Gaming as a Situated Collaborative Practice

Björn Sjöblom, Linköping University

This article uses video data of gaming sessions at an Internet café in order to explore the situated and embodied resources that players use for doing collaborative gaming. Prior studies of gaming have, to a great extent, not accounted for many of the semiotic resources that players utilize for conducting gaming as a collaborative enterprise. From an ethnomethodological and interaction analytical perspective, drawing on Goodwin’s concepts of semiotic fields and contextual configurations, this study shows how both on- and off-screen semiotic resources structure the gaming interaction, and how the use of these resources relate to the players’ different interactional projects, such as issuing instructions or orienting themselves in an on-screen space. These embodied semiotic resources are situated in the on- and off-screen spaces in which the players interact. They provide ways of playing computer games in collaboration that are specific for gaming where the players are co-present.

Keywords: action, collaboration, computer games, embodiment, interactivity, video analysis

Playing computer games in collaboration is done in a wide variety of ways, ranging from friends playing console games in the living room, to LAN-parties, Internet cafés and MMORPGs. Contexts of gaming are known to be varied and diverse, and gaming has been understood as a complex phenomenon (see for example Aarsand 2007a; Jansz & Martens 2005; Johansson 2000 or Linderoth 2004, for examples of the diversity of settings in which gaming has been studied). Gaming is studied as a
social and collaborative enterprise, conducted in a variety of contexts, but the collaborativity of the gaming sessions themselves is seldom examined in its interactional detail. This article is based on a short-term fieldwork in an Internet café in Stockholm, and it tries to explore some features of collaborativity of gaming in this setting.

The data is drawn from video ethnographic recordings of computer gaming at an Internet café. Collaborative action was something that was striking when analyzing the gameplaying. Most players were seated in front of a screen of their own, yet they were recurrently engaged in the gaming of other co-present players, asking questions, providing “online commentary” (Heritage & Stivers 1999) or using “response cries” (Aarsand & Aronsson forthcoming Goffman 1981, ch. 2) as ways of involving themselves in the games. As the games that are played in the café are designed to be playable over the Internet, and the present players have access to the necessary technology at home, there must be something about this type of gaming that differs from playing the games online. What is it, then, about the gaming in this place that is uniquely situated and dependent on the players’ co-presence? Understanding how collaborativity in co-present gaming comes about is a key issue in this article. This is important for understanding complexity of gaming interaction, especially as a situated and embodied phenomenon.

Playing games in this café means, in a variety of ways, to play together with other people. However, playing together does not necessarily mean to log onto the same network or MMORPG-realm or to occupy the same space in the games. Rather, playing together may mean to stand behind someone else playing, watching and commenting on choices, progress or other relevant features of the game play. The roles of participants in the gaming session are varied and negotiable (cf. Goffman 1974). This leads to the question of how various gaming actions are constructed and accomplished as collaborative activities. How collaboration is achieved and maintained, in episodes of naturally occurring gaming activities, is the focus for this study. Here, collaborativity is seen as a phenomenon performed and observed in interaction and studied through the analysis of video recordings of gaming. Collaborativity and “togetherness” are understood as accountable phenomena. In the ethnomethodological definition, the concept of accountable phenomena means that actions are pro-
duced in order to be “observable-and-reportable” (Garfinkel 1967, 1), i.e., to be recognizable for what they are. This means that, in order for the gaming sessions to be recognized as collaborative events, they have to be constructed and oriented to as such by the players themselves. The findings of this study are not primarily theoretical findings; instead they have the nature of a *praxeology* of gaming: an explorative foregrounding of the details of how collaborativity in gaming is constructed from embodied actions (see for example Reeves, Brown & Laurier 2006, for an account of the practice of playing *Counter-Strike*).

The findings will be presented in terms of excerpts and analyses of video recorded gaming sessions. Each excerpt shows some aspect of how collaborativity and togetherness appear, for the players themselves, in and as embodied interaction, on and around the screen. The focus of this article is on embodied actions and practices of gaming. It is not an ethnographic account of the Internet café in which the gaming takes place.

**Prior Studies**

There are several types of studies to which this article is indebted. LAN parties and Internet cafés have been studied with various survey techniques and ethnographic methods (Beavis, Nixon & Atkinson 2005; Jansz & Martens 2005; Jonsson 2006; Légrain & Stewart 2003), exploring for example the players’ motives for playing in these specific locations. These kinds of studies, while instrumental in providing an understanding of co-present gaming as a social phenomenon, rarely analyze how the games themselves are played. In contrast, the present article foregrounds these very issues. In line with ethnomethodological ideas, it treats collaboration and sociability as members’ concerns, that is, issues to be studied through the members’ own demonstrated orientations in their gaming activities. Playing computer games must somehow, for the participants, be understandable as “playing together”, i.e., be *accountable* as playing together.

Studies that broadly fit within CSCW and ethnomethodological studies of work (Goodwin & Goodwin 1996; Heath & Luff 2000; Suchman 1987) are useful for the ways in which they show how collaborativity in computer mediated activities is a members’ concern and something which they *achieve* in those activities. They demonstrate how em-
bodied action and the use of (screen based) artifacts in specific settings cannot be studied as separate types of actions, but as part and parcel of the same situated activity.

Gaming has been studied using various forms of interaction analysis, and some of these studies have been illustrative in showing how concepts and methods of CSCW can be applied to activities that are not considered work in the traditional sense of the word (Crabtree, Rodden & Benford 2005; see also Brown & Barkhuus 2007). Crabtree, Rodden and Benford (2005) demonstrate how playing a pervasive game is accomplished through embodied engagement with the game’s artifacts.3

Games studies informed by interaction analysis have demonstrated that, for example, players’ understanding of the game varies depending on their interaction, and that players often focus on the functionality of in-game features rather than their “theme”, i.e., what they represent (Linderoth 2004, 136ff). Also, studies have shown how children playing computer games deploy response cries and thereby conduct the gaming in coordinated ways (Aarsand & Aronson forthcoming) and demonstrate how computer gaming is used as a resource in inter-generational interaction in families (Aarsand 2007b). These studies are important in that they show how gaming is performed both as a situated phenomenon and as sequences of embodied actions.

