

Coordinating Joint Activity in Avatar-mediated Interaction

¹Robert J. Moore, ²E. Cabell Hankinson Gathman, ¹Nicolas Ducheneaut, ¹Eric Nickell

¹Palo Alto Research Center

3333 Coyote Hill Road

Palo Alto, CA 94304

{bobmoore, nicolas, nickell}@parc.com

²University of Wisconsin-Madison

1180 Observatory Drive

Madison, WI 53706

cgathman@ssc.wisc.edu

ABSTRACT

Massively multiplayer online games (MMOGs) currently represent the most widely used type of social 3D virtual worlds with millions of users worldwide. Although MMOGs take face-to-face conversation as their metaphor for user-to-user interaction, avatars currently give off much less information about what users are doing than real human bodies. Consequently, users routinely encounter slippages in coordination when engaging in joint courses of action. In this study, we analyze screen-capture video of user-to-user interaction in the game, *City of Heroes*, under two conditions: one with the game's standard awareness cues and the other with enhanced cues. We use conversation analysis to demonstrate interactional slippages caused by the absence of awareness cues, user practices that circumvent such limitations and ways in which enhanced cues can enable tighter coordination.

Author Keywords

Avatar-mediated interaction, massively multiplayer online games, conversation analysis.

ACM Classification Keywords

H.5.3. [Collaborative Computing]: online games

INTRODUCTION

Since the text-based multi-user dungeons (MUDs) and MOOs of the 1980s and early 1990s [1, 8] and the collaborative virtual environments (CVEs) of the late 1990s [3, 4, 5, 13, 14], computer-mediated virtual worlds have gone mainstream today in the area of online computer gaming. Millions of people worldwide are spending an average of twenty hours a week [21] playing together in massively multiplayer online games (MMOGs). For example, *World of Warcraft* alone has over 7 million active users worldwide. MMOGs have come to flourish since their inception in the late 1990s with the proliferation of

broadband internet access and powerful PC graphics cards.

While MMOGs are primarily designed and used for recreation or play, they nonetheless constitute a new communication medium that is beginning to be used for non-recreational purposes as well. For example, Forterra has adapted the MMOG, *There*, for training the U.S. military, while Education Arcade's *Revolution* and Numedon's *Whyville* use MMOGs as a platform for education. And with new generic MMOG platforms like *Multiverse*, *Big World*, *Croquet* and even *Second Life*, we can expect to see an explosion of new kinds of virtual 3D social worlds - beyond *Dungeons-&-Dragons*-inspired combat games - in the very near future.

What makes MMOGs distinctive is the kind of social interaction they enable. Thousands of simultaneous users navigate persistent 3D worlds via characters or "avatars." In real time, they can approach and "face" each other, activate preprogrammed gestures and "talk" using text chat and sometimes voice. Avatar-mediated interaction in MMOGs is thus based on the metaphor of face-to-face conversation. It relies on users' knowledge of how to interact face-to-face and at the same time it departs from the mechanics of face-to-face in interesting ways. Past studies show that users heavily utilize the ability to position their avatars near and "facing" each other when interacting in much the same way we do in face-to-face encounters [7, 2, 22]; however, they use discrete hand and body gestures less frequently than in face-to-face [19].

While the graphical sophistication of MMOGs has greatly increased over the past decade (e.g., *EverQuest II*) in terms of texture maps, character models, physics and lighting simulation, motion capture of real human bodies, etc., the mechanics of avatar-mediated interaction have evolved more slowly. Furthermore, across today's games there is relatively little agreement among game designers regarding what the best features of avatar interaction systems are. Beyond basic avatar movement and orientation, IRC-style text chat and preprogrammed gestures, avatar interaction systems vary widely. In some systems, avatars display facial expressions; in others, they do not. In some systems, avatar orientation determines the user's view of the world; in others, it does not. In some systems, avatars can turn their heads semi-independently of their bodies to "look at" each other, maintain "eye contact," or glance automatically when in close proximity; in many other systems, avatars

cannot turn their heads, only their entire bodies. In some systems, users give some outward indication that they are engaged in composing a turn-at-chat, either the fact *that* they are typing or *what* they are typing, although in the majority, turn composition is private. Most systems implement preprogrammed gestures and animations (e.g., waves, shrugs, nods, thumbs up, dance) but they use different libraries of such animations and may or may not cause the avatar to face the recipient, couple avatar animations with text descriptions, or enable looping animations. Thus, the MMOG industry still seems to be discovering which features of avatar interaction systems are the most effective.

Prior work on avatar-mediated interaction in CVEs [3, 13] and MMOGs [15] shows how lapses in user-to-user coordination occur when various kinds of user activity is hidden from public view by design. One important component of co-presence in the face-to-face world is that actions are “accountable” or observable-reportable [10, 15]; that is, we can see at a glance what our co-participants are doing through verbal and nonverbal cues and can make use of this information when designing our own actions. For example, we can account for delays in speaking when we can see that a co-participant is rummaging through a purse, and we approach others very differently if we can see that they are on the phone versus visibly unoccupied. People “give off” such behavioral cues as a consequence of engaging in embodied activity, although they may at times manage them more or less consciously [12].

