

1 **Co-Constructing Tele-Presence by Embodying Avatars:**
2 **Evidence from Let's Play Videos**

3 *Axel Schmidt & Konstanze Marx*

4 **1 Introduction**

5 The research object of our contribution are so-called *Let's*
6 *Plays*, in which video games are played for an internet
7 audience (cf. Hale 2013; Ackermann 2016a). To make
8 watching a video game attractive for viewers, the gamer(s)
9 usually produce verbal as well as embodied comments on
10 their game. In addition, their face often appears in an extra
11 video embedded within the feed of the video game (a so-
12 called facecam) (see section 2, fig. 1). This allows spectators to
13 have access to the verbal and embodied reactions of the
14 players during the game.

15 Essentially, gamers take on an additional role of moderator
16 not only playing but also mediating their game play activities
17 for an audience. By doing this they try to make their game
18 play 'watchable' (cf. Schmidt/Marx 2020). One crucial aspect
19 that enhances the pleasure of watching a video game is to
20 make moves in the game more transparent and
21 understandable for viewers. For this purpose, gamers verbally
22 formulate their game moves (e.g. now I knock at a door). In
23 addition, they produce exclamations or following Goffman
24 (1981b) response cries during or after their actions like oh if
25 they are surprised or ahhh if they are shocked or frightened
26 by game events. In this way they animate their avatars. We
27 understand such practices of either formulating actions or

28 animating avatars as an embodiment of avatars by gamers for
29 viewers.

30 By doing so, gamers construct a specific kind of
31 participation framework (Goffman 1981a; Goodwin/Goodwin
32 2004) consisting of a) a human-machine-interaction (playing
33 the game) which results b) in represented virtual activities
34 (avatars 'doing' something within the game) which are c)
35 presented for viewers. By animating their avatars and making
36 their game moves more understandable by verbally
37 formulating their actions, gamers co-construct (tele-)presence
38 as they allow the spectators to participate in their immediate
39 experience.

40 Our contribution focuses on the types of practices gamers
41 employ to embody their avatars in video games for an
42 internet audience. We investigate two main practices of Let's
43 Players, *formulating actions* and *animating avatars via*
44 *response cries*, which interact in the embodiment of avatars.

45 2. Computer Games and Let's Plays

46 Our paper aims to contribute to a growing body of studies
47 dealing with computer gaming from an interaction-
48 theoretical perspective (cf. summarizing
49 Reeves/Greiffenhagen/Laurier 2016). Most studies emphasize
50 that playing a computer game extends interaction to a virtual
51 world and, by that, creates a different and – in the case of
52 several players – a more complex *participation framework*
53 (cf. Keating/Sunakawa 2010; Laurier/Reeves 2014; Mondada
54 2012; Piirainen-Marsh 2012; Tekin/Reeves 2017) as well as
55 different time layers (at least the time prescribed by the game
56 play and time in terms of interaction) which have to be
57 temporally coordinated (cf. Mondada 2013).

58 In the simplest case of gaming, which is playing alone, two
59 levels of (inter)action arise: First, there is an interaction
60 between the player and the computer, that is, a kind of
61 human-machine-interaction, whose base is a *control action*
62 by the player (e.g. moving the mouse). Secondly, this creates a
63 *represented (inter)action* within the game (e.g. an avatar's
64 knocking at a door or an avatar's fight with a non-play
65 character). We call this *game action*. Control actions happen
66 in the *real world*, game actions in the *game world*.

67 Computer games can be played alone or together. Playing
 68 together creates additional levels of interaction in the real
 69 world as well as in the game world. When several players
 70 control different avatars in a joint game, interactions between
 71 the players (either because they are co-present or because
 72 they are connected via technology) and/or between their
 73 avatars on screen may occur (cf. Mondada 2012; Baldauf-
 74 Quilliatre/Colón de Carvajal 2019; Marx/Schmidt 2019).
 75 However, in the following we focus mainly on single players
 76 who present their gaming activities to an internet audience.

77 In *Let's Plays*, video games are not only played but
 78 presented to an internet audience at the same time. This is
 79 usually done by commenting verbally on the game during
 80 playing via a facecam. Figure 1 shows the typical
 81 representation of *Let's Plays*.



82

83 **Figure 1:** Typical representation of a *Let's Play*

84 The verbal comments add an extra level of communication, a
 85 kind of para-social interaction (cf. Horton/Wohl 1956) or – as
 86 Ayaß (1993) has suggested renaming the term – social para-
 87 interaction. This has two effects: First, gamers produce talk
 88 and embodied conduct for an audience to which they have
 89 no direct access. Similar to mass media products such as
 90 television, direct address is frequently used to create an
 91 intimate interaction situation, even though there is no

92 immediate feedback from the audience, and so the audience-
93 centric talk is only based on assumptions about typical
94 viewers (cf. Ayaß 2005, Hausendorf 2001). Secondly, viewers
95 of *Let's Plays* are not able to participate, and importantly,
96 they are not able to influence the game. All they can do is
97 watch and listen to the comments of those who are actively
98 involved in playing. Interactivity and immersion are
99 considered to be computer games' most attractive features
100 (cf. Freyermuth 2015). Although the term *Let's Play* promises
101 a joint gaming experience, both of these features are lost for
102 viewers of *Let's Plays*, as joining a *Let's Play* merely means to
103 be in the role of a spectator (and, therefore, 'Watch Me Play'
104 may actually be a more appropriate name, as suggested by
105 Ackermann (2016a)). As our examples will show, this *inter-*
106 *passivity* or *de-interactivization* (cf. Pfaller 2008 and
107 Ackermann 2016b) is faced by the *Let's Players* by a kind of
108 *interactivity by proxy* (cf. Ligman 2011). This means that
109 players convey their own interactive immersion by letting
110 viewers participate in their immediate experiences. A
111 prevalent technique to achieve this is by various practices of
112 *embodying avatars*.

113 The first *Let's Plays* were created in the year 2006 on the
114 platform *Something Awful*.¹ These videos are still very
115 popular, ranking among the most popular videos on
116 YouTube, and representing almost the entire content of the
117 online live streaming platform *Twitch*.² The German Let's
118 Player *Gronkh*, for instance, has 4.8 million subscribers for
119 his YouTube-channel. *PerwDiePie*, a well-known global Let's
120 Player from Sweden, can even record 60 million.

121 There are two different ways in which *Let's Plays* can be
122 presented. One possibility is to record the gameplay as well
123 as the video feed including the facecam and upload this video
124 to websites like YouTube. Another possibility is to livestream
125 the gameplay and the video feed on special platforms like
126 *Twitch*. In this case, spectators usually have the possibility to
127 participate directly via live chat tools, creating additional
128 opportunities for interaction (cf. Schmidt/Marx/Neise 2020;
129 Recktenwald 2017). In the following we focus on recorded

1 See <https://www.somethingawful.com/>.

2 Twitch is a live streaming video platform that specializes on live streams of Let's Plays.

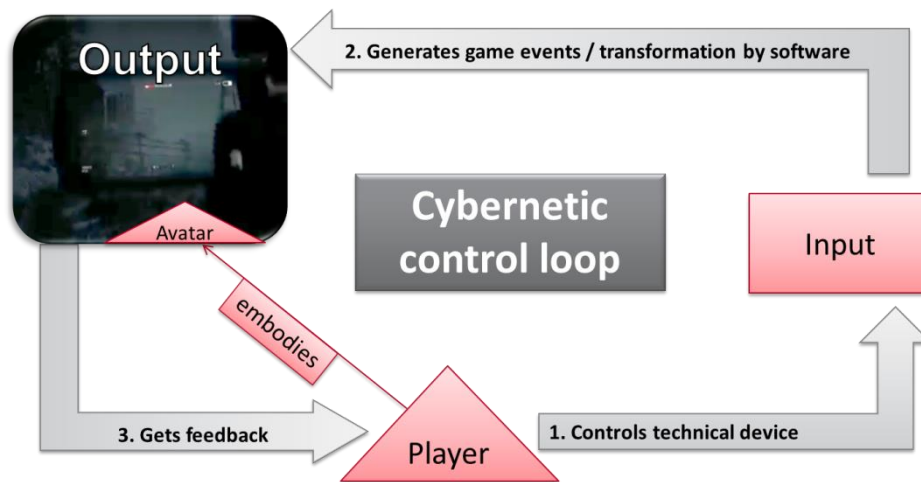
130 *Let's Plays*. In contrast to live streams, no interaction via chat
131 with an audience *during* gameplay is possible.

132 In the next section we introduce the notions of
133 embodiment and tele-presence and argue that they play a
134 crucial role when video games are presented to an audience.

