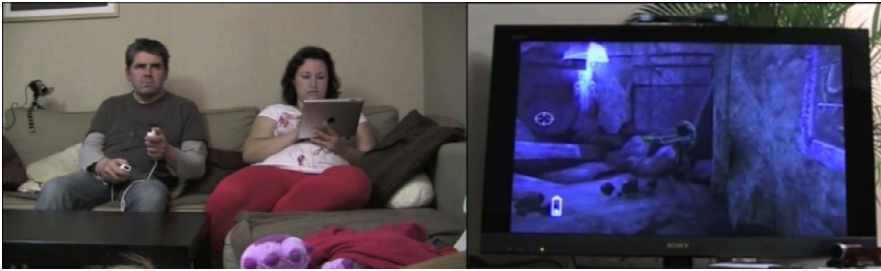
**Figure 2. Non-player's gaze to the screen and sequence organisation**

Three aspects became immediately visible: First, 52.8% ( $n = 103$ ) of the gaze shifts are not followed by a verbal action, neither from Greg nor from Lucie. In this case, the gaze is very short (appr. 1.91 sec.). Second, 22.7% ( $n = 42$ ) of the gaze shifts are followed only by one single verbal turn from Lucie that is not taken up verbally by Greg. Of course, this does not provide any indication about the way Greg eventually adjusts his gaming action, but it might suggest that both participants do not orient to (verbal) interaction. Third, less than one quarter of the gaze shifts (23.8%,  $n = 46$ ) actually precede an interactional sequence with at least one turn from each of the two participants. In this case, Lucie's gaze at Greg's screen lasts rather long (appr. 16.57 sec.). Generally speaking, throughout the whole session we can distinguish a short gazing to the screen with a gaze shift that is either not followed by any turn or visible responsive action or that is only followed by one single verbal turn, and gaze shifts that lead to a longer gazing at the gaming screen – and that are rather often followed by sequences of interaction. This observation allows us to distinguish two types of participation and maybe of engagement: a “noticing” in the sense of Kidwell and Reynolds (2022) that does not project interaction, but displays a minimal form of engagement, and a “spectating”, close to Kidwell and Reynold's “watching”.

### Silent gazing at the screen as display of minimal engagement

In the following section we analyse a segment of about two minutes that represents a moment of silent gaming with multiple gaze shifts. We will look at the way these gaze shifts contribute to display engaging in/disengaging from the gaming activity and show how they are used to display availability for interaction as a core element for the construction of togetherness. The segment highlights how gaze shifts participate in the construction of what we call “minimal engagement” by briefly, but consistently orienting to the screen as the object of potential joint actions and by therefore displaying availability for collaboration.

Lucie is playing on her tablet, Greg is playing on the Wii (Figure 3). Lucie is completely silent during the whole segment; Greg produces one turn that is not taken up (see Extract 1 further below in the text).

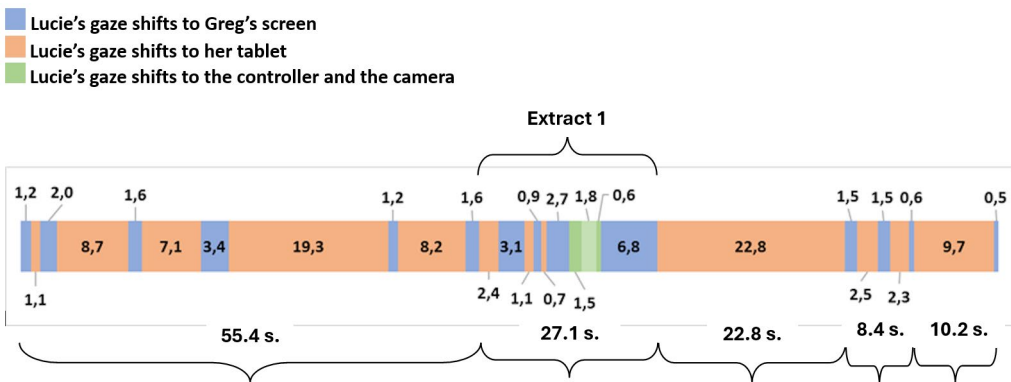


**Figure 3. Greg and Lucie’s postures at the beginning of segment 1 (00:47:57)**

The timeline (Figure 4) that indicates the direction of Lucie’s gaze as well as its duration, allows us to focus on our main point. Going back to Peräkylä and colleagues’ (2022) components of engagement, the timeline shows the non-player’s bodily orientation to the co-participant through gazing at the gaming screen, but collaboration and a shared moral order cannot be investigated, since the player never gazes at the non-player and reacts only once. He neither changes nor adapts his gaming. The non-player herself gazes only momentarily at the gaming

screen without producing any verbal turns. The collaboration in a joint activity can therefore only be established unilaterally with Lucie’s potential noticing of what happens on the screen.

One single gaze shift cannot be interpreted as display of engagement. However, the multiple gaze shifts to the gaming screen, despite a continuing engagement in another activity, allow us to consider a minimal engagement of the non-player in the gaming activity.



**Figure 4. Timeline 1 (00:47:57–00:49:57)**

Looking more closely at what Lucie is doing in this segment, we will divide it into five moments with regard to gaze organization. During the first 55.4 seconds of this segment (the first 55.4 seconds in Timeline 1, Figure. 4), Lucie al-

ternates between longer gazing at her tablet and very short glances to Greg’s gaming screen. She hereby displays some engagement but not an unconditional availability for interaction, since the longer focus on the tablet signals an engage-

ment in another activity than Greg's gaming. And this activity competes with a continuous monitoring of the gaming screen.

During the next 27.1 seconds, Lucie gazes nearly exclusively at the screen and the controller. She moves her tablet down, displaying that she is, momentarily, pausing from her own activity, and gazes at the gaming screen. Lucie's pausing might be related more to her activity on the tablet rather than to Greg's gaming since nothing has changed in the game or in the gaming. Nevertheless, she gazes at the gaming screen and the controller instead of doing something else.