In contrast, avatars tend to “give off” far fewer cues about what users are doing than real human bodies. While avatars “do nothing,” they often conceal many kinds of work that the user is engaged in, such as browsing inventories, consulting maps, chatting with other users and more [15]. The experience of virtual co-presence can break down when avatars remain inactive for too long.

In prior work [15], we show, on a micro-interactional level, how private player activities *inhibit* tight coordination in MMOGs and how MUD-style text emotes can mitigate these effects. In the current study, we build on our prior work by showing how different types of awareness cues - avatar animations, chat bubbles, and variable-duration cues - as well as player-invented workarounds can *enhance* player-to-player coordination. This enables us to offer fine-grained design guidelines.

DATA AND METHODS

For this study, we have adapted the approaches of ethnology and conversation analysis. These fields are concerned with how people organize and achieve recognizable social activities, such as “taking turns in conversation” or “standing in a service line,” through their concrete embodied actions, and especially through their talk [10, 16]. The aim of analysis is twofold: to identify the generic systematic features of social action and interaction

and to explain how these features are uniquely accomplished on particular occasions [18].

The focus of conversation analysis is on how interactional practices work as systematic, repeatable sequences. While comparison of cases is a central part of the method, measuring the frequency or distribution of a practice is not usually the goal of analysis. On the contrary, conversation analytic and ethnomethodological studies tend to show how attempts to quantify interactional structures gloss over the details of how particular cases are locally and uniquely achieved [18]. Instead of attempting exhaustive coding of our rich video data, we apply the Grounded Theory procedure of building collections of phenomena until “theoretical saturation” [11] is reached, i.e., the point at which we no longer discover new types of cases which require us to modify our analytic categories. Similar to this, we also follow the conversation analytic procedure of “deviant case analysis” [17]: searching specifically for potential counter examples to our main findings and including them in the analysis rather than throwing them out as “outliers.”

Both ethnology and conversation analysis use detailed transcripts of recordings of naturally occurring activities in order to discover their endogenous, local organizations. The transcription conventions capture many details of talk, but focus especially on the timings of turns-at-talk relative to one another. We adapt these conventions to avatar-mediated interaction by annotating system logs. As conversation analysts have pointed out for IRC [9], system logs fail to capture the temporal dimension of composing a chat message. The log merely records when the entire message is posted to the server. Furthermore, current chat logs do not capture most avatar actions, especially movement through space or user interface actions. To capture these details, we manually annotated the system logs with additional detail by using screen-capture video recordings.



Figure 1. City of Heroes

For this study, we used City of Heroes (CoH) as a quasi-experimental test-bed. CoH is a popular superhero-themed MMOG in which users band together in teams to cooperatively engage in combat missions that involve battling system-controlled villains. We created two conditions: one with the default system settings and the other with a set of enhanced awareness cues. We recruited 6 users to play under one or both conditions for at least one hour. Users formed teams of three and completed missions together. We recorded 12 hours of screen-capture video, 6 hours for each condition, from two users' perspectives. We then searched the video data manually for instances of joint activities. Finally we collected nearly 70 such cases, although we still did not entirely exhaust the data set.

Like most avatar-interaction systems, CoH does not provide public cues for several user activities such as managing ability enhancements, consulting the map, reviewing the mission log, trading with another player, composing a turn-at-chat, chatting with a remote character and more. It does provide public cues for a few actions; for example, avatars kneel when players activate the "rest" command and perform appropriate animations when players activate combat abilities.

One feature of the CoH system, as with many games, is that it enables users to define custom macros and key bindings. We used this functionality to create a special set of enhanced awareness cues for users.

- *Manage enhancements* - When the user opens the "manage enhancements" window, which takes up the entire screen and cuts the user off from the game world, the user's avatar automatically says "Adding (or Managing) enhancements" and initiates an animation of working on a laptop computer on a small table. (See Figure 1, left; Figure 2, left.)
- *Consulting map* - When the user opens the in-game map, which obscures only a small portion of the screen, the avatar automatically says "Consulting map" and initiates an animation of reading a newspaper (Figure 2, center).
- *Start chat* - When the user begins to type a chat message, a small chat bubble with an "uh..." appears over the avatar's head (Figure 2, right). This reveals *when* a user is taking a turn-at-chat, although it does not reveal *what* the user is typing in real time.¹
- *Trading* - When the user invites another user to trade, the avatar automatically says "Trade with me, [name of recipient]" and initiates an animation of waving its hand in a summoning gesture.
- *Chat with NPC* - When the user initiates interaction with a non-player character, the avatar automatically says

"Starting mind-meld with [target's name]" and initiates an animation of a military-style salute.

- *Going AFK* - When the user activates the AFK command, the avatar automatically says "Going AFK" and initiates an animation of sleeping, i.e., closing eyes, tilting head and emitting a string of Z's.



Figure 2: Visual Cues

With the exception of the hand waving animation for trading, these animations are automatically looped by the system until the player moves the avatar. Thus, the animations are "held" for the duration of the actions they represent and broken only when the player moves after completing those actions. Although the macros required users to use keystrokes different from the defaults, subjects quickly learned them. Although we created macros for several functions, there were some functions for which we could not create an enhanced cue due to limitations of the system. For example, we could not create a macro for opening the mission log.