135 3. Presented Gaming, Embodiment and (Tele-)Presence

136 As mentioned above, when computer gaming is presented for
137 an audience, a specific kind of *participation framework* arises
138 involving three levels: First, a human-machine interaction
139 (HMI) between gamer(s) and game software; second,
140 interactions between avatars within a represented virtual
141 reality (VR); and finally, social para-interaction (SPI) between
142 gamer(s) and viewers. The three levels are intertwined as the
143 first level (HMI) generates the second level (VR) which in turn
144 provides the content for the third level (SPI). It is this specific
145 relationship that creates the affordances to embody avatars
146 for viewers and, by that, to co-construct (tele-)presence. In
147 the following we would like to take a closer look at these
148 relationships.

149 Actions in computer games are usually mediated by
150 movements of avatars. Players control most of the
151 movements of their avatars via interfaces using technical
152 devices such as a game controller or a mouse and a keyboard.
153 The result of the gamers' control actions are displayed on
154 screen as movements of their avatars. Like this, players can
155 monitor the effect of their control actions immediately on
156 screen. Thus, players and avatars are connected to each other
157 by means of a cybernetic control loop in which inputs
158 generate immediate outputs that are, in turn, the basis for
159 further inputs as shown in Figure 2.



160

161 **Figure 2:** Gaming as cybernetic control loop

162 The close cycle between the player's control actions and the
 163 avatar's movements on screen forms the basis for an
 164 interactive and immersive game experience. When an avatar,
 165 for instance, is knocking at a door in the game world, this
 166 'action' is generated by a player's control action (e.g. by
 167 pressing a certain key on the keyboard). The avatar's action
 168 therefore has two sides: First, it is triggered by a technical
 169 operation outside the game world; second, it is at the same
 170 time a representation of an action in the game world (in this
 171 case a 'door-knocking-action') that is visually displayed on
 172 the screen. Consequently, immersion is understood as a
 173 (partial) occupation of our senses by the events within virtual
 174 reality (Lombard/Ditton 2006).

175 This connection between controlling the game (HMI) and
 176 representation of the game (VR) is presented in *Let's Plays* in
 177 a certain way for viewers. Only the latter, the represented
 178 action, is visualized in *Let's Plays* and thus accessible for
 179 viewers. The real world action of manipulating the controller
 180 remains invisible for spectators (cf. Schemer-Reinhard 2016).
 181 For spectators, the technical control action (pressing a key)
 182 and the representation of that action on screen (avatar
 183 knocking at a door) therefore merge into one action. As a
 184 result, the experience (and pleasure) of playing a computer
 185 game is emulated for viewers as the representation acts as if
 186 there is only one relevant level of action, namely the events
 187 in the virtual world. As we will see in the data, this is
 188 additionally enforced by *Let's Players* as they often comment
 189 on such actions with expressions such as *I'm going to knock*
 190 *at that door*, thereby referring to the avatar's action as their

191 own action. Using this kind of footing (cf. Goffman 1981a), the
192 identities of players and avatars seem to overlap. By saying
193 *I'm going to knock at that door*, players refer to both
194 themselves and their avatars controlled by them as agents of
195 the announced action. Such a construction of *player-avatar-*
196 *hybrids* (cf. Baldauf-Quilliatre/Colón de Carvajal 2015, 2019)
197 as agents in/of the game through the use of language enables
198 and enhances illusion and immersion for viewers. It creates
199 the impression that players are *directly active in the play*
200 *world*.

201 These affordances enable and constrain a specific kind of
202 participation framework which is, in turn, the basis of a
203 specific kind of embodiment of avatars conducted by gamers
204 for viewers. The notion of embodiment, roughly speaking,
205 emphasizes the reflexivity of cognition and situated behavior,
206 above all senso-motoric coordination (cf. Clark 2001; Gibbs
207 1995; Rohrer 2007; Suchman 1987; Wachsmuth/Lenzen/
208 Knoblich 2008). Most importantly, cognition is seen as rooted
209 in the body so that perception is only possible within in a
210 functional cycle of sensing, kinesthetic and movement (cf.
211 Lakoff/Johnson 2011; Streeck 2008; Streeck/Goodwin/
212 LeBaron 2011).

213 What occurs as a unit in real life is separated and
214 reconnected via media technology in computer games. Acting
215 in computer games is often mediated by *avatars* generating
216 *two poles of agency*, the gamer and the avatar. In addition,
217 gamers are involved in two situations simultaneously, the
218 situation of playing a game in the real world and the situation
219 of being in a game in the virtual world. For most gamers, this
220 creates a very enjoyable experience because they feel as if
221 they really are part of the virtual world when they play, and
222 so for them the real world can become temporarily replaced
223 by the virtual world. Players are perceptively and
224 psychologically immersed (cf. Lombard/Ditton 2006). When
225 gamers embody their avatars for viewers, they basically try to
226 convey their own experience of acting in a virtual world.
227 They reveal their thoughts and allow access to their
228 (emotional) experiences during their actions. As outlined
229 above, gamers and their avatars are closely connected. With
230 respect to agency, they merge into player-avatar-hybrids.
231 Taking this into account, access to their inner states
232 (thoughts, feelings etc.), which players grant viewers, will be

233 transferred to their avatars. By that, avatars are equipped
234 with human features (such as being sensitive, rational etc.).

235 Virtual reality in computer games, especially if presented in
236 a point-of-view-perspective, “provide media users with an
237 illusion that a mediated experience is not mediated”
238 (Lombard/Ditton 2006, 1) which “creates for the user a strong
239 sense of presence” (ibid.). In view of recent developments in
240 the field of media technology to convey interaction more and
241 more realistically, traditional notions of presence are
242 questioned (cf. Licoppe 2015; Spagnolli/Gamberini 2005).
243 Traditionally, presence was tied to the spatial concept of
244 situation, which means that two (or more) people are
245 required to be in the same place at the same time (cf.
246 Goffman 1963, Gumbrecht 2012). With the help of
247 technology, however, it is possible to create a realistic
248 representation of interaction partners who then appear to be
249 present or at least tele-present (cf. Höflich 2005, Meyrowitz
250 1990, Shanyang 2005). In addition, virtual worlds promote
251 immersive experiences that replace (or at least superimpose)
252 the perception of a real world by that of the virtual world
253 creating a strong sense of presence. When users consider
254 something mediated to be present, they tend to “respond
255 directly to what they see and hear in a mediated experience,
256 as if what they see and hear was physically present in their
257 viewing environment (...)” (Lombard/Ditton 1997: 10). This is
258 exactly what *Let’s Players* do to entertain their viewers: They
259 treat the virtual world and react to it as if it were physically
260 present. In the following we are interested in *Let’s Players*’
261 verbal and embodied practices of creating such a sense of
262 presence.

263 4. Data and Method

264 Our contribution follows an interactional perspective, or
265 more precisely, a multimodally extended EMCA approach,
266 which asks how participants create social reality. EMCA
267 stands for Ethnomethodology and Conversation Analysis (cf.
268 Heritage 1984a). By multimodal extension, we mean that the
269 focus is not only on talk, but also on embodied actions,
270 including the use of objects, media technology and space (cf.
271 Deppermann 2013, Mondada 2008, Streeck/Goodwin/
272 LeBaron 2011). Specifically, our analysis focuses on the

273 strategies or practices that players employ to make their *Let's*
 274 *Plays* attractive and engaging for potential viewers (and
 275 therefore “watchable”).

276 To illustrate the core strategies and practices for making
 277 *Let's Plays* watchable we draw on an example of a so-called
 278 ‘blind play’ (where the game has not been played before by
 279 the players) by the popular German player Pan³, who
 280 presents a current computer game from the adventure-
 281 action-genre. A few examples are taken also from multiplayer
 282 *Let's Plays*, in which several well-known German *Let's*
 283 *Players* participate in a joint adventure-horror game. The
 284 combination of blind play and the adventure/action/horror
 285 genre promises situations that are potentially unpredictable
 286 and/or surprising and therefore both require explanation by
 287 the player(s) and provoke spontaneous reactions. Selecting
 288 popular *Let's Players* offers us a chance to pin-point more
 289 typical, well-established practices in this community of *Let's*
 290 *Players*.

291 5. Analysis

292 5.1 Hypothesis

293 In the following analysis we investigate two highly frequent
 294 practices used by *Let's Players* to embody their avatars for
 295 viewers. The first practice relates to *formulating one's own*
 296 *actions*, while the second one deals with *animating avatars*
 297 *via response cries*.

298 A very simple example for *formulating actions* is shown in
 299 the following transcript: The gamer introduced above, *Pan*,
 300 plays a demo version of the action-adventure game *Outlast 2*.
 301 The transcript shows her talk (original German with an
 302 English translation in bold) as well as important game events
 303 (GE) and game sounds (GS). Special characters (such as * or ~)
 304 align non-verbal events with talk.⁴ Still images are

3 Pan is the short form for the pseudonym Pandorya, a well-known German
 Let's Player, and at the same time her nickname as Let's Player.

