Journal of Virtual Worlds Research

Volume 2, Number 1.

Pedagogy, Education and Innovation in 3-D Virtual Worlds.

ISSN 1941-8477

Vol. 2. No.1
"Pedagogy, Education and Innovation in 3-D Virtual Worlds"
April 2009

This issue was sponsored, in part, by the Singapore Internet Research Centre, the Department of Radio, TV & Film at the University of Texas at Austin, and the Texas Digital Library Consortium.

The Journal of Virtual Worlds Research is owned and published by the Virtual Worlds Research Consortium, aTexas non-profit corporation. (http://vwrc.org)

Journal of Virtual Worlds Research
jvwresearch.org ISSN: 1941-8477

Vol. 2. No.1
ISSN: 1941-8477
"Pedagogy, Education and Innovation in 3-D Virtual Worlds"
April 2009

Editor-in-Chief's Corner
Leveraging Game-Playing Skills, Expectations and Behaviors of Digital Natives to Improve Visual Analytic Tools

By Theresa A. O'Connell, National Institute of Standards and Technology;
John Grantham, Systems Plus;
Kevin Workman, Millersville University
Wyatt Wong, Forterra Systems Inc.

Abstract

We report gaming research aimed at improving innovative visual analytic (VA) tools. Digital natives are entering the information analysis workplace. There are similarities between innovative visual analytic (VA) tools used by information analysts and the video games digital natives play. These similarities provide an opportunity to leverage the game-playing skills, expectations and behaviors of digital natives in the design of VA tools that will help them perform information analysis. To this end, we performed a user-centered usability engineering (UE) study of digital native video game players' interactions with each other and with a collaborative virtual environment (CVE). We measured player success (in terms of efficiency and effectiveness) and satisfaction with PanelPuzzle, a puzzle-solving game. Players were digital natives, having grown up surrounded by technology. The study showed that roleplaying positively impacted their success and satisfaction. Despite an expectation for immediate feedback, digital natives found workarounds to mitigate impacts of delayed feedback. Teammate communication was essential to collaboration and thus to success. In particular, players used communication records to build collaborative knowledge. This work provided data for a future study on discourse during gameplay. Findings will provide user-centered feedback to improve the design of innovative visual analytic (VA) tools.

Keywords: avatar; collaboration; digital natives; immersion; virtual worlds; video game; visual analytic tools.

Editor-in-Chief's Corner

Leveraging Game-Playing Skills, Expectations and Behaviors of Digital Natives to Improve Visual Analytic Tools

By Theresa A. O’Connell, National Institute of Standards and Technology;
John Grantham, Systems Plus;
Kevin Workman, Millersville University
Wyatt Wong, Forterra Systems Inc.

Our research derives from similarities between video games and VA tools, software information analysts use to visualize and understand massive data sets (O’Connell, Choong, Grantham, Moriarty & Wong, 2008). Our ultimate goal is to improve the design of innovative VA tools to leverage the game-playing skills, behaviors and expectations of digital natives while helping them perform information analysis. Our prior work on riddle solving in a CVE (O’Connell et al, 2008) showed that gameplay collaboration in VEs leads to higher success in gameplay and might be transferable to collaboration among information analysts. Information analysis resembles assembling puzzle pieces. Building on this metaphor, we constructed a game that required players to collaborate in finding and assembling puzzle pieces.

A rich literature addresses games in education (e.g., de Freitas, 2006). Our interest is an aspect of e-learning not often studied: analytical skills digital natives develop when playing games, whether the games be serious or for enjoyment. Squire (2005) offers that gamers develop skills for decision making and problem solving and posits that gaming provides practice for transferring these skills to out-of-game domains. The line between business software and the gaming paradigm is diminishing (ESA, 2008; Chao, 2001; Chao, 2004, Malone, 1982). Gaming approaches are fusing into software designed for digital native defense workers (Capps, McDowell & Zyda, 2001; Hendrick, Knight, Menaker, O'Connor & Robbins, 2008). The merger of VE work and gaming is amply documented (e.g., Zyda, 2005), as is CVE use in analytic workplaces (e.g., Maybury, 2001). However, there is a dearth of literature on studying gameplay to improve VA tool design. Thus, an innovative approach was required. Our user-centered approach adapted UE best practices developed for VA tool studies (Choong & O’Connell, 2008), examining players’ skills, expectations and behaviors and measuring their gameplay success and satisfaction. This work differs from classic UE because its goal was not to assess or improve the game’s usability. Instead, we examined three factors common to gameplay and information analysis: team dynamics relating to collaboration in a CVE; the impact of roles on collaboration in a CVE; and the impact of the timing of feedback delivery on players’ success and satisfaction in solving puzzles.

Digital Natives

Prensky (2001a) established the term, digital natives, to describe a generation that grew up surrounded by electronics. They are comfortable using text messages, social networking sites, video game user interfaces (UI), and other recent technologies. Prensky’s work started in the educational sector, but an inflow of digital natives also affects business, research, and government (Prensky, 2005).

Digital natives who play action video games have higher visual acuity and the cognitive ability to handle larger amounts of information than their predecessors (Green, & Bavelier, 2003, 2007). Because of their immersion in technology, their brains differ from those of people who grew up without technologies such as personal computers, the internet, and video games (Prensky, 2001b). They have developed new communication models, using cell phones and the Internet to communicate quickly and frequently. They are very collaborative and comfortably communicate with several parties simultaneously utilizing several forms of communication (Prensky, 2004).

Digital natives blur the lines between education and fun, between work and play. Thanks to their rapid-fire style of interacting with digital devices and UIs, they expect immediate feedback from technology and other people (Prensky, 2005). False or unshared expectations can lead to misunderstandings. People we call digital foreigners often mistake digital natives' impatience for inaptitude. Digital natives expect digital foreigners to follow the protocols of new communication models. Such opposing viewpoints impact the workplace causing it to evolve as digital natives gain employment.

Digital natives defy definition because they are the constantly evolving products of technology and social interaction. As digital foreigners retire, digital natives will replace them, bringing evolved forms of working and communicating. The first step to helping digital natives perform to the best of their ability in the VA workplace is to understand how they use technology and how to leverage their technology-induced adaptations.

Hypotheses

VA tools are applied to problems so complex that analysts must collaborate to solve them (O'Connell & Choong, 2008). Thus, our studies focus on collaboration skills. We are also interested in how digital native gamers' expectations and behaviors affect their gameplay. To understand the collaborative aspects of gameplay as manifested in communication and as affected by the imposition of roles and the timing of feedback and to understand the experiential dimensions of teams' interactions with PanelPuzzle, we formulated three hypotheses.

H1: Communication among teammates will facilitate collaboration, resulting in higher scores for teams that communicate most effectively. The social aspects of gaming are widely recognized as motivational and integral (e.g., Chao, 2001; Zubek & Khoo, 2000; Whang & Chang, 2003). We considered communication to be the essential aspect of society among players as it is among analysts. We expected teammate communication to facilitate collaboration. Studies show that voice communication and text chat impact collaboration (e.g., Jensen, Farnham, Drucker & Kollock, 2000), but give little attention to differences among voice chat, text chat and forum communication. We investigated whether differences existed and if so, how they impacted engagement and scores. Because digital natives are frequent digital communicators, we expected constant streams of voice and text chat to accommodate tactical communication. We expected the players forum to be reserved for strategic communication on building collaborative knowledge.

Although we did not find literature on transferring gaming language to VA tools, an emerging body of literature discusses transferring gaming language to the workplace. Chao (2001) reasons that using gaming language and gaming metaphors in workplace applications will facilitate the experience of users who grew up digital. He posits that communication among children and even adults who are not digitally savvy has incorporated gaming slang. Khoo and Zubek (Khoo & Zubek, 2002; Zubek & Khoo, 2000) identified characteristics of chat during competitive gameplay. They observed disconnected discourse with poor spelling and grammar. Topics change frequently. Simultaneous threads cause players to miss parts of discourse. We expected to see these trends in player discourse.