PLAYER COORDINATION

In this study, we examine how users coordinate text chat and various game commands with those of other players on a turn-by-turn basis. As conversation analysts have shown, real-life interaction, both verbal [16, 17] and embodied [20] is coordinated remarkably tightly. People can interweave their actions with minimal gap and overlap [16]. This tight coordination is made possible by the fact that turns-at-talk and embodied actions are largely "projectable" and therefore speakers can anticipate their completion with surprising precision.

In avatar-mediated interaction, looser coordination (i.e., greater gaps and overlaps) is possible than in face-to-face. This is due in part to the fact that players actions are not instantly visible since they must first travel through a central server: server lag prevents tight coordination. However, in addition, in current MMOGs, tight coordination is also impacted by the fact that many player activities are hidden from public view and therefore are not projectable. As we have shown in prior work [15], such private activities further inhibit players' ability to achieve tight coordination jointly. Players may not be able to produce particular actions "on time" because part of the

¹ A better solution is to post messages on a word-by-word basis, like in There [6], so players can also see *what* is being said in near real-time.

interactional context is unavailable to them. We call such lapses in timing and coordination simply “slippages.” Specifically, we focus on situations in which users compose turns-at-chat, manage enhancements, consult the map or the mission log and observe if engagement in these activities causes users to “fall behind” their teammates in terms of travel or combat.

In analyzing interactions from the two conditions of gameplay in *City of Heroes*, we discovered four rough categories of coordination sequences (70 cases total): (1) those in which a private activity causes no slippage in coordination (17 cases; 24%), (2) those in which it does (11 cases; 16%), (3) those in which users avoid slippage by making a special effort to reveal private activity (18 cases; 26%), and (4) those in which an awareness cue enables tighter coordination (23 cases; 33%). We found one case in which the use of an awareness cue initially appears to result in a coordination slippage (see #10); however, we show how it in fact affirms the general pattern.

No public cues, no slippage

Although the CoH avatar interaction system, like many systems, does not provide public cues for many user activities, users can often nonetheless use it successfully to coordinate their actions.

The following case from condition 1 is an instance in which a user, Elizabot, engages in a private activity--reviewing the mission list--while in the midst of a conversation with her teammates, Fuyuonna and Napaul, but it causes no problems in coordination. The three have just exited city hall, where they picked up new missions, and are standing near and facing each other discussing which mission to do first.

1. Reviewing Mission Log - No slippage (Condition 1)

```
01→ ((E opens mission log))
02   (4.6)
03 E: which mish?
04   (1.7)
05 F: Should we do N's door mish and try
06   to get my 5 Hellions on the way?
07   (2.0)
08 N: lead on, i'm just here to help
09   (6.0)
10 N: sounds good
11 E: k
12   (3.9)
13   ((E sets team mission))
14→ (4.3) ((E closes mission log))
15 F: cool.
16   ((F runs down steps toward mission
17   marker))
18   ((E follows))
```

In this case, Elizabot privately opens her mission log (line 01), examines it, polls the team about which one to do first (line 03), selects one (line 13) and closes the menu (line 14) all while chatting with her teammates. Although Fuyuonna

and Napaul cannot see exactly when Elizabot opens and closes her mission log, they nonetheless successfully coordinate leaving the scene together (lines 16-18). They can infer from Elizabot's question (line 03) that she is probably reviewing her mission log at that point and can infer from her setting of the mission (line 13), which is indicated to them by a red star appearing on their compasses, that she is likely done reviewing missions by that point. Furthermore, Fuyuonna is already typing her extended turn (lines 05-06), a proposal to do Napaul's “door mish,” when Elizabot asks the team “which mish” they want to do (line 03). Because the composition of turns-at-chat is private, Elizabot cannot see that her teammate is already in the course of answering her question before she asks it. Therefore, there is some minor overlap in turn-taking in this case, but it does not impact the team's ability to coordinate their joint exit. In many cases in our data, even though users are acting without awareness of what their teammates are doing, they can nonetheless achieve adequate coordination.

No public cues, slippage

Although users can often achieve joint coordination without public cues of activity, there are situations in which even experienced users encounter interactional slippage. For example, in the case below, while Jain Reaction is reviewing her mission log (for which we could not create a macro), her teammates run ahead without waiting for her to finish.

2. Reviewing Mission Log - Slippage (Condition 2)

```
01   ((JR & D exit building))
02   ((JR & D run forward & stop))
03   ((D begins to type))
04→  ((J opens her mission log))
05 D: can we get 5 hellions?
06   (4.6)
07 JR: sure
08   (3.5)
09→  ((D runs ahead))
10   (0.4)
11 ED: yep
12   (1.5)
13→  ((ED runs ahead))
14   (2.9)
15   ((JR expands mission description))
16   ((JR resets team mission))
17→  ((JR closes mission log))
17   (1.6)
18   ((JR joins others))
```

After exiting a building simultaneously with Jain (line 01), Derowen begins to compose a message (line 03), as indicated by the “uh...” which appears over her head. A moment later, Jain opens her mission log (line 04) for which there is no public cue. Derowen's question appears (line 05) and Jain's responds shortly after (line 07). Derowen then runs ahead (line 09) while Jain is still reviewing her mission log. Earth Dude responds and runs

ahead as well (lines 11 and 13). Jain continues using her mission log by expanding a mission description (line 15) and resetting the team's mission (line 16) before joining them (line 18). Thus Jain falls behind briefly while she manages the mission log for the team who cannot see what she is doing.