Moreover, interaction analysis has been used in gaming studies of on-screen, in-game actions. Reeves, Brown and Laurier (2006) have developed a phenomenological and ethnomethodological account of gaming, demonstrating how expertise in playing Counter-Strike can in part be understood as an embodied phenomenon in the sense that gaming competence is linked to skillful movements of the avatar in the game space. Accountability in online gaming and virtual worlds has been explored in detail, examining the discrepancies between avatar-mediated interaction and face-to-face conversation (Moore, Ducheneaut & Nickell 2007). Manninen (2003a; 2003b) offers accounts of the various interactional possibilities of contemporary multiplayer gaming. While these studies all discuss interaction in computer games, however, they tend to overlook the embodied resources that the players have at hand: only their on-screen resources are analyzed.
Embodied Action

The collaborativity that this study examines is, in contrast to much prior research, explored through analyses of sequences of embodied and situated action (Dourish 2001; Goodwin 2000; Suchman 1987). Such sequential analyses of the detailed features of actions, both on- and off-screen, facilitate an account of collaborative gaming that adheres to the participants’ own common-sensical understanding of their situation in the course of doing those actions. Collaborative gaming is seen as a reflexively accountable phenomenon (cf. Dourish 2001, 78-81; also Garfinkel 1967, ch. 1; Garfinkel 2002): the actions that constitute the gaming, both on- and off-screen are the very same actions by which the gaming gets understood, and the situation (context) in which this is done is constructed from those very actions (Dourish 2004).

Gaming actions (actions performed on or around the screen in the gaming session) can be seen as embodied actions. They are actions that occur in real space and time, and that are performed through physical manifestations in those dimensions (Dourish 2001, 100ff.). The body is understood as a locus for producing and understanding the situation that the participants inhabit (Goodwin 2000), that is, for making it meaningful. However,

‘embodiment’ does not simply mean ‘physical manifestation’. Rather, it means being grounded in and emerging out of everyday, mundane experience. [...] embodiment is a foundational property out of which meaning, theory, and action arise [...] embodiment is a participative status, a way of being, rather than a physical property. (Dourish 2001, 125)

Embodied action is “the creation, manipulation, and sharing of meaning through engaged interaction with actifats” (Dourish 2001, 126). Playing computer games should not be understood primarily as “perception”, but as skilled bodily engagement with the game-world (Walkerdine 2007).

Exploring the achievement of collaborativity, moment by moment, is done here using Goodwin’s (2000, 1490) notion of semiotic fields and contextual configurations. In this view, accountable action is produced through members’ assemblage of resources from various semiotic fields. In
the gaming studied, these include talk, gestures, posture as well as on-screen action. A locally (in precise space and time) relevant assemblage of semiotic fields, to which the participants demonstrably orient, constitutes a contextual configuration. The exact configuration therefore varies moment by moment. Which one is relevant at any particular moment is a negotiable matter. These semiotic fields elaborate each other so that, for example, pointing becomes meaningful for participants in its temporal and spatial relation to other means of communication. Using contextual configurations as a framework for analysis allows for the simultaneous exploration of a variety of semiotic resources, such as talk, gestures, postures, and locomotion, both on- and off-screen. This is something that is of paramount importance when investigating a highly visual domain such as computer gaming in order to avoid focusing the analysis on spoken interaction only.

Even though there have been studies of the use of semiotic resources in computer mediated interaction (e.g., Heath & Hindmarsh 2000; Klerfelt 2007), or in pervasive games (Crabtree, Rodden & Benford 2005), the contextual configurations in the playing of off-the-shelf commercial computer games are still in need of further exploration, especially the interplay between on- and off-screen semiotic resources.

Method
In this study, gaming is examined as actions-in-interaction through a careful analysis of video recorded gaming sessions. These were obtained in an Internet café in Stockholm during a period of 10 weeks in 2007, and are all examples of collaborative gaming in one sense or another. Around 10 hours of gaming were recorded, in a total of 12 different sessions, ranging in duration from 10 minutes to over 5 hours. Moreover, the café was visited, in between recording sessions, by the analyst on at least 5 other occasions. These visits were deployed to participate in the regular activities at the café, building rapport with its regular attendants. Informal interviews were conducted in order to better understand the setting in which computer gaming took place, as well as the people participating in it.

Recordings were made during various times of the day, both in the early afternoon, when there were not many visitors, as well as later in the afternoon and evening, when the café filled up with young people drop-
ping by after school hours or after dinner. Also, one whole night was spent recording a “night gibb”, where players play computer games at the café for up to 10 hours straight.

On average, each recording is around one hour in length. The recordings were made using one or two video cameras and external microphones. Of the 12 sessions recorded, 5 were shot using two cameras. The players filmed were all between the ages of 11 and 21 (the majority of players in the café are in this age bracket), all of them male except for one female player. This is a representative sample of the regular players at the café. Even though there are often female players present in the café, they only rarely play computer games there. More frequently, they engage in collectible card games such as *Magic: The Gathering* (1993).

The players in the recordings were not arranged in any way so as to facilitate the collection of data. Instead, they were approached in the configuration that they had already established, so that the ways in which they conducted their activities were not the product of the researcher’s arrangements. It should be noted, however, that the constellations of players that were recorded were all deemed to be filmable, i.e., at least one camera angle framing both the player and screen could be obtained.

Relying on video recordings of interaction allows for the close examination of how gaming actions construct the collaborativity of the gaming, that is, how both on- and off-screen semiotic resources are assembled to make up the gaming session, but field notes were also taken, noting such details as who the players were and what game they were playing.

The cameras were not arranged as a result of a fixed formula, but were set up in order to be able to capture interaction features deemed as most relevant. Since collaborativity in gaming is the focus of the study, a variety of situations in which collaborative gaming occurred were recorded. This includes:

- single players playing multiplayer games over the Internet,
- multiple players playing multiplayer games together over the local network or on the Internet, as well as
- multiple players playing different games side by side while still being oriented towards each others’ gaming.
In most cases, one camera was set up so that it framed the players’ screens, their upper bodies and a side view of their faces, allowing for analyses of gaze and facial expressions. When playing on two computers side by side, this meant that one player is closer to the camera, showing more detail on his screen, and sometimes obscuring the other player altogether. The other camera (where this procedure was deployed) was set up to capture either

- a view symmetrical to the one described above, but more focused on the other player’s screen, or
- a general overview of the situation, e.g. the two players seen from a distance.