4 Talk is transcribed according to GAT2 (cf. Selting et al. 2009), embodied
 conduct according to multimodal conventions following Mondada (2014). For a
 more comprehensive overview of the transcription convention used, see
 Appendix.

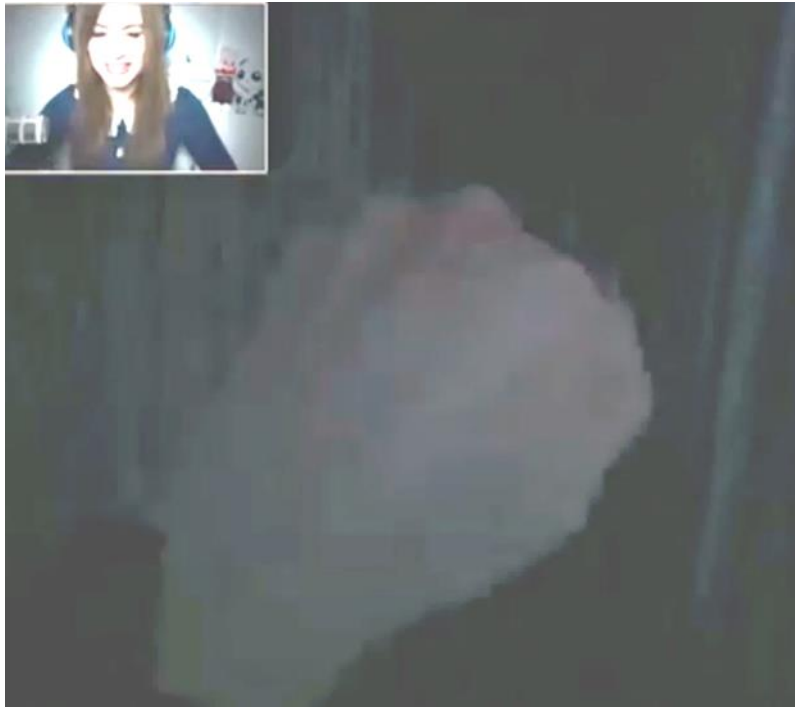
305 represented by an extra line termed *Fig*, and their exact
 306 position in relation to talk is indicated by a hashtag (#).

307 **Transcript (1):** knocking (video 1)⁵

```

01 P   ich klopfe hier einmal *~AN.#
       i'm gonna knock here now
      GE                               *avatar's
      GS   fist/knocking
      Fig                               ~knocking sound
                                               #fig. 3
  
```

308 In this example, *Pan* is performing a *knocking-action* with
 309 her avatar which is represented in a point-of-view-
 310 perspective (cf. Figure 3).⁶



311

312 **Figure 3:** Knocking action in point-of-view-perspective

313 The action is not only carried out by manipulating the avatar,
 314 but is also *verbalized* by saying “*Ich klopfe hier einmal*
 315 *an*”/“*I’m gonna knock here now*” (line 1). With this
 316 verbalization, *Pan* is letting us know what she is doing. We
 317 call such practices *formulating actions*.

318 A simple example for *animating avatars via response cries*
 319 is the following case: *Pan*’s avatar is hiding in a locker in a

⁵ You can examine video 1 [here](#).

⁶ Video games in point-of-view-perspective enhance the impression, for players and viewers alike, to be directly active in the game (cf. Neitzel 2013).

320 deserted hallway. As she is about to open the door in order to
 321 get out again, a monster suddenly appears in front of the
 322 locker. Her strong physical reaction is accessible via the
 323 integrated facecam (FC).

324 **Transcript (2):** being scared (video 2)⁷

01 Pan kannst du mal BITte *wieder;
 can you please again
 GE *starts to open the locker-
 -->*

02 Pan DANke schön das wär super*
 thanks that would be great
 GE -----*

03 **\$(0.5)**
 GE \$monster appears suddenly in front of the
 locker

04 Pan *eeeeehhhhhh
 eeeeehhhhhh
 FC *distorts the face, tears both hands upwards,
 jerks back from the screen
 Fig #fig. 4

05 ***(0.5)**
 FC *grabs her head with both hands, eyes and mouth
 torn open

06 Pan *huuuuaaa
 huuuuaaa
 Fc *takes hands down, gets closer to the screen
 again

325 In this example, *Pan* shows strong *embodied reactions* to an
 326 unexpected game event, the sudden appearance of a monster.
 327 The extract begins with *Pan* talking to ‘the game’ by
 328 requesting for a possibility to get out of the locker (line 1),
 329 which she thanks for when she figures out how to reopen the
 330 door (line 2). At this point the monster appears, and after a
 331 half-second pause (line 3), *Pan* moves her upper part of the
 332 body quickly away from the screen and produces a fright
 333 sound accompanied by a corresponding facial expression
 334 (line 4; cf. Fig. 4).

⁷ You can examine video 2 [here](#).



335

336 **Figure 4:** Appearance of a monster and Pan's scared reaction

337 With this reaction, she displays a stance (she is shocked)
 338 towards an emergent result in the game (the appearance of a
 339 monster). At the same time, the game event is supplemented
 340 by an immediate and physical reaction of the gamer. Such
 341 reactions not only enrich game events with additional
 342 meaning (cf. also Recktenwald 2017), but give avatars a lively,
 343 audio-visually accessible experience. We call such practices
 344 *animating avatars*.

345 In the following we argue that there is an *inherent relation*
 346 between *formulating actions* as an attempt to make
 347 intentions accessible and *animating avatars* as spontaneous
 348 reactions which add an emotional dimension to represented
 349 game events. That is, in *Let's Plays*, both forms of comments,
 350 formulating actions and animating avatars, can be understood
 351 as different but *complementary practices of embodying*
 352 *avatars*.

353 In the following we aim to support this hypothesis by first
 354 discussing practices of *formulating actions* and, secondly,
 355 practices of *animating avatars* in order to show their
 356 *capability to embody avatars*.

357 5.2 Formulations

358 Especially in *Let's Plays* which are dedicated to an audience,
 359 *Let's Players* are forced to make their gaming activities
 360 attractive for viewers. This is done by making them more

361 *transparent* with the help of *verbal moderation* like for
 362 example when players tell viewers what they plan to do next
 363 or which problems they may expect (cf. Marx/Schmidt 2019).
 364 Within the moderation a frequent practice is to *verbalize*
 365 *one's own actions*. A case in point is the formulation of the
 366 knocking-action as discussed above (“*I’m gonna knock here*
 367 *now*”). The way in which *Pan* formulates this action is
 368 designed to *ascribe intentions* to the visible actions of her
 369 avatar. Importantly, saying *I’m gonna knock* is not formulating
 370 *her* action of manipulating the game controller (like pressing a
 371 key to trigger the knocking action in the game); rather it
 372 formulates the result of her control action, which is an *audio-*
 373 *visual representation* of an avatar knocking at a door. By this,
 374 her *intention* to knock in the game by pressing a certain key
 375 is *transferred* by the game technology to a visible and
 376 *accountable action* of an avatar in the game world.
 377 Nevertheless she is using the pronoun *I*, thereby conflating
 378 her own actions with those of the avatar. Therefore,
 379 formulating one’s own actions in this way equips avatars with
 380 plans, intentions and, in the long run, with rationality. We
 381 seemingly get to know why an avatar is acting in a certain
 382 way within the game.

383 If we have a closer look at the knocking-example with
 384 respect to its *temporal structure*, we realize that the verbal
 385 part precedes the represented action on the screen, which
 386 immediately follows in slight overlap (Transcript 1 is
 387 presented here again for convenience):

388 **Transcript (3):** knocking (video 1)⁸

```

0 P      ich klopfe hier einmal *~AN.#
1       i'm gonna knock here now
      GE      *avatar's fist/knocking
      GS      ~knocking sound
      Fig     #fig. 3
  
```

389 This is the typical ordering as it occurs in most of our
 390 instances. A *verbal formulation* (line 1: *I’m gonna knock here*
 391 *now*) is followed by a *corresponding action* in the game.
 392 Thereby both parts are tied together in a *reflexive* way, in
 393 that the verbal part appears as a projection and the avatar’s
 394 action itself as its fulfilment. This is made possible by the
 395 fine-grained temporal coordination of verbal projection and

⁸ You can examine video 1 [here](#).

396 embodied implementation. At the same time, game actions
 397 get a verbal label (here: *knocking*) whereby visually accessible
 398 conduct is unambiguously categorized (here as *knocking*).

399 Sometimes next actions are not only verbally projected,
 400 but specific *expectations* concerning the results of next
 401 actions may also be explicitly anticipated. This is done by
 402 embodied conduct indicating expectations. In the following
 403 example *Pan* is exploring a room with several doors. One of
 404 these doors is half open, which motivates her to take a closer
 405 look at the door. Her exploration is accompanied by facial
 406 expressions and shifts in body posture which indicates
 407 anticipated trouble.

