

ORIGINAL ARTICLE

Can You Hear Me Now? The Impact of Voice in an Online Gaming Community

Dmitri Williams¹, Scott Caplan², & Li Xiong¹

1 Annenberg School for Communication, University of Southern California, Los Angeles, CA 90089-0281

2 Department of Communication, University of Delaware, Newark, DE 19716

This paper reports the results of a controlled field experiment in which voice communication was introduced into an existing online community (online gaming guilds within the popular game “World of Warcraft”), comparing a mix of voice and text with text only. Quantitative results suggest increases in liking and trust due to the addition of voice, as well as insulation from unexpected negative impacts of text-only play. The findings are discussed with respect to social capital, cyberbalkanization, and the general computer-mediated communication literature, with special attention paid to social information processing theory.

doi:10.1111/j.1468-2958.2007.00306.x

Voice has several functions in human communication beyond the transmission of content. For evolutionary reasons, humans have come to use voice to determine gender, personality, and intentions (Nass & Gong, 2000), and when voice is absent, we rely on other cues or interact differently (Walther, 1992). The current moment in communication technology can therefore be seen as an aberration in our long history as communicators; our telecommunications systems have given us mastery over voice across large distances for over a century, yet our most vibrant modern communications tool—the Internet—is mostly text- and graphics-based with no voice.

The diminished nonverbal cues in text-based computer-mediated communication (CMC) have given rise to a variety of theories seeking to explain how reduced relational cues may affect interpersonal interactions online (for reviews, see Walther, 1996; Walther & Parks, 2002). For example, the cues-filtered-out model (Culnan & Markus, 1987) hypothesized that the relative lack of nonverbal cues in text-only, computer-mediated interaction stifles the development of interpersonal relationships and hinders impression formation. However, most researchers have dismissed the cues-filtered-out model in light of evidence indicating that CMC interactions are

Corresponding author: Dmitri Williams; e-mail: dcwillia@usc.edu

A previous version of this paper was presented at the 2006 meeting of the International Communication Association.

not lacking in cues but rather require more time than face-to-face (FtF) conversations to attain similar levels of intimacy and depth of impression formation (Walther, 1992, 1996; Walther, Anderson, & Park, 1994). As a result, researchers have studied a range of social phenomena online as people have sought, avoided, and managed communication in text-based online environments. Yet, this text-only period may well be drawing to a close with the advent of an increasingly popular technology known as Voice-over Internet Protocol (VoIP).

VoIP is the use of computer networks, including the global Internet, to carry a digitized voice signal. People engaging in CMC over the Internet can thus interact with one another via voice, instead of or in addition to text. VoIP is inexpensive, can connect with existing phone service, and allows for many-to-many communication. For groups that rely on text-only communication such as game players, and even some corporations (Olsen, 2005), the move to VoIP represents a fundamental shift in the quality of that communication.

With the introduction of vocal cues, the norm of relative anonymity in games and other Internet activities may change. The now-famous *New Yorker* cartoon featuring a canine at a keyboard, which reads "On the Internet, nobody knows you're a dog," has been a clever way to analogize these issues. The dog, unforced to speak or show itself, can happily pass as a human. Likewise, the dog is not discriminated against. This analogy provokes questions about the social networks that result from such anonymity: Although the lack of vocal and expression-based cues represents a paucity of information, it provides an opportunity for different kinds of people to connect and interact. It is the very lack of cues that enables teenagers and elders, housewives and construction workers, rich and poor, and countless other less likely offline interactions to "talk" with one another online. Unable to see one another and likely less aware of their differences, the parties are less hampered by restrictive social norms based purely on their demographics. Conversations can instead be based on their mutual interests or activities. This horizon-broadening function is one of the facets that optimists regularly cite about the impact of the Internet (Negroponte, 1995; Wellman et al., 2003). Nevertheless, that same lack of cues means that the horizons are seldom broadened in practice. If the two parties are of radically different religions or political ideologies, but never know this about each other, the practical effect of their diversity is negated. But technology is advancing and may well soon reintroduce those cues. When voice and video return, what happens to the conversations and networks?

As previously text-only communications between individuals and groups add a voice component, what will the social and psychological impacts be? Specifically, how will voice impact so-called virtual communities? VoIP technology relates to a channel effects question raised by contemporary communication theorists: What role does channel (i.e., voice and text vs. text only) play in how interpersonal processes differ online compared to FtF contexts? Even as VoIP systems become more common (Fitzgerald, 2004), the research on their social uses has been limited to interface studies (Clarke, 2003), with the result that the systems have not been

immediately appealing to consumers (Burke, 2002), which is typical in the early stages of diffusion (Rogers, 1995). What usability studies do not consider are traditional CMC questions. What will happen to the quality of the communication and relationships as media change? What will happen to communities? This paper presents an experimental test of these questions by examining the social impact that VoIP systems have on preexisting communities found in large, vibrant online video games.

One popular example of online communities: Massive multiplayer online games and player guilds

Video games have long been a site for technological innovation and early adoption (Burnham, 2001). Game makers have invented new technologies, whereas game players have from the start innovated “hacks” to improve and alter the game code and its social affordances (Levy, 1994). It should not be surprising, then, that a social video game would be the site of a social technological adoption. These games are massively multiplayer online (MMO) games. They are the graphical descendents of “multiuser dungeons,” the text-based networked games in which players compete and collaborate in fantastical imagined settings (Trubshaw & Bartle, 1978). Those text-only environments have been recognized as important sites of social activity (Kendall, 2002); rather than a compelling narrative or a pleasing aesthetic, it is the existence of the other players that drives these games’ popularity (Turkle, 1995). In an MMO, players create and direct an avatar on the screen. In 3-D space, players see their and other players’ avatars in the virtual world. They can chat with them via text messages and engage in adventures together—slaying dragons, searching for treasures, exploring fantastic areas, and so forth.

The most typical activity in MMOs involves a player trying to advance their avatar from an initial level of relative powerlessness to more advanced levels imbued with superhuman abilities and skills. This “leveling” process takes place by completing quests and gaining experience and items. Notably, the process has been engineered in most MMOs to work more easily when done in groups. Thus, groups of players regularly band together to explore, fight monsters, exchange information—and along the way, simply talk—often in groups of two to five players. Once the players reach the highest level possible, they are incentivized to continue their adventures in what is known as the “end game.” At this stage, players confront larger and more complex tasks and monsters that can only be defeated with groups of 10–60 people. Thus, there is a need to group with a steadily larger number of others as the game goes on in order to access the most glamorous content.

An important form of group in MMOs is known as a “guild.” Unlike the ad hoc groups that can form between strangers in the game world, guilds are more permanent associations with rosters and hierarchies. Compared to solo players, guild members can more easily communicate with more people and can reduce the risks of dealing with potentially troubling players who may not have the same goals or play styles. Guilds then often form around like-minded players or players with a

connection that predates the game. Membership in a guild bestows a degree of status, but most importantly, it gives the player a readily available roster of vetted playmates. Research on World of Warcraft (WoW) guilds (Williams et al., 2006) reveals that guilds are formed for both social and strategic support. They range in size from a small clutch of four or five players to many hundreds of players, sometimes even extending over more than one game. Within guilds, many players receive social, emotional, and psychological support (Taylor, 2006), although other guilds function purely as task-oriented groups with no social component. A guildmate may be an unknown stranger, a strong friend, or something in between. Because game guilds must facilitate communication between large numbers of people within hierarchies, they have adapted communication modalities to support themselves. Most notably, guilds have begun to use out-of-game voice systems to supplement the text-based systems built into the game. Guild players then have the option to communicate via text or also with voice and often suit the medium to the task at hand (Williams et al., 2006); private and group conversations requiring fewer cues and less speed take place via text, and more complex and time-sensitive tasks move to voice.