The data material is therefore not uniform, and is not suitable for an analysis of, e.g., the frequency of a certain event, as it would not have been observable in all the camera views. Instead, the analyses developed in this article should be seen as examples of embodied action in gaming. The relevance of an embodied resource is in part determined by the way in which the data was recorded. By setting up the camera(s) in these ways, an analysis of which modes of interaction that are relevant to gaming has already begun (cf. Mondada 2006). Positioning the camera in such a way as to not include the whole room, for example, displays an analytical disposition to regard interaction happening on and in near proximity to the screen as being “the most relevant interactions” for the gaming session.

The videotapes were scrutinized for episodes of interest for this article, and relevant episodes were then transcribed, using conventions of conversation analysis (Atkinson & Heritage 1984), but with extra care taken to preserve details of various forms of bodily actions as well as the organization of speech (see appendix 1 for a summary of transcript symbols). The transcripts were used as means for facilitating analysis of the data (the video recordings), and were not considered data in and by themselves, even though this relationship is a complex one (Ten Have 2004). Transcripts were used in conjunction with viewing the video recordings, thereby allowing the transcript to illuminate the complexity of the multimodal interaction. In this article, the transcripts are presented as translations, with the original Swedish transcript included for reference.
The excerpts were taken from a collection of transcripts. They have been chosen among a multitude of possible excerpts in order to exhibit common properties of collaboration and embodiment in co-present computer gaming and to provide illustrative examples of these phenomena. The analysis revolves around two very similar games, both based on Warcraft III: The Frozen Throne (2003). These are commonly played in the café, as is World of Warcraft (2004) and Counter-Strike (2000), but the extent to which phenomena of collaboration and embodiment are dependent on the precise game played is not discussed further in the article. The four excerpts in this article are taken from three different gaming sessions, involving six different players.

Analyzing Video Recorded Computer Gaming
Details of on-screen events turned out to be a somewhat problematic issue in the recording of gaming sessions. The researcher attempted to record gaming in a natural setting, and did not arrange the players to suit the analytical needs of the study. Instead, the methods used had to be able to capture gaming interaction without being intrusive to either the players or the interests of the Internet café in which the recordings were made. This meant that video-capture software could not be installed, and that the study has had to rely on video cameras to record on-screen events.\(^1\)

Obviously, the players, in all the games played in the recordings, are presented with a lot more information than is visible in the videos. These details therefore become unavailable for analysis. However, the interaction analytical focus of the study enables analyses of visual details that are not directly visible in a recording, but are observable in the ways they are made interactionally relevant. The semiotic fields (Goodwin 2000) that are used as interactional resources by the participants are not already in the situation, prior to any engagement by the parties into that situation. The relevance of any contextual feature is produced as a matter of course, moment by moment, rather than being omnirelevant factors ready to be discovered and used in analyses (cf. Dourish 2004; Garfinkel 2002). For example, consider this gaming episode, involving the players G and K:
Excerpt 1

Recording 2 03:15  
Players G and K  
Game: WC III, Angel Arena  
1 K ((walks with his hero towards some enemies and attacks))  
2 K ↑Wu::h check out how it crits >500 500 500 500 500<  
↑Wu::h Kolla va den critta >500 500 500 500 500<  
3 G and it happens often too  
4 (the enemies die, and K examines the loot using the mouse cursor))

This episode is analyzable for its visuality even though the actual numbers are not visible to the analyst examining the video data (the fact that numbers do show up on the screen each time a critical hit is scored is known from experience of the game). However, the way in which Player K produces his utterance as an exclamation and a notification of something out of the ordinary is available for analysis. Player K, on line 2 (the exclamatory response cry “↑Wu::h”) shows that the on-screen events are something unusual and worth noticing. He specifies what it is that is worth noticing in two ways: first, by an invitation to “check out how it crits” and then, by specifying that it is the amount of damage done (500) and the speed by which this happens that is worth noticing. K’s speedily delivered “>500 500 500 500 500<” mimics the way in which numbers show up on his screen, and serve to highlight the rapid succession of critical hits (for further discussion on mimicking on-screen actions, see Aarsand & Aronsson forthcoming). What the participants see in their visual field is available for analysis through their practices of accounting for it.

Player G’s response cry and answer underscore that he takes Player K’s exclamation to be about the amplitude of the hits, i.e., the large amount of damage (500 per hit) that the critical hits do. To this, Player G adds more information regarding what it is about the event that is perceived as unusual (“and it happens often too”). This utterance is a further specification through the use of “and”: player G is providing additional important information about the event, thereby specifying rather than refuting Player K’s remarks. This excerpt is interesting as it demonstrates how
Players K and G’s cooperative gaming can be seen as an instructive sequence where the players demonstrate different levels of competence in the game. While Player G approves of Player K’s comprehension of the event, he apparently finds it in some ways insufficient and in need of a “more accurate” account. How the game should be seen and understood in a locally competent way is not something taken for granted, but something that the players themselves negotiate. Through instructing Player K on how to see the game, Player G is demonstrating something that could be understood as a professional vision of playing this game (Goodwin 1994).

When examining this event, it can be seen that the participants are indeed formulating the relevant visual details on the screen into features of verbal interaction. These details are not just some objectively present part of the user interface, but are actively lived, experienced and referred parts of the gaming session. The sequential analysis of gaming interaction allows for the analyst to grasp what the players are talking about through the “proof criterion” that subsequent actions provide interaction (Sacks, Schegloff & Jefferson 1974, 702). Interaction is seen as reflexive, whereby is meant that each action is added to the prior chain of events. The meaning of each action or utterance is shown in the actions or utterances that follow it (ibid.). In cases where there are misunderstandings, these are demonstrated in, and as various forms of, repair sequences (Schegloff, Jefferson & Sacks 1977). The sequential order of actions is both a resource for the participants of the interaction as well as for the analyst (Sacks, Schegloff & Jefferson 1974, 702).

Analyzing what the participants themselves demonstrate as the relevant semiotic events is useful for research as

we can use the visible orientation of the participants as a spotlight to show us just those features of context that we have to come to terms with if we are to adequately describe the organization of their action (Goodwin 2000, 1509).