### Extract 1 (49.02–49.15)

```

1 luc looks at screen
2 luc #looks at controller in Greg's hands
  scr #change of perspective
3 luc looks at camera
4 ava starts swimming around
5 luc #looks at screen ->9
  ava #swims around
6 ava blocked by a wall, changes direction
7 GRE #xx où est la sortie
   xx where is the exit
  ava #swims alongside the wall
8 ava stops and changes direction
9 luc looks at tablet

```

But Lucie does not offer help: in the third part of this segment, during the following 22.8 seconds, she turns away and focuses again on her tablet, displaying that she is no longer available for interaction and that she is no longer collaborating. This part gives an insight into the complex relationship between constructing togetherness, showing availability for interaction and responding to visible displays of trouble. Not offering assistance is not treated as problematic and does not question the principal construction of togetherness; Greg does not insist and continues the gaming.

We consider this choice as “doing togetherness” and showing a minimal participation in Greg's activity. Interestingly, Lucie not only gazes at the screen, but also moves her gaze around, monitoring Greg's technical manipulations as well as the recording device (Extract 1, l. 2–4). At this very moment, Greg answers by an interrogative as a verbal display of searching (l. 7) which could be interpreted as display of trouble and recruitment for assistance (Kendrick & Drew, 2016) and therefore as uptake of the gaze to the gaming screen as display of availability for interaction.

In the fourth part, during 8.4 seconds, Lucie rapidly alternates between her tablet and the gaming screen. Again, she moves her tablet down, but only briefly. For a very short time, she continues pursuing two activities quasi-simultaneously, moving down and up the tablet, continuing writing and making short pauses, gazing at the gaming screen. The visual display of multi-activity as “dual involvement” (Raymond & Lerner, 2014) shows her possible re-availability for interaction with regard to Greg's gaming as well as her ongoing engagement in her own activity. She does not transit to focused interac-

tion but indicates that she might be available at this very moment. Greg, on his part, continues moving around his avatar without responding to Lucie's potential availability.

During the last 10.2 seconds, Lucie focuses on her tablet, before gazing again very shortly at the gaming screen. Each of the two participants momentarily pursues their own activity, without any collaboration or taking into account the co-participant's activity. The segment ends with Lucie gazing shortly at the screen.

During the whole time, both participants remain focused on their own activities and never publicly co-construct joint attention (Kidwell & Zimmerman, 2007): when Lucie gazes at the gaming screen, Greg never monitors her gaze (Stukenbrock, 2020) to gather evidence of emerging joint attention<sup>8</sup>, he only once addresses her indirectly as co-present participant by verbally displaying his trouble – but at a moment when she did not gaze at his screen and did not display availability for interaction. Lucie, on her side, gazes silently at the gaming screen, shifting her gaze constantly, but not in even intervals, between her and Greg's screen. She displays only very briefly by her gaze to the controller in Greg's hands that she is following his gaming. She never announces changes in the participation framework, putting Greg's gaming in the foreground and her own activity

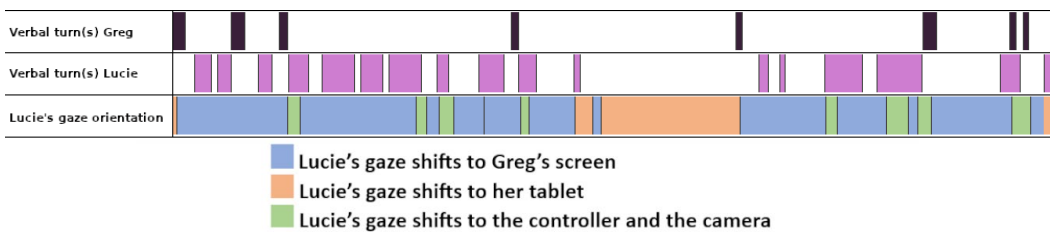
in the background – in other words she does not “watch” the gaming (Kidwell & Reynolds, 2022). Nevertheless, by sitting together on the sofa and displaying a general availability for interaction (Lucie through gaze shifts, Greg through his recruitment of assistance), they display their presence to each other.

## Multimodal practices of engagement

In contrast to the first example, the following two segments show a verbal exchange between the two participants and allow us to identify practices of engagement through the analysis of two extracts (Extracts 2 and 3).

### Responding to a recruitment of assistance

In the second segment (Extract 2), the player requests the non-player's assistance. We will focus on three aspects: the design of the requesting turn, the sequence organisation and the non-player's sensorial engagement in the gaming activity through touching the controller. The timeline of this segment of 55 seconds (Figure 5) shows a different structural gaze organization compared to the silent gaming in segment 1 (Extract 1).



**Figure 5: Timeline 2 (00:21:12–00:22:19)**

<sup>8</sup> But his turn (Extract 1, l. 7) can be considered as possible proof of his awareness of her monitoring him.

Lucie gazes more frequently and much longer at the gaming screen, also shifting her gaze to the player. Besides an initiation of the interaction by Greg, most of the turns are produced by Lucie.

Looking more closely at the transcription, Greg explicitly solicits Lucie's assistance through an

open question (l. 1), and Lucie not only helps with suggestions, explanations and instructions but also engages physically in the gaming activity. Lucie is thus monitoring Greg's gaming - watching it - and actively collaborating in the gaming, which becomes the joint activity of both participants.