Game players, especially those with preexisting friendships or guild memberships, are forming the classic early adopter group of VoIP technology. This timing makes the current state of diffusion ideal for an experiment examining the impact of vocal cues on interpersonal and social variables. There is no previous systematic research on game guilds, but there is a growing body of research within organizational communication about virtual group dynamics. That work suggests that longer-term groups, for example, guilds compared to ad hoc teams, are more likely to develop relationships but that the shift from one type of media to another in such teams has surprisingly little research on it (Poole & Zhang, 2005).

Theories converging on voice and gaming

How will voice impact the online social experience? The sociological literature on online communities as well as the CMC literature offer suggestions. The modern debate began with Rheingold (1993), Postman (1992), and Beniger (1987) but has its roots in the postindustrial tensions first identified by Tönnies (1957). Whereas modern scholars worry about whether online communities are “pseudocommunities,” in which the connections do not matter in any substantive way (Beniger, 1987), Tönnies first identified the tensions between the small, insular, and supportive folk-based community and the impersonal, modern state-imposed social order. The desire for this sense of supportive familial community in the face of a steady decline in civic life (Putnam, 2000) and communal spaces (Oldenburg, 1997) may be one factor pushing people into online communities, including complex games (Steinkuehler & Williams, 2006). Nie and others (Nie & Erbring, 2002; Nie & Hillygus, 2002) are concerned that vacuous online communities will *displace* existing communities and relationships that matter more and that CMC will never suffice to support real communities and relationships. Increases in time online combined with decreases in existing communications media are therefore often seen in a negative light.

Another concern is continued exposure to diversity. The new fear is that individuals will be so enabled in their ability to customize connectedness that they will fashion an experience that exposes them to only like-minded people and ideas—total control of their exposure could mean perfect isolation and lack of connection to demographic and ideological diversity. The awful downside of Negroponte's (1995) utopian "Daily Me" is a perfectly isolated interest group (Sunstein, 2001) in a system of "cyberbalkanization" (VanAlstyne & Brynjolffson, 1996) and an ultimate drop in social diversity for users. A useful conceptual framework in deciphering such online community dynamics is Putnam's (2000) adaptation of "social capital," which he breaks down into two component parts: bridging, or the ability to reach new ideas and people, and bonding, or the ability to secure practical and emotional support. The bonding function matches closely with the concerns over "pseudocommunity," and the bridging function matches closely with the concerns over cyberbalkanization. A test of these two indicators would address the original theoretical issues.

The existing CMC and human-computer interaction (HCI) literature on voice chat suggests that, when compared to text, it may improve the bridging function through the expansion of social cues, bringing communication closer to the presumed gold standard of FtF communication. In particular, prior research on VoIP and gaming has focused on usability issues and found that sociability improves when the players know precisely who is talking (Halloran, Fitzpatrick, Rogers, & Marshall, 2004), players are on the same team, and background noise is minimized (Hew, Gibbs, & Wadley, 2004). When these factors are not controlled, voice appeared to do little for sociability. In contrast, general research on audio has found enhanced communication and collaboration as compared to visual cues (Jensen, Farnham, Drucker, & Kollock, 2000) and has shown the ability to increase trust (Burgoon et al., 2002).

More general research and theory on CMC offer predictive frameworks about what will happen with the introduction of a richer communication medium. As more cues are supported by a medium, media richness theory (Daft & Lengel, 1984) suggests that it is better suited to maintaining relationships. The theory suggests a general ordering of media from rich to poor in this order: FtF, videoconference, audioconference, e-mail, written documents, and then numerical information (Rice & Gattiker, 2001). In keeping with the initial research on voice use in game groups, more complex tasks such as group organization and dealing with dynamic real-time situations require richer media (Connaughton & Daly, 2004). A test of text versus voice in a virtual team found that the voice group had a larger volume of communication (Salinas, 2005). More cues should lead to more information, both enabling more complex tasks and also helping the group members reduce uncertainty about each other. Text-based communications have frequently been noted to contain fewer cues, yet fewer cues can lead to both positive and negative outcomes (Culnan & Markus, 1987). Fewer cues lead to more antisocial flaming and longer decision processes but also to more egalitarian participation and less reliance on status (Garton, Haythornthwaite, & Wellman, 1999).

If the cues that go missing in text communication produce positive and negative social outcomes, reintroducing them ought to reverse those effects, albeit moderated by the setting. The handful of CMC-focused voice studies to date offer general support for this idea. The first functions of voice are to fill in those missing cues. The listener wants to know first if the speaker is human, then what gender, and then what personality (Nass & Gong, 2000). Listeners then use this information to make judgments about the person's intelligence (Reeves & Nass, 1996) and to decide if the person has a compatible personality to the listener (Nass & Lee, 2000). More information may help individuals decide who they *do not* like as much as who they do. Thus, the true outcome of moving from voice to text in an online community is likely complex. One study suggests that voice will improve trust and cooperation (Jensen et al., 2000), but the impacts on liking and the social variables associated with bridging and bonding remain less clear. The present study hopes to replicate and clarify those associations.

Other theories suggest that it is not the richness of the medium but the intent of the users or the influence of group norms that play the biggest role. In particular, social information processing (SIP) theory suggests that CMC users will adapt to a medium over time and will work with it to develop strategies, which compensate for any lack of richness (Walther, 1994; Walther & Parks, 2002). The support for this approach is strong and growing but is also marked by subject groups with no prior history and for relatively short periods of time. Alternatively, the social identity model of deindividuation effects (SIDE) theory suggests that group structures and norms drive the CMC outcomes (Postmes, Spears, & Lea, 2000), with conformity to those norms increasing generally over time (Postmes, Spears, Sakhel, & deGroot, 2001). But although the support for SIDE is also strong, the theory cannot account for individual-level changes and everything must stem from group-based processes (Walther, 2006).

Hypotheses and research questions

The assumptions offered by the CMC research are that text-only communication has fewer cues and is less rich in information than FtF or voice communication. However, SIP suggests that any differences in richness will reconverge if they are given enough time. Thus, voice will lead to richer communication and more information, and potentially more social connectedness, but only in the short term. Is there a truly qualitative difference beyond that? Introducing cues allows the presence of true boundary crossing as individuals learn the demographic backgrounds and personalities of the other players, yet those same characteristics may offer them reasons to segment themselves into groups of similar people. In the research detailed below, the richness of the available communications media mix was manipulated to test these issues. The following research questions and hypotheses were tested in an experimental design among text-communicating game players in which preexisting WoW player guilds were given access to voice communication (in addition to the existing text chat) or remained in a text-only communication environment. In other words, the richness of the available media was varied to test questions of social and

relationship impact within the competing frameworks of uncertainty reduction, media richness, SIP, and SIDE.

The first question involves how voice will affect the composition of online groups.

RQ1: Will voice lead to more balkanization or more bridging?

The CMC literature otherwise predicts more positive outcomes:

H1: Voice will lead to stronger relationships.

H2: Voice will lead to greater levels of bonding.

H3: Voice will lead to greater levels of trust.

Nie's question about social displacement can be examined as well, although there is evidence on both sides of the debate:

RQ2: Will VoIP use displace local real-world social capital?

Putnam suggests that this displacement of social capital can lead to psychological distance from friends, coworkers, and relatives; replacing rich FtF social connections with poorer, mediated ones could damage relationships. The resulting decline in solid relationships could manifest in declining happiness and an increase in loneliness.

RQ3: Will voice use have an impact on affective state?