The contextual configuration should in this view also be a delimiting framework for the analyst.
Games Played

The games played in all these excerpts are based on Warcraft III: The Frozen Throne (2003; hereafter referred to as WC III). The variants played are Defence of the Ancients (2002; hereafter referred to as DotA, in excerpts 1 and 4), and Angel Arena (2005; in excerpts 2 and 3). They are both modifications, or mods, of the standard off-the-shelf game. While utilizing the game engine, graphics and themes of WC III (and sharing its interface), these mods forfeit the resource management system of the original game, and each player only controls one avatar (called a hero) instead of an entire army. Each hero has different capabilities and possible developmental trajectories. The hero is upgradable through various items found on the playing field, through upgrades bought using gold gained as loot from killed enemies, and through experiences gained when killing enemies. The games last for about one to two hours, and it is not a persistent game: the game starts anew every time it is played, although players may develop a proficiency in playing a particular hero.

DotA is a highly popular WC III mod, with international tournaments being held at regular intervals. The game is played in teams of up to five players on each side, who try to work their way across a battlefield and, with the aid of computer controlled troops, destroy an “ancient” (an important building) located in the enemy base. The enemy can be another team of players or it can be computer controlled. Angel Arena exists in many variants, generally consisting of a battlefield, littered with computer controlled enemies, with which the players (in teams or one-on-one) battle in order to develop their hero and defeat the opponent a number of times. Enemies are, to some extent, modeled on Christian characters, including the Devil, Baby Jesus and God.

The interfaces of the two games are similar, as they are both modifications of the WC III-platform. The battlefield is several times larger than the screen, and therefore only a small proportion of it is visible to the player at any given moment. The game features full 3D graphics, with the possibility of positioning the camera at any angle. The default setting, and the one used in all the examples in this study, is a bird’s-eye view, showing the battlefield from above. Players regularly scroll the battlefield in order to be aware of the geography and activities in more distant on-screen locations, using the keyboard or mouse.
The hero is controlled by clicking on the ground, making the hero walk there, or on an enemy, making the hero walk there and attacking it, using its default attack. Special abilities are available by pressing specific buttons, either in the “inventory” or “abilities” panel (see appendix 2 for an example of the interface in WC III). These special manoeuvres are more powerful, but can either be used only seldom (once the ability has been used it is on “cooldown” and cannot be used again until some time has passed) or when the magical energy (“mana”) is expended, without which the ability cannot be used, limiting the frequency of use. Each hero can, after progressing a certain number of levels, learn the “ulti”, an ability specific to each hero. This ability is very powerful, but generally also has a very long cooldown.

Other major features of the interface are the minimap and the stat-panel. The minimap displays the whole battlefield in a miniaturized version. Clicking on the minimap instantly moves the view of the battlefield to that point (but does not move the hero). This can be used as an instant scrolling, in order to be aware of enemies or battlefield geography in distant locations. The stat-panel displays the hero’s characteristics (strength, intelligence and agility), how many experience points he has accumulated, and a bar showing how many more experience points the hero has to gather before reaching the next level. The top left corner of the screen is home to a portrait of the hero, flashing red every time he or she takes damage, as well as two bars that are measures of health-points and mana respectively.

**Resources for Demonstrating Collaboration**

The excerpts presented here all demonstrate different aspects of collaboration in co-present computer gaming. The analyses focus on the various semiotic resources that the players assemble in order to play the games together, and what kinds of interactions the use of these resources make possible.

**Posture as a Resource for Collaboration**

In local multiplayer games, the players are often distributed in the room in accordance with available computers. This calls for ways of displaying collaborativity that accounts for the layout of the room and the distribu-
tion of players. In the example below, eight persons in the room are playing DotA over the local network. Player G is seated about five meters away from the other players, but is still involved in the on-screen action via his hero. The seating arrangements in this gaming session are shown in appendix 3. Leaning to the side and back, Player G displays bodily orientation towards the other players, frequently turning his head away from the screen in order to be able to monitor what the others are saying and doing. This excerpt demonstrates how Player G’s interaction with the other players is structured by spatial relations in the room where they are situated, as well as by the ongoing on-screen actions.

**Excerpt 2**

Recording 8 52:53
Players: G, in view, and Player D, as well as six other players (A, B, C, E, F, H)
Game: WC III, DotA

1. G \((\text{engages enemies with his hero})\) (5 s.)
2. G \((\text{turns head towards the other players})\) But what’ll happen later
ne va händer sen när
3. when you get ganged \(=\) yeah
du blir gangad \(=\) a:
4. now< for example?
nu< till exempel?
5. D Yea but why would I do that cause
A men va skulle ja de för
‘cause I used the \(\text{ulti?} \) \(=\) [really]
för att ja lagt \(\text{ulti?}\)
6. \[\text{[egentligen]}\]
7. \[\text{[Yaboj]}\]
8. \[\text{[A men]} \] om de
9. it was let’s say only two heroes \(=.5\)
va bara två heroes säger vi \(=.5\)
10. use cyclone when you can
använd cyclone när du kan
11. (2 s.)

(Cont’d.)
In this excerpt, Player G has positioned his body so that he can take part in the ongoing gaming not just on-screen, but also in the ongoing activity in the room where the game is played. Leaning towards the majority of other players, who are seated with their backs to him, he displays an orientation to them which affords engagement and shows willingness to take part in their conversation. His bodily position is thereby used as a semiotic resource in constructing a participatory framework in which he is participating in the gaming interaction. His distal spatial position in the room requires him to position himself so that he is able to monitor and show the others he is monitoring the gaming session. In other words, he is attempting to be observable and reportable (i.e. accountable) as being a player in the gaming session as a whole, not just as a participant in the on-screen interactions.

In his leaning position, he minimizes the bodily work required for sur-reptitious monitoring (Hearth & Luff 1992) of the ongoing activity. This monitoring is however also structured by his own engagement in the on-screen activities. He cannot leave his hero completely unattended (apart from when being inside his base), in which case he could get attacked by enemy heroes (player characters) and computer controlled troops. Therefore, it is only after he has engaged some enemies (on line 1, letting his hero auto-attack, i.e., continue until he or the enemies are dead), and made sure that this fight will continue for some time, that he can turn his attention away from the on-screen action of his own screen and direct to activities that the other co-present players are part of.

The interactional project that Player G is pursuing in this excerpt can be said to be one of ensuring his place in the playing of the game, even though he is not placed in an ideal spatial position to do this. He does so by displaying that he is aware not only of what goes on in the game, but of what happens specifically on other players’ screens (lines 4 – 6)
and in the surrounding conversation. This excerpt also shows how the placement in a gaming environment is not arbitrary or something describable merely using objective measurements, such as showing a map of the room (see appendix 3). Rather, in the participants’ interactions, Player G’s placement as peripheral becomes a demonstrable phenomenon by way of his specific interactional work to display himself as a player in the gaming session. This can be seen as an example of the inherent meaningfulness that members orient to in relation to spatial positionings (Dourish 2006). As in this episode, the contextual configuration differs depending on the distance between players. In other words, the semiotic resources that are available to them are contingent upon spatial relations.