H2. Roles positively affect group dynamics in a CVE. We expected digital natives to collaborate and build collaborative knowledge to solve puzzles. Collaboration is the key to resolve VA problems. Usually, analysts' workplace culture strictly defines roles, e.g., senior analysts supervise junior analysts. The potential of CVEs to promote collaboration is a rich area of research (e.g., Benfield, et al, 2001), but few discuss the role of roles in CVEs. Maybury (2001) identified the need for role-based access, with defined rights, to communication in a CVE. Our work aligned with this in defining roles based on access to communication with a game master (GM). Maybury (2001) also identified the need for a leadership role in a CVE, noting that without a leader, group interactions are not successful.

Our definition of role was narrower than gaming literature usage. For example, Yee (2006), in his definition of roleplaying, includes story improvisation by the role player. Because PanelPuzzle was straightforward in its goals, we did not expect role improvisation. We expected higher efficiency and higher satisfaction in mandatory-role conditions because they defined responsibilities and restricted communication with the GM to one player. We expected roles to facilitate collaboration, with voluntary-role players refusing roles and collaborating less successfully than mandatory-role players.

H3. Deferred feedback will impact gameplay strategy. Digital natives often expect immediate feedback, but VA tools give both immediate and deferred feedback. We expected delayed feedback to reduce satisfaction. Gergle, Kraut & Fussell (2006) demonstrated that, in a shared workspace, millisecond-long delays in visual feedback impaired communication, negatively impacting the performance of two collaborators solving puzzles on computers. We expected their findings to extend to longer delays. We wanted to see whether players modified gameplay strategies to remedy problems caused by delayed feedback.

Independent Variables

PanelPuzzle was designed to force player collaboration to foster investigation of two sets of independent variables, roles and feedback. For **roles**, there were two variables. The **mandatory role (RM)** variable compelled players to select a role. Under the **voluntary role (RV)** variable, players decided whether or not to play one or more roles. Under both RM and RV variables, players had autonomy in choosing among four roles. Under RM conditions, the GM responded only to a designated player. Under RV conditions, the GM responded to any player.

Feedback pertained to the GM's responses to players' requests for insertion of puzzle pieces into puzzles. A change in the state of a panel, i.e., the filling of a section, was considered visual feedback. For feedback, there were two variables, **feedback deferred (FD)** and **feedback immediate (FI)**. The FD variable constrained players to request piece insertion through the GM forum. The GM waited up to five minutes to insert pieces. With the FI variable, the GM received messages through text chat and inserted pieces immediately upon request. Players using text chat sometimes issued so many simultaneous requests that the GM needed a moment to catch up. Despite occasional brief intervals between requests and feedback, we called the second variable *immediate* because the GM inserted pieces with no intentional or purposeful delays.

Players

Before gameplay, players took a demographic survey. All questions were optional. Fourteen males and two females below the age of 29 participated in the study as players. All fit the definition of digital natives, e.g., all had grown up surrounded by technology. Eleven played interactive video games weekly: seven for one to five hours; three for 6-15 hours, and one for 16-25 hours. Five were not gamers. To protect anonymity, players received gameplay names, e.g., Zulu_Delta. Because communication was an integral aspect of PanelPuzzle, we surveyed players about their communication behaviors and preferences. Players were familiar users of telecommunication devices and technologies. Chat rooms and forums, both important means of communication in PanelPuzzle, were among the least frequently used and least popular means for communication with friends.

PanelPuzzle Platform

To examine how digital natives interacted and collaborated in a VE, we required a platform on which to organize a multiplayer game. To monitor how digital natives use different communication forms, the platform had to accommodate text and voice chat and support their recording. We chose Forterra's On-Line Interactive Virtual Environment (OLIVE) platform which provided a rich collaborative experience in a persistent 3D (three-dimensional) VE. OLIVE supports capabilities essential to communication and inter-player interaction, including avatars, text and Voice-over-Internet Protocol (VOIP) communication. Its session record and playback capabilities save and replay the entire VE simulation from any viewpoint.

OLIVE is a complex system with various interfaces, controls, and synchronous communication channels. To this, we introduced PanelPuzzle, an application with uniquely different interfaces, controls, and asynchronous communication channels. PanelPuzzle required players to operate OLIVE as well as Web-based forums. The default OLIVE graphical UI and virtual city filled many PanelPuzzle requirements, requiring few customizations to support game design and analysis. To accommodate gameplay, we changed the layout of UI elements; removed unnecessary controls; and added controls for new functionality. Avatar customization and establishing the avatar-player bond familiarize players with the environment and the conditions and representation under which they will play (de Freitas, 2006). So, we modified the OLIVE client to launch in face view, empowering players to see and customize their avatars. We modified the OLIVE client to log text chat and the OLIVE server to log voice chat.

For in-game communication, binaural headsets with microphones (mics) enabled VOIP proximal communication, projecting players' voices from their avatars using 3D-audio techniques. Consequently, as in the real world, players' voices faded as distance increased between their avatars. We provided the text forum for distal exchanges.

The Projection Screen (ProjScreen) in OLIVE resembles a real world projection screen, where custom content can be loaded and displayed. The GM's computer contained a folder of PowerPoint slides named by piece identification number. The GM loaded these onto ProjScreens to display puzzles.

Playing PanelPuzzle

Solving puzzles was analogous to building collaborative knowledge by coordinating distributed knowledge, e.g., when one player found one piece of a puzzle, and teammates found others. We wanted to see how players handled knowledge that is useless by itself, but solves a problem when combined with other players' knowledge.

Four sessions each accommodated four players. Players worked on whichever puzzle they wanted at any point in the 75 minutes allotted. Puzzle pieces were placed throughout the city, inside and outside of buildings. This caused players to move throughout the city and to use the communication mechanisms. Each puzzle piece had a random numerical identifier to facilitate discussion. Pieces also indicated the size of the panels to which they belonged. Enough pieces were scattered throughout the CVE to solve nine puzzles. After finding a piece, players asked the GM to insert it into a specific section on a panel. The center of Peninsula City contained three panels: each with sections to accommodate one three-piece, four-piece, or five-piece puzzle. There was no area large enough to display nine panels. Thus, only one puzzle of each size was visible at any time. Maybury (2001) notes the importance of context, i.e., a focal point, in a workplace CVE. The city center where the puzzle panels were located provided this context. It served as an assembly point where teammates received visual feedback on progress.

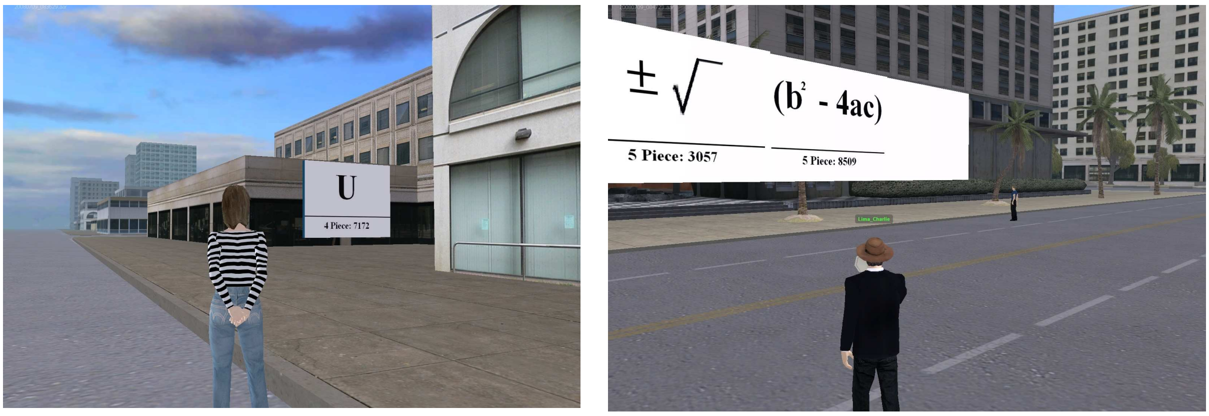

Figure 1. (Left) A player finds a puzzle piece labeled "4 Piece" and numbered (7172). (Right) A player (center) directs the GM (right, in distance) to insert a piece into the rightmost section of a panel.