Method

This study used a month-long, three-wave, field-based panel study with a control group to test the hypotheses and research questions. The method tested the transition of preexisting groups from text to voice rather than the effect of the game itself. Those data were reflected in the control group baseline. Participants were current players of *WoW* who had never used voice communication in the game. Participants were recruited and assigned randomly to a treatment group that received free hardware, software, and a VoIP service called "Teamspeak" or to a control group that continued to play the game using only in-game text to communicate. Surveys with repeated measures were completed via secured Web sites, beginning with a pretest measure before group assignment. Treatment groups were mailed the voice hardware and start-up instructions. Based on pretesting, the delivery time was an average of 2 days. The study then lasted for 1 month, with a second measure 2 weeks in and a final measure after the month. The design offers naturalism at the expense of perfect control. Players in the open-ended game environment completed varying tasks with varying levels of play style, yet they did so in ways made equivalent through random assignment. The mean number of hours played by participants in each condition over the month of study was 57, although as is typical of field settings, the exposure was not uniform (control $M = 53.00$, $SD = 41.02$; treatment $M = 61.12$, $SD = 46.35$). All pretest measures were collected before the voice systems

were distributed, and all posttest measures were collected after the month ended. In each case, these were collected within 1 week of the beginning and end of the test period, with the majority collected within 2 days of the survey openings.

Participants

Participants were solicited via the game's official online message boards in forums labeled "General World of Warcraft," "Off-topic," and "Guild Recruitment." With the assistance of the game company, these threads were made semipermanent and were endorsed as legitimate. This addressed the suspicious nature of a free service, which is a major issue for game players in what are often cantankerous forums. Interested subjects were told that they were invited to a general opinion survey and that small-sized guilds of 8–15 members were sought. Interested guild members then had to organize their membership enough to send in e-mails and screen names. All participants were promised \$20 for their participation, although that was raised quickly to \$40 when recruitment stalled early in the process.

To reduce the confounds discovered by prior HCI voice research, the design sought to minimize problems associated with knowing who was talking, eliminating background noise, and being on different teams. ("Latency," or the lag between speaking and hearing, may prompt players to play with others within similar time zones to reduce overlapping chatter.) Thus, whole preexisting guilds were recruited so that the players would have a prior association and team orientation with their fellow communicators. Also, players were instructed in the "push to talk" method of VoIP, that is, sounds only went out to others when the player held down a keyboard key. Pretests found that this practice greatly minimized background sounds such as breathing, eating, music, traffic, pets, and so forth.

The sample collected here was 83% male and 85% White, with a mean age of 26.48 years. This is representative of MMO players (Griffiths, Davies, & Chappell, 2003; Yee, 2006). The mean guild size was 10.18 ($SD = 2.63$, low = 7, high = 17), and members showed substantial preexisting ties: 92.4% were described by at least one other guild member as a "friend," and 51.33% had a substantial offline connection with a guild member (relative, coworker, neighbor, roommate, classmate, boyfriend/girlfriend, or spouse). There were 103 complete cases in the control (text) condition and 109 in the treatment (voice) condition. The retention rate was 74.15% for members of the treatment group and 81.10% for the control group. The groups were checked for equivalence at Time 1. Multilevel modeling (MLM) tests (alpha set at .05) of the individuals across the two conditions (with group membership used as a random effect in the model) showed that random assignment achieved equivalence on most but not all of the variables under study.

Yet, the assignment was of groups, not individuals, meaning that the chances of errors rise; if the members of any one group can be assumed to be relatively homogeneous, then assigning whole groups to condition weakens the power of randomization to the number of groups rather than the larger number of individuals. That was the case here, in that the people assigned to the treatment condition made more

money, voice group $M = \$42,700/\text{year}$; text group $M = \$31,700/\text{year}$; $t(240) = 3.37$, $p < .001$, worked more, voice group $M = 36.81$ hours/week; text group $M = 27.52$ hours/week; $t(231) = 3.20$, $p < .001$, had better social skills, Riggio's Social Control Scale: voice group $M = 32.32$; text group $M = 36.54$; $t(226) = 3.37$, $p < .001$, and were less likely to argue with their guildmates, 5-point scale: voice group $M = 1.87$; text group $M = 2.30$; $t(246) = 3.32$, $p < .001$. The groups were not different on any of the 80 other measures taken, including age, gender, and many other demographic and psychographic measures. In short, the individuals in the treatment conditions were more successful and mature. These variables were therefore used as controls where theoretically appropriate.

Measurement

Self-reported questionnaires were completed pre- and posttest online via a secure Web site and included a range of demographic, behavioral, and personality variables. To avoid suspicions about the intent of the research, the measures used in this study were scattered within a larger questionnaire and comprised less than 30% of the total questions. To avoid completion fatigue, the measures were pre-tested and found to take 10–15 minutes without complaint from a pool of 15 existing game players.

Balkanization and bridging were measured with two scales. First was the Saguaro Seminar's 11-question diversity index (sample items: Do you have a personal friend who is a manual worker; owns a vacation home; is Asian? etc.; Yes/No answers; scale alpha = .74) ("The Saguaro Seminar: Civic Engagement in America," 2004). Second was the bridging subscale of Williams' (2006b) Internet Social Capital Scales (ISCS) adapted for guilds. This additive subscale covers dimensions related to reaching out for new experiences and people through social outlets (sample items: The people in my guild make me curious about other places in the world; In my guild I have new people to talk to; 7-point range; scale alpha = .85). Relationship strength was measured with a standard 0–100 feeling thermometer (How close would you say that you feel to <guildmate name>?). This measure was customized for each guild to use the in-game names of their fellows, allowing for the construction of group-based means. Bonding was measured with the bonding subscale of the ISCS (sample items: In my guild there are several people I trust to help solve my problems, ... there are several people I can talk to when I feel lonely; 7-point range; scale alpha = .87), whereas trust was measured with adapted standard National Election Study (NES) question forms (How much do you trust <guildmate name>? single-item, 4-point range.).

To test real-world social capital, a series of questions were asked regarding preexisting local communities. These covered subjects' sense of community in their neighborhoods, in their city, at work and school, and with people online (e.g., The people I have met online give me a strong sense of community; 5-point scale) adapted from Kraut et al. (1996). A second series of questions, also adapted from Kraut et al. (1996), examined relationships within families. These covered the frequency (How many minutes a day do you communicate with <family member>?)

and quality of conversations with family members (Rate the quality of your communication with this person; 5-point scale).

Psychosocial well-being was assessed with two scales. First, a 2-item scale adopted from Kraut et al.'s (1996) Internet Paradox series measured happiness (About how many days would you say the following statements were true? I was happy; I enjoyed life; 14-point summative scale; $\alpha = .93$). Additionally, loneliness was operationalized with the short form, six-question summative version of the University of California at Los Angeles Loneliness Scale (Russell, Peplau, & Cutrona, 1980) (sample item: I feel isolated from others; 5-point scale; $\alpha = .74$). Other social and psychological measures were included for exploratory reasons, including Riggio's (1989) Social Control Scale for social maturity (sample item: I sometimes say the wrong thing when starting a conversation with a stranger; $\alpha = .83$).

Lastly, the issues of usability and modality preference were seen as relatively exploratory. A series of open-ended questions were asked at the end of the final survey to discover whether subjects enjoyed using VoIP or not while playing, whether they had any particularly good or bad experiences, whether they thought it affected their interactions with their guildmates, and how they compared it to the text-only experience.

Analyses: Interdependent data

The data collected here were sufficiently nonindependent to suggest the use of MLM (Hayes, 2006; Kenny, Kashy, & Cook, 2006). All participants were nested within groups (i.e., guilds), suggesting that intraclass correlations be tested to see if the scores for individuals could in part be explained by group means (Southwell & Doyle, 2004). Using the protocols outlined by West, Welch, and Galecki (2007), intraclass correlations on the major study variables were found ranging from .01 to .14. In other words, a small portion of an individual's score could be explained by group membership. Therefore, following Hayes' recommendation for situations where the intraclass correlations exceed .05, MLM was used.

The data collected for the study reported here contained three levels: First-level units were the three overtime survey administrations within each player. Second-level units were the game players themselves, who were in turn nested within the guilds. Guilds represented the third level in the analysis. Each model presented below was implemented through the SPSS MIXED procedure. In the hypothesized model, random effects were assumed to be controlled by random assignment to condition. However, as noted above, members of the treatment condition were more mature and successful. In particular, they had higher income, more hours worked, and had higher scores on the Social Control Scale and were less prone to arguing. Because income and hours worked were highly correlated, they could not both be included in the model as controls. Likewise, the maturity scale and arguing were correlated. Therefore, only income and Social Control were used as fixed-effects covariates in each model. Condition was also included as a fixed effect throughout.