### Extract 2 (00:21:12–00:22:19)<sup>9</sup>

01 GRE #lqu'est-ce que je fais avec \*le grappin/ x  
**what do I do with the grapple x**  
 lu\_ga \*screen->13




fig. #1

02 (0.5)  
 03 LUC \*t` sais tu::: \*  
**you know you**  
 lu\_ge \*opens hand with outflung arm\*

04 (0.5)  
 05 LUC comme \*spiderman là \*  
**like spiderman here**  
 lu\_ge \*opens hand with outflung arm\*

06 GRE ouais et alors  
**yeah so what**  
 07 (0.2)  
 08 GRE et après  
**and then**  
 09 (1.1)  
 10 LUC ah \*bah là tire// \*  
**ah well here shoot**  
 lu\_ge \*pulls hand towards herself\*

11 (0.6)  
 12 GRE et là je tire/  
**and here I shoot**  
 13 LUC \*c'est pas b/ \*ouais ouais tire/  
**it's not b yeah yeah shoot**  
 lu\_ga \*controller \*screen->19

14 (1.2)  
 15 LUC \*t` as a- ou: rappui:e sur b/\* j` sais pas/  
**you have p- or press again on b I don't know**  
 lu\_ge \*pulls hand towards herself \*  
 16 (0.5)  
 17 LUC i` t'a pas dit d` rappuyer sur un truc/  
**it didn't tell you to press again on something**  
 18 (0.6)  
 19 LUC \*en fait c'est pour euh \*faire baisser le pont-\*levis là\\*  
**in fact it's to uh let down the drawbridge here**  
 lu\_ge \*pulls hand towards herself\*lowers hand with open palm \*  
 lu-ga \*controller

20 \*(1.2)  
 lu\_ga \*screen

<sup>9</sup> We adopted the multimodal transcription convention from Lorenza Mondada (2018) and from Groupe ICOR (2013).

21 LUC **\*\*la flèche là/  
the arrow here**  
lu-ga \*controller  
lu-ge \*touches button on controller->23

22 **\*(2.4)**  
lu-ga \*screen->25

23 LUC et c'est sur quoi qu` t` as fait le grappin/  
**and on what did you put the grapple**  
(omitted transcript lines not mentioned in the analysis)

38 LUC ça doit être un autre \*bouton\ ou l` \*bouton d` derrière nan/  
**it must be another button or the button behind no**  
lu-ga \*controller  
lu-ge \*touches controller

39 y a pas un \*bouton \*derrière/  
**isn't there a button behind**  
lu-ga \*screen  
lu-ge ->\*

40 **\*(0.3)**  
lu-ga \*controller

41 GRE **\*atten- -ten -ten \*-tion \***  
**atten- -ten -ten -tion**  
lu-ge \*touches behind controller\*  
lu-ga \*screen

42 **(1.2)**

43 **@(0.4) @ (3.0) @# (0.3)**  
av\_ge @launches grapple @pulls grapple  
gr\_ge @searches the button for shooting



*fig.*

44 LUC voi:\*[là\ t` as fait] quoi/  
**that's it what did you do**

45 GRE [ah d'accord ]  
**ah okay**  
lu-ga \*controller

46 GRE (merci) munchuck  
**(thank you) munchuck**  
47 **(0.3) \*(0.6) \*(0.3)**  
lu-ga \*screen \*tablet

48 LUC munchuck  
**munchuck**

The extract starts with Greg's request for help, formulated as a question and designed for a co-orienting participant (l. 1): despite a longer phase of silent gaming, neither the turn gives evidence for re-opening or attention-getting, nor does the speaker situate or explain his problem. Lucie, though, gazes at her tablet, and only at the end of Greg's turn, she does withdraw her gaze from the tablet and shifts it to the gaming screen. In other words, (1) Greg

displays that he considers the situation as "incipient talk" (Schegloff & Sacks, 1973, p. 325) where openings or closings are not necessary, (2) he shows that he takes for granted that Lucie is available for and willing to offer assistance and (3) he indicates that he takes for granted that Lucie sufficiently follows his gaming so that she understands what he is talking about. Greg's suppositions concerning Lucie's availability and attention are only possible if we consider a sit-

uation where both participants are “together”, even if they are principally not engaged in the same activity, and if Greg has taken into account Lucie’s preceding gaze shifts and considered them as at least minimal engagement.

After Lucie’s embodied demonstration (l. 3–5; Keevallik, 2014), Greg specifies his problem and thus confirms his request: he needs help to find out what he should do next (l. 8). This is the starting point for a longer collaborative searching sequence where player and non-player search together for a way to open the drawbridge (l. 10–48). The non-player participates in this searching through instructing the player, asking questions, and monitoring the gaming. Lucie thereby shows her knowledge of the game (e.g., “you need to press the B button to pull something”) and affirms her “epistemic status” (Heritage, 2012), allowing her to instruct the player. Greg mostly accomplishes the instructed actions silently. While the instructions show and insist on the speaker’s knowledge (K+, or more knowledgeable) and project an accomplishment of the instructed action, questions (as in l. 17) initiate a verbal exchange with the player.

All questions in this extract (except the first one by Greg, l. 1) are asked by the non-player Lucie. They occur when the instructed action could not be realized or when the targeted gaming action does not show effects. They bring in a new aspect in order to resolve the problem: an eventual specific instruction as part of the game (l. 17), the target of the grappling iron (l. 23), another still unidentified button (l. 38–39). In all cases, Lucie manifests a higher epistemic status by proposing possible solutions, but the action of “asking” displays less certainty than “instructing”, and the turn design mostly foregrounds technical or game design-related aspects that put the player’s agency into

the background. The non-player’s monitoring of the game as well as of the player’s physical actions on the controllers allows offering temporally finely tuned assistance. The non-player follows the player’s gaming actions to “gather information” (Stukenbrock, 2020, p. 20): The monitoring starts after Greg’s request for assistance where the avatar still does not move on. When Greg insists that he is accomplishing the instructions (l. 12), but still nothing happens in the game, Lucie suggests that the problem is related to the manipulation of the controller. This suggestion is accompanied by a short withdrawal of gaze, to the controller, allowing Lucie to see which button Greg actually activates. She then focuses again on the screen where the avatar still does not move. Lucie adjusts her assistance with another proposal (l. 17–19). But again, her suggestions do not resolve the problem; Lucie gazes once more to the controller (l. 19) and to the screen (l. 20), silently watching the consequences of Greg’s technical manipulations. These observations lead to a new instruction, an instruction that she accomplishes herself (l. 21), gazing first at the controller and then at the screen to monitor the success or failure of her action.