A variety of styles reflected the variety of cognitive abilities characteristic of digital natives and information analysis. For each puzzle size, there were three types of puzzle. Math puzzles contained common equations e.g., the Pythagorean Theorem. Word puzzles contained common words or sentences. Image puzzles contained pictures, e.g., a scene from a familiar video game.

The GM was a human who provided feedback according to strict rules, e.g., requiring players to identify the puzzle piece, the panel size and the exact section in a panel where they wanted the piece inserted. Within such constraints, the GM followed players' instructions, even when players requested piece insertion into the wrong panel or section. To earn points, a player informed the GM that a puzzle was complete. Teams earned 50 points for submitting a correct three-piece puzzle, 100 points for a four-piece puzzle, and 200 points for a five-piece puzzle. Assembling pieces incorrectly and then reporting the puzzle as solved resulted in a 25 point penalty. Upon completion of a puzzle, the GM informed players of their success and the points won, using the forum or text chat.

PanelPuzzle resided in the virtual Peninsula City. Its 36 city blocks covered 1,300 scaled kilometers, with over 100 architectural models. Ten models had unique interiors. Avatars walked and ran at a pace that scaled to real life walking or running; they teleported to major destinations.

Figure 2. An overhead view of Peninsula City shows much of the downtown area. Four white square puzzle pieces are visible.

Players had an enormous amount of information to process, and their choice of communication tools could impact puzzle-solving. Players could communicate with each other and with the GM synchronously using text chat, or asynchronously through the forum. Players could use a mic to communicate with each other, but not with the GM.

Players in RV and RM conditions had the same choice of four roles. A GM Coordinator communicated with the GM through either text chat or the GM forum. A Communications Coordinator oversaw player forum communications. A Map Coordinator used the map to coordinate player activities. A Puzzle Piece Coordinator directed puzzle assembly. We expected the GM Coordinator to be the team leader, although the rules did not designate this responsibility. In RV conditions, players had the option of adopting an unofficial leader. Prior to gameplay, each team had ten minutes in a virtual conference room to discuss strategy, roles, and teamwork. Seventy-five minutes after players left the conference room, gameplay stopped.

Experimental Environment

The experiment ran on four desktop computers, each with an Intel Xeon 3.0 GHz processor; two GB of memory; and an nVidia Quadro FX 1400 128 MB 3D graphics card. To support multi-tasking, each player had two monitors. Peripherals included a standard 101/102 keyboard; a three-button click/scroll-wheel mouse; two monitors (19”, 20” or 21”) set to maximum resolution; headphones and a noise-cancelling, free-standing desktop PC mic. Peripherals varied slightly, but were functionally equivalent, e.g., display areas were roughly equal. The operating system on each computer was Microsoft Windows XP Professional, Service Pack 2. Other software was limited to the Microsoft Internet Explorer 7 Web browser and OLIVE 2.0.1.

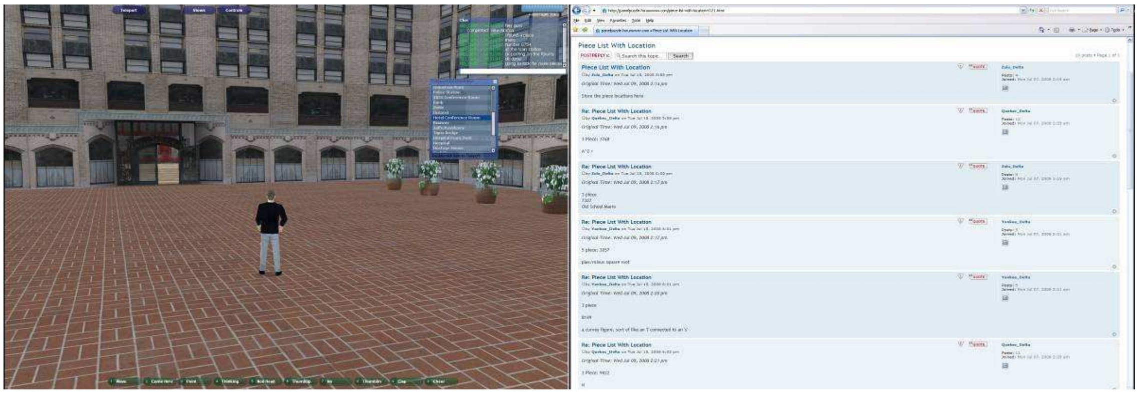

Figure 3. On their left monitor, players viewed Peninsula City, the text chat box and a teleportation destinations box. On their right, they viewed the forums.

Sessions

Four three-hour sessions each accommodated one condition, i.e., one combination of the variables.

Bravo Feedback Deferred (FD) Roles Voluntary (RV)	**Charlie** Feedback Immediate (FI) Roles Mandatory (RM)
Delta Feedback Deferred (FD) Roles Mandatory (RM)	**Echo** Feedback Immediate (FI) Roles Voluntary (RV)

Figure 4. Four conditions each accommodated two variables.

Each session presented the same activities in the same sequence. After a demographic survey, players received written gameplay instructions specific to the session’s conditions. A self-paced fifteen-minute tutorial showed how to customize an avatar; navigate the CVE; communicate with teammates and the GM; find pieces; instruct the GM to insert pieces; and teleport. Players customized avatars and explored the world before gameplay. A competency test verified that players could exercise PanelPuzzle’s basic functionality. Then, players received paper maps of Peninsula City. Out-of-game communication was prohibited. An exit survey followed gameplay. A discussion completed each session.

Measures And Metrics

Measures of players' interactions with PanelPuzzle derived from the three attributes of usability defined by the International Standards Organization (ISO, 1998): efficiency, effectiveness and satisfaction. These attributes are not discrete; each impacts the others.

Efficiency and Effectiveness

Efficiency and effectiveness equate with player success. Measuring **efficiency** in gameplay studies has been discouraged because efficiency is often defined in terms of the game, not player success. An example is to define efficiency in terms of resources a game provides to players (Federoff, 2002). We defined efficiency in user-centered terms of player success. We measured nine dimensions of efficiency:

- Number of requests to insert a piece into the correct section
- Number of requests to insert a piece into the wrong section
- Minutes taken to correctly solve a puzzle
- Ratio of puzzles started to puzzles completed
- Number of duplicate requests to GM
- Number of sections in each puzzle solved
- Number of requests for piece insertion into a wrong panel
- Number of instances of incorrect request formatting
- Number of penalty points for incorrect puzzle solutions.

To measure efficiency of communication among teammates, we counted text chat, voice chat and player forum messages. We looked at the number of messages in terms of the independent variables and success. We examined the relationship between the number of text chat messages and satisfaction with roles.