Results

This section reports the results of tests of time by condition interactions on various dependent variables relevant to our hypotheses and research questions. With regard to RQ₁, there was no significant time by condition effect either on diversity, $F(2, 158.32) = .263, p > .05$, or on bridging social capital, $F(2, 191.49) = .286, p > .05$. However, there was a significant main effect of condition on bridging where those in the voice condition exhibited higher levels of bridging than those in the text condition, $F(1, 210.65) = 7.57, p < .01, d = .36$.¹

Next, H₁ predicted that voice would lead to stronger relationships, increased bonding, and trust. There was a significant time by condition interaction effect on the feeling thermometer scores, $F(2, 180.50) = 3.945, p < .05, d = .30$ (see Figure 1). However, the post hoc comparison² of estimated marginal means did not reveal any significant differences among means across time or between conditions. There was not a significant time by condition interaction effect on the bonding subscale of the ISCS, $F(2, 177.36) = .42, p > .05$. There was a significant time by condition interaction on using the Internet to keep in touch with local friends, $F(2, 185.50) = 5.17, p < .01, d = .36$. A post hoc comparison revealed that the estimated marginal mean score for text-only users at Time 3 was significantly less than either groups' mean scores at Time 1 or Time 2. Finally, there was a significant time by condition interaction effect on trust $F(2, 180.50) = 3.93, p < .05, d = .30$. A post hoc comparison did not reveal any meaningful pattern of significant differences between groups across time (see Figure 2).

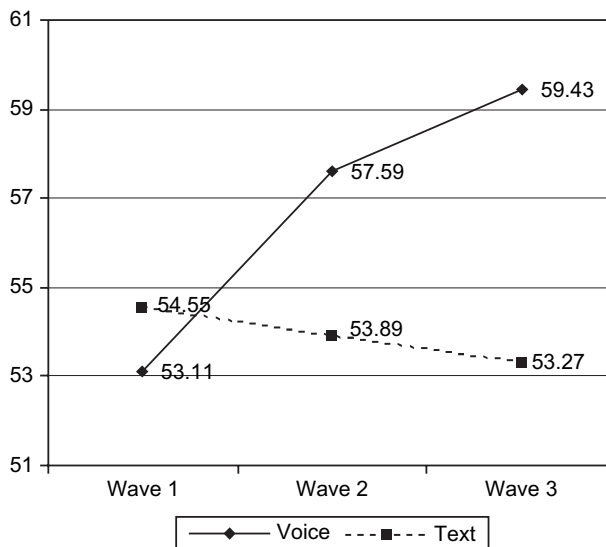


Figure 1 Feeling thermometer outcomes for treatment versus control.

Note: Values come from means created for each respondent after they had assessed every guildmate individually.

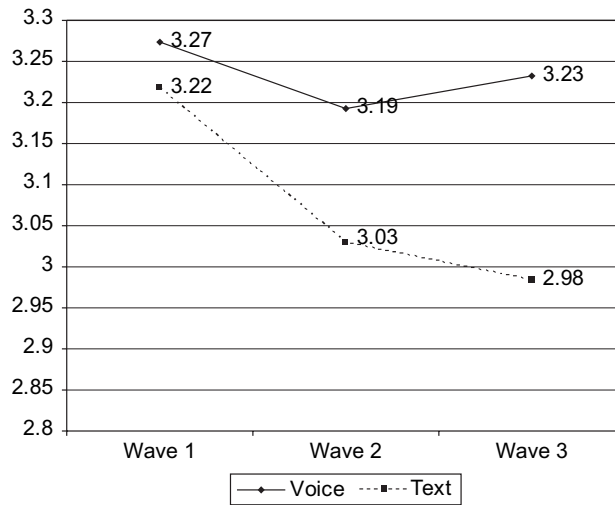


Figure 2 Mean trust in guildmates.

Note: Values come from means created for each respondent after they assessed each guildmate.

Local community as measured by subjects’ sense of community at the neighborhood, city, or online level was unaffected by the introduction of voice. However, there was a significant time by condition interaction effect on the sense of community derived from coworkers and schoolmates, $F(2, 182.89) = 3.24, p < .05, d = .22$. A post hoc comparison did not reveal any meaningful pattern of significant differences among each condition’s scores over time (see Figure 3). A second series of questions

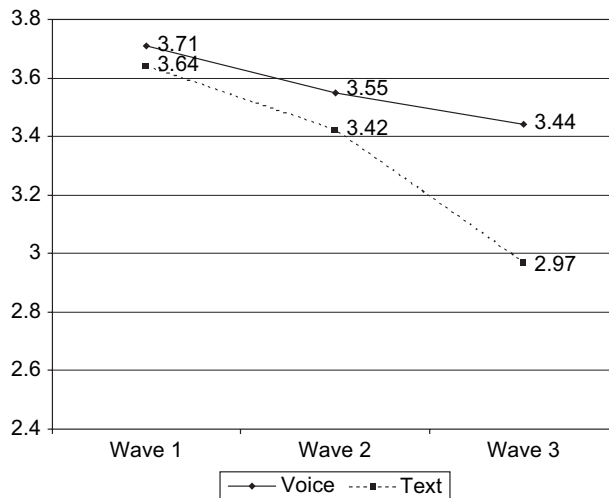


Figure 3 Players’ sense of community coming from coworkers and schoolmate.

examining family connections yielded no significant outcomes, with one marginal finding. The quality of conversation with family members was marginally higher among voice users, $F(2, 183.57) = 2.93, p < .10, d = .24$.

The next question asked whether the use of voice would have an impact on affective state. There was a significant time by condition interaction on happiness, $F(2, 180.02) = 5.23, p < .01, d = .27$. The post hoc comparison did not reveal any meaningful pattern of significant differences among each condition's scores over time. The analyses also revealed that the time by condition interaction on loneliness was marginally significant, $F(2, 168.36) = 2.95, p \leq .10, d = .22$.

Discussion

The study sought to test the impact of adding voice to text-based online communities, exploring diversity, social capital, trust, and affective state. A three-level multi-level model analysis examined the impact of using VoIP within gaming guilds over a month in an experimental design. The use of voice had no impact on balkanization, bridging, or bonding social capital, but interaction effects suggest significantly higher levels of relationship strength and trust between voice-based guildmates when compared to the text condition over time. Local social capital was largely unaffected. The only exception was a positive impact on community between coworkers and schoolmates, indicating that game play between preexisting friends likely serves as a medium for maintaining relationships. Affective state impacts were also positive, with significantly higher levels of happiness and lower levels of loneliness among the voice users over time. The patterns and their relation to theory will be discussed, followed by a brief discussion of online diversity in groups, concluding with implications for other uses of voice and future research to address the theoretical implications of the findings.

There were two major patterns in the results. First, there were consistent patterns in the results suggesting differences between the treatment and the control groups over time. A visual inspection of the graphs suggests that the difference between voice and text becomes larger over time. Many of the post hoc tests do not allow certainty about precisely when these differences are significant. Second, and unexpectedly, there were significant drops for several dependent variables within the text group only. What is the theoretical reason for the significant divergence between the conditions? SIP and SIDE offer some explanatory power, but in the end, the best fit comes from uncertainty reduction and media richness theories.