It is significant that Player G launches an instruction sequence when he takes the opportunity to enter into the other players’ conversation, claiming that a substandard tactic has been used (using up the “ulti”) which leaves the player vulnerable to attacks from other players (lines 4-6). The specific instructiveness of Player G’s comments is shown in how he, when leaning closer to Player D, spots on-screen “trouble”: player D’s game play stands in need of improvement. When not requested, issuing advice can in this situation be seen as an attempt by Player G to position himself as more competent in playing the game than Player D. This piece of advice is redesigned in relation to specific on-screen events, apparent by the way Player G reformulates the advice from a general query about recurrent game states (lines 4-5) to a specific one related to what is happening currently (“yeah now for example”, line 6). This is done in order to show that Player G’s advice is relevant “right now”, making Player D accountable not only to his general skills as a player, but to his ability of being able to handle this specific situation.

Player D, however, resists taking instructions in his game play by questioning the reason for Player G giving the advice in the first place. Apparently, Player G is implying that there is a relationship between what spells are “on cooldown” and the risk of being attacked by other players (lines 4-6). Yet to Player D, using the “ulti” does not cause players to attack you (lines 7-8). Player G attempts to remedy his failed attempt at giving advice by specifying a condition under which his previous utterance can still be seen as good advice (lines 9-10). It is interesting that he does not persist in this specification of his earlier utterance.
Clearly, he finds he is better served by cutting to the chase, namely aborting his attempt to clarify the necessary conditions and instead just instructing Player D what to do (“use cyclone when you can”, line 11). Player D’s response following this (“I can’t”, line 12) is designed once again so that the condition for following Player G’s instruction is not fulfilled. The instruction to use cyclone “when you can” is denounced in and through Player D claiming that “I can’t”. This can be seen as a demonstration of how Player D considers Player G’s advice invalid for his game play at this time. The episode ends with Player G once again orienting to his own screen and Player D interacting with the players closer to him.

Even though, as shown below in excerpts 3 and 4, interactional resources have to be deployed in order to secure mutual attention between players even when seated next to each other, the resources are of a different order compared to when players are spatially separated. Leaning is a resource that is more visible in the room, but also one that, to a greater extent, changes what persons are available for the player’s attention and interaction. It opens up a new path of interaction in the activity which thereby contributes to the sense of collaborativity in the gaming. The players are neither isolated from each other nor collaborating solely by virtue of co-presence. It is their interactional work that makes this possible, and Player G’s peripheral spatial position necessitates work on his part in order to secure his place in the activity. This can be seen in the way he, when he receives attention from the more centrally placed player, tries to secure additional influence by attempting to demonstrate higher gaming competence than Player D. Player G’s attempt to secure interactional leverage in the gaming session by attempting to display better game-playing skills fails when Player D simply says “I can’t” (line 13) in response to G’s suggestion to use another of the hero’s abilities. Player D’s way of dismissing G’s suggestions demonstrates how spatial relations in the room affect participation in the gaming. Player D is not obliged to specify or further formulate why he cannot follow G’s advice: instead he can simply focus on the on-screen actions of his game. The placement of Player D’s screen makes it easy to focus in a direction facing away from Player G. By dismissing the premises presented by Player G altogether, Player D makes further comments unnecessary.
This is not to say that the spatial relation between the other players (e.g. between A and E, which is at least as far as between D and G) makes interaction easy. However, players being seated right next to each other have semiotic resources for interacting that are not available to Player G at this point. The contextual configuration of semiotic resources more readily accommodates players seated next to each other. The next two excerpts demonstrate some of the semiotic resources that players use when seated in close proximity.

**Resources for Creating a Shared Game Space**

In spatial configurations where the players are seated side by side, as in the next excerpt, the players can monitor each others’ on-screen activities more easily than in the previous example. In these cases, being observant of the other player’s actions can be done by glancing at the other person’s screen. But looking at the other player’s screen is not something that “just happens”. It is related to locally defined interactional projects and trajectories of action, and cannot be understood isolated from its sequential context. The next excerpt shows how Player T uses glancing to keep track of what Player S does, in order to make sure that he does not die unnecessarily and that he understands and learns how to play the game.9

In the episode (and all through this gaming session) Player T is actively instructing Player S on how to play DotA. Player T functions as an informal, self-appointed “teacher” in their gaming session, guiding Player S through interface, game mechanisms, and tactics in order for them to make progress and win the game. Player T is experienced while Player S has never played the game before.

**Excerpt 3**

Recording 3 09:23, using two cameras
Players: T and S
Game: WC III, DotA

1

S  

((gets attacked by many enemies, tries to escape by scrolling the screen and clicking on an empty part of the ground. He scrolls back to his character’s location, but is unable to find him))

2

T  

This is unplayable look >it’s like< all.

De här går ju inte å lira kolla >de e så här< helt.

(Cont’d.)
(Cont’d.)

5 (3 s.)
6 S ((looks around the screen))
7 S ↓Damn.
   ↓Fan.
8 T ((quickly looks at Player S’s screen)) Did (. ) did you die?
    Do- (. ) dog du?
9 [no (. ) you’re alive< ((leans over closer to Player S))
   [ne] (. ) >du lever<]
10 S ((moves cursor down to stats<))
11 S ((scrolls the screen down, hero comes into view))
12 T There you are ((looks on his own screen)) [then just to go down
   här e du]
13 =
   här e du bara å gå ner
14 S [[(looks at Player
15 T’s screen)]]
16 T =( (. ) Down here where you were buying]
   =( (. ) här e du handla)
17 =
18 T [[[moves cursor to minimap, clicks on the home base and the
   screen scrolls there]]
19 =
20 =
   ((Player T continues showing Player S how he can heal the damage
   he has suffered))

In this episode, Player T, being the more experienced player, exhibits a watchful eye over Player S. While first formulating an account of why the gaming environment is out of the ordinary (on line 4, a continuing topic in their conversation, and something that amounts to “working with constant interruption”; see Crabtree, Rodden & Benford 2005), he nevertheless continues playing. When Player S cannot find his character (lines 1-3) and says “damn” (line 7), Player T quickly reacts in response (line 8). Such response cries (Goffman 1981, ch. 2) have been shown to elicit co-players’ attention to key events in the game (Aarsand & Aronsson forthcoming). This swiftness shows an active orientation to the progress of Player S’s game, and that they are cooperating in the game session. In many other instances of the data, such response cries by the players are not immediately reacted to, but instead the next turns by participants are delayed until the game allows for verbal interaction. Here,
however, the swiftness of Player T’s response demonstrates engagement in the events on the other player’s screen.