Similarly late communications can cause slippages in coordination as well. In the following case, the team has just completed one battle and has paused before engaging in the next.

3. Composing a Turn - Slippage (Condition 1)

```

01 ((E runs ahead))
02 (2.2)
03 [((E turns around))]
04 [((F runs ahead ))]
05 (1.7)
06 ((E turns around))
07 (2.5)
08 ((E turns halfway))
09 (0.8)
10→ ((E begins typing))
11 (3.5)
12→ ((F runs ahead))
13 ((E still typing))
14 (2.1)
15→ ((N runs ahead))
16 ((E still typing))
17 (2.1)
18 E: why don't you guys attack first
19 (1.0)
20 ((E turns around))
21 (0.8)
22→ ((E begins typing))
23 (0.7)
24 ((F jumps down, ahead))
25 ((E still typing))
26 (2.0)
27 ((N runs behind F))
28 ((E still typing))
29 (2.4)
30→ [((F attacks mobs2)]
31 E: [then I'll try to take aggro]
32 ((E runs ahead))
33 (5.5)
34 ((E attacks))
35 (0.9)
36 F: I really don't think that works.
37 ((more fighting))

```

In the lull between combat encounters, Elizabot turns to Fuyuonna and begins to compose privately a proposal regarding tactics (line 10). While in the midst of Elizabot's

typing, Fuyuonna runs ahead (line 12) and is joined by Napaul (line 15). Elizabot finally completes her message (line 18) and turns to see where the others went. Before receiving any response, she then begins to expand her proposal (line 22). While she is still typing, Fuyuonna initiates the next combat encounter (line 30). Elizabot completes her second turn (line 31) and joins the battle (line 33) and a response to the proposal from Fuyuonna appears (line 35). Thus, the team experiences trouble in coordinating a conversation about tactics with the next battle. Because Fuyuonna and Napaul cannot see that Elizabot is composing turns-at-chat, they cannot wait for her (lines 12, 15, 24, 27 and 30).

No public cues, workarounds, no slippage

Although avatar systems tend not to offer many awareness cues for user activity, players sometimes employ special practices that circumvent the interactional effects of private activities. When tight coordination really matters for gameplay, for example in difficult combat situations, users sometimes enact the "readiness check" sequence before initiating next attacks.³ This is no doubt a generic coordination device that players have imported from real-life situations, yet it works effectively in virtual worlds as well. Regardless of the cause of the lack of readiness - engagement in competing (private) activities, distracted attention, fatigue, etc. - the readiness check can enable a team to act in unison. The following is a routine case of a readiness check in CoH.

4. Readiness Check - Compact (Condition 1)

```

01 ((E passes F))
02 ((F cycles through targets))
03 (0.8)
04 ((E approaches group of mobs))
05 (0.8)
06 ((E targets a mob))
07 (4.6)
08→E: rdy?
09 (3.2)
10→F: ready
11 (3.0)
12→ ((E runs into aggro range))
13 (1.5)
14 ((F attacks mob))
15 (1.0)
16 ((E attacks mob))

```

In this case, Elizabot passes Fuyuonna (line 01) whom she can see is stopped and therefore possibly engaged in a private activity. Indeed Fuyuonna is cycling through a series of potential next targets (line 02). Elizabot then approaches (line 04) and targets a member of the same group of mobs (line 06), likely preparing to start the next attack. However, before doing so, she initiates a readiness

² "Mob" is short for *mobile object*, a legacy term from text-based MUDs that is still commonly used in MMOGs to refer to non-player combat opponents.

³ They may even abandon text chat entirely in favor of real-time voice through third-party voice applications.

check: she asks “rdy?” (short for “ready?”) (line 08) and waits for a go-ahead signal from Fuyonna *before* she attacks. After a brief pause (line 09), Fuyonna gives the go-ahead with “ready” (line 10), and Elizabot only then initiates the attack by running into “aggro range” (line 12), the proximity within which hostile system-controlled opponents will attack a player character. The two then begin coordinated attacks (lines 14 and 16). Thus the readiness check (line 08) enables Elizabot to avoid initiating the next battle while her fellow teammates are possibly are busy with a private activity or simply not paying attention. If Fuyonna were engaged in such an activity, she would have the opportunity to finish it. This is precisely what happens in the following case.