SIP theory suggests that people adapt to the media at hand to reach affective and social goals (Walther, Loh, & Granka, 2005). One of the key findings of SIP research has been that users will use those media to maintain relationships similarly to how they would in FtF interactions (Walther & Parks, 2002). This has been studied with zero-history task-based groups, who have experienced initial divergence in outcomes as a result of using different media but then later experienced reconvergence when the CMC groups caught back up to the FtF groups (Walther, 1993). Despite relatively

similar goals for the groups there and here, that model is difficult to apply to the findings for a number of reasons. First, the groups in this study were not zero history. More than half of the guild members know each other before playing the game, and those who did not knew each other for periods of time before the study began considered each other friends. Second, the durations of the studies differ greatly. Walther's groups met once every 2 weeks for 2 hours over 6 weeks, for a cumulative time together of only a handful of hours. The WoW players in this study reported time in the game consistent with what prior research has estimated for MMO players: over 50 hours across the 4 weeks. There are therefore two explanations for the difference between the studies suggested by SIP theory. One is that the original CMC study found an effect that is short term and that was swallowed by the longer term and coarser instrument used here; the divergence and reconvergence of the two conditions could have occurred within that same 6 hours and then diverged again before the midpoint test at 2 weeks. Such a result would cast doubt on shorter CMC experiments on communication modality and socioemotional connections reconverging over time. The second explanation is that time limits are the significant predictor in CMC group relations experiments. A meta-analysis (Walther et al., 1994) suggested that the tests without time limits—as was the case here—found more positive socioemotional communication than studies with a known end point. This would explain the positive outcomes for the voice condition but not the negative outcomes from the text condition. The sheer difference in media remains a more powerful explanation than user intent.

SIDE (Lea & Spears, 1992; Spears & Lea, 1992) would also seem to be a useful theory to explain the results. With its emphasis on group-based norms, SIDE would predict that the changes occurring were an outcome of group norms established in different ways through the different media and there is a pattern here that suggests this (Postmes et al., 2001). The voice group appears to have a temporary drop in some outcomes before reverting to a more neutral or positive pattern. This could have been the result of a group norm slowly forming around the use of voice. Once established, the norm would then influence the members to become comfortable with the new group practice.

Yet, SIDE cannot fully account for all of the findings. For example, SIDE suggests that colocated groups have better shared norms and understandings than dispersed groups (Cramton, 2001, 2002). In this case, the colocation between conditions was equivalent, yet one condition had more positive outcomes. Intriguingly, this suggests that medium is a slightly more powerful factor than distance for groups that are not wholly colocated. Further research focusing on the degrees of distance and the level of offline FtF contact would help explain these findings with more nuance. It is still possible that group-based processes were in effect over the month of the study, but SIDE is unable to explain the individual-level findings offered by the feeling thermometer scores. These were calculated by asking each guild member to assess a score for each other group member. Thus, although group means were reported and used for analysis, there was a clear trend of voice users gaining liking on an interpersonal, not group-based, level.

The most powerful explanations for the findings are two simpler ones and an adaptation of SIP: media richness and uncertainty reduction. The reduction of uncertainty is a major theme in the organizational literature on teams, virtual, or otherwise (Poole & Zhang, 2005), and this study appears to be a good intersection between that approach and media richness. If the spectrum of richness and cues indeed moves from text to voice to video to FtF, then moving along that spectrum introduces more cues. Those cues are then used by community members to reduce their uncertainty about others. Communities found across the Internet are consistently plagued and blessed with quasi-anonymous media. And although that anonymity has allowed for unlikely offline connections (breadth), it does so at the expense of true human connections (depth). The findings here suggest that reintroducing those richer cues allows for more genuine connections in the way suggested by Galston (1999); without surrendering some of the anonymity that online connections give us, we cannot expect to have quality relationships. When that anonymity is removed—even by simply reintroducing voice—it is apparently enough to improve relationships and insulate members from losses that might occur without intervention.

We are thus left unexpectedly with media-based theories. Even though players likely used text in nonstandard ways to communicate affective cues (Utz, 2000) and make truly personal ties (Tidwell & Walther, 2002), the richness of the text-only information could not compete with voice plus text. McLuhan gets his moment. This combining and mixing of media suggests a hybrid theory for media richness and SIP. Certainly, the relatively cue-rich media of text plus voice was superior to text only, but it is important to remember that the game players were not given a totally binary choice. Instead, they were given a richer *mix* of media versus a poorer one. As many players attested, they made active choices about when and how they would add that second medium into the mix. For example, one player noted that he/she selectively used voice to avoid conflict:

There is one member of the guild whom I get frustrated with his play and with voice it is a little too easy to say something to that effect. Without voice, I am better at holding my tongue when he does something detrimental.

And another, who was more comfortable using it frequently, noted that the richness of the medium had a direct affect on group affinity:

If for nothing else, it has added another dimension to the bad jokes we tell. Seriously, though, I'd say that it has brought us closer together as a group.

Thus, SIP's contentions about users taking an active role remain a useful one if we consider not just one medium versus another but users' considerations of how and when to select from an available mix. Walther's (1999) suggestion therefore sounds particularly apt if we alter the singular "medium" to the plural "media" to allow for mixed media use:

The greater the number of cue systems (along with a nonspecific concurrency with natural language potential, immediate feedback, and personalization potential, resulting in “richer” media), the more beneficial the medium is [media are] when the topic is complex (p. 1).

This emphasis on richer mixes of media is not a purely academic thought exercise: As the Internet acquires voice and video functions, the elements of choice, availability, and skill (Katz, Rice, & Aspden, 2001) may offer strong predictive power for social and affective outcomes. And as users mix and multipurpose their media, new uses, patterns, and effects are likely (Jones & LeBaron, 2002). New avenues of research should consider the way in which different new media are used simultaneously among actors given a choice.

Depressing worlds or depressing media?

It is unusual for the group that is considered theoretically to be the control to experience change, so the general visual patterns of decreases in the text condition merit consideration. The most obvious explanation is that the base state of MMO game worlds depresses social capital and group dynamics. This explanation would extend Williams' (2006a) findings from a previous MMO (*Asheron's Call 2*), which, although much less popular (and populated) than WoW, was found to depress offline social capital. In keeping with media richness and uncertainty reduction theory, the data suggest that the social tools necessary to maintain positive and trusting groups of relative strangers are not provided by either game. Predicting this, Steinkhueler (2004) has called upon game developers to provide stronger support tools for the complex social and managerial task of organizing and maintaining a guild. Nevertheless, the groups studied here were likely not part of a steady linear decline, or else the 2 million North American WoW players would be noticeably depressed in their daily offline lives. It is more likely that MMO games do indeed have a baseline depressive effect on social capital but that deviations from that baseline will vary widely from game to game and at different points within a game's life cycles. The timing of the study is a reminder that unlike most CMC research, these were not zero-history groups and were instead being examined at a point in their development.

Thus, a contributor to the possible downward trend in the study here likely involves the life span of guilds within this particular game world. As described earlier in the paper, guilds within WoW follow a general trajectory as they approach what is known as the “end game,” in which players reach the maximum level and then turn their attention to content that suddenly necessitates large groups. Small guilds, such as the ones used here, typically experience social tensions as the members realize that their group size is not optimal for completing the most glamorous tasks (Williams et al., 2006). It is therefore typical for these guilds to experience periods of political instability as they consider merging and/or splitting up in order to accomplish these goals. The small-group bonds formed in the run-up to this period are threatened by

the possibility of joining a larger and often less personal guild. This instability may offer an explanation for some of the decreases exhibited by the text groups. Given the launch timing of the game (November 2004) and the timing of the data collection (October and November 2005), it is very likely that the small guilds used in the study were undergoing this period of uncertainty. Data show that the general WoW population was experiencing guild mergers in the study time period as the general population approached the end game.³

Nevertheless, the large disparity between the conditions shows that the voice-based guilds were better equipped to handle such socially disruptive events. This is especially noteworthy given the preexisting ties in these groups. As noted earlier, nearly every participant (92%) was considered the friend of at least one other guildmate. In other words, the guild members had ties to each other before the study started, with over half (52%) of those extending to some kind of offline relationship. Voice made these groups stronger, even though they had these prior connections. In some cases, this was likely the strengthening of a relatively weak tie. This prior connection has implications for the text-only group as well: that group suffered *despite* the prior and offline connections. The weak ties that existed were not enough to insulate the guilds from social losses, and the offline (assumedly stronger) ties were not enough either. The social impact of playing WoW with only text was clearly negative, even for groups with genuine offline connections.