408 **Transcript (4):** anticipation (video 3)⁹

```

01 Pan      +(xxx) ^was ist HIER?
            (xxx)  what is here
            GE      +approaches a half-open door--->+
            FC      ^tilts head to side--->^

02          + (0.6) ^# (2.23)
            GE      +starts to open the door--->+
            FC      ^pulls tilted head back, twists mouth-->^
            Fig      #fig. 5

03 Pan      YE::^%AH,
            yeah
            FC      ^bends her head forward--->>
            GE      %turns on Night Vision--->>

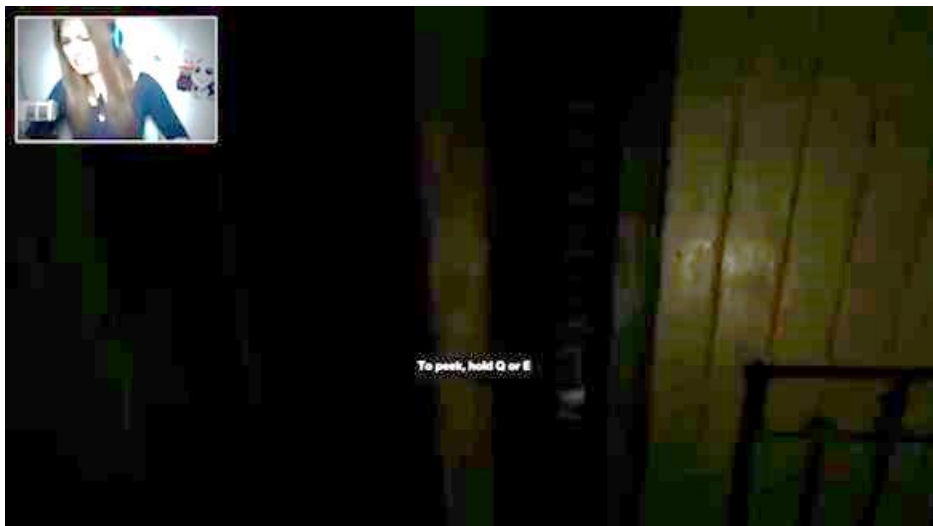
04          (0.4)

05 Pan      +halLO?
            hello
            GE      +moves towards/through the door
  
```

409 When Pan is approaching the half-open door, she produces
 410 the question *what is here?* (line 1) which serves as an
 411 *exploratory announcement* indicating her (following) action as
 412 *'exploring something'*.¹⁰ She tilts her head slightly to the side
 413 as if she is peeking carefully at what is behind the door. When
 414 she starts to open the door (line 2) she pulls her tilted head
 415 back and twists her mouth (s. Fig. 5).

⁹ You can examine video 3 [here](#).

¹⁰ For a more extensive analysis of exploratory announcements in *Let's Plays* cf. Marx/Schmidt (2019).



416

417 **Figure 5:** Facial expression anticipates trouble

418 By her embodied conduct, especially by her posture shift and
 419 her facial expression (s. Fig. 5), Pan indicates that she *expects*
 420 to be scared by a sudden event (as it was the case above with
 421 the monster). But in this case, she makes her *expectation* of a
 422 specific possible result of her exploring action publicly
 423 available before being confronted with the actual outcome.
 424 She maintains her strained (facial) expression for several
 425 seconds (line 2), and only relaxes it when the potentially
 426 dangerous situation is resolved, accompanied by a drawn out
 427 *Yeah* indicating relief (line 3). At the same time she turns on
 428 the night vision to have a better view. Once the danger is
 429 over, she starts to explore what is behind the door (line 4).

430 Action formulations and embodied expectations (like in the
 431 case before) are not always used in an anticipating function
 432 preceding the next action. They can occur at *different*
 433 *temporal positions* within an action process. Besides their
 434 occurrence as projections in initial position (like in the two
 435 cases above) they can also appear *during* and *after* action
 436 processes. As we will see in the following examples, they can
 437 mark actions as relevant next actions, as ongoing actions or as
 438 completed actions, depending on their temporal positioning.

439 In the next example, *Pan's* avatar moves quickly through a
 440 cornfield. As she performs that action, she repeats the phrase
 441 *I'm looking for something*:

442 **Transcript (5):** I'm looking for something (video 4)¹¹

01 Pan ^oKAY,
 oKAY
 GE >>moves through gaming world--->>
 ^leans slightly forward --->>
02 Pan ich SUche was,
 i'm looking for something
03 Pan ich SUche was,
 i'm looking for something
04 Pan ich SUche was;
 i'm looking for something

443 In this example Pan repeats the formulation *I'm looking for*
444 *something* three times. All three instances are very similar.
445 They use the same words in the same order, and prosodically
446 they are produced in a nearly identical manner. This
447 redundancy, achieved by almost exact repetition, is used to
448 signal that the *action is still underway* and not yet completed.
449 As Stivers (2006) has shown, the main function of multiple
450 sayings is displaying that the turn is addressing an in progress
451 course of action. Furthermore, the repetitions here indicate
452 that there is no change with regard to the ongoing action
453 process, and especially that it has not yet been successful.

454 However, by using the expression *looking for something*, a
455 specific kind of outcome is anticipated, which is *finding*
456 *something*. This means that the current action is not only
457 categorized (as *looking for something*) and marked as ongoing
458 by accompanying verbalizations while it is underway, but at
459 the same time it is prospectively limited by implying a
460 possible end point (*finding something*).

461 Formulations are not only used before and during actions
462 processes, as in the examples above. They can also be used
463 *after* the completion of an action process. However, usually
464 they follow projections, so that they hardly ever occur alone.
465 Thus, they are embedded in *sequences* of projecting a next
466 action, conducting that action and finally evaluating it
467 afterwards.

468 In the next example, two such sequences of action are
469 concluded by evaluative formulations (line 3 and line 6). We
470 join the action when Pan is exploring a room she has just
471 entered:

¹¹ You can examine video 4 [here](#).

472 **Transcript (6):** action sequence (video 5)¹²

01 **(0.2)**
 GE >>focus on a book on a table in a dark room ->*

02 Pan kann ich das hier LEsen?*

can i read this

03 Pan *nein kann +ich NICH.

no i can`t

 *defocus book

 +focus cupboard/attempt to open--->+

04 **(0.29)**

05 Pan kann ich den SCHRANK-

can i the cupboard

06 Pan +°h ich kann den SCHRANK a-

°h i can the cupboard o

 +zooms in on cupboard >>

07 Pan °h (.) A:A:H.

°h (.) aah

473 In this case, two action processes are projected (lines 1 and 5)
 474 and respectively completed by statements about the action's
 475 outcome (lines 3 and 6) which can be understood as
 476 evaluations as they assess the success of the projected actions
 477 retrospectively. The first one starts with an exploratory
 478 announcement in form of a question (line 2: *can I read this?*)
 479 followed by a negative evaluation of the action's outcome
 480 (line 3: *no, I can't*). The second one is launched in a similar
 481 way by an aborted question (line 5: *can I the cupboard...*) and
 482 followed by a positive evaluation (line 6: *I can the*
 483 *cupboard...*). Between projection (questions: *can I...*) and
 484 evaluation (*I can't/I can*) the projected actions are tested
 485 (*reading the titles of a book | opening a cupboard*). In both
 486 cases the final evaluation signals a *completion* of the action
 487 process and is bracketed by initial and terminal formulations.
 488 As the second action sequence (cupboard) shows, verbal
 489 formulations are adapted to the pace of visually conveyed
 490 action processes (and not vice versa). Once Pan has managed
 491 to open the cupboard, she aborts her question (line 5) and
 492 proceeds seamlessly to an evaluation (line 6) which she also
 493 aborts in favor for a change-of-state-token¹³ (a drawn out *ah*
 494 in line 7) conveying that she has learned something. The fact
 495 that the action sequence *try to open the cupboard* is treated as
 496 completed at this point is also indicated by a camera action,

¹² You can examine video 5 [here](#).

¹³ Cf. Heritage (1984b).

497 the *zooming in* on the cupboard (line 6). By doing this, Pan
498 switches her focus from opening to exploring the cupboard.

499 Interestingly, by formulating their actions before, during
500 and after its conduction, *Let's Players* not only reveal their
501 plans, but sometimes they *project an expected result* on the
502 basis of which the action can be evaluated afterwards. This
503 becomes particularly apparent in the case of *failures*. In the
504 following example *Pan* explores a building from the outside.