The physical intervention of the non-player constitutes a step further: the collaboration of players and non-players has been considered as offering assistance through verbal or embodied demonstrations, instructions, questions, etc. (Olbertz-Siitonen et al., 2021), but not as sensorial engagement through touching and manipulating the controller and acting on their own in the game. This sensorial engagement is interesting insofar as it enlarges the collaboration from the non-player’s side from simple noticing to interfering in the game and momentarily inverts the participation framework: the non-player plays and the player observes. The player’s acceptance of this intervention (l. 46)

ratifies this local change in participation frameworks as well as the shared moral order. The following extract (Extract 3) will nevertheless show that this acceptance has its limits.

### Offering assistance and seeking for understanding

In the third example, the non-player offers assistance. Offering assistance without any

request implies not only that the non-player “watches” the game but also “sees” it as a potential player. The timeline (Figure 6) shows that Lucie maintains her gaze to the gaming screen. Both participants produce turns, and the turns are approximately equal in length. This is remarkable with regard to Extract 2 where nearly all turns are produced by the non-player.

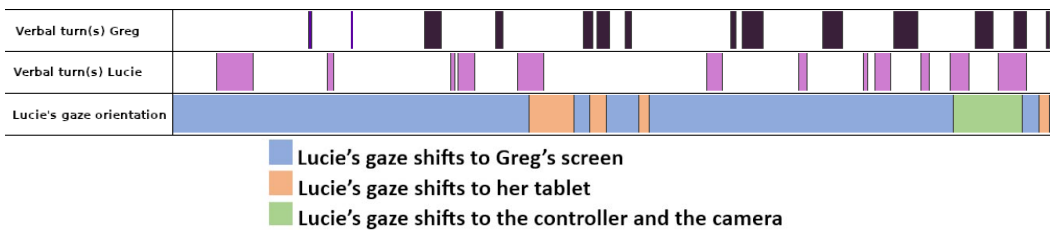


Figure 6. Timeline 3 (00:11:09–00:12:29)

Lucie's turns are followed by rather long periods of silence and/or different interventions by Greg. Looking more closely at the transcript (Extract 3), we will see how Lucie engages actively in the gaming, despite the moments of

silence. Our analysis focuses especially on the sequence organisation as well as on the local negotiation of a shared moral order with regard to the non-player's sensorial engagement and interfering with the player's gaming.

### Extract 3 (00:11:09–00:12:29)

```

01      *@(3.3)           @
    lu_ga *screen->18
    av_ge @jumps on a rock@
02      *(0.3)
    lu_ge *hand gesture from up to down*
03 LUC  t` es comme ça/* t- t-           *#t` es un peu/ en vue/*
    you are like this y- y- you are you're a little bit visible
    lu_ge           *hand gesture up *hand gesture up down up*
    scr           #shift camera perspective
04      *comme ça/ ah:*
    like that
    lu_ge *hand gesture *
05      (4.5)
06 GRE  .h
07      (1.2)
08 LUC  *uhm::\           *
    lu_ge *points to screen*
09      *#(6.3)           #
    lu_ge *head upwards->14
    scr   #camera perspective turns horizontally#

```



((omitted transcript lines not mentioned in the analysis))  
 18 LUC m\*ais tu viens \*d'où/ \*j` l'ai pas vu/ en fait j` suivais pas  
**but where do you come from/ \*I didn't see it/ in fact I  
 wasn't following**  
 lu\_ge \*hand gesture \*  
 lu\_ga \*tablet  
 19 (2.3) # (0.6) \* (0.3)  
 scr #shift camera perspective->22  
 lu\_ga \*screen->20  
 20 GRE (inaud.)\*  
 lu\_ga \*tablet->22  
 21 (0.3)  
 22 GRE #moi je viens d` là normalement/\*  
**I come from there normally**  
 lu\_ga \*screen->25



fig.

#3

23 (1.2)  
 24 GRE #je viens d` là/  
**I come from there**



fig.

#4

25 (0.5) \* (1.0) \* @ (4.6)  
 lu\_ga ->\*tablet \*screen ->43  
 av\_ge @jumps down and hangs on with hands  
 26 LUC tu viens d'en bas:/ ou tu viens d'en haut\@  
**you come from downstairs or you come from upstairs**  
 av\_ge ->@  
 27 (0.6)  
 28 GRE j` viens d'en bas\  
**I come from downstairs**  
 29 (0.5)  
 30 GRE ah ouais (inaud.)  
**ah yeah (inaud.)**  
 31 (1.1) @ (1.8)  
 av\_ge @jumps on the rock by doing a salto  
 32 LUC wa:@o:\  
**wow**  
 av\_ge ->@  
 33 (1.3)  
 34 GRE c'est balaise hein main`enant j` fais #ça/  
**that's cool isn't it now I'm doing this/** #  
 scr @camera perspective  
 turns horizontally #

35 (0.6) @ (0.1) \* (0.9)  
 av\_ge @jumps down  
 lu\_ge \*points to screen->38

36 LUC tiens/  
**look**  
 (0.6)

37 AH:: \*t` es passé \*@au-dessus/#  
**OH you did go above it**  
 lu\_ge ->\*hand gesture up\*  
 av\_ge @runs to the door  
 scr #instructional window