Network makeup

The cyberbalkanization hypothesis illustrates a central tension in online community formation. Are online communities good because they broaden our horizons and enable us to meet disparate groups, or are they bad because they limit our contact to others who hold our same demographic and ideological traits? As with many binaries, this one appears to be oversimplified. Other research suggests that players within guilds did in fact meet others from a wide range of backgrounds while using voice (Williams et al., 2006). Yet in this test, they simultaneously ceased meeting additional new players. Thus, there are two countervailing forces at work: the establishing of connections that can form long-term closeness versus the insularity of forming an in-group. Such forms of in-group insularity and out-group ambivalence are forecasted by much of the sociology literature and are reinforced by Putnam's (2000) work on social capital. The upside in this context is that those in-game ambivalences are unlikely to translate to offline antipathies, where they make little sense.

The results also speak to the complex interplay of online and offline relationships and show that this barrier is flimsy at best. One notable dimension of this positive pattern was geographical. Even though prior research has suggested that only one third of WoW guild members know each other before playing (Williams et al., 2006), and are thus likely to be geographically nearby, there was a significant gain (or insulation from loss) in players' keeping in touch with local friends. If more of the other players were local, this effect would theoretically be larger as more local friends

would use the medium (as SIP suggests) as a means of maintaining relationships. This finding matches other recent work on online communities, that is, that Internet use is increasingly a supplement to offline life rather than a substitute (Horrigan & Rainee, 2002; Pew Internet and American Life Project, 2004; Wellman & Haythornthwaite, 2002). People simply migrate their existing connections online, where they are moderated by new forces. For MMO players, play with nearby friends with a voice component appears to protect them from losing local social capital and substitutes for previous media as a means of everyday connection. Future research might specifically investigate locality as a factor in these issues, perhaps drawing on the Netville research (Hampton & Wellman, 2001).

Theoretical implications and applications in other contexts

SIP suggests that the voice plus text combination allows users to adapt the media at hand to specific communication tasks. For some tasks, one medium might be superior to another, but the choice involved allows for the two channels to complement rather than conflict (Dutta-Bergman, 2004). A key example here would be the adaptation of mixed media choices for a mix of synchronous and asynchronous tasks. Players needing to communicate immediately, or with little time delay, might opt for voice, whereas players wanting to communicate without immediacy may opt for text. This element of choice and synchronicity is an avenue for future research.

This test of VoIP in gaming is only one of several future applications that may appear. As VoIP becomes cheaper and easier, it will likely transfer to new contexts and applications. Given enough bandwidth, users could participate in social networking sites, shopping, and myriad other online activities with a voice component. Or, voice could become an element of the content itself in the form of participatory site content. Video games show an early example of how the true content of interest is often the other users.

This voice-over text-only outcome might have implications for the growing body of research on e-mail versus FtF or telephone use. Research has demonstrated that e-mail can be used to maintain personal relationships and support collaboration (Haythornthwaite, 2002; McCormick & McCormick, 1992; Sproull & Kiesler, 1992), but the results here suggest that it is still inferior to voice plus text, at least in the context of team-based collaboration. Yet, the open-ended comments suggest that voice was superior for joint task coordination, problem solving, and dealing collectively with dynamic situations (however fantastical they may have been). This is not to say that SIP is not still at play. Users clearly adapted to whatever media were available. Moreover, the subjects here were previously unaware of the potential for using voice while playing. A true test of the interaction between intentionality and information richness by medium would include some sense of user choice rather than the artificial introduction used in this experimental design.

Usability was a factor cited in many of the open-ended comments and is a reminder that there is a body of literature available in the HCI domain related to the technical and performance-based aspects of voice systems. These are fruitful

variables to consider in future research, especially when considering how voice systems may be more useful in some contexts than others. For example, drivers using voice systems show reduced driving performance (Ranney, Harbluk, & Noy, 2005). Military uses of voice range from speech recognition in cockpits to communications between individuals and groups (Paper, Rodger, & Simon, 2004). It is therefore worth exploring how voice might influence task performance within online environments. Conversely, the tasks carried out in such environments might require a level of focus that inhibits simultaneous social interaction. One initial study found that voice employed poorly enough can actually hinder social interaction (Hew et al., 2004). Another found that voice could be disruptive to social interaction in war game settings when people were not sure who was talking (Halloran et al., 2004). Aside from technical issues, it is feasible that voice used for task completion could inhibit social interaction if it occupies a large amount of players' mental capacity. Given the SIP approach, users may therefore select text over audio in some contexts (Scholl, McCarthy, & Harr, 2006). In a more nuanced sense than the current findings can detect, this suggests that the way voice is employed—how easily the speaker is identified, push-to-talk versus always on, trash-talking versus collaborative, and so forth—can have major social impacts. Future research should therefore be careful in examining how VoIP is used within a given context.

In sum, the mix of CMC options appears to have a strong impact on community and friendships. Those with more and richer options do comparatively better than those relegated to fewer choices or poorer media. And, as an unexpected finding, poorer media in an online play-based setting appears to be potentially harmful to groups. The challenge for future research is now to delve into these applications with more nuance and power. Online systems are a complex and moving target as social architectures and interfaces change and new community and entertainment forms emerge. Last decade's couch potato is this decade's active creator of content and interface. Regardless, communication theory can be adapted and used to make sense of these strange new worlds and applications.

Notes

- 1 All multilevel analysis effect sizes were estimated using the d coefficient recommended by Kenny et al. (2006, p. 57). However, because the current study employed group data rather than dyadic data, the equation was slightly modified. Instead of computing d_D , we computed d_G . According to D. A. Kenny (personal communication, April 19, 2007), $d_G = 2 \times \text{square root of } F / \text{square root of } k$, where k is the number of groups. Then to get d for individuals, we used the following equation: $d_G / (\text{square root of } n / (1 + (n - 1)r))$, where n is the average group size and r the intraclass correlation.
- 2 All post hoc analyses involved comparing 95% confidence intervals for the estimated marginal mean scores of each condition across time.
- 3 See data from Palo Alto Research Center's PlayOn project at http://blogs.parc.com/playon/archives/data/wow_data/guilds/index.html. Retrieved December 17, 2006.