505 **Transcript (7): failure (video 6)¹⁴**

01 (1.13)
GE >>focus on a window--->+
02 Pan ich kann hier REINGucken;*
i can look inside
03 Pan +°h oke wahrscheinlich kann ich auch da
REINGehn;
°h okay probably i can also enter there
+panning movement, starts moving--->+
04 Pan das werd ich ja wo_ma direkt MACHEN,
i will directly do that now
05 +(1.89)
GE +moves tw. entrance, tries to open the door-->+
06 Pan +HÄÄh;
Huh
07 (0.39)
08 Pan halLO?
Hello
09 (0.46)
10 Pan KANN ich nich;
i can't
11 Pan (.) OH;
(.) oh

506 In this example, visual evidence (line 2: *I can look inside*)
507 leads to an inference (line 3: *I can also enter*) which serves
508 not only as a projection for a specific kind of next action
509 (*entering the house*) but at the same time as a projection of an
510 expected (successful) outcome (*being able to enter the house*).
511 After announcing the intention to do so (line 4: *I will directly*
512 *do that now*), the following conducted action of opening the
513 door of the house fails (line 5) and is commented on by a
514 response cry-like surprise sound (*hääh/huh* in line 6),
515 followed by a summon (*hello* in line 8). Both reactions convey
516 her disbelief to have failed. Finally, the action sequence is
517 evaluated with *I can't* (line 10) and a change-of-state-token (a
518 freestanding *oh* in line 11) retrospectively contextualizing the

¹⁴ You can examine video 6 [here](#).

519 action's outcome as surprising and the information sequence
520 as complete (cf. Heritage 1984b).

521 By *projecting* a possible result of her action (*being able to*
522 *enter the house*), Pan makes the *conditions of a failure*
523 *explicit* (in this case not being able to enter) before actually
524 conducting the action. By this, visually conveyed action
525 processes in the game get both a *projected course* by marking
526 start and end points and a normative structure by *projecting*
527 *expected outcomes*. Both contribute to making Let's Plays
528 more *transparent*.

529 Making action processes transparent in order to involve
530 viewers is particularly important when it comes to *less*
531 *obvious actions*. The next example is an extended version of
532 transcript 5 where Pan is moving quickly through a cornfield:

533 **Transcript (8):** I'm looking for something (extended) (video 7)¹⁵

01 Pan ^oKAY,
 oKAY
 GE >>moves through gaming world--->>
 FC ^leans slightly forward --->>
02 Pan ich SUche was,
 i'm looking for something
03 Pan ich SUche was,
 i'm looking for something
04 Pan ich SUche was;
 i'm looking for something
05 **(1.08) & (2.0)**
 GE &Night Vision
06 Pan *DAS suche ich.
 that's what i'm looking for
 GE *water container visible, moves towards it
07 ***(0.92)**
 GE *climbs into water container

534 In this case, if only the images were available, the viewer
535 would see a flow of movements that is hardly recognizable as
536 a certain activity. Only the verbal descriptions of this flow of
537 images as *looking for something* enables the viewer to
538 comprehend what is happening on the screen. This example
539 highlights how otherwise apparently *random movements* on
540 the screen are framed by the *verbal formulations* as *rational*
541 *and accountable* activities, in this case as a process of
542 *searching*.

543 Furthermore, the flow of visual representations is
544 *packaged* into comprehensible *action units*, in this example

¹⁵ You can examine video 7 [here](#).

545 starting with a search lasting a while (marked by the
546 repetitions in lines 2-5) and ending successfully, in this case
547 by finding a specific place to hide (a water container), which
548 is also verbally announced (line 6: *that's what I'm looking*
549 *for*). Note that the found 'object' is not explicitly named, but
550 referred to with the demonstrative pronoun *that*. By using
551 *that* (instead of a referential noun like *water container*), the
552 visually conveyed game world and the talk about it are
553 reflexively tied together as we have to scan the images in
554 order to detect the reference of *that*. In turn, only by the
555 verbal formulations, parts of the images get the status of a
556 searched/found object. Moreover, by indexically referring to
557 the game world, Pan makes her perception relevant. To
558 understand the meaning of *that*, we have to see the game
559 world with her eyes or – as Hausendorf (2003) has put it –
560 we are invited to perceive her perception.

561 5.3. Animations via response cries

562 Game events are not only verbally formulated, but also
563 commented on via *exclamations* or (in Goffman's words) via
564 *response cries* which convey a *bodily involvement* and a
565 *player's stance* towards actual game events. Having access to
566 *spontaneous reactions* of the player (transmitted via facecam)
567 makes the game more transparent and attractive for viewers.
568 A case in point is the *shock reaction* of the player Pan
569 described above, when suddenly a monster appeared (see
570 transcript 2).

571 Reactions of this kind are very frequent. In our selected
572 30-minutes *Let's Play* from the player *Pan*, response cries
573 occur every 30 seconds on average; altogether we found
574 more than *105 instances* (for a compilation see video 8¹⁶).
575 They occur in lexicalized forms (such as *shit, oh my god* etc.)
576 and non-lexicalized forms (e.g. shock cries as in the monster
577 example, or pain cries, cf. below). They are related to various
578 events, such as sudden game events (as in transcript 2),
579 anticipated game events caused by own actions (as in
580 transcript 6), the game control, status displays or simply the
581 development of the story within the game. Depending on
582 how demanding the game is at a certain point, forms,
583 functions and density of response cries may vary. Especially

16 You can examine video 8 [here](#).

584 when the game gets particularly thrilling, the use of response
585 cries increases and tends to replace verbal comments overall.
586 Furthermore, response cries play a crucial role in structuring
587 and constructing action processes and thereby assign
588 meaning to the images conveyed by the computer game (cf.
589 section 5.2 above). Sometimes they animate avatars in terms
590 of reacting or speaking on their behalf (see below the
591 comments on *pain cries* and *interaction with non-play-*
592 *characters*).

593 What all cases have in common is that the use of response
594 cries connects game events with the player's emotions and
595 thus roots them in physical experience. Often the results of
596 players' announced actions are qualified affectively
597 afterwards by response cries (as above in transcript 6).
598 Goffman (1981b) has argued that response cries are designed
599 to convey an internal state without doing it in an explicit
600 communicative way (cf. also Baldauf 2002, Heritage 1984b).
601 Saying *Ooops* in a public place, e.g. if you are stumbling over
602 a step, signals to all those present within earshot that the
603 stumbling person realized that it was an accident. In this way,
604 it is framed as an exception and, by that, individuals show
605 that they are normal members of society who are aware of
606 their misconduct. At the same time, uttering *Ooops* does not
607 oblige anybody to engage in an interaction or conversation.
608 Response cries can be registered without any comment. One
609 reason for that is that they are read, partly because of their
610 non-lexicalized sound structure, as direct reactions of the
611 body, which are then more of an indicator or an indexical
612 sign than a full-fledged symbol or sign such as lexical units
613 like words. Response cries are therefore seen as more rooted
614 in autonomous reactions of the body than verbal assertions,
615 for instance, which may refer to the same or similar
616 circumstances.

617 As research on board games (cf. Hofstetter 2020) and
618 computer gaming has shown (cf. Aarsand/Aronsson 2009;
619 Baldauf-Quilliatre 2014; Piirainen-Marsh 2012), response
620 cries are used frequently when playing (video) games. They
621 are not only an expression of involvement in the game, but
622 are also used continuously to convey to fellow players,
623 spectators and recipients how the players experience
624 individual game events.

625 Interestingly, although the gaming experience is conveyed
626 through a virtual character, the avatar, which has no real

627 body and no feelings, the player's reactions also include
 628 response cries that suggest direct contact with the "material"
 629 world in the game. This holds for instance for *pain cries*, like
 630 in the following collection of examples from a multiplayer-
 631 *Let's Plays* in which four participants (A, B, C, D) play the
 632 adventure-horror-game *Dead by Daylight*.¹⁷

633 **Transcript (10):** collection of pain cries (video 9)¹⁸

634 (a) whiny voice

01A D's name hör mal auf zu STRUGgeln jetzt;
 name of D stop struggling now
 GE >>D on shoulder of A; A moves towards place of
 execution>>
 02D °hh° <<weinerliche Stimme> lass mich runTE:R->
 °h h° <<**whiny voice**> **let me down** >

635 In this extract, gamer A, who plays a killer, carries the injured
 636 avatar of D on his shoulders in order to execute him. While
 637 being carried, D tries to escape by making his avatar struggle
 638 hard (which A complains about jokingly in line 1). While D is
 639 carried by the killer, he whines quietly and pleads with a
 640 whiny voice to release him. He therefore lets his avatar talk
 641 to the killer-figure played by A.

642 (b) ouch (simplified)

01C ihr müsst natürlich LEIse *~sein;
 you need to be quiet of course
 GE >>C fixes generator *C gets electric shock
 GS ~banging
 02 **(0.43)**
 03C AUa SCHEIße-
 ouch shit

643 In this extract, C tries to fix a generator while talking about
 644 something else (line 1). He fails and gets an electric shock (line
 645 1). In reaction to this, C produces a pain cry and a swearword
 646 (line 3: *ouch shit*).