On line 8, Player T produces a first interpretation of Player S’s response cry (line 7) by formulating a query (“di- did you die?”). This formulation of a possible course of events shows that Player T found that Player S’s exclamation was well founded. Such a response cry is something that Player T shows should be done in relation to a major event in the game, such as dying. Player S at the same time moves his cursor down to the hero’s stats (line 10), in order to find out more about its current state (i.e. if he in fact has died or not). This movement displays Player S’s orientation to the fact that this is a prime on-screen location for finding out more about the hero’s current status. At the same time, Player T leans closer to get a better view, and simultaneously responds to his own question (“no, you’re alive”, line 9). At this point, Player S understands that hope has not vanished, and he scrolls down to find his hero standing where he had originally intended it to go. The utterance produced by Player T on line 12 (“You’re right there”) demonstrates understanding, in conjunction with his prior utterance that (i) the player’s stats are incompatible with the death of the character and (ii) that the reason for Player S’s original response cry was that he incorrectly thought that he was dead when he had really just lost track of his hero. Saying “there you are” (“där är du” in Swedish) also performs a specific task: it shows that Player S’s first exclamation was uncalled for.

In the continuing sequentially ordered interactions, Player T also shows, as an instruction, that the situation is accountably “severe” in that the hero has sustained too much damage to be able to continue fighting. When Player T says “then just to go down here” (lines 12-13), he shows that he foresees the next action by Player S in order to remedy his problems. That Player T projects some action by Player S demonstrates a competent understanding of the situation in the game (a “professional vision”, see Goodwin 1994). Player T articulates what Player S should be doing and seeing, and how he should be doing it, i.e., Player T is issuing instructions. Player S clearly recognizes this in the way he observably orients to what Player T is doing on his screen. On line 12, Player T looks back at his own screen, and Player S follows his gaze (on lines 14-15), accountably “listening to instructions”.

-437-
In this example, gazing is a demonstrably relevant resource for the gaming interaction. By gazing and leaning, Player T is responding to Player S’s exclamation. By examining such adjacency paired structures (cf. Sacks, Schegloff & Jefferson 1974), analysis can show how collaborativity and intersubjectivity are constructed in a specific situation. Mutual attention, and thereby collaborativity, is demonstrated in how the players construct, first, Player S’s screen, and later, Player T’s screen as the locus for the relevant gaming actions. By looking back at his own screen, and also by assuming a regular non-leaning posture, Player T’s actions serve to elicit attention to his own screen, where a solution to the problem can be found (showing Player S where his hero can be healed). They are also orienting to, and thereby producing as interactionally relevant, the larger framework of game playing, such as the rules of the game. Player S’s confusion and suspicion that he has died, and Player T’s instruction for a solution to his low-health problem, are only intelligible for the participants within the situated activity system (Goffman 1961; see also Heath & Hindmarsh 2000) of this game, where the hero’s low health can be seen as a problem to be solved.

**Pointing and Scrolling**

In the episode discussed next, two of the players are using computers, seated next to each other, while the third player G actively comments on the ongoing action. Just prior to this excerpt, K has announced himself as the eventual winner of the game, something that both Player L and Player G refute. Player K picks up on a prior topic, where Player L has complained of how badly his hero is performing. In the pictures, Player L is seated closest to the camera, followed by Players K and G. The pictures in the excerpt have been edited in order to show the pointing more clearly: the original recording shows more of the players.
Excerpt 4

Recording 2 09:00
Players: L, K and G (from left to right)
Game: WC III, "Angel Arena"

1 K But why won’t you change character men varför byter du inte gubbe
   then? ((turns his palms upwards)) där?
2 G ((coughs)) Now you’re buying life Nu köper du liv
3 L You can’t De går inte
4 K Yea you can= (((moves his right hand Jo du kan-
   towards the screen)) )
6 L (((image 1: clicks on his hero and moves the cursor towards
    the exit of the base)) )
10 K =enter this one (((image 2: waves hand
    går in i den här downwards several times, brings his
    hand down again)) )
13 L (((moves cursor down towards the bottom of the screen)) )

(Cont’d.)

Image 1, excerpt 4: Player L is about to exit on the left side, K moves his hand towards the screen.

Image 2, excerpt 4: K points downwards several times.
In lines 1 and 2, Player K suggests a way for Player L to improve what he does not like about his hero. Player K does this with a distinct gesture and while looking at L’s screen. With his palm turned upwards, he points in the general direction of L’s screen. It is a pointing gesture without a deictic function, but it is used as an accentuation of K’s utterance as a suggestion, i.e. as a possible course of action. Player L, however, looking intently at his screen, claims in response to Player K that this is impossible (line 4), accounting for the mechanisms and rules of the game. Player K knows this is wrong, and shows this to Player L by directly opposing his version of the game rules, pointing with his right hand on the screen (lines 5-6). This pointing is an assemblage of several different embodied actions, the meaning of which is seen in how they are appropriated in interaction. Through varying uses of pointing, Player K tries to ensure that Player L is aware of how to change his hero and thereby enjoy more of his gaming.

Starting on lines 5-6, Player K’s pointings in combination (i.e. as a contextual configuration of semiotical fields) with his utterance show (i) the opposing opinion to what Player L has claimed in line 4, and (ii) that the solution of the problem at hand is an on-screen physical location. With a general pointing he establishes the solution as a feature of the on-
However, as the building is not in view, and the players use no other vocabulary for talking about it, it cannot be interactively handled in any concrete way yet. Player K therefore has to instruct Player L to bring this building into view so that it can be handled and talked about. Player L, in response to Player K’s suggestion to “enter this one”, (together with his hand suggesting to scroll the screen down by waving it downwards, lines 10-12) moves his cursor towards the bottom of the screen, preparing to scroll down (lines 13-14). This is also a display of gaming geography knowledge, although not very specific, as he cannot adequately project the place Player K is talking about. Player K specifies this in the next action (lines 10-12) where he specifies the direction of the scroll by pointing downwards several times. He performs this pointing in response to Player L’s demonstrated lack of knowledge of the gaming geography: Player L is demonstrably unable to find the building himself. Player L clearly understands Player K as instructing him to scroll down (line 16), which brings the building into view. Still, Player K has to make sure that Player L does not scroll the screen too far, as the building has to be kept as an interactable object, that is, within the coordinated co-orientations of the players. He does so by pointing at a specific building, at which point B stops scrolling. In this contextual configuration of semiotic resources, the current view of the screen delimits what gaming actions can be taken.