5. Readiness Check - Expanded (Condition 1)

```
01 GE: im not really getting hurt Bad
02     (3.0)
03     ((F & E run upstairs))
04     (6.3)
05     ((E faces & targets mob))
06     (2.1)
07→E: rdy?
08     (1.5)
09 F: yeah, but always nice to have
10     one on hand.
11     (1.2)
12→F: ready
13     ((E turns to face teammates))
14     (4.6)
15→GE: yup
16     (6.0)
17→   ((E initiates attack))
18     ((F attacks))
19     (4.1)
20     ((N attacks))
21     (0.4)
22     ((GE attacks))
```

Here the team has been having a conversation between battles (line 01), and then Elizabot runs ahead and finds the next group of mobs (lines 03-05). Before attacking, she initiates a readiness check sequence (line 07). After a short pause (line 08), an extended turn-at-chat appears from Fuyonna (lines 09-10) indicating that she was indeed not ready. After another short pause (line 11), Fuyonna gives a go-ahead signal (line 12) followed by GE (line 15). Elizabot then initiates the attack (line 17) only after two teammates have indicated readiness and each teammate joins in (lines 18-22). The readiness check sequence is a practice we have observed across several combat MMOGs. While it is effective, it is somewhat cumbersome and players in fact do not use it before *every* encounter (for example in excerpt 3).

Another example of player devices that circumvent slippages in coordination involve extended text chat messages. As we saw in excerpt 3, when a user types an extended chat message privately, teammates may initiate alternate joint courses of action, such as initiating the next

attack, exiting the scene or introducing a new topic of conversation. However, in the following case, Elizabot combats this kind of slippage by projecting a turn-at-chat with a short preface turn.

6. Prefacing an Extended Turn (Condition 1)

```
01     ((E exits city hall))
02     (5.4)
03→E: btw
04     (15.4)
05→E: running with mouse will go to
06     autorun after several seconds
07     (9.4)
08 F: I never run with the mouse.
09     (2.2)
10 F: Weird.
11     (8.2)
12 E: so be careful about running into
13     mobs accidentally
14     (4.1)
15 E: drives me crazy
```

At line 01, Elizabot exits city hall and joins Fuyonna outside the entrance. Napaul is still inside city hall, but the team chat channel is visible to all members regardless of spatial location. Elizabot types a short turn, “btw” (line 03), a common abbreviation for “by the way,” and then begins typing an extended turn, which appears in team chat over 15 seconds later (line 04-06). Hence, the “btw” projects further chat and thereby accounts for the long a gap (line 04) as being due to typing. This preface device then succeeds in preventing Fuyonna and Napaul from initiating new courses of action while she is composing her turn. This is very similar to a technique of breaking long turns-at-chat into shorter ones that has also been identified in internet relay chat [9]. We see then that players have ways of making a system work for them despite its limitations.

Public cues, no slippage

In order to reduce the kinds of slippages in coordination that result from private activities, we created several macros to automatically give off public cues when users are engaged in certain activities. For example, when a user begins to type a chat message, a small bubble saying “uh...” appears over their avatars’ heads. This awareness cue certainly would have helped in the previous case and indeed helps in the following one in which the team begins a mission in the sewers.

7. Waiting for a Turn-at-chat (Condition 2)

```
01     ((D turns toward ED & JR))
02     ((JR runs ahead))
03     ((D runs ahead & passes JR))
04     (6.0)
05     ((D turns around))
06→   ((ED is typing))
07     (1.0)
08→   ((D runs back to teammates))
09     (3.7)
```

10 ED: yes, the odor is from me too
 11 (6.1)
 12 JR: eeek
 13 (3.3)
 14 JR: cool music
 15 D: ready? :p
 16 (6.8)
 17 D: yeah, I like the whispering
 18 effects.
 19 (10.6)
 20 ((D runs ahead))
 21 (4.2)
 22 ((D turns around))
 23 (1.5)
 24 ((D runs ahead))

Immediately upon entering the sewers, Derowen turns to her left and can see that both of her teammates are inside (line 01). Jain runs forward (line 02) and Derowen does the same and passes her (line 03). Derowen continues on for some distance until she reaches a small doorway and then turns around (line 05) no doubt to see if her teammates are following. When she turns, she can see that they are still back at the entrance and that Earth Dude is typing a message as indicated by the “uh...” directly over his head (line 06). Derowen then returns to her teammates (line 08) in time for Earth Dude’s utterance to appear (line 10). Jain responds (line 12) and adds her own comment. Derowen then initiates a readiness check sarcastically, as indicated by the sticking-out-tongue emoticon (line 15) and responds to Jain’s comment (lines 17-18). Derowen waits for over 10 seconds (line 19), during which time she can see that no one is composing a turn-at-chat, and then runs ahead into the sewer (line 20). When she reaches the same doorway as before (line 05), she again turns back toward her teammates (line 22). This time she can see that they are indeed following her and she continues into the sewer (line 24).

The typing cue thus enables Derowen to see why her teammates are not following her and to better predict when they might be ready to move on. Without such a cue, while a player may have an idea of several possible reasons for the failure to follow, they have few clues for deciding which is most likely. The sense of co-presence in CoH is improved by the availability of an account of teammates’ actions, and this availability also allows Derowen to participate more fully in the interaction by returning to engage in the conversation.

Similarly, automatic cues for other user activities help users interpret each others’ actions (or lack of actions) and better coordinate joint courses of action. For example, in the following case, the teammates use cues for managing enhancements. Adding enhancements to one’s character’s abilities improves one’s performance but also renders one temporarily “blind” to the environment and unavailable for other action, since the menu takes up the entire screen.