¹⁷ In Multiplayer *Let's Plays* the participants are in different places and play an online game together. They are connected via voice conference software (cf. Marx/Schmidt 2019).

¹⁸ You can examine video 9 [here](#).

647 (c) ahh

```

01B   glaube wir %~*könn da hier ^#[(xxx)]
      believe we can here xxx
02C   [!AH! ]
      ah
      >>C runs in crouching position
GE    %A hits C with machete from behind
GS    ~stroking sound
GE    *C falls over
FCc   ^moves back, eyes
      wide open
      Fig. #Fig.6
03A   du HUM[peilst nich mehr,]
      you're not limping any more
04D   [NEI:N NEI:N; ]
      no no

```

648 In this example, the killer (A) hits C with a machete from
649 behind (line 2). C shrugs back from the screen with his eyes
650 wide open and cries out loud (line 2), drowning out the end of
651 B's prior utterance (line 1). Partly overlapping with A's ironic
652 comment (line 3), C comments on his situation by saying *no*
653 twice in a modulated voice (line 4).

654 In this collection of examples, represented physical states
655 and pains of avatars caused by virtual events are embodied by
656 the players producing both vocal sounds such as pain cries,
657 whining and moaning, and embodied conduct such as bodily
658 position changes and facial expressions accessible via the
659 integrated facecam. Figure 6 shows C's reaction when hit by
660 the killer from behind.



661

662 **Figure 6:** C's reaction when hit by the killer

663 The physical expressiveness indicates not only their
 664 involvement in the game but adds more liveliness to the
 665 overall experience of watching the gameplay. Pain cries in
 666 particular appear to be directly connected to physical
 667 experiences (cf. Ehlich 1985, Mannheim 1980).

668 The animating effect of response cries gets especially
 669 obvious when gamers speak as avatars with non-play
 670 characters (NPC)¹⁹ as in the next collection of examples:

671 **Transcripts (11):** collection of interaction with NPC (video 10)²⁰

672 (a) leave me in peace

```

01          ** (0.81)
    GS      >>threatening music>>
    GE      *NPC reaches for her and tries to hold her>>
    GE      +tries to escape>>
02 Pan     !FUCK!,
           Fuck
03          * (0.22)
    GE      *fast camera movements
04 Pan     <<lachend> WOAHA SCHEIÙe >
           <<laughing> wow shit >
05          ~ (1.78)
    GS      ~hittig sound
06 Pan     !NEIN!,
           No

```

¹⁹ Non-play characters (NPC) are characters in the game that are not controlled by a human player, but by the software of the game.

²⁰ You can examine video 10 [here](#).

07 (0.29)
 08 Pan LASS mich;
leave me in peace;
 09 (0.42)
 10 Pan LASS ~mich;
leave me in peace
 GS ~hitting sound
 11 (0.92)
 12 Pan lass mich RAUS,
let me out
 13 (0.55)
 14 Pan LASS mich einfach raus;
just let me out
 15 Pan *lass mich einfach RAUS,
just let me out
 *escapes und flees>>

673 In this extract, Pan is suddenly attacked by an NPC (line 1)
 674 which tries to catch her. During the whole extract she tries to
 675 escape and get rid of him. First, she comments on her
 676 situation via response cries (*fuck* in line 2, *wow shit* in line 4,
 677 *no* in line 6), later she addresses the NPC directly telling him
 678 to stop (lines 8-15). As she slowly manages to free herself, her
 679 voice gets calmer and softer (lines 14/15).

680 (b) no no no no no

01 (1.0) * (0.2) ~ (1.0)
 GE *NPC approaches, reaches for her
 GS ~hitting sound
 02 Pan *NEI:N nein nein nein nein ~nein nein
 nein nein nein nein nein nein nein nein
 nein nein,
no no no no no no no no no no no no no no no
no no no
 GE *escapes and flees>>
 ~hitting sound
 03 Pan lass mich in RUhe,
leave me in peace
 04 Pan lass mich in RUhe.
leave me in peace

681 As in the extract before, Pan is attacked and pursued by an
 682 NPC (line 1). This time she is able to escape a scuffle and flees
 683 directly. Her escape is accompanied first by a series of
 684 response cries (multiple *no*'s in line 2), later by direct calls to
 685 the NPC to leave her alone (lines 3/4). Furthermore, her
 686 multiple sayings display that her turn is addressing an ongoing
 687 'action' (cf. Stivers 2006), in this case of an NPC.

688 As Aarsand/Aronsson (2009) have argued, animating
689 avatars is a common way of engaging in joint gaming. In the
690 above examples, Pan uses several response cry-like
691 expressions (like outcries, negation particles etc.) when she is
692 attacked by a non-play character. She not only reacts to game
693 events but stages them as an interaction with a real
694 counterpart. By using imperatives, for instance (like *get away*
695 *from me, leave me in peace* etc.), and a stronger voice, she
696 treats the non-play character as a person-like figure that can
697 change his behavior through interaction. This results in a
698 more animated nature of her gameplay and the presented
699 actions of her avatar.

700 6. Conclusion

701 Using the example of *Let's Plays*, we have shown the close
702 connection between affordances, participation framework
703 and practices of Let's Players to make their gaming
704 'watchable'. *Let's Plays* are video games in which players get
705 involved in a virtual world by using software. Involvement
706 basically means that the control actions of players are
707 translated into represented movements of avatars on a screen
708 which are interpreted as actions in a virtual world. As players
709 and game/avatars are closely interconnected through a
710 cybernetic control loop, actions of players get immediate
711 feedback. This is the precondition of being immersed in a
712 virtual world. Immersion means that our sense of presence
713 has shifted from the real world to a virtual world. One
714 indicator for immersion in a virtual world is that players sense
715 and react to virtual rather than real world events. The
716 pleasure and attractiveness of video games is largely the
717 result of this possibility to be active and immersed in a virtual
718 world.

719 That said, video games are not, at least not in the first
720 place, made to be presented to an audience. The sense of
721 being present in a virtual world that video games enable is
722 first off only a gratification for active players, not for those
723 who watch how others play. To make watching video gaming
724 attractive, Let's Players try to convey this sense of presence
725 to the audience.

726 Two aspects, as we have seen, are crucial for achieving
727 this: First off, the screen representation of *Let's Plays*

728 involves the game play itself and the player appearing in a
729 facecam. By simultaneously representing the game and those
730 who play the game in a split-screen mode, the most
731 important feature of video games is highlighted: its close
732 connection between control actions and actions in a play
733 world in form of a cybernetic control loop. Like this, for
734 viewers, game actions and reactions of the player(s) are
735 directly related. Based on this tele-presence, the viewers are
736 able to vicariously experience the players' sense of
737 presence.²¹

738 Secondly, as we have seen, Let's Players have to invest a
739 lot of work to actively achieve this kind of illusion. They are
740 continuously oriented towards their viewers providing them
741 with comments on their actions in the running game. Thereby
742 they co-construct tele-presence with regard to viewers. For
743 instance, saying *I'm now going to knock at that door* while
744 performing this action in the game world is constructed to
745 make the move in the game tangible for viewers. Its
746 construction systematically takes viewers into account (ergo
747 co-constructs) with the goal to create a lively representation
748 of their sense of presence (which is then, given the mediated
749 character of *Let's Plays*, a kind of tele-presence).

750 As we have shown, Let's Players basically draw on two
751 kinds of practices to convey their sense of presence when
752 playing the game. On the one hand, they formulate their
753 game actions in order to give action processes a recognizable
754 and rational structure, thus making them more
755 comprehensible. On the other hand, they produce response
756 cries in reaction to game events. Producing response cries
757 during the game not only indicates high involvement and
758 grants viewers access to players' emotions, but also
759 contributes to an animation of avatars. Both formulating
760 actions and animating avatars via response cries are part of
761 *embodying* avatars. Formulating actions equips avatars with
762 cognition (e.g. intentions), animating with sensibility (e.g. pain
763 sensation). Both practices interact to make the visual events
764 on screen understandable as actions. By bringing intentions
765 (or more general: inner processes of consciousness) and
766 external behavior together through spoken discourse (mainly
767 by formulating actions) and embodied conduct (mainly by

21 Basically, this rests on mechanisms of vicarious role-taking, which is always relevant when we observe the conduct of others (cf. Ellis/Streeter/Engelbrecht 1983).

768 animating avatars via response cries) gamers facilitate the
 769 perception of avatars as full-fledged persons/characters. This
 770 is an important precondition to *understanding* and thus to
 771 *enjoying* what is happening on the screen.