Effectiveness is not a typical measure of human interaction with games although points and progression to higher levels are reported to players as engaging feedback. An exception is to define effectiveness as following the ideal path to a game's end goal (e.g., Federoff, 2002). However, PanelPuzzle offers many equally appropriate paths to its goals. A user-centered approach assessed effectiveness in terms of players' success, i.e., results. Our metrics addressed five dimensions of effectiveness. We counted

- Puzzles solved
- Points earned
- Unique puzzle pieces found
- Unique pieces inserted correctly into panel sections
- Sections in correctly populated panels.

Satisfaction

Satisfaction surveys that quantify subjective user feedback are common in UE. An exit survey used 1-to-7 ascending Likert scales with open-ended questions for players to explain their ratings. In gaming, satisfaction indices reflect ratings of game aspects e.g., sound, scenario or graphics (e.g., Ham & Lee, 2006). We measured satisfaction across dimensions of the player experience rather than game attributes. We measured enjoyment, engagement and comfort, factors associated with human interaction with games (e.g., Federoff, 2002), although not always considered measures of satisfaction.

Enjoyment. Jegers (2008) epitomizes the literature on enjoyment in saying that a game becomes pointless if players do not enjoy it. Enjoyment is sometimes measured in terms of time spent in gameplay (e.g., Malone, 1982). Imposing a time limit required us to rely on other metrics. For example, we assessed roleplay enjoyment and enjoyment in general.

During gaming, digital native analysts develop an expectation for immediate feedback that they transfer to their use of VA tools (O'Connell & Choong, 2008). In this study, we considered messages from the GM to players as system feedback. Because these messages pertained to notification of points won, we equated system feedback with the performance feedback that is considered integral to game enjoyment (e.g., Malone, 1982).

Engagement. Van Eck (2006) stresses the need to engage digital natives. We anticipated engagement to be impacted by communicating with other players; customizing avatars; exploring the world; interacting and collaborating with teammates; and roleplay. Challenge contributes to enjoyment of video games (Malone, 1982; Von Ahn & Dabbish, 2008). We expected challenge to also contribute to engagement, so we surveyed players on how engaging they found four dimensions of challenge: accumulating points; hunting for puzzle pieces; assembling puzzle pieces; and trying to finish as fast as possible.

Data Collection

Recordings of all in-game activities in the CVE, but not in the forums, were manually started and stopped by the GM, stored on the OLIVE server with automatically date- and time-stamped filenames, and converted into MPG format via video-out to a camcorder during recording playback. **Surveys** were Web-based.

Figure 5. The chat message (C) represents chat-window content. (A) and (B) appeared in the logs.

The OLIVE client produced **text chat logs** in time-stamped text files. **Forum logs** automatically collected data on all messages posted to the player and GM forums. At the end of each session, forum logs were manually archived in individual MS Word documents for each thread. Their format preserved the title of each thread/message, player/poster identification, the date-and-time stamp, and the message contents. The OLIVE server generated **voice chat logs** from its VE recordings during playback of recorded sessions.

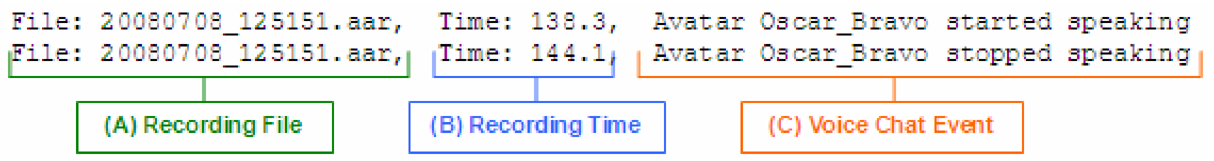

Figure 6. Voice-chat recording files (A) registered speaking events (C), capturing start and stop times (B) relative to the recording start time, whenever the mic was open.

The GM kept an **electronic record** of puzzles solved and incorrect submissions. Usability engineers took electronic, time-stamped **notes** on players' activities, using a fly-on-the-wall protocol, having no interactions with the players. **Video and audio recordings** of sessions facilitated investigation of incidents flagged in observers' notes. During post-gameplay discussions, observers noted comments about participants' experiences playing PanelPuzzle.

Results

Bravo (FD, RV) ranked fourth (last) in points (0), fourth (last) in the number of puzzles solved (0), and third in the number of unique pieces found (23). **Charlie** (FI, RM) ranked third in points earned (200), third in the number of puzzles solved (1), and first in the number of unique pieces found (27). **Delta** (FD, RM) ranked first (tied) in points earned (250), first (tied) in the number of puzzles solved (2), and fourth (last) in the number of unique pieces found (18). **Echo** (FI, RV) ranked first (tied) in points earned, first (tied) in the number of puzzles solved (2), and second in the number of unique pieces found (24).

Communication during Gameplay

We consider discourse essential to collaboration. Players' comments reinforced this. When asked how they collaborated, all players except one in Bravo (0 points) discussed collaboration in terms of communication. Players rated the helpfulness of collaboration: Bravo 6.7; Charlie (200) 6.8; Delta (250) 6.8; Echo (250) 5.5.

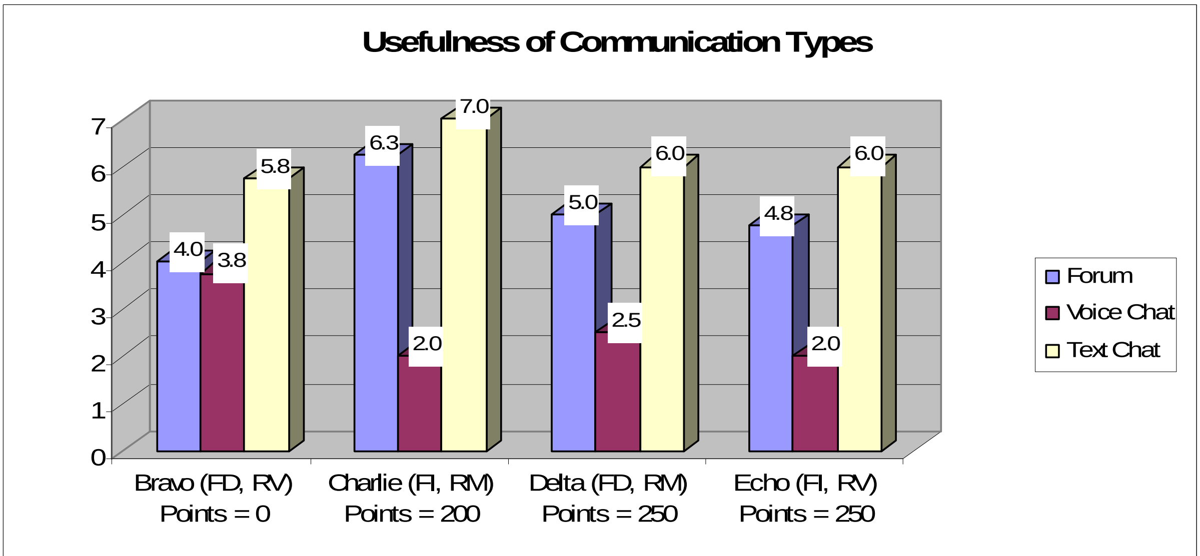

Figure 7. All teams found text chat the most useful communication type and voice chat the least.

Bravo (0 points), was fourth (last) in the number of text-chat messages sent (177), and second in forum messages posted (50). Charlie (200) ranked second in the number of text-chat messages sent (273), and fourth in forum messages (20). Delta (250) was first in text-chat messages (395), and third in forum messages (23). Echo (250) was third in text-chat messages (221), and first in forum messages (51).

Moments after gameplay started, joking ceased and communication focused on strategy and finding pieces. We asked players what percentage of in-game time they spent communicating with teammates and then averaged the responses to understand teams' perceptions of time spent on team communication.

Table 1. Players communicated with each other via text chat, a forum or the mic. They communicated with the GM via a forum.