8. Adding Enhancements Together (Condition 2)

01→ED: ((offscreen)) Adding enhancements
 02 (6.3) ((turns around to see ED))
 03 ((D interacts with glowie))
 04 (4.0) ((J follows D))
 05 *KEY FOUND!*
 06 (0.9) ((D turns around))
 07→JR: ((laptop)) Adding enhancements
 08 (2.7)
 09→D: ((laptop)) Managing enhancements.
 10 (7.8) ((D adds enhances rest))
 11 ((D runs ahead))

In between combat encounters, Earth Dude opens his enhancements menu. When he does so, the macro automatically generates the text message, “Adding enhancements” and a laptop computer on a small table appears in front of his avatar. In this transcript from Derowen’s perspective, Derowen cannot see the laptop animation initially because she is not facing Earth Dude, but can see the chat bubble (line 01). She then turns around and can see Earth Dude at his laptop (line 02) and continues to search the immediate vicinity. After finding a mission-related item (line 05), she turns back toward Earth Dude. As she turns, she can see that Earth Dude is done adding enhancements since his laptop animation is gone but can also see Jain now open her enhancement menu. Derowen then follows suit and opens her own enhancements menu (line 09). The 3D world is then blocked for several seconds as Derowen adds an enhancement to her “rest” ability (line 10). When she closes the menu, she can see that Jain is also done managing her enhancements—although the “Adding enhancements” chat bubble is still visible on the screen, the laptop animation has disappeared, indicating that the action is finished. After seeing that her teammates appear unoccupied, Derowen runs ahead toward the next combat encounter.

Similarly, in the following case, the adding enhancements macro proves critical in coordinating a next attack.

9. Waiting for Enhancements & Rest (Condition 2)

01 ((D approaches mobs))
 02 (1.0)
 03→ED: ((laptop)) Adding enhancements
 04 (1.4)
 05→ ((D retreats))
 06 (2.0)
 07 ((ED closes enhancements))
 08 (2.0)
 09 D: ready
 10 (2.0)
 11 JR: running out of tape
 12 ((D advances & retreats))
 13 (3.1)
 14 D: yeah, one min
 15 ((turns toward ED))
 16 (2.4)
 17 ((D turns her view back & forth))

18 between JR & ED, approaches mob))
 19 (1.0)
 20→ ((ED is kneeling))
 21 ((D turns away from ED and
 22 back toward him))
 23→ ((ED stands up))
 24 (0.8)
 25→ ((D approaches & attacks mob))
 26 (2.6)
 27 ((JR attacks mob))

In this case, Derowen approaches a group of mobs (line 01) to initiate the team's next combat encounter. However, just before she enters aggro range, she can see Earth Dude open his enhancement management window (line 03) both by his avatar's "laptop" animation and the text alert in the chat bubble. In response, Derowen retreats from the group of mobs (line 05) and waits. When Earth Dude's laptop animation disappears (line 07), indicating that he has closed his enhancements window, Derowen then indicates that she is ready (line 09). Jain and Derowen then discuss the state of their video recording (lines 11-14). Derowen glances back and forth between her two teammates, whom she cannot view simultaneously because they are positioned too far apart (lines 15-18). When she glances back at Earth Dude (line 20), she can see that he is kneeling, a default animation in CoH which indicates that the user is "resting" to regenerate power. Derowen then waits further until she can see that Earth Dude is done resting (line 23) and only then she initiates the next attack successfully (line 25) and without the use of a readiness check.

Thus, the automatic text messages and avatar animations enable teammates to better see what the others are doing without anyone doing extra work to make their actions more transparent. Furthermore, in this case, knowing that one teammate is taking a particular kind of maintenance action both transforms that action into a kind of group activity as others can choose to do it at the same time, and also increases efficiency, since no one's time is wasted waiting for a teammate to verbally indicate completion of a maintenance action.

A potential deviant case

Overall, we found that the enhanced awareness cues provided by the macros enabled users to coordinate their actions more tightly and avoid slippages. In most cases, we see this in terms of waiting for a fellow teammate to complete an activity before initiating a next joint course of action. However we have only one case (out of 70) in which a user fails to wait, despite seeing an awareness cue. In Excerpt 10, the team has just completed a battle and mission.

10. Not Waiting for Map Reading (Condition 2)

01 D: cool
 02 (2.5)
 03 D: Train?
 04 (4.0)
 05→D: ((newspaper)) Consulting my map.

06 (4.3)
 07 JR: ok
 08 ED: FEAR THE EARTH DUDE, MWAHAHA
 09 (3.4)
 10 JR: haha
 11 (3.6)
 12 ((D laughs))
 13 (3.8)
 14→ ((D runs off screen))
 15→JR: ((newspaper)) Consulting my map
 16 (1.8)
 17→ ((ED runs off screen))
 18 (5.6) ((JR sets compass point
 19 & closes map))
 20 ((JR follows teammates))

At line 03, Derowen proposes that the team return to the NPC trainer as its next joint course of action. She then opens her map and sets the compass point for the trainer (line 05). Jain confirms the proposal (line 07) and Earth Dude makes a joke (line 08) to which Jain laughs through chat (line 10) and Derowen through avatar emote (line 12). After a moment, Derowen then heads off toward the NPC trainer (line 14). Just as Derowen passes by Jain, the latter opens her own map (line 15). Derowen cannot see the reading-the-newspaper animation because Jain is no longer in her view, but she *can* see the chat bubble saying, "Consulting my map" (line 15). Similarly Earth Dude, who can see both Jain's automatic text message and her avatar's animation, fails to wait as he follows Derowen (line 17). Jain then sets her compass point (line 18), closes her map (line 19) and follows behind her teammates (line 20).