772 Our contribution focused on practices of Let's Players
 773 which enhance an 'illusive' (Rapp 1973) experience of
 774 viewers. Such practices are designed to create the impression
 775 that players are directly active in the play world and that their
 776 avatars are capable of acting. However, Let's Players are not
 777 only deploying practices which are designed to create illusion
 778 (for viewers). There are instances in which they talk with
 779 their avatars, with the game itself (as in transcript 2) or in
 780 which they meta-communicate the fact of playing a video
 781 game (e.g. by discussing the production of the video game, its
 782 narration or the game control). Instead of concealing the
 783 process of mediation (as practices of embodying avatars do),
 784 meta-communicative practices disclose the impression of
 785 being present in a virtual world. Such practices do not create
 786 illusion, rather they destroy them. They are not illusive, they
 787 are 'inclusiv', that is they are creating distance instead of
 788 immersion (cf. Rapp 1973). Further research is needed to
 789 explore this intricate relationship between opaqueness and
 790 transparency or illusion and inclusion in re-mediations such as
 791 *Let's Plays*.²²

792 7. Appendix

793 Conventions for the notation of physical activities (cf. Mondada
 794 2014)

795 *Nonlinguistic events and activities*

- 796 • appear after the abbreviations GE, GS, FC, and Fig
- 797 • in lines following pauses or conversation activities
- 798 (without own number)

22 Following Bolter/Grusin (2000), *Let's Plays* are understandable as a form of re-mediation as they mediate an already existing medium, the video game (see also Ackermann 2016b). Re-mediation can be opaque, then they promote the illusion of the re-mediated medium, or they can be transparent, then they expose the illusion of the re-mediated medium. In our contribution, we focused on practices which enhance an opaque re-mediation of video games in *Let's Plays*.

- 799 • are aligned with conversation/pauses with the help of
800 special characters (like *, ~, + etc.) indicating the
801 beginning and (if relevant) the end of events

802 *Further conventions for the notation of physical movement*

- 803 ---> movement continues
804 --->\$ movement continues after the line until reaching
805 \$ the same sign
806 >> continues after transcript ends
807 >> starts before transcript

808 8. Bibliography

- 809 Aarsand, André/Aronsson, Karin (2009): Response cries and
810 other gaming moves – Building intersubjectivity in gaming.
811 In: *Journal of Pragmatics*, 41, 1557–1575.
- 812 Ackermann, Judith (2016a) (Ed.): *Phänomen Let's Play-Video:
813 Entstehung, Ästhetik, Aneignung und Faszination
814 aufgezeichneten Computerhandelns*. Wiesbaden: Springer
815 VS.
- 816 Ackermann, Judith (2016b): Einleitung. Entstehung und
817 wissenschaftliche Relevanz eines
818 Remediatisierungsphänomens. In: Ackermann, Judith (Ed.):
819 *Phänomen Let's Play: Entstehung, Ästhetik, Aneignung
820 und Faszination aufgezeichneten Computerhandelns*.
821 Wiesbaden: Springer VS, 1–15.
- 822 Ayaß, Ruth (1993): Auf der Suche nach dem verlorenen
823 Zuschauer. In Holly, Werner/Püschel, Ulrich (Eds.):
824 *Medienrezeption als Aneignung*. Opladen: Westdeutscher
825 Verlag, 27–41.
- 826 Ayaß, Ruth (2005): Interaktion ohne Gegenüber? In Jäckel,
827 Michael/Mai, Michael (Eds.): *Online-Vergesellschaftung?
828 Mediensoziologische Perspektiven auf neue
829 Kommunikationstechnologien*. Wiesbaden: VS, 33–49.
- 830 Baldauf, Heike (2002): *Knappes Sprechen*. Tübingen:
831 Niemeyer.
- 832 Baldauf-Quilliatre, Heike (2014): Formen knapper
833 Bewertungen beim Fußballspielen an der Playstation. In:
834 Schwarze, Cordula/Konzett, Carmen (Eds.):
835 *Interaktionsforschung: Gesprächsanalytische Fallstudien
836 und Forschungspraxis*. Berlin, 107–130.

- 837 Baldauf-Quilliatre, Heike/Colón de Carvajal, Isabel (2015): Is
838 the avatar considered as a participant by the players? A
839 conversational analysis of multi-player videogames
840 interactions. In: *PsychNology Journal*, 13, 2-3, 127–147.
- 841 Baldauf-Quilliatre, Heike/Colón de Carvajal, Isabel (2019):
842 Interaktionen bei Videospiele-Sessions: Interagieren in
843 einem hybriden Raum. In: Marx, Konstanze/Schmidt, Axel
844 (Eds.): *Interaktion und Medien – Interaktionsanalytische*
845 *Zugänge zu medienvermittelter Kommunikation*.
846 Heidelberg: Universitätsverlag, 219–254.
- 847 Bolter, Jay David/Grusin, Richard (2000): *Remediation:*
848 *understanding new media*. Cambridge: MIT Press.
- 849 Clark, Andy (2001): *Being there: putting brain, body, and*
850 *world together again*. Cambridge: MIT Press.
- 851 Deppermann, Arnulf (2013): Multimodal interaction from a
852 conversation analytic perspective. In: *Journal of*
853 *Pragmatics*, 46, 1–7.
- 854 Ehlich, Konrad (1985): The Language of Pain. In: *Theoretical*
855 *Medicine*, 6, 177–187.
- 856 Ellis, Godfrey J./Streeter, Sandra K./Engelbrecht, Joann D.
857 (1983): Television characters as Significant Others and the
858 Process of Vicarious Role Taking. In: *Journal of Family*
859 *Issues*, 4, 2, 367–384.
- 860 Freyermuth, Gundolf S. (2015): *Games, game design, game*
861 *studies: eine Einführung*. Bielefeld: transcript.
- 862 Gibbs, Raymond W. (2007): *Embodiment and cognitive*
863 *science*. Cambridge: Cambridge Univ. Press.
- 864 Goffman, Erving (1963): *Behavior in public places: notes on*
865 *the social organization of gatherings*. London: Free Press.
- 866 Goffman, Erving (1981a): Footing. In: Goffman, Erving: *Forms*
867 *of Talk*. Philadelphia: University of Pennsylvania Press,
868 124–159.
- 869 Goffman, Erving (1981b): Response Cries. In: Goffman,
870 Erving: *Forms of Talk*. Philadelphia: University of
871 Pennsylvania Press, 78–122.
- 872 Goodwin, Charles/Goodwin, Marjorie Harness (2004):
873 Participation. In Duranti, Alessandro (Eds.): *A companion*
874 *to linguistic anthropology*. Vol.1, Malden: Blackwell, 222–
875 244.
- 876 Gumbrecht, Hans Ulrich (2012): *Präsenz*. Berlin: Suhrkamp.
- 877 Hale, Thomas (2013): *From Jackasses to Superstars: A Case*
878 *for the Study of 'Let's Play'*. Online resource:
879 https://www.academia.edu/5260639/From_Jackasses_to_

- 880 Superstars_A_Case_for_the_Study_of_Let_s_Play_Sept
881 ember_2013_
- 882 Hausendorf, Heiko (2001): Warum wir im Fernsehen so
883 häufig begrüßt und angeredet werden. In Sutter,
884 Tilmann/Charlton, Michael (Eds.): *Massenkommunikation,
885 Interaktion und soziales Handeln*. Wiesbaden:
886 Westdeutscher Verlag, 185–213.
- 887 Heritage, John (1984a): *Garfinkel and Ethnomethodology*.
888 Cambridge: Polity Press.
- 889 Heritage, John (1984b): A change-of-state token and aspects
890 of its sequential placement. In Atkinson, Maxwell/Heritage,
891 John (Eds.): *Structures of social action: Studies in
892 Conversation Analysis*. Cambridge: Cambridge University
893 Press, 299–345.
- 894 Höflich, Joachim R. (2005): Medien und interpersonale
895 Kommunikation. In Jäckel, Michael (Eds.):
896 *Mediensoziologie: Grundfragen und Forschungsfelder*.
897 Wiesbaden: VS, 69–90.
- 898 Hofstetter, Emily (2020): Nonlexical "Moans": Response Cries
899 in Board Game Interactions. In: *Research on Language and
900 Social Interaction*, 53, 1, 42–65.
- 901 Horton, Donald/Wohl, Richard R. (1956): Mass
902 Communication and para-social interaction: Observations
903 on intimacy at a distance. In: *Psychiatry*, 19, 215–229.
- 904 Keating, Elizabeth/Sunakawa, Chiho (2010): Participation
905 cues: Coordinating activity and collaboration in complex
906 online gaming worlds. In: *Language in Society*, 39, 331–
907 356.
- 908 Lakoff, George/Johnson, Mark (2011): *Metaphors we live by*.
909 Chicago: Univ. of Chicago Press.
- 910 Licoppe, Christian (2015): Contested norms of Presence. In:
911 Hahn, Cornelia/Stempfhuber, Martin (Eds.): *Präsenzen 2.0:
912 Körperinszenierung in Medienkulturen*. Wiesbaden:
913 Springer VS, 97–112.
- 914 Ligman, Kris (2011): Let's Play: Interactivity by Proxy in a
915 Web 2.0 Culture. In: *PopMatters*. URL:
916 [https://www.popmatters.com/lets-play-interactivity-by-
917 proxy-in-a-web-2-0-culture-part-1-2496052960.html](https://www.popmatters.com/lets-play-interactivity-by-proxy-in-a-web-2-0-culture-part-1-2496052960.html),
918 06.02.2017.
- 919 Lombard, Matthew/Ditton, Theresa (1997): At the Heart of It
920 All: The Concept of Presence. In: *Journal of Computer
921 Mediated Communication Computer-Mediated*