References

- Beniger, J. (1987). Personalization of mass media and the growth of pseudo-community. *Communication Research*, *14*, 352–371.
- Burgoon, J., Bonito, J., Ramirez, A., Dunbar, N., Kam, K., & Fischer, J. (2002). Testing the interactivity principle: Effects of mediation, propinquity, and verbal and nonverbal modalities in interpersonal interaction. *Journal of Communication*, *52*, 657–677.
- Burke, R. R. (2002). Technology and the customer interface: What consumers want in the physical and virtual store. *Journal of the Academy of Marketing Science*, *30*, 411–432.
- Burnham, V. (2001). *Supercade, a visual history of the videogame age 1971–1984*. Cambridge, MA: The MIT Press.
- Clarke, R. (2003). ENUM: A case study in social irresponsibility. *Privacy Law & Policy Reporter*, *9*, 181–187.
- Connaughton, S., & Daly, J. (2004). Long distance leadership: Communicative strategies for leading virtual teams. In D. Pauleen (Ed.), *Virtual teams: Projects, protocols, and processes* (pp. 116–144). Hershey, PA: Idea Group.
- Cramton, C. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization Science*, *12*, 346–371.
- Cramton, C. (2002). Attribution in distributed work groups. In P. Hinds & S. Kiesler (Eds.), *Distributed work* (pp. 137–162). Cambridge, MA: MIT Press.
- Culnan, M., & Markus, M. (1987). Information technologies. In F. Jablin, L. Putnam, K. Roberts, & L. Porter (Eds.), *Handbook of organizational communication: An interdisciplinary perspective* (pp. 420–443). Newbury Park, CA: Sage.
- Daft, R., & Lengel, R. (1984). Information richness: A new approach to managerial behavior. In B. Straub & L. L. Cummings (Eds.), *Research in organizational behavior* (pp. 191–223). Homewood, IL: JAI Press.
- Dutta-Bergman, M. (2004). Interpersonal communication after 9/11 via telephone and internet: A theory of channel complementarity. *New Media & Society*, *6*, 659–673.
- Fitzgerald, T. (2004, December 2). Should your PC be your telephone? *New York Times*, E6.
- Galston, W. (1999). *Does the Internet strengthen community?* Retrieved November 20, 2001, from http://www.puaf.umd.edu/IPPP/fall1999/internet_community.htm
- Garton, L., Haythornthwaite, C., & Wellman, B. (1999). Studying on-line social networks. In S. Jones (Ed.), *Doing Internet research* (pp. 75–105). Thousand Oaks, CA: Sage.
- Griffiths, M., Davies, M. N., & Chappell, D. (2003). Breaking the stereotype: The case of online gaming. *CyberPsychology & Behavior*, *6*, 81–91.
- Halloran, J., Fitzpatrick, G., Rogers, Y., & Marshall, P. (2004, April). *Does it matter if you don't know who's talking? Multiplayer gaming with voice over IP*. Paper presented at the CHI 2004, Vienna, Austria.
- Hampton, K., & Wellman, B. (2001). Long distance community in the network society: Contact and support beyond Netville. *American Behavioral Scientist*, *45*, 476–495.
- Hayes, A. (2006). A primer on multilevel modeling. *Human Communication Research*, *32*, 385–410.
- Haythornthwaite, C. (2002). Strong, weak, and latent ties and the impact of new media. *Information Society*, *18*, 385–401.
- Hew, K., Gibbs, M., & Wadley, G. (2004, February). *Usability and sociability of the Xbox Live voice channel*. Paper presented at the Australian Workshop on Interactive Entertainment, Sydney, Australia.

- Horrigan, J., & Rainee, L. (2002, October). *Getting serious online*. Paper presented at the Association of Internet Research 3.0, Maastricht, The Netherlands.
- Jensen, C., Farnham, S., Drucker, S., & Kollock, P. (2000). The effect of communication modality on cooperation in online environments. In *Proceedings of CHI'00* (pp. 470–477). The Hague, Amsterdam.
- Jones, S., & LeBaron, C. (2002). Guest editor's introduction. *Journal of Communication*, 52, 499–521.
- Katz, J. E., Rice, R. E., & Aspden, P. (2001). The Internet, 1995–2000: Access, civic involvement, and social interaction. *American Behavioral Scientist*, 45, 405–419.
- Kendall, L. (2002). *Hanging out in the virtual pub: Masculinities and relationships online*. Berkeley: University of California Press.
- Kenny, D., Kashy, D., & Cook, W. (2006). *Dyadic data analysis*. New York: Guilford Press.
- Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukhopadhyay, T., & Scherlis, W. (1996). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist*, 53, 1011–1031.
- Lea, M., & Spears, R. (1992). Paralanguage and social perception in computer-mediated communication. *Journal of Organizational Computing*, 2, 321–341.
- Levy, S. (1994). *Hackers: Heroes of the computer revolution*. New York: Penguin Books.
- McCormick, N., & McCormick, J. (1992). Computer friends and foes: Content of undergraduates' electronic mail. *Computers in Human Behavior*, 8, 379–405.
- Nass, C., & Gong, L. (2000). Speech interfaces from an evolutionary perspective: Social psychological research and design implications. *Communications of the ACM*, 43(9), 36–43.
- Nass, C., & Lee, K. (2000, April). *Does computer-generated speech manifest personality? An experimental test of similarity-attraction*. Paper presented at the CHI 2000, The Hague, Netherlands.
- Negroponste, N. (1995). *Being digital*. New York: Vintage Books.
- Nie, N. H., & Erbring, L. (2002). Internet and society: A preliminary report. *IT & Society*, 1, 275–283.
- Nie, N. H., & Hillygus, D. S. (2002). The impact of internet use on sociability: Time-diary findings. *IT & Society*, 1, 1–20.
- Oldenburg, R. (1997). *The great good place: Cafés, coffee shops, community centers, beauty parlors, general stores, bars, hangouts, and how they get you through the day*. New York: Marlowe.
- Olsen, S. (2005). *VoIP wants to cut the computer cord*. Retrieved September 26, 2005, from http://www.nytimes.com/cnet/CNET_2100-7352_3-5879421.html
- Paper, D., Rodger, J., & Simon, S. (2004). Voice says it all in the Navy. *Communications of the ACM*, 47(8), 97–101.
- Pew Internet and American Life Project. (2004). *Percent of American adults online, 1995–2004*. Washington, D.C.: Pew Internet and American Life Project Surveys.
- Poole, M. S., & Zhang, H. (2005). Virtual teams. In S. Wheelan (Ed.), *Handbook of group research* (pp. 363–386). Thousand Oaks, CA: Sage.
- Postman, N. (1992). *Technopoly: The surrender of culture to technology*. New York: Knopf.
- Postmes, T., Spears, R., & Lea, M. (2000). The formation of group norms in computer-mediated communication. *Human Communication Research*, 26, 341–371.
- Postmes, T., Spears, R., Sakhel, K., & deGroot, D. (2001). Social influence in computer mediated communication: The effects of anonymity on group behavior. *Personality and Social Psychology Bulletin*, 27, 1243–1254.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

- Ranney, T., Harbluk, J., & Noy, I. (2005). Effects of voice technology on test track driving performance: Implications for driver distraction. *Human Factors*, *47*, 439–454.
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press/CSLI.
- Rheingold, H. (1993). *Virtual communities*. Reading, MA: Addison-Wesley Publishing.
- Rice, R., & Gattiker, U. (2001). New media and organizational structuring. In F. Jablin & L. Putnam (Eds.), *The new handbook of organizational communication: Advances in theory, research, and methods* (pp. 544–581). Thousand Oaks, CA: Sage.
- Riggio, R. (1989). *The social skills inventory manual: Research edition*. Palo Alto, CA: Consulting Psychologists Press.
- Rogers, E. (1995). *Diffusion of innovations*. New York: The Free Press.
- Russell, D., Peplau, L. A., & Cutrona, C. E. (1980). The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*, *39*, 472–480.
- The Saguaro Seminar: Civic Engagement in America. (2004). Retrieved October 26, 2004, from <http://www.ksg.harvard.edu/saguaro/>
- Salinas, E.-L. (2005). Effects of communication mode on social presence, virtual presence, and performance in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *14*, 434–449.
- Scholl, J., McCarthy, J., & Harr, R. (2006, November). *A comparison of chat and audio in media rich environments*. Paper presented at the CSCW, Banff, Alberta, Canada.
- Southwell, B., & Doyle, K. (2004). The good, the bad, or the ugly? A multilevel perspective on electronic game effects. *American Behavioral Scientist*, *48*, 391–401.
- Spears, R., & Lea, M. (1992). Social influence and the influence of the ‘social’ in computer-mediated communication. In M. Lea (Ed.), *Contexts of computer-mediated communication* (pp. 30–65). New York: Harvester Wheatsheaf.
- Sproull, L., & Kiesler, S. (1992). *Connections: New ways of working in the networked organization*. Cambridge, MA: MIT Press.
- Steinkuehler, C. (2004, March). *Providing resources for MMOG guild leaders*. Paper presented at the MUD Developers Conference, San Jose, CA.
- Steinkuehler, C., & Williams, D. (2006). Where everybody knows your (screen) name: Online games as “third places”. *Journal of Computer-Mediated Communication*, *11*. Retrieved March 22, 2007, from <http://jcmc.indiana.edu/vol11/issue4/steinkuehler.html>
- Sunstein, C. (2001). *Republic.com*. Princeton, NJ: Princeton University Press.
- Taylor, T. L. (2006). *Play between worlds: Exploring online game culture*. Cambridge, MA: The MIT Press.
- Tidwell, L., & Walther, J. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research*, *28*, 317–348.
- Tönnies, F. (1957). *Gemeinschaft und gessellschaft [Community and society]* (G. Simpson, Trans.). East Lansing: The Michigan State University Press.
- Trubshaw, R., & Bartle, R. (1978). *MUD1*. Essex, United Kingdom.
- Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet*. New York: Touchstone.
- Utz, S. (2000). Social information processing in MUDs: The development of friendships in virtual worlds. *Journal of Online Behavior*. Retrieved July 7, 2006, from <http://www.behavior.net/JOB/v1n1/utz.html>