Formally, this case resembles excerpts 2 and 3 in which two teammates move on to a next joint course of action without waiting for the third member to complete a current activity. The main difference in excerpt 10 is that the two teammates can actually see what the third is doing and see that she is not finished yet. So why don't they wait? Although awareness cues enable users to wait for each other, they do not constrain users to do so. Users may choose not to wait. The other important difference between excerpt 10 and excerpts 2 and 3 is the nature of the next joint course of action. In 2 and 3 it is combat, an activity that greatly benefits from teamwork and tight coordination. In 10 it is returning to the NPC trainer, which is located in area that acts as a kind of home base between missions. This activity *does not* typically benefit from teamwork nor tight coordination. Compass points show each individual user where to go and each can outrun any opponents along the way, so sticking together for protection is not necessary. In fact, in CoH fighting jointly but traveling independently are routine practices. Furthermore, although Derowen and Earth Dude do not wait for Jain to finish consulting her map, they do so knowingly. That is, they likely know why she is falling behind and therefore are in a position to make an informed judgment. Without awareness cues, they would not be in a position to as accurately account for her delay.

DISCUSSION

We see then that the ability to see what activities teammates are currently engaged in enables players to achieve tighter coordination in interaction than they can without such awareness information. Although experienced players employ various strategies, such as the pre-attack “readiness check” sequence, to enhance coordination within the limits of existing systems, these are somewhat limited devices. Even capable players are not always able to predict when others may be engaged in action-stopping activities such as producing a turn-at-chat, as we demonstrated in excerpt 3 above.

Where the system does provide visual cues for activities that render characters unavailable for action, such as “resting,” it is much easier for players to quickly notice and accommodate others’ activities. The visual cues that we provided through macros broadened the range of accountable, observable-reportable actions undertaken by users and correspondingly enhanced the coordination of joint action. Furthermore, even where a lack of visual cues does not actually lead to visible breakdowns in coordination, such as uncoordinated attacks, their deployment contributes to a stronger sense of co-presence than is currently available in MMOGs.

Although the chat bubble component of our awareness cue macros was obviously important for enhancing player awareness of each other’s activities, as seen in excerpt 10, where only an off-screen avatar’s chat bubble is visible, the use of linked animations provided additional, more finely detailed information about those activities. As demonstrated in excerpt 8, the way that the system allows animations to loop/hold until players break them with movement lets them communicate automatically when they have completed the associated action. Integrating such animations into the system itself would make it possible for *all* users to communicate and access information about their currently invisible in-game actions, not just users with enough technical knowledge to create and use macros. As other researchers have found, experienced users tend to communicate through the channels that are the least trouble [7]; by integrating visual cues into the system itself as City of Heroes has already done with “rest,” virtual-world designers could ensure maximum availability of information for minimum user effort.

CONCLUSION

It may indeed never be possible to enable avatars to give off as much rich information as real human bodies do, nonetheless our data suggest that even crude awareness cues can go a long way to improve user-to-user coordination. By simply revealing *that* fellow users are engaged in an activity of a particular type - managing enhancements, consulting a map, typing a chat message - and *when* they are done, users have enough feedback to wait until such activities are complete.

In past work [15], we examined the impact of MUD-style text emotes on player-to-player coordination. While the text descriptions of player activities (e.g., “Player rummages through her bags” or “Player checks her To-Do list”), did indeed enable tighter coordination, they were problematic in practice because at times they can become lost in players’ text boxes, which they share with chat and system messages, while at other times (especially in large groups) they can flood the text box.

In the current study, we examine alternative types of activity awareness cues - avatar animations and chat bubbles - which enable us to draw the following design guidelines:

- Animating the avatar enables other players to see at a glance who is doing what when; however, they can be missed if one is not looking directly at the avatar or cannot see it due to a limited field of view.
- Avatar animations can be supplemented with cues that are noticeable even when other players are not looking directly at one’s avatar. These may include the chat bubbles in our study, which are visible even when the avatar is not, as well as sounds which are likewise noticeable even when one is facing the wrong direction.
- Looping animations make the *duration* of activities publicly observable, and not just their onsets. This enables other players to see at a glance *when* a user is done and therefore facilitates the initiation of next joint actions (City of Heroes is one of the few MMOGs that provides such looping animations).

While many MMOGs provide this kind of feedback for *combat* activities, most fail to provide it for non-combat activities. Yet our data shows that joint coordination is important for even mundane activities like leaving the scene together. Furthermore, it shows that coordination in mundane activities can also impact combat activities especially in the case of timing the initiation of an attack.