Condition	Bravo (FD, RV)	Charlie (FI, RM)	Delta (FD, RM)	Echo (FI, RV)
Points earned	0	200	250	250
Total messages	398	307	506	279
Total messages among players	338	301	499	276
Gameplay time spent communicating with teammates	50.5%	60.5%	36.8%	45.5%
Text Chat				
Text chat messages to GM	0	8	1	2
Text chat messages to teammates	177	273	395	221
Total chat Messages	177	281	396	223
Forums				
Forum messages to teammates	19	20	18	48
Player forum threads	6	6	5	7
Forum messages to GM	60	0	7	3
GM forum threads	24	0	3	2
Total forum messages	79	26	25	51
Voice Chat				
Total time in minutes spent in voice chat	17.2	.87	6.2	.29
Average length in seconds of messages	4.42	.10	2.74	.97
Messages with human speech	142	0	85	5
Button presses not followed by voice chat	93	7	51	13

Despite complaints about its usability, players favored text chat over voice and forums. Zulu_Delta explained why, "[The] forum takes too long for a game that is timed and the voice chat was only for talking locally."

Efficiency

Players were aware of a defined end game condition with a 75 minute time span from start to end of gameplay. Efficiency metrics were collected within this time span. Speed of gameplay was important and, in this way, efficiency impacted effectiveness, e.g., Delta (250 points) posted a correct solution just after time ran out, gaining no more points.

Table 2. For gameplay speed and errors during requests for piece insertion, the lower the value, the higher the team's efficiency. Gray areas indicate no puzzle was solved.

Condition	Bravo (FD, RV)	Charlie (FI, RM)	Delta (FD, RM)	Echo (FI, RV)
Points earned	0	200	250	250
Minutes to solve first puzzle		62	48	40
Sections in first puzzle solved		5	3	5
Minutes to solve second puzzle			11	25
Sections in second puzzle solved			5	3
Requested panel and section correct	14	8	8	13
Requested section wrong	13	12	0	14
Requested panel wrong	3	0	0	0
Requested undo, cancel, or clear	0	0	0	4
Request formatting wrong	7	1	0	0
Total requests (includes inserts and moves, no duplicates)	37	21	8	31
Duplicate insertion requests	1	2	0	2

No team submitted an incorrectly assembled puzzle. In the ratio of puzzles started to puzzles completed *puzzles started* refers to puzzles for which at least one piece was inserted into a panel, regardless of whether the panel or section were correct. Delta (FD, RM, 250 points) had a 2:2 ratio; Echo (FI, RV, 250) had 5:2; Charlie (FI, RM, 200) 6:1; and Bravo (FD, RV, 0) 9:0.

Effectiveness

Effectiveness assessed outcome success, and was impacted by communication. Charlie (200 points) found all the pieces for a four-piece puzzle, and communicated about them within the time limit, but only discussed inserting one. Completing this puzzle would have resulted in the highest point total.

Table 3. In assessing effectiveness during gameplay, pieces inserted refers to pieces the GM was asked to insert at least once.

Condition	Bravo (FD, RV)	Charlie FI, RM)	Delta (FD, RM)	Echo (FI, RV)
Points earned	0	200	250	250
Puzzles solved	0	1	2	2
Unique pieces found	23	27	18	24
Requests for unique (not necessarily correct) piece insertion	21	12	8	17
Pieces inserted or moved correctly into panels	14	8	8	13
Sections in 1st puzzle solved		5	3	5
Sections in 2nd puzzle solved			5	3

Bravo (FD, RV) used panels as a visual aid to puzzle solving by inserting and moving pieces frequently to obtain different views of puzzles-in-progress. Bravo inserted more pieces (21) than any other team, but solved no puzzles.

Satisfaction

Table 4. Team averages for overall satisfaction and comfort were positive to high.

Condition	Bravo (FD, RV)	Charlie (FI, RM)	Delta (FD, RM)	Echo (FI, RV)
Points earned	0	200	250	250
Overall Satisfaction	6.0	5.3	5.0	5.3
Comfort	6.3	6.3	5.0	5.3

Foxtrot_Charlie based his rating (6) for overall satisfaction on the fact that he found PanelPuzzle to be engaging. One Bravo (FD, RV), two Charlie (FI, RM), one Delta (FD, RM) and one Echo (FI, RV) player reported that ease of playing motivated their high comfort ratings. Zulu_Echo said lack of team organization caused discomfort, but still gave a high comfort rating (6). Bravo's satisfaction scores were high because they were enjoying the environment and remained engaged.

Players perceived challenge as a dimension of both enjoyment and engagement which, in turn, are dimensions of satisfaction. Two players reported challenge as the motivation for their very high overall satisfaction ratings (6). Hotel_Bravo said, "… it was challenging to distinguish what the puzzle actually was and actually putting the pieces in the correct position … I love a challenge."

Table 5. Team averages for enjoyment were positive to high.

Condition	Bravo (FD, RV)	Charlie (FI, RM)	Delta (FD, RM)	Echo (FI, RV)
Points earned	0	200	250	250
Enjoyment	6.3	6.3	5	5.7

Players' comments supported our assertion that **enjoyment** is a dimension of satisfaction. One player from each team attributed their very high overall satisfaction ratings to the fact that PanelPuzzle was fun. Some players factored collaboration into their satisfaction ratings, e.g., Tango_Bravo cited the "team element" as motivation for his enjoyment rating (7). Higher point totals did not coincide with higher enjoyment ratings. Bravo (0 points) and Charlie (200) gave ratings of 6.3, but the teams that tied for highest points (250) gave lower ratings Delta, 5 and Echo, 5.7.

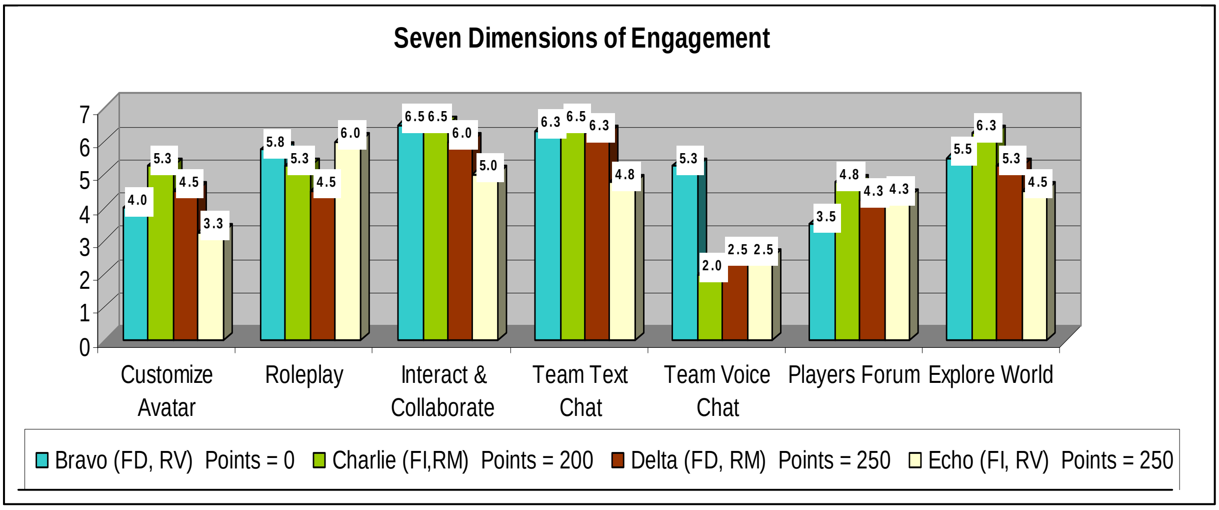

Figure 8. Players reported the degree to which seven factors impacted engagement during gameplay.