It has been demonstrated how collaborative game play is an interplay of semiotic resources in different modalities. The meaning of an action in one modality (such as Player K’s pointing on line 10) is shown as a response in a different modality (Player L scrolling the screen downwards). It is of no concern whether or not this resource is grounded on-screen or off-screen. Adjacency pairs (Sacks, Schegloff & Jefferson 1974) are constructed utilizing various embodied semiotic resources, such as pointing (of different kinds), in conjunction with talk. In producing relevant second pair-parts of adjacency pairs, the players establish their gaming as a collaborative activity. When Player K waves his hand downwards and says “enter this one” (line 10) he is issuing an instruction on how to solve the current problem of trying to find the relevant building. This instruction is accountably followed by Player L when he moves the cursor towards the bottom of the screen (lines 13-14) and starts scrolling down.
(line 16). Player L’s on-screen actions can only be understood as following Player K’s instruction through their sequential placement. Producing the pair “instruction” and “following instruction” is a way in which the players’ use of the game can be seen to be collaborative. This is altogether separate from the fact that they are in a competitive situation, in which one will eventually be the winner.

**Concluding Discussion**

This last part of the article provides a broader and more “lateral” analysis of some phenomena that are present in the excerpts. Common themes are discussed, together with what these reveal about collaboration and embodiment in computer gaming.

**On-Screen and Off-Screen Gaming Actions**

Coordination in and through gaming actions happens on different levels and in relation to different interactional contexts. Mutual orientation to the contextual configurations can be considered as a prerequisite for collaboration in gaming. A feature of all the episodes explored in this article is the participants’ simultaneous orientation to on-screen and off-screen space. For example, in excerpt 4, it is through the coordination of off-screen pointing and on-screen scrolling that the previously only talked-about feature of the on-screen geography (the specific building) can be brought into view, thereby allowing other types of on-screen actions (such as entering the building), and off-screen actions (such as pointing directly at it). As the participants do not orient to on-screen and off-screen resources as in any way qualitatively different from each other, the analyst should reconsider whether or not to stipulate a major distinction between them. In these examples there is nothing about gaming that is “virtual”. Players are neither outside the game nor inside it, no more than they are “outside” or “inside” other embodied actions, be it writing, drawing or playing soccer.

Secondly, it is important to see how these gaming actions relate to broader communicative projects in the sessions. In excerpt 2, Player G’s actions serve to ensure him a place in the (spatially) more centrally placed players’ gaming. Even though they are playing over a local network and have symmetrical access to the on-screen actions, their modes of off-
screen interaction matter just as much as what happens on the screen. Understanding the contextual configurations in the gaming situation involves seeing how gaming actions are composed of both on-screen and off-screen elements. The interactions, composed of the locally relevant semiotic fields that make up the gaming actions are not separated from each other, but are sequentially intertwined, regardless of whether or not they are produced using on-screen or off-screen resources. In the present analyses, therefore, on-screen and off-screen resources cannot be separated. Gaming actions are accountable embodied actions, produced in and as a sequential order on the screen and around it. Instructions are, for example, often produced as assemblages of talk and gesture (such as pointing) followed by non-verbal uptake, actions like avatar movement or scrolling of the screen (this is what Player L does in order to follow Player K’s instructions in excerpt 4).

Moreover, it is striking (but perhaps coincidental) that all the present examples contain instructional sequences in one form or another, and that the players assist each other through advice, pointing and explaining. What is noticeable is that this also happens when they act in opposing teams (such as in excerpt 4). Even competitive gaming in this setting thus contains elements of collaboration. Instructive sequences therefore indicate, among other things, that players recurrently demonstrate an orientation to the overall long term enjoyment of all players, including opponents.

**Professional Vision in Gaming**

The semiotic fields that assemble gaming sessions involve both general resources used in everyday interaction and gaming specific competencies. General resources include bodily interaction such as gazing, posture and pointing, but also talk and prosodic and sequential features of talk. However, in order to produce and understand gaming interaction, the players also utilize a set of resources specific to gaming and games. The use of these resources serves both to demonstrate the participants as competent gamers and to render the activity opaque (Goodwin 2000) to the outside observer.11 It is the use of these esoteric (to the outsider) language games (Wittgenstein 1953, §2) that provides the rationale for studying gaming in situated gaming settings rather than, say, in a laboratory.
In such specific situational and embodied language games we might find a professional vision of gaming, the methods used for perceiving a phenomenon in institutionally relevant ways (Goodwin 1994), and thereby how it affords specific courses of action. Players are continuously seeing and analyzing the ongoing on-screen and off-screen action for possible upcoming actions. Being a competent gamer means to be able to foresee events in the game; to be able to demonstrably understand what will happen on the screen; what courses of action are suitable for solving upcoming situations; as well as being able to actually perform the necessary actions. Such gaming competencies are demonstrated in various ways in the present excerpts. In excerpt 2, the way in which Player G attempts to make himself part of the gaming session involves seeing and demonstrating how Player D’s tactics are substandard in a possible upcoming event (getting attacked by a group of other players). Excerpt 3 shows how Player T, through the use of embodied semiotic resources, guides Player S through a difficult part of the game, and shows him what actions need to be performed in order to progress in the game. Excerpt 4 explores the finely tuned interactions that go into demonstrating knowledge of on-screen geographies, and how on-screen actions such as scrolling, moving the cursor and the avatar, as well as off-screen actions such as pointing, make the players’ actions understandable to each other.

Pointing and scrolling towards something not yet seen on the screen demonstrates the players’ competence in the game: they are able to foretell what will happen next on the screen. Being able to project future actions and events, and thereby to plan future actions to be taken, is a central characteristic of a competent player. Such projections are shown in several of the excerpts. In excerpt 2, Player G issued advice and instructions based on future and present trouble. In excerpts 3 and 4, the players demonstrated competence in the game by being able to locate specific places in the on-screen geography, and by being able to instruct their co-players on how to find these places as well.