As we argued in past work regarding text-emote cues [15], increasing the public transparency of player activities facilitates coordination but also increases the visibility of player (in)competence. Players who are accustomed to the particular kind of *public privacy* afforded by current systems might dislike a higher level of public exposure. Designers can always give players the option of turning activity cues on and off. Indeed there may be some occasions on which players may wish to hide their activities. Yet we believe these occasions are the exceptions and that players overall will value greater coordination over privacy. (All of the players in our study responded positively to the new awareness cues.) Furthermore, MMOGs already make much of players’ *combat* activities, and therefore competence in combat, transparent. We propose being more systematic and fine-grained in the ways player activity is revealed.

In this study, we have examined avatar-mediated interaction in the context of a combat game, City of Heroes, yet we argue that the coordination issues we identify apply to non-game-oriented persistent 3D virtual worlds as well. Indeed from our observations in Second Life and There, we see that users in those environments likewise encounter slippages in coordinating private activities with conversation, travel, and other joint actions (although less so in There [6] since it provides much better awareness cues than most systems). As virtual worlds expand to include a wider variety of social activities (e.g., education and commerce), there will be a growing need to develop avatar systems that enable richer and richer social interactions.

REFERENCES

1. Bartle, Richard. A Voice from the Dungeon. *Practical Computing*, December, (1983), 126-130.
2. Becker, Barbara and Gloria Mark. Social Conventions in Computer-mediated Communication: A Comparison of Three Online Shared Virtual Environments. In *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*, ed. Ralph Schroeder. Great Britain: Springer-Verlag London Limited (2002), 19 – 39.
3. Benford, S., J. Bowers, L.E. Fahlén and C. Greenhalgh. Managing mutual awareness in collaborative virtual environments. In *Proceedings of Virtual Reality Software and Technology*, ed. G. Singh, S.K. Feiner and D. Thalman. Singapore: World Scientific Publishing Co. Pte. Ltd. (1994).
4. Benford, S. D., Bowers, J. M., Fahlen, L. E., Greenhalgh, C. M., Snowdon, D. N. Embodiments, Avatars, Clones and Agents for Multi-user, Multi-sensory Virtual Worlds, *Multimedia Systems*, 5(2) Springer-Verlag, (1997), 93-104.
5. Bowers, John, James Pycock and Jon O'Brien. Talk and Embodiment in Collaborative Virtual Environments. In *Proceedings of SIGCHI 1996*. ACM Press: New York, (1996), 58 – 65.
6. Brown, Barry and Marek Bell. CSCW at play: 'there' as a collaborative virtual environment." *Proceedings of CSCW 2004*. ACM Press New York, NY, USA, (2004).
7. Cheng, Lili, Shelly Farnham and Linda Stone. Lessons Learned: Building and Deploying Shared Virtual Environments. In *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*, ed. Ralph Schroeder. Great Britain: Springer-Verlag London Limited (2002), 90 – 111.
8. Curtis, Pavel. Mudding: Social Phenomena in Text-Based Virtual Realities. Proceedings of the 1992 Conference on the Directions and Implications of Advanced Computing, (1992).
9. Garcia, Angela and J. Jacobs. The Eyes of the Beholder: Understanding the Turntaking System in Quasi-synchronous CMC. *Research on Language and Social Interaction*, vol. 32, no. 4, (1999), 337–367.
10. Garfinkel, Harold. *Studies in ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall. (1967).
11. Glaser, Barney G., & Strauss, Anselm L. *The Discovery of Grounded Theory*. Chicago: Aldine. (1967).
12. Goffman, Erving. *The Presentation of Self in Everyday Life*. Doubleday: New York, (1959).
13. Hindmarsh, Jon, Mike Fraser, Christian Heath, Steve Benford and Chris Greenhalgh. Fragmented interaction: establishing mutual orientation in virtual environments. In *Proceedings of CSCW 1998*. ACM Press: New York, (1998), 217 – 226.
14. Hindmarsh, Jon, Mike Fraser, Christian Heath and Steve Benford. Virtually Missing the Point: Configuring CVEs for Object-Focused Interaction. In *Collaborative Virtual Environments*, Springer, September (2001). 115-133.
15. Moore, Robert J., Nicolas Ducheneaut, and Eric Nickell. Doing virtually nothing: Awareness and accountability in massively multiplayer online worlds. *Computer Supported Cooperative Work(Online)*. (2006).
16. Sacks, Harvey, Schegloff, Emanuel A., and Jefferson, Gail. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50: (1974) 696-735.
17. Schegloff, Emanuel A. Sequencing in Conversational Openings. *American Anthropologist* 70:(1968)1075-1095.
18. Schegloff, Emanuel A. Analyzing single episodes of interaction: an exercise in conversation analysis, *Social Psychology Quarterly* 50: (1987), 101-14.
19. Smith, Marc, Shelly Farnham and Steven Drucker. The Social Life of Small Graphical Chat Spaces. In *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*, ed. Ralph Schroeder. Great Britain: Springer-Verlag London Limited (2002), 205 - 220.
20. Streeck, Jurgen: On Projection. In E. Goody (ed.). *Interaction and Social Intelligence* (pp.84-110). Cambridge: Cambridge University Press (1995).
21. Yee, Nick. The Daedalus Gateway. <http://www.nickyee.com/daedalus>.
22. Yee, N., Bailenson, J.N., Urbanek, M., Chang, F., Merget, D. The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. *The Journal of CyberPsychology and Behavior*, (in press).