Challenge fostered **engagement**. Four players responded to a question asking for the top factor that kept them engaged. Of these, three cited challenges, including puzzle assembly and inter-team competition. Ranking engagement factors over all the dimensions of engagement and challenge combined, ten players reported dimensions of challenge as the top factor that engaged them: four cited hunting for pieces; four, assembling pieces; one, accumulating points; and one, trying to finish as fast as possible.

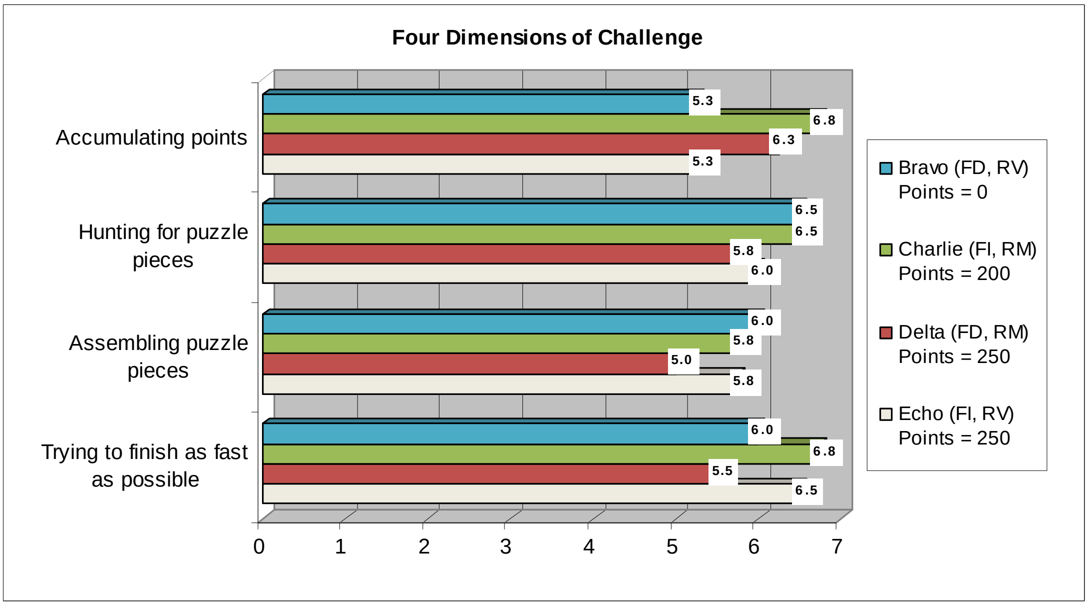

Figure 9. Ratings for four dimensions of challenge as measures of engagement were positive to high.

Findings

H1: Communication among teammates will facilitate collaboration, resulting in higher scores for teams that communicate most effectively. The principal factor impacting success was the effectiveness of collaboration strategies. Communication was the key to collaboration. Usually, one player stayed near the panels; others checked them, then dispersed to seek pieces. Thus, teammates rarely shared visual feedback. Teams depended on communication to solve puzzles. Bravo (0 points), which had little collaboration, sent more messages than Charlie (200) or Echo (250), but earned no points because of ineffective communication. Delta (250), which exhibited the most intra-team communication tied for the highest score.

True to Prensky's (2004) claim, digital native players preferred text over voice. All teams except Bravo (FD, RV) gave voice chat very low engagement ratings. Early in the game, players used the mic to organize before commencing their tasks, then abandoned it. Players preferred text chat for tactical communication and the forum for strategic communication. Mic abandonment may be attributable to several factors. Players were in the same room and could hear each other. They sometimes did not understand or forgot that mic activation required pressing a button. As players dispersed in the CVE, mic capacity diminished and it was not possible to hear other players. Possibly, having three communication mechanisms was counter-efficient; most games provide no more than two. Human factors also impacted, e.g., Lima_Charlie said, “I like typing to people while I play. It felt weird to use the microphone, so I didn't.” Text chat and forums provided histories that proved integral to executing team strategies; voice chat did not.

The most successful team, Delta (250 points), were the most active and effective communicators with 506 messages. Delta solved two puzzles faster than Echo (250), manipulating only those pieces that they inserted. As gameplay ended, Delta were on the brink of solving a five-piece puzzle. They made no requests for incorrect insertions. They sent the most intra-team chat messages. They were second for number of player forum messages.

Teams gave their highest engagement ratings for communication media to text-chat messages. RM conditions sent more text messages than RV conditions. Charlie, (FI, RM, 200 points) sent 281; Delta (FD, RM, 250) sent the most text messages, 506. Bravo (FD, RV, 0), sent the fewest text chat messages, 177. Echo (FI, RV, 250), sent the lowest total number of messages, 279, and was third in text chat messages, 223.

During text chat, we expected players to import linguistic traits from emailing and text messaging. Both were popular in the demographic survey. Chat analysis disclosed brevity, quick topic changes and deviations from Standard English, aligning with the observations of Zubek, Khoo (2000, 2002) and Chao (2001).

PanelPuzzle success depended on building collaborative knowledge. Teams approached this differently, depending on their assigned condition and their success at communicating and collaborating. Different strategies arose for teammate communication. The players’ forum proved a locus for strategic communication. Each team stored piece information in the forum. Each used text-chat for discussions. Teams developed collaboration strategies at different points in gameplay. Echo (FI, RV, 250 points) developed their forum strategy quickly, in the eighth text-chat message, approximately five minutes before gameplay started; Delta (FD, RM, 250) approximately seven minutes into gameplay (in the 102nd text-chat message); and Charlie (FI, RM, 200) approximately 16 minutes into gameplay (in the 110th text-chat message). Only one Bravo (FD, RV, 0) player posted information on pieces to the forum (13 minutes into gameplay). Bravo evidenced no strategy for collaboration, instead devoting planning time to a strategy for finding as many pieces as possible by distributing avatars geographically.

All teams gave neutral ratings to the engagement of team communication via the forum, but all used the players forum. Organizing pieces in the players forum varied among teams. Delta (250 points) chose to post all information regarding pieces in one thread allowing teammates to view all pieces found in one location without needing to navigate multiple threads. However, this approach may have increased the time taken to find information about pieces belonging to a particular panel. Charlie (200) posted piece information in two separate threads, one for piece numbers with descriptions and another for assembling pieces. Echo (250) organized piece information in three separate threads, one for each puzzle size. This approach facilitated puzzle solving by making it easier to view pieces of the same size puzzle together; but made it more difficult to view all of the pieces found at once (e.g., in order to avoid duplicate postings).

H2: Roles positively affect group dynamics in a CVE. H2 proved true. Roles helped teams organize and collaborate and made them more effective. If teams did not adopt roles at the outset, they experienced little collaboration and no success. Even when lines between roles blurred, the advantages of having roles from the outset carried teams forward to earning points. Both RV and RM players were self-organizing in determining roles and sometimes changed roles in-game. Contrary to our expectations, most RV players assumed roles. Players who abandoned their roles or did not play roles said after gameplay that they wished they had done so. Roles provide structure. Lacking agreement on roles, Echo (FI, RV) had two competing solvers, resulting in conflicts. Successful teams took roles seriously; they did not consider roles to foster enjoyment. Conversely, Bravo (FD, RV), which had the least organized roles gave the highest rating (6.8) for enjoying roles and the highest rating (6.0) for impact of roles on progress. RM conditions did not experience higher satisfaction. Successful teams did not perceive roles' positive impact on progress, giving this factor neutral to negative ratings. Players focused on winning rather than roleplay. Indeed, across roles, high point totals did not coincide with high enjoyment ratings. Bravo (RV, 0 points) and Charlie (RM, 200) gave ratings of 6.3, but the teams that tied for highest points (250) gave lower ratings Delta (RM) gave 5 and Echo (RV) 4.5. Delta, the highest achiever, gave the lowest rating (4.5) to the engagement of roleplay.