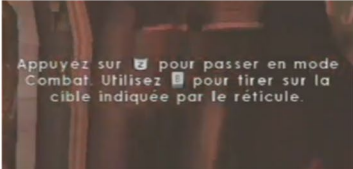


fig. #5

39 GRE a- appuyer sur (inaud.)  
**p- press on (inaud.)**  
 (0.3)

41 LUC \*su:r/ \*  
**on**  
 luc\_ge \*body to the right ->50  
 luc\_mi \*raises eyebrows\*

42 (0.2) @ (0.2) @ (1.2)  
 gr\_ge @press controller@

43 LUC c'est\* z c'est derr\*ière non/ \*  
**it's z it's behind isn't it**  
 lu\_ga \*controller Greg->50  
 lu\_ge \*touches controller\*

44 \*(0.5)  
 lu\_ge \*lifts up controller->

45 GRE on peut pas appu\*yer sur le sur \*@z @  
**you can't press on the on z**  
 lu\_ge ->\* \*lifts up second controller  
 gr\_ga @controller@

46 (0.4)

47 LUC @mais z @ c'est ce que je te  
**but z that's what I'm**  
 av\_ge @move towards the door@

48 cher[che\ c'est ça/\* Z#] \*  
**looking for that's it z**  
 lu\_ge ->\*presses controller\*  
 scr #weapon activated  
 lu\_ga \*screen->51

49 GRE [nan mais (inaud.)]  
**no but (inaud.)**

50 (1.0) \*(0.6)  
 lu\_ga \*tablet

51 GRE (inaud.)  
 52 @ (8.9)  
 av\_ge @takes weapon and opens door

When Greg makes the avatar jump on a rock, the view of the gaming space changes (l. 1): the player now only sees the upper part of the hall where the avatar is progressing and from this position the way out is unclear. For Lucie, this change of perspective is even more inconvenient since she has been gazing at the tablet before this extract and therefore did not follow the avatar's movements.

Lucie starts gazing at the gaming screen when the avatar jumps on the rock – maybe alerted by the increasingly tense music. She first displays the problem of the camera perspective through embodied demonstrations, indicating the direction in which Greg needs to shift the camera according to her (l. 2–9). With these demonstrations, she draws attention to a potential trouble source which she has identified and which could be related to Greg's problems progressing in the game: shifting the camera perspective could allow them to get a better overview of the space and therefore to find the exit. She uses hand and head gestures and always keeps her tablet in her hands, showing that she might assist but not give up her activity (“dual involvement”).

Despite Greg's shifts of the camera perspective, the avatar does not find the exit and for Lucie, the problem still remains the same. She then initiates three repair sequences, partially in overlap with the demonstrations, asking where the avatar comes from (l. 10–16, 18–24 and 24–28). While Greg shows precise knowledge concerning the avatar's movements, Lucie seems to be lost. This is an additional hint that Greg's camera shifts are his own strategy of searching for clues of an eventual exit and do not indicate an orientation problem for himself. Only when the avatar makes a back flip and jumps again (l. 31), which leads to a new change of the camera perspective (l. 34), Lucie does display un-

derstanding, indicated by the change-of-state token “ah” and the formulation of her understanding (l. 38).

In the previous segment (Extract 2), Lucie commented on the game or on gaming actions according to her understanding, and she offered help once she was asked to do so. Both participants displayed the same “seeing” of the game and the occurring problems, allowing the non-player to rapidly provide assistance and instruct the player. In Extract 3, the non-player provides self-initiated assistance based on her understanding. But rapidly she realises that her understanding has failed and she tries to understand by “seeing with the player's eyes”, through repair sequences and silent monitoring. What to see is “embedded in the local history of the activity in progress” but this local history needs to be “activated through specifically designed talk” (Nishizaka, 2000, p. 111). To achieve joint seeing, both participants need to explicitly collaborate, and what started to be a sequence of offering help becomes a joint construction of understanding.

After Greg's answer to Lucie's last question (l. 28), the trouble seems to be repaired and Lucie's understanding problem seems to be resolved (l. 38). Both can now fully engage in the avatar's search for the exit, which allows it to move on. An instructional pop-up-window appears on the screen and they read parts of the written text (l. 39, 41) which indicates the technical manipulations the player has to accomplish. From the beginning of the reading, Lucie positions herself as “co-player” (Olbertz-Siitonen et al., 2021) and displays a sensorial participation in the gaming through touching and manipulating: she orients her body towards Greg (l. 41), she gazes at the first controller and then touches it (l. 43), she lifts up both controllers (l. 44–45), etc. These embodied actions show that she

closely monitors what Greg is doing (or not). As in the previous extract, she intervenes on her own initiative, interpreting Greg's absence of successful gaming actions as display of trouble and recruitment of assistance: he only starts in line 42 to press a button on the controller and since nothing happens in the game, it was apparently not the correct one.

But in contrast to Extract 2, Greg does not affiliate with this type of joint collaboration; he explicitly contests Lucie's proposals (l. 45) as well as her interfering in the gaming (l. 49).

By touching the controllers (l. 43, 45) and finally pressing the "Z" button (l. 48), Lucie meshes with the technological device as part of the player's space (Meyer & Jucker, 2022) and challenges "the active player's primary rights to perform and make decisions about gameplay actions" (Olbertz-Siitonen et al., 2021, p. 116). The taken-for-granted agreement concerning a shared moral order - especially the rights and limits of non-playing participants - is challenged and needs to be re-established before the gaming can continue. Whereas the non-player's intervention in Extract 2 was temporally and gesturally very limited - Lucie only touched the controller once to show the arrow button - and accompanied by a verbal explanation allowing the player to accomplish the action himself, the participation in Extract 3 is longer, more complex and culminates in the non-player's physical accomplishment of a gaming action. Even if Lucie finally activates the weapon which allows her to open the door and to progress in the game, her sensorial engagement seems to go too far. The player re-orientes to her status as non-player, indicating that player and non-player collaborate in the joint activity through different practices. The non-player aligns with this disapproval through her follow-up actions: once the button and the weapon activated, she shows

disengagement in Greg's gaming by gazing back at her tablet (l. 50).