- VanAlstyne, M., & Brynjolffson, E. (2005). *Global village or cyberbalkans*. Modeling and measuring the integration of electronic communities. *Management Science*, 51(6), 851–868.
- Walther, J. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 19, 52–90.
- Walther, J. (1993). Impression development in computer-mediated communication. *Western Journal of Communication*, 57, 381–398.
- Walther, J. (1994). Anticipated ongoing interaction versus channel effects on relational communication in computer-mediated interaction. *Human Communication Research*, 20, 473–501.
- Walther, J. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal. *Communication Research*, 23, 3–43.
- Walther, J. (1999, May). *Visual cues and computer mediated communication: Don't look before you leap*. Paper presented at the International Communication Association, San Francisco, CA.
- Walther, J. (2006). Nonverbal dynamics in computer-mediated communication, or :(and the net :(s with you, :) and you :) alone. In V. Manusov & M. Patterson (Eds.), *The Sage handbook of nonverbal communication* (pp. 461–480). Thousand Oaks, CA: Sage.
- Walther, J., Anderson, J., & Park, D. (1994). Interpersonal effects in computer-mediated interaction: A meta-analysis of social and anti-social communication. *Communication Research*, 21, 460–487.
- Walther, J., Loh, T., & Granka, L. (2005). Let me count the ways: The interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of Language and Social Psychology*, 24, 36–65.
- Walther, J., & Parks, M. (2002). Cues filtered out, cues filtered in: Computer-mediated communication and relationships. In M. Knapp & J. Daly (Eds.), *Handbook of interpersonal communication* (3rd ed.). Thousand Oaks, CA: Sage.
- Wellman, B., & Haythornthwaite, C. (2002). *The Internet in everyday life*. Malden, MA: Blackwell Publishing.
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Diaz, I. I. d, et al. (2003). The social affordances of the Internet for networked individualism. *Journal of Computer-Mediated Communication*, 8(3).
- West, B., Welch, K., & Galecki, A. (2007). *Linear mixed models: A practical guide using statistical software*. Boca Raton, FL: Chapman & Hall/CRC.
- Williams, D. (2006a). Groups and goblins: The social and civic impact an online game. *Journal of Broadcasting and Electronic Media*, 50(4), 651–670.
- Williams, D. (2006b). On and off the 'Net: Scales for social capital in an online era. *Journal of Computer-Mediated Communication*, 11. Retrieved March 22, 2007, from <http://jcmc.indiana.edu/vol11/issue2/williams.html>
- Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., & Nickell, E. (2006). From tree house to barracks: The social life of guilds in World of Warcraft. *Games & Culture*, 1, 338–361.
- Yee, N. (2006). The demographics, motivations and derived experiences of users of massively-multiuser online graphical environments. *PRESENCE: Teleoperators and Virtual Environments*, 15, 309–329.

Maintenant, peux-tu m'entendre? L'impact de la voix dans une communauté de jeu en ligne

Dmitri Williams

University of Southern California

Scott Caplan

University of Delaware

Li Xiong

University of Southern California

Résumé

Cet article fait état des résultats d'une expérience de terrain contrôlée dans laquelle la communication par la voix fut incorporée à une communauté en ligne existante (une guilde de jeu en ligne dans le jeu populaire « World of Warcraft »), en comparant un mélange de voix et de texte avec du seul texte. Les résultats quantitatifs suggèrent une hausse de la sympathie et de la confiance due à l'ajout de la voix, ainsi qu'une isolation des impacts négatifs imprévus du jeu contenant exclusivement du texte. Les résultats sont discutés en regard du capital social, de la cyberbalkanisation et de la littérature générale sur la communication assistée par ordinateur, avec une attention particulière accordée à la théorie du traitement social de l'information (*Social Information Processing, SIP*).

Kannst du mich jetzt hören? Der Einfluss von Stimme in Online-Spielegemeinschaften

Dieser Beitrag erörtert die Ergebnisse eines kontrollierten Feldexperiments, in dem Stimmkommunikation in eine existierende Online-Gemeinschaft (eine Online-Gilde des beliebten Spiels „World of Warcraft“) eingeführt wurde. Eine Kombination von Stimme und Text wurde mit einer Nur-Text-Variante verglichen. Quantitative Ergebnisse zeigen einen Zuwachs an Gefallen und Vertrauen aufgrund der ergänzenden Stimme sowie eine Minderung von unerwarteten negativen Einflüssen des textbasierten Spiels. Die Ergebnisse werden im Zusammenhang mit Sozialkapital, Cyberbalkanisierung und der allgemeinen CvK-Literatur mit besonderem Augenmerk auf die SIP Theorie diskutiert.

¿Puedes Escucharme Ahora? El impacto de la Voz en una Comunidad de Juego Online

Dmitri Williams

University of Southern California

Scott Caplan

University of Delaware

Li Xiong

University of Southern California

Resumen

Este ensayo reporta los resultados de un experimento de campo controlado en el cual una comunicación de voz fue introducida en una comunidad online existente (una sociedad de jugadores del popular juego “World of Warcraft”), comparando una mezcla de voz y de texto solamente. Los resultados cuantitativos sugieren un incremento en la preferencia y la confianza debido a la adición de la voz, así como el aislamiento de impactos negativos inesperados del juego de texto sólo. Los resultados son discutidos con respecto al capital social, la balcanización cibernética, y la literatura general de la comunicación mediatizada por la computadora, prestando especial atención a la teoría SIP.

听到我吗？声音在线游戏社区的作用

Dmitri Williams

南加州大学

Scott Caplan

特拉华大学

Li Xiong

南加州大学

[摘要]

本论文报告控制性实地实验的结果，声音传播引入了已有的在线社区（一流行的游戏社区：战舰世界）和混合声音和文字和只有文字的社区的区别。数据表明声音的引入引起更高的喜欢度和信任度，只有文字的游戏由于出乎意料的负面效果造成了隔绝效应。我们将这些结果与社会资本，在线化和通常的电脑中介传播文献结合论述，特别讨论了对 SIP 理论的含义。

지금은 잘 들리는가: 온라인 게임 커뮤니티에서의 음성 영향에 대한 연구

Dmitri Williams
University of Southern California

Scott Caplan
University of Delaware

Li Xiong
University of Southern California

요약

본 연구는 음성 커뮤니케이션이 기존의 온라인 커뮤니티 (인기게임인 WoW 내 온라인게임 길드)에 도입 되었을 경우의—음성과 텍스트의 혼합과 텍스트의 경우만을 비교함으로써—통제된 현장실험의 결과들에 관한 것이다. 양적 결과들은 텍스트만의 게임에서 기대하지 않은 부정적인 영향들로부터의 단절뿐만 아니라 음성을 더하였기 때문에 좋아함과 신뢰가 증가한다는 것을 보여주고 있다. 연구의 결과들이 사회적 자본, 사이버발칸화, 그리고 일반적인 컴퓨터 매개 커뮤니케이션 문헌, 특별히 SIP 이론에 주목하여, 논의되었다.