- 922 *Communication*, 3, 2, DOI: 10.1111/j.1083-
923 6101.1997.tb00072.x].
- 924 Mannheim, Karl (1980): *Strukturen des Denkens*. Darmstadt:
925 Luchterhand.
- 926 Marx, Konstanze/Schmidt, Axel (2019): Making Let's Plays
927 watchable – Praktiken des stellvertretenden Erlebbar-
928 Machens von Interaktivität in vorgeführten Videospielen.
929 In: Marx, Konstanze/Schmidt, Axel (Eds.): *Interaktion und*
930 *Medien – Interaktionsanalytische Zugänge zu*
931 *medienvermittelter Kommunikation*. Heidelberg:
932 Universitätsverlag, 319–351.
- 933 Meyrowitz, Joshua (1990): Redefining the Situation:
934 Extending Dramaturgy into a theory of social change and
935 media effects. In: Riggins, Stephen H. (Eds.): *Beyond*
936 *Goffman: studies on communication, institution, and social*
937 *interaction*. Berlin: de Gruyter, 65–97.
- 938 Mondada, Lorenza (2008): Using Video for a Sequential and
939 Multimodal Analysis of Social Interaction: Videotaping
940 Institutional Telephone Calls [88 paragraphs]. In: *Forum*
941 *Qualitative Sozialforschung / Forum: Qualitative Social*
942 *Research*, 9, 3, Art. 39.
- 943 Mondada, Lorenza (2012): Coordinating action and talk-in-
944 interaction in and out of video games. In: Ayaß,
945 Ruth/Gerhardt, Cornelia (Eds.): *The appropriation of media*
946 *in everyday life*. Philadelphia: Benjamins, 231–270.
- 947 Mondada, Lorenza (2013): Coordinating mobile action in real
948 time: The timely organization of directives in video games.
949 In: Haddington, Pentti/Mondada, Lorenza/Nevile, Maurice
950 (Eds.): *Interaction and mobility: language and the body in*
951 *motion*. Berlin: de Gruyter, 300–341.
- 952 Mondada, Lorenza (2014): *Conventions for multimodal*
953 *transcription*. Online-Dokument:
954 [https://franzoesistik.philhist.unibas.ch/fileadmin/user_uplo](https://franzoesistik.philhist.unibas.ch/fileadmin/user_upload/franzoesistik/mondada_multimodal_conventions.pdf)
955 [ad/franzoesistik/mondada_multimodal_conventions.pdf](https://franzoesistik.philhist.unibas.ch/fileadmin/user_upload/franzoesistik/mondada_multimodal_conventions.pdf).
- 956 Neitzel, Britta (2013): Point of View and Point of Action –
957 Eine Perspektive auf die Perspektive von
958 Computerspielen. In: *RMFK. Repitorium*
959 *Medienkulturforschung*, 4, online:
960 <http://repositorium.medienkulturforschung.de/?p=360>.
- 961 Pfaller, Robert (2008): *Ästhetik der Interpassivität*. Hamburg:
962 Philo Fine Arts.
- 963 Piirainen-Marsh, Arja (2012): Organizing participation in
964 video gaming activities. In: Ayaß, Ruth/Gerhardt, Cornelia

- 965 (Eds.): The appropriation of media in everyday life.
 966 Philadelphia: Benjamins, 197–230.
- 967 Rapp, Uri (1973): *Handeln und Zuschauen: Untersuchungen*
 968 *über den theatersoziologischen Aspekt in der menschlichen*
 969 *Interaktion*. Darmstadt: Luchterhand.
- 970 Recktenwald, Daniel (2017): Toward a transcription and
 971 analysis of live streaming on Twitch. In: *Journal of*
 972 *Pragmatics*, 115, 68–81.
- 973 Reeves, Stuart/Greiffenhagen, Christian/Laurier, Eric (2016):
 974 Video gaming as practical accomplishment:
 975 Ethnomethodology, conversation analysis, and play. In:
 976 *Topics in Cognitive Science*, 9, 2, 308–342.
- 977 Rohrer, Tim (2007): The Body in Space: Dimensions of
 978 embodiment. In: Zlatev, Jordan; Ziemke, Tom; Frank, Roz;
 979 Dirven, René (Eds.): *Body, Language and Mind*, vol. 2.
 980 Berlin: de Gruyter, 339–378.
- 981 Selting, Magret et al. (2009): Gesprächsanalytisches
 982 Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung -*
 983 *Online-Zeitschrift zur verbalen Interaktion*, 10, 353–402.
- 984 Schemer-Reinhard, Timo (2016): Let's play without
 985 Controller. Zu den Effekten des Verschwindens der
 986 Steuerung im remediatisierten Game. In: Ackermann,
 987 Judith (Eds.): *Phänomen Let's play-Video: Entstehung,*
 988 *Ästhetik, Aneignung und Faszination aufgezeichneten*
 989 *Computerhandelns*. Wiesbaden: Springer VS, 55–83.
- 990 Schmidt, Axel/Marx, Konstanze (2020): Making Let's Plays
 991 watchable: An interactional approach to multimodality. In:
 992 Crispin Thurlow/Christa Dürscheid/Diémóz, Federica
 993 (Eds.): *Visualizing (in) the New Media*. London: John
 994 Benjamins, 131–150.
- 995 Schmidt, Axel/Marx, Konstanze/Neise, Isabell (2020):
 996 Produktion - Produkt - Rezeption? Medienketten in
 997 audiovisuellen Webformaten am Beispiel von Let's Plays.
 998 In: Marx, Konstanze/Lobin, Henning/Schmidt, Axel (Eds.):
 999 *Deutsch in Sozialen Medien: interaktiv, multimodal,*
 1000 *vielfältig*. Berlin: de Gruyter, 265–289.
- 1001 Spagnolli, Anna/Gamberini, Luciano (2005): A Place for
 1002 Presence. Understanding the Human Involvement in
 1003 Mediated Interactive Environments. In: *PsychNology*
 1004 *Journal* 3/1, 6–15.
- 1005 Stivers, Tanya (2004): "No no no" and Other Types of
 1006 Multiple Sayings in Social Interaction. In: *Human*
 1007 *Communication Research*, 30, 2, 260–293.

- 1008 Streeck, Jürgen (2008): *Gesturecraft. The manu-facture of*
1009 *meaning*. Amsterdam: Benjamins.
- 1010 Streeck, Jürgen/Goodwin, Charles/LeBaron, Curtis D. (2011):
1011 Embodied interaction in the material world: an
1012 introduction. In: Streeck, Jürgen/Goodwin,
1013 Charles/LeBaron, Curtis D. (Eds.): *Embodied interaction -*
1014 *language and body in the material world*. New York:
1015 Cambridge Univ. Press, 1–26.
- 1016 Suchman, Lucy A. (1987): *Plans and situated actions: the*
1017 *problem of human-machine communication*. Cambridge:
1018 Cambridge Univ. Press.
- 1019 Tekin, Burak S./Reeves, Stuart (2017): Ways of spectating:
1020 unravelling spectator participation in Kinect play. In:
1021 *Proceedings of the 2017 CHI Conference on Human Factors*
1022 *in Computing Systems*, 1558–1570.
- 1023 Utterback, Camille (2004): Unusual positions – Embodied
1024 interaction with symbolic spaces. In: Wardrip-Fruin,
1025 Noah/Harrigan, Pat/Crumpton, Michael (Eds.): *First*
1026 *Person. New Media as Story, Performance, and Game*.
1027 Cambridge: MIT Press, 218–235.
- 1028 Wachsmuth, Ipke/Lenzen, Manuela/Knoblich, Günther (Eds.)
1029 (2008): *Embodied communication in humans and*
1030 *machines*. Oxford: Oxford University Press.
- 1031 Zhao, Shanyang (2005). The Digital Self: Through the
1032 Looking Glass of Telecopresent Others. In: *Symbolic*
1033 *Interaction*, 28, 3, 387–405.
- 1034 *Video Games*
- 1035 Red Barrels (2017): *Outlast 2*.
- 1036 Starbreeze Studios (2016): *Dead by Daylight*.