## Conclusion

In the data we analysed in this study, the two participants are engaged in two different activities: Greg plays a single person videogame on the large TV screen, Lucie is acting on her tablet. Whereas Greg never displays engagement in Lucie's activity, Lucie takes more or less part in Greg's gaming. We therefore consider the gaming as the principal activity and we asked how both participants collaborate and co-construct the non-player's engagement in the gaming activity. This question is related to our interest in a broader understanding of technology use in everyday situations where participants are engaged more or less simultaneously in different activities.

In our data, the non-player's engagement in the gaming is closely related to her other activity: she is not always available for interaction with the player and shares her focus between the two screens. The non-player thus constantly transits between different activities, pausing one activity to engage in another or even following the activities simultaneously. Our analysis has shown that gaze orientation plays a central role in the organization of these transitions and the construction of engagement in the other's activity and availability for interaction. This observation echoes Kamunen and Haddington's (2020, p. 82) observation on the role of "gaze behaviour [...] to project an imminent activity transition".

The coding of Lucie's gaze during the whole session points to phases of clear engagement in or a disengagement from Greg's gaming as well as phases where the two activities compete and

the gaze constantly shifts between the gaming screen and the tablet. This alternation between engagement/disengagement not only allows her to follow two activities simultaneously over a longer time span, but also constructs togetherness by regularly signalling potential availability for interaction. The focus on the 45 minutes of gaming highlights the global trajectory of engagement during the whole session. The identification of the development of engagement over a larger timescale allows identifying fluctuations, withdraws and local practices of displaying availability for interaction and engagement in the gaming as “doing togetherness”.

The detailed analysis of a first, mostly silent segment (Timeline and Extract 1) shows how the non-player’s gaze to the gaming screen and the gaze duration display “noticing” (Kidwell & Reynolds, 2022) and a potential availability to interact and participate in a joint activity that we describe as “minimal engagement”. The non-player uses gaze as a resource to switch between two activities and two participation frameworks. The player eventually notices the gaze as a display of the non-player’s availability for interaction.

The sequential and multimodal analysis as well as the analysis of the timeline of two other segments (Extracts 2 and 3) detail the temporal unfolding of the non-player’s engaging in the gaming activity as co-construction of both participants. We focused on two practices: the player recruiting the non-player’s help and the non-player offering assistance on her own. In both cases, gaze duration and gaze shifts play an important role. The non-player follows the player’s gaming through longer gazing at the gaming screen and gaze shifts between the screen and the controllers to gather information on the actions in the game as well as the accomplished technical manipulations, similar

to what Stukenbrock (2020) describes for “gaze following”.

The two extracts show a different sequence organisation: while in Extract 2 we mostly find embodied demonstrations, instructions, proposals or questions from the non-player, nearly without any verbal turns from the player, Extract 3 is characterised by verbal turns from both participants. The analysis revealed that this is not only due to the sequence initiation (request for help vs. self-initiated offer to help), but also to the problems that did occur, more precisely the collaboration in the non-player’s search for understanding and the negotiation of the non-player’s limits concerning the participation in the gaming. Further research is needed to investigate the co-construction of engagement when the participants are involved in different activities and, more precisely, to observe how different types of sequences are initiated and which resources the participants use to engage in the other participants’ activities.

The negotiation of limits in term of participant roles is particularly interesting with regard to technology use. It illustrates that non-players can collaborate even physically in the gaming as joint activity, in other words they can show sensorial engagement – but only up to a certain point. On the one hand, this supports our claim that “player” and “non-player” are not clearly separated categories but need to be co-constructed locally through specific practices of playing and spectating/watching/co-playing. On the other hand, the player’s follow-up actions in Extract 3 display that the non-player’s interventions are considered as intrusions and the non-player, additionally, retracts her engagement in the gaming thereafter. This points to a complementary distribution of the players’ and non-players’ rights and obligations that eventually need to be renegotiated locally. A

shared moral order is not only a display of engagement; the negotiation of moral norms also participates in the co-construction of engagement.<sup>10</sup>