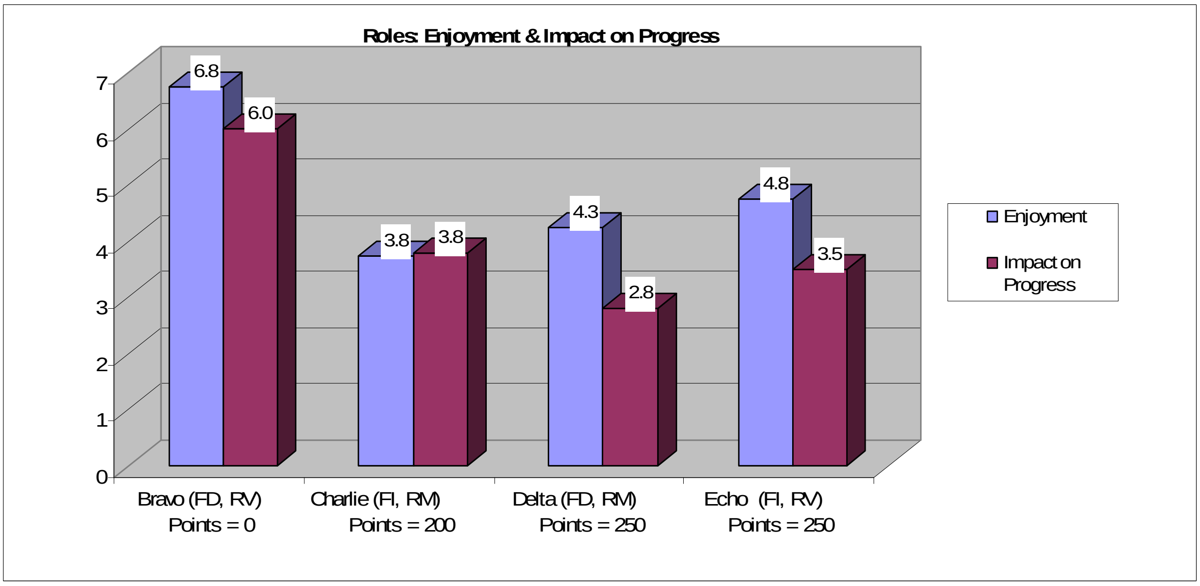

Figure 10. Other than Bravo, teams gave neutral ratings to roles' impact on progress.

As Yee (2006) observed, players changed the story. Roles evolved during gameplay; players switched roles and assumed multiple roles, even in RM conditions. When the Delta (RM) Puzzle Coordinator became confused, a teammate solved puzzles alone on a sheet of paper and posted them to the forum. When two Echo (RV) players fought over the role of GM Communicator, their teammates redefined their own roles to maintain efficiency in finding pieces. Collaboration on finding pieces became the priority goal for most players; sometimes even the GM coordinator sought pieces. Zulu_Delta observed, “Everyone was working together. It seemed like everything was working smoothly without people being sticklers for their roles.”

As expected, players perceived the GM Communicator role to be the leader role. Leadership emerged differently in each condition with at least one player becoming the GM Communicator. In RV conditions, players either adopted an unofficial leader or later indicated that they would elect a leader if given another chance to play PanelPuzzle. We expected roles to facilitate collaboration, resulting in superior collaboration in RM conditions. Positing that the leadership role empowers group interaction as Maybury (2001) asserted, a survey question asked how helpful collaboration was in solving puzzles. RM teams Charlie and Delta gave 6.8 ratings. These were higher than RV teams, Bravo. 5.7 and Echo, 5.5. Echo, where there was competition for leadership, gave the lowest score. The RM teams’ very high ratings support the expectation of more efficiency in RM conditions.

A one-way analysis of variance (ANOVA) test was performed between the number of chat messages sent by each player and that player's answer to the question, "How enjoyable was your role?" The analysis was significant, $F(1, 14) = 9.488062$, $p < .01$. A regression test indicated that the number of chat messages sent and a player’s enjoyment of roles were negatively correlated with a p value of 0.008145. We interpreted this to mean that while effective communication was crucial to a team's success, ineffective communication diminished role enjoyment. Players who were forced to repeat themselves or ask teammates for clarification enjoyed their roles less than those who communicated effectively with their teammates. This showed us that although communication is a key to success in a CVE, more is not always better. A CVE must give its users the ability to communicate effectively. Otherwise, enjoyment diminishes as players are forced to repeat themselves or ask teammates to do so.

H3: Deferred feedback will impact gameplay strategy. Observational evidence supported H3. After experiencing deferred feedback, Delta (FD, RM) adopted an efficient and effective strategy for working under this condition. Teammates from FD conditions Bravo and Delta complained about delayed feedback. One instance supported the observations of Gergle et al (2006) that delayed visual feedback impairs communication. A Charlie (FI, RM) player's impatience with the GM's response time caused a duplicate insertion request.

There were two principal strategies for communicating with the GM. We called these *piece-at-a-time* and *panel-at-a-time*, referring to how the teams communicated piece movement instructions to the GM. Bravo (FD, RV), Charlie (FI, RM), and Echo (FI, RV) requested insertion as pieces were found. Bravo correctly expressed 21 requests for single-piece insertions, waiting an average of 2.5 minutes for feedback. Once, Bravo requested insertion of eight pieces simultaneously, waiting seven minutes from the time of the request until the last piece was inserted. The least successful team experienced the longest feedback delay, yet did not change its strategy. Delta (FD, RM) alone waited until they were confident they had all the pieces assembled, then requested insertion of an entire panel. This proved the most effective strategy. Delta made only two requests, one to insert three pieces and one for five pieces, each time completing a puzzle. Although Delta waited an average of 4.5 minutes between request and insertion of the last piece, they minimized their waiting time by minimizing requests. In both FD conditions, searches continued during delays.

Players' comments uncovered an issue related to delayed feedback which aligned with our expectations for digital natives: a desire for fast gameplay. Anticipating this, we had provided teleportation and the ability to run, but some players found these insufficient. Golf_Echo (overall satisfaction, 4) explained, "I want to fly to travel faster, what is the point of having the same constraints of the real world in the virtual one?" Three players complained about the avatar's speed. Others were impatient with the responsiveness of PanelPuzzle. FI players complained about the GM's speed. Players wanted keyboard shortcuts.

Conclusions

Our UE approach investigated efficiency, effectiveness and satisfaction during gameplay. We examined factors common to gameplay and information analysis: communication, roleplay and feedback. Although there were only four players in each of four sessions, the study uncovered many aspects of team dynamics during gameplay. The dimensions of satisfaction we identified, engagement and enjoyment, with challenge as an aspect of each, coincided with players' mental models of what makes gameplay satisfying. Several factors influenced the ability to accrue points in PanelPuzzle. Good communication strategies promoted collaborative play and building collaborative knowledge. Satisfaction kept players engaged. No players abandoned the game and none gave low overall satisfaction ratings. For satisfaction, engagement is more important than winning. Leadership promotes team success. Digital natives developed strategies to prevent deferred feedback from impeding success.

To leverage digital natives' skills and expectations, the next generation of VA tools must promote engagement. Analysis is tedious and answers can remain elusive. Engagement is integral to the satisfaction that fosters perseverance. To leverage digital native behaviors, it is essential that VA tools facilitate collaboration by providing effective communication mechanisms and access to a history of these communications. Players consulted communication histories to organize collaborative knowledge. Indeed, the most organized team, Delta (FD, RM) was the most successful. Digital natives can handle delayed feedback in VA tools. Obligating roleplaying is not necessary for VA tools, but they should accommodate a leadership role because it is likely that leaders will emerge.