Local Competence
In all the excerpts, players variously prove themselves competent players, and instruct each other on how to see the game in the way it should be seen and played at that specific time. Showing yourself as a competent
player is not always done in order to dominate or defeat other players in the
game. Rather, displayed competence fulfills a multitude of functions.
In excerpt 2, issuing advice and attempting to demonstrate a correct and
better understanding of the game was done in order to secure a more
central position in the gaming interaction. In excerpt 3, Player T moni-
tored Player S’s gaming actions, and demonstrated a higher competence
by being able to sort out and instruct Player S on the next relevant gam-
ing action. In this episode, Players T and S position themselves as teacher
and student by issuing and following instructions. It is also another ex-
ample of how to establish collaboration in computer gaming: the more
competent gamer being vigilant and caring for what happens to another
player in the game. This kind of monitoring is accomplished through the
simultaneous invoking of a multitude of embodied semiotic resources:
gazes, postures, talks and on-screen navigations. In excerpt 4, Players K
and L are discussing the placement and function of a specific building in
the game. Through distinctively embodied resources, such as pointing,
scrolling and avatar movement, they are negotiating each other’s pro-
fessional vision and gaming competence. This episode shows players
demonstrating collaboration, but also that such collaboration involves
tensions when the players also strive to display themselves as competent
gamers. The examples presented in this article show how the gaming
practices in this Internet café are social and collaborative affairs, and fur-
ther that collaborating in this location also means presenting yourself as a
competent gamer, something that at times involves competitive gaming
actions.

Last of all, let us return to one of this article’s initial observations: the
players play games that “gamewise” are equally playable online at home,
with no other players co-present. This study can partly be seen as an ex-
ploration of what it is that makes playing these games collaborative
events. Even though gazes, postures and gestures might be supported in
online play (through e.g. emotes in MMORPGs), that kind of inter-
action is nevertheless different from co-present face-to-face conversation
(Moore, Ducheneaut & Nickell 2007). Online gaming interaction
should be studied on its own premises and not be seen as “flawed” face-
to-face interaction (cf. Bennerstedt 2007), as something in need of com-
pensatory mechanisms. Yet, the ways in which the players in the present
excerpts make use of situated and embodied resources in order to structure their gaming as collaborative is interesting. In a pure online context, achieving joint attention to a spatial feature on the screen cannot be done via pointing-and-scrolling (as in excerpts 3 and 4), but are resources that are uniquely situated and embodied in the off-screen context. The ways in which these local semiotic resources are used to construct collaborative play might carry some of the appeal that playing games co-presently has for players.

Björn Sjöblom is a doctoral student at the Department of Child Studies, Linköping University. He is interested in gaming as a collaborative activity in a range of settings, such as Internet Cafés, LAN-parties, gaming associations and private homes, using video recording and other ethnographic methods. His theoretical perspectives include ethnomethodology, conversation analysis and interaction analysis. In relation to children and gaming he is interested in how gaming changes the life-worlds of children, as well as views of children and childhood. This study is part of the research project “Computer gaming as an aesthetic practice and informal learning”, funded by the Swedish Research Council and the Department of Child Studies, Linköping University, project number 721-2007-3208.
E-mail: bjorn.sjoblom@tema.liu.se
Notes

2. Computer Supported Cooperative Work.
3. A pervasive game is one in which the player’s real-worldly spatial position is incorporated into gameplay, using for example PDAs, cellphones and GPSs (Björk et al. 2002).
4. Video capture software, such as Fraps (Beepa 2008), while being able to record every detail of on-screen action (including audio), requires pre-installing, start-up management, as well as large amounts of space on the hard drive. It was deemed infeasible to install such software on every computer in the Internet café, and the management of this software would have been intrusive to both the players and the operations of the café. Furthermore, it would not have eliminated the need to video-record off-screen interaction.
5. Critical hits are, in the present game, attacks that do more damage than usual hits, but occur randomly (say 1 in 10 hits). The chance to do a critical hit (English “to crit”, Swedish “att critta”) can be increased as the player’s hero progresses in the game.
6. A mod is a modification of a standard off-the-shelf computer game, often developed by gamers themselves rather than by a commercial company. A well known example is *Counter-Strike* (2000), which is a mod of *HalfLife* (1998). Mods in WC III are generally called custom maps.
7. It should be noted that the other players are not in view of the camera, and whether or not they have eye contact or are bodily oriented towards Player G is unavailable to the analysis.
8. Most of the other players are seated next to each other (see appendix 3), and therefore have other interactional means of accountably participating in the game.
9. Players T (Teacher) and S (Student) are so named because of the marked difference in competence in playing the game, with T constantly instructing and giving advice to S.
10. The part of the interface showing the hero’s current attributes.
11. To some degree it is therefore necessary for the analyst to have some knowledge of the games played (c.f. Aarseth 2007) but also, more crucially, of the setting and the local cultures in which they are played.
References


Appendix 1: Transcription Legend

[ ] Overlapping talk, bodily interaction or on-screen action

= Continuous utterance

wo- Interrupted utterance

wo: Elongated sound

↑ Raising intonation

> < Higher paced speech

damn Emphasis

( ) Micropause

(2 s.) Longer pause (in seconds)

moves cursor On-screen actions

((  ))) Non-spoken interaction
Appendix 2: The WC III Interface

Important features of the WC III interface, used in DotA and Angel Arena. The white frames are added for clarity.

1. The player's avatar, the hero
2. Minimap: showing a minitaturized version of the whole battlefield
3. Hero's portrait: showing health-points and remaining mana (magical energy)
4. Hero's stats: its level, how much strength, agility and intelligence the hero has, how much damage it does and how much armour it has
5. Inventory: special items that the hero is carrying, currently empty
6. Abilities: the special spells and skills that the hero can use

The text on the left hand side is a chat window and combat log that displays major events in the game, and in which players can also communicate with each other.
Appendix 3: Seating Arrangements

Seating arrangements in excerpt 2. The letters represent players seated in front of computers. The picture is not to scale, but meant to show the players' spatial relations to each other. The picture only includes those players and computers that are involved in this gaming session, and only those parts of the room in which it takes place.