## References

- Aarsand, P. A., & Aronsson, K. (2009). Gaming and territorial negotiations in family life. *Childhood, 16*(4), 497–517. <https://doi.org/10.1177/0907568209343879>
- Baldauf-Quilliatre, H., & Colón de Carvajal, I. (2021). Spectating. How non-players participate in videogaming. *Journal für Medienlinguistik, 4*(2), 123–161. <https://doi.org/10.21248/jfml.2021.33>
- Cekaite, A., & Kvist Holm, M. (2017). The Comforting Touch: Tactile intimacy and talk in managing children's distress? *Research on Language and Social Interaction, 50*(2), 109–127. <https://doi.org/10.1080/08351813.2017.1301293>
- De Stefani, E., & Mondada, L. (2018). Encounters in public space: How acquainted versus unacquainted persons establish social and spatial arrangements. *Research on Language and Social Interaction, 51*(3), 248–270. <https://doi.org/10.1080/08351813.2018.1485230>
- Gehle, R., Pitsch, K., Dankert, T., & Wrede, S. (2017). How to open an interaction between robot and museum visitor? Strategies to establish a focused encounter in HRI. *Proceedings of the 2017 ACM/IEEE International Conference on Human Robot Interaction, 187–195*. <https://doi.org/10.1145/2909824.3020219>
- Goffman, E. (1957). Alienation from Interaction. *Human Relations, 10*(1), 47–60. <https://doi.org/10.1177/2F001872675701000103>
- Goffman, E. (1967). *Interaction ritual: Essays on face-to-face interaction*. Aldine Publishing Company. <https://doi.org/10.4324/9780203788387>
- Goodwin, M. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry, 50*(3–4), 303–317. <https://doi.org/10.1111/j.1475-682X.1980.tb00024.x>
- Goodwin, C. (1981). *Conversation organization: Interaction between speakers and hearers*. Academic Press.
- Goodwin, C., & Goodwin, M. H. (1996). Seeing as a situated activity: Formulating planes. In Y. Engeström & D. Middleton (Eds.), *Cognition and communication at work* (pp. 61–95). Cambridge.
- Goodwin, C., & Goodwin, M. H. (2004). Participation. In A. Duranti (Ed.), *A companion to linguistic anthropology* (pp. 222–244). Blackwell. <https://doi.org/10.1002/9780470996522.ch10>
- Groupe ICOR (2013). Conventions ICOR. Laboratoire ICAR, UMR 5191, accessed on 12/20/2021. <http://icar.cnrs.fr/corinte/conventions-de-transcription/>
- Haddington, P., Keisanen, T., Mondada, L., & Nevile, M. (2014). Towards multiactivity as a social and interactional phenomenon. In P. Haddington, T. Keisanen, L. Mondada, & M. Nevile (Eds.), *Multiactivity in Social Interaction* (pp. 3–32). Benjamins. <https://doi.org/10.1075/z.187.01had>
- Hausendorf, H. (2013). On the interactive achievement of space - and its possible meanings. In P. Auer, M. Hilpert, A. Stukenbrock, & B. Szmrecsanyi (Eds.), *Space in language and linguistics: Geographical, interactional and cognitive perspectives* (pp. 276–303). de Gruyter. <https://doi.org/10.1515/9783110312027.276>
- Heath, C. C. (1982). The display of reciprocity: An instance of a sequential relationship in speech and body movement. *Semiotica, 42*(2–4), 147–168. <https://doi.org/10.1515/semi.1982.42.2-4.147>
- Heath, C., & Luff, P. (1996). Convergent activities: Line control and passenger information on the London Underground. In Y. Engeström & D. Middleton (Eds.), *Cognition and communication at work* (pp. 96–129). Cambridge.
- Heritage, J. (2012). The epistemic engine: Sequence organization and territories of knowledge. *Research on Language & Social Interaction, 45*(1), 30–52. <https://doi.org/10.1080/08351813.2012.646685>
- Holly, W., & Baldauf, H. (2002). Grundlagen des fernsehbegleitenden Sprechens. In W. Holly, U. Püschel, & J. R. Bergmann (Eds.), *Der sprechende Zuschauer. Wie wir uns Fernsehen kommunikativ aneignen* (pp. 41–60). Westdeutscher Verlag.
- Kamunen, A., & Haddington, A. (2020). From monitoring to co-monitoring: Projecting and prompting activity transitions at the workplace. *Gesprächsforschung - Online Zeitschrift zur verbalen Interaktion, 21*, 82–122. <http://www.gespraechsforschung-online.de/fileadmin/dateien/heft2020/ga-kamunen.pdf>
- Keevallik, L. (2014). Turn organization and bodily-vocal demonstrations. *Journal of Pragmatics, 65*, 103–120. <https://doi.org/10.1016/j.pragma.2014.01.008>

<sup>10</sup> The authors are grateful to the ASLAN project (ANR-10-LABX-0081) of the Université de Lyon, for its financial support within the French program “Investments for the Future” operated by the National Research Agency (ANR) as well as to the editors of the special issues as well as to two anonymous reviewers for insightful comments on previous versions.