Future Work

Digital natives' gameplay behaviors and the potential to reflect them in VA software combine to offer a rich area to investigate from a UE perspective. We found collaboration among digital natives building collaborative knowledge during a video game to be a complex area where findings generated questions opening opportunities for research as richer CVEs become part of VA tools.

Our next step is analysis of linguistic data from PanelPuzzle. Khoo and Zubek (Zubek & Khoo, 2000; Khoo & Zubek 2002) observed that emotional involvement is key to gameplay enjoyment. They cited emotional involvement and verbal posturing as key to social interactions during gameplay. Does dialog during PanelPuzzle gameplay indicate these factors and their impacts on team success and satisfaction? PanelPuzzle required the analytical skills of decision making and problem solving. Are there are discourse patterns that evidence collaborative decisions? Are there patterns that facilitate puzzle solving and information analysis?

Will collaboration change in a more controlled collaborative environment? If players are geographically dispersed and voice chat facilitated, will communication change? We identified innovative metrics for human interaction with CVEs. Can we expand and apply them to VA tools? For example, we want to further study measuring engagement. Will players who establish a bond with their avatar become more engaged in the CVE and their task?

We designed the PanelPuzzle game to be played in a "mirror world" environment, a literal representation of the real world in digital form. What if we design the PanelPuzzle game in a true virtual world, where the environment, simulation, and physics can be modified to place the player in a figurative space that better matches the cognitive model required by the game? Will the game fundamentally change – will the CVE be the actual game? Will adapting to the laws of the CVE differ significantly for each individual as the frames of reference to the laws of the literal world no longer apply? The possibilities for further research are intriguing.

Acknowledgements

This work was supported in part by the Intelligence Advanced Research Projects Activity, the National Institute of Standards and Technology (NIST) and Forterra Systems Inc. The authors would like to thank Yee-Yin Choong at NIST for contributing to the experiment and Alex Wang for supporting this work during his NIST internship.

Bibliography

Capps, M., McDowell, P., & Zyda, M. (2001). A future for entertainment-defense research collaboration. I*EEE Computer Graphics and Applications*, 21(1), 37–43.

Chao, D. (2001). Doom as an interface for process management. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 152–157). March 31 – April 5, 2001, Seattle, Washington.

Chao, D. (2004). Computer games as interfaces. *Interactions* 11(5), 71–72.

Choong, Y. and O'Connell, T. (2008). Planning user-centered evaluations for interactive information visualizations. *Proceedings of Usability Professionals Association (UPA) Annual International Conference*. (CD-ROM). June 16 – 20, 2008, Baltimore, MD.

de Freitas, S. (2006). *Learning in immersive worlds: A review of game-based learning.* JISC (Joint Informational Systems Committee) Report. Retrieved February 5, 2009 from http://www.jisc.ac.uk/media/documents/programmes/elearninginnovation/gamingreport_v3.pdf

ESA (Entertainment Software Association). (2008). *Video games and the workplace*. Retrieved January 2, 2009 from http://www.theesa.com/gamesindailylife/workplace.asp

Federoff, M. (2002). *Heuristics and usability guidelines for the creation and evaluation of fun in video games.* Unpublished master's thesis, Department of Telecommunications, Indiana University, Bloomington, IN.

Gergle, D., Kraut, R., & Fussell, S. (2006). The impact of delayed visual feedback on collaborative performance. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1303–1312). April 22 - 27, 2006, Montréal, Québec, Canada.

Green, C., & Bavelier, D. (2003, May). Action video game modifies visual selective attention. [Letters to Nature.] *Nature*, *423*, 534–537.

Green, C., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18 (1),* 88–94.

Ham, H. & Lee, Y. (2006). An empirical study for quantitative evaluation of game satisfaction. *Proceedings of ICHIT '06. International Conference on Hybrid Information Technology* (Volume 2, pp. 724–729). November 9 – 11, 2006, Cheju Island, Korea.

Hendrick, A., Knight, J., Menaker, E., O'Connor, D., & Robbins, C. (2008). Extension of SCORM based learning content into game based, multiplayer training environments. (JADL N61339-07-C-0046 Final Report). Retrieved January 12, 2009 from January 30, 2009 from https://acc.dau.mil/GetAttachment.aspx?id=229915&pname=file&aid=37133&lang=en-US

ISO (International Organization for Standardization). (1998). Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability.

Jegers, K. (2008). Investigating the applicability of usability and playability heuristics for evaluation of pervasive games. *Proceedings of the Third International Conference on Internet and Web Applications and Services, ICIW '08.(pp.* 656–661). June 8 – 13, 2008, Athens, Greece.

Jensen, C., Farnham, S., Drucker, S., & Kollock, P. (2000). The effect of communication modality on cooperation in online environments. *Proceedings of the SIGCHI Conference on*

Human Factors in Computing Systems (pp. 470–477). April 01 - 06, 2000, The Hague, The Netherlands.

Khoo, A. & Zubek, R. (2002). Applying inexpensive AI techniques to computer games. *IEEE Intelligent Systems*, 17(4), 48–53.

Malone, T. (1982). Heuristics for designing enjoyable user interfaces: Lessons from computer games. *Proceedings of the 1982 Conference on Human Factors in Computing Systems* (pp. 63–68). March 15 - 17, 1982, Gaithersburg, Maryland.

Maybury, M. (2001). Collaborative virtual environments for analysis and decision support. *Commun. ACM*, 44(12), 51–54.

O'Connell, T., & Choong, Y. (2008). Metrics for measuring human interaction with interactive visualizations for information analysis. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1493–1496). April 5 - 10, 2008, Florence, Italy.

O'Connell, T., Choong, Y., Grantham, J., Moriarty, M. & Wong, W. (2008). From a video game in a virtual world to collaborative visual analytic tools. *Journal of Virtual World Research*, 1(1). Retrieved February 5, 2009 from http://journals.tdl.org/jvwr/article/view/292/246

Prensky, M. (2001a). Digital natives, digital immigrants, part I. *On the Horizon*, 9(5), 1–6.

Prensky, M. (2001b). Digital natives, digital immigrants, part II: Do they really think differently? *On the Horizon*, 9(6), 15–24.

Prensky, M. (2004). The emerging online life of the digital native. Retrieved Jan 13, 2009 from http://www.marcprensky.com/writing/Prensky-The_Emerging_Online_Life_of_the_Digital_Native-03.pdf

Prensky, M. (2005). Listen to the natives. *Educational Leadership*, 63(4), 8–13.

Squire, K. (2005). *Game-based learning: Present and future state of the field. Report to the Masie Consortium*. Retrieved January 13, 2009 from http://www.masieweb.com/dmdocuments/Game-Based_Learning.pdf

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review*, 41(2) pp. 16–30.

von Ahn, L. & Dabbish, L. (2008). Designing games with a purpose. *Commun. ACM*, 51(8), 58-67.

Whang, L. & Chang, G. (2004). Lifestyles of virtual world residents: Living in the on-line game "Lineage". *CyberPsychology & Behavior*, 7(5) 592–600.

Yee, N. (2006). Motivations for play in online games. *CyberPsychology & Behavior*, 9(6), 772–775.

Zubek, R., & Khoo, A. (2002). Making the human care: On building engaging bots. *AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment.* AAAI Technical Report SS-02-01.

Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, 38(9), 25–32.

Certain commercial equipment, instruments, materials, services or companies are identified in this paper in order to specify the experimental procedure. This in no way implies endorsement or recommendation by NIST.