- Kendrick, K. H., & Drew, P. (2016). Recruitment: Offers, requests, and the organization of assistance in interaction. *Research on Language and Social Interaction*, 49(1), 1–19. <https://doi.org/10.1080/08351813.2016.1126436>
- Kidwell, M., & Zimmerman, D. H. (2007). Joint attention as action. *Journal of pragmatics*, 39(3), 592–611. <https://doi.org/10.1016/j.pragma.2006.07.012>
- Kidwell, M., & Reynolds, E. (2022). Gaze and the organization of participation in collective visual conduct. *Social Interaction: Video-Based Studies of Human Sociality*, 5(2). <https://doi.org/10.7146/si.v5i2.119332>
- Kohonen-Aho, L., & Vatanen, A. (2021). (Re-) Opening an encounter in the virtual world of Second Life: On types of joint presence in avatar interaction. *Journal für Medienlinguistik*, 4(2), 14–51. <https://doi.org/10.21248/jfml.2021.30>
- Kremer-Sadlik, T., Fatigante, M., & Fasulo, A. (2008). Discourses on family time: The cultural interpretation of family togetherness in Los Angeles and Rome. *Ethos*, 36(3), 283–309. <https://doi.org/10.1111/j.1548-1352.2008.00015.x>
- Luff, P., & Heath, C. (2001). Naturalistic analysis of control room activities. In J. Noyes & M. Bransby (Eds.), *People in Control. Human factors in control room design* (pp. 151–165). Institute of Engineering and Technology. [https://doi.org/10.1049/PBCE060E\\_ch10](https://doi.org/10.1049/PBCE060E_ch10)
- Merleau-Ponty, M. (1960). *Signes*. Gallimard.
- Meyer, N., & Jucker, A. H. (2022). Co-presence and beyond: Spatial configurations of communication in virtual environments. In A. H. Jucker & H. Hausendorf (Eds.), *Pragmatics of Space* (pp. 579–608). De Gruyter. <https://doi.org/10.1515/9783110693713-018>
- Mondada, L. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of pragmatics*, 41(10), 1977–1997. <https://doi.org/10.1016/j.pragma.2008.09.019>
- Mondada, L. (2012). Coordinating action and talk-in-interaction in and out of video games. In R. Ayass & C. Gerhardt (Eds.), *The appropriation of media in everyday life* (pp. 231–270). Benjamins.
- Mondada, L. (2018). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Mondada, L. (2022). Appealing to the senses: Approaching, sensing and interacting at the market's stall. *Discourse & Communication*, 16(2), pp. 160–199.
- Nishizaka, A. (2000). Seeing what one sees: Perception, emotion, and activity. *Mind, culture, and activity*, 7(1–2), 105–123. <https://doi.org/10.1080/10749039.2000.9677650>
- Nishiaki, A. (2014). Sustained orientation to one activity in multiactivity during prenatal ultrasound examinations. In P. Haddington, T. Keisanen, L. Mondada, & M. Nevile (Eds.) *Multiactivity in Social Interaction* (pp. 79–107). Benjamins. <https://doi.org/10.1075/z.187.03nis>
- Olbertz-Siitonen, M., Piirainen-Marsh, A., & Siitonen, M. (2021). Constructing co-presence through shared VR gameplay. *Journal für Medienlinguistik*, 4(2), 85–122. <https://doi.org/10.21248/jfml.2021.31>
- Peräkylä, A., Voutilainen, L., Lehtinen, M., & Wuolijoki, M. (2022). From engagement to disengagement in a psychiatric assessment process. *Symbolic Interaction*, 45(2), 257–296. <https://doi.org/10.1002/symb.574>
- Piirainen-Marsh, A. (2010). Bilingual practices and the social organisation of video gaming activities. *Journal of Pragmatics*, 42(11), 3012–3030. <https://doi.org/10.1016/j.pragma.2010.04.020>
- Raymond, G., & Lerner, G. H. (2014). A body and its involvements. In P. Haddington, T. Keisanen, L. Mondada, & M. Nevile (Eds.), *Multiactivity in Social Interaction* (pp. 227–246). Benjamins.
- Reeves, S., Greiffenhagen, C., & Laurier, E. (2017). Video gaming as practical accomplishment: Ethnomethodology, conversation analysis, and play. *Topics in Cognitive Science*, 9(2), 308–342. <https://doi.org/10.1111/tops.12234>
- Robinson, J. D. (2020). Revisiting preference organization in context: A qualitative and quantitative examination of responses to information seeking. *Research on Language and Social Interaction*, 53(2), 197–222. <https://doi.org/10.1080/08351813.2020.1739398>
- Schegloff, E., & Sacks, H. (1973). Opening up closings. *Semiotica*, 8(4), 289–327. <https://doi.org/10.1515/semi.1973.8.4.289>
- Sidnell, J., & Stivers, T. (Eds.). (2013). *The handbook of conversation analysis*. Wiley & Sons. <https://doi.org/10.1002/9781118325001>
- Stivers, T. (2015). Coding social interaction: A heretical approach in conversation analysis? *Research on Language and Social Interaction*, 48(1), 1–19. <https://doi.org/10.1080/08351813.2015.993837>

Stukenbrock, A. (2020). Deixis, meta-perceptive gaze practices, and the interactional achievement of joint attention. *Frontiers in psychology, 11*, 1779, 1–23. <https://doi.org/10.3389/fpsyg.2020.01779>

Tekin, B. S. (2021). Quasi-instructions: Orienting to the projectable trajectories of imminent bodily movements with instruction-like utterances. *Journal of Pragmatics, 186*, 341–357. <https://doi.org/10.1016/j.pragma.2021.10.018>

Tekin, B. S., & Reeves, S. (2017). Ways of spectating: Unravelling spectator participation in kinect play. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1558–1570. <https://doi.org/10.1145/3025453.3025813>

Voida, A., & Greenberg, S. (2009). Wii all play: The console game as a computational meeting place. *Proceedings of the SIGCHI conference on human factors in computing systems*, 1559–1568. <https://doi.org/10.1145/1518701.1518940>

vom Lehn, D. (2013). Withdrawing from exhibits: The interactional organisation of museum visits. In P. Haddington, L. Mondada, & M. Nevile (Eds.), *Interaction and mobility: Language and the body in motion* (pp. 65–90). De Gruyter. <https://doi.org/10.1515/9783110291278.65>

Weilenmann, A. (2003). “I Can’t talk now, I’m in a fitting room”: Formulating availability and location in mobile-phone conversations. *Environment and Planning A: Economy and Space, 35*(9), 1589–1605. <https://doi.org/10.1068/a34234>

---

OTSIKKO JA AVAINSANAT SUOMEKSI:

**Ei-pelaajien keholliset videopelihin osallistumisen käytännöt**

KEYWORDS: ei-pelaaja, keholliset käytännöt, multimodaalinen keskustelunanalyysi, seuraaminen, sitoutuminen, videopelit



## Appendix 1. ICOR Convention (Groupe ICOR 2013)

<http://icar.cnrs.fr/corinte/conventions-de-transcription/>

### Text in bold: translation

Text in grey: information concerning events on the screen (scr), avatars' (ava) or players' (player's pseudonym) actions

[ ]	Overlapping talk
/ \	Rising or falling intonation
°°	Lower voice
:::	Lengthening of the sound or the syllable
p`tit	Elision
trouv-	Truncation
xxx	Incomprehensible syllable
=	Latching
( )	Uncertain transcription
(( ))	Comments
&	Turn of the same speaker interrupted by an overlap
(.)	Micro-pause
(0.6)	Timed pause

## Appendix 2. Multimodal convention (Mondada 2018)

<https://www.lorenzamondada.net/multimodal-transcription>

*	Gestures and descriptions of embodied actions are delimited between two identical symbols (one symbol per participant) and are synchronized with corresponding stretches of talk
#	Screen events, is indicated with a specific symbol showing its position within the turn at talk
->	The action described continues across subsequent lines