

Mode Choice in Multimodal Comment Threads: Effects on Participation, Sociability, and Attitude

Susan C. Herring
Bradford Demarest

Indiana University, Bloomington

Abstract

This study investigates how participation, sociability, and attitude differ according to commenting mode and commenter gender in Voicethread.com, an interactive Internet platform that allows commenting on multimodal slideshows via text, audio, or video. In such platforms, the question arises as to whether “richer” modes such as video and audio are preferred for social communication and “leaner” text for more impersonal, contentious communication, as previous research has claimed (e.g., Daft & Lengel, 1984; Sproull & Kiesler, 1991). Comments in all three modes were analyzed for the frequency of metadiscourse terms that express social awareness and for the frequency and polarity of attitude categories in three extended public VoiceThreads. Video comments were made mostly by males, and male comments were more negative, consistent with previous research on gender and computer-mediated communication (CMC). Mode differences were also found that provide support for the claims of early CMC theorists associating richer media with more social, interpersonal communication and text with critical judgment and negativity. These findings lead us to revisit claims in the literature that have previously been largely discredited, including technological determinism itself, and to suggest an alternative account that reconciles the VoiceThread findings with those of previous research (e.g., Walther, 1996) regarding the positive sociability of textual CMC.

Introduction

Interactive multimodal platforms that combine video and user-generated textual comments represent a growing trend in convergent media computer-mediated communication. Typically in such environments, visual content serves as a prompt to which text comments respond. Examples range from comments posted below YouTube videos to text annotations inserted directly into videos and, perhaps less familiarly, into dynamic visual displays of audio files, as on SoundCloud.com. The platform that is the focus of this study, VoiceThread.com, takes mixed-modality computer-mediated communication (CMC) to a different level, however, by allowing commenting on multimodal slideshows *in multiple modes*¹—text, audio, or video—and in displaying all comments together in a single format.

VoiceThread.com raises questions about why, how, and to what effect users choose to participate in CMC in a given mode. In particular, the question arises as to whether “richer” modes such as video and audio are preferred for social communication and “leaner” text for more impersonal, contentious communication, as previous research has claimed (e.g., Daft & Lengel, 1984; Sproull & Kiesler, 1992). Most previous research in support of such claims has been experimental; an advantage of studying VoiceThread comments is that they are generated in authentic contexts of use. Moreover, in most experimental studies that compare communication modes, subjects are assigned to one

mode and do not have the option to select another, whereas VoiceThread offers users a choice of commenting modes within a single platform. All that is required to post audio and video comments is a microphone and web cam, which most laptop computers come equipped with these days; audio comments can also be made through a telephone or uploaded as sound files.

This study investigates how style and content of participation differ according to commenting mode in three public VoiceThreads, employing language-focused content analysis. To determine if audio and video are associated with greater sociability, all comments were analyzed for the frequency of metadiscourse—hedges, boosters, personal pronouns, and other language forms that indicate a degree of social awareness and an orientation on the part of the commenter to the addressees (Hyland, 2005). To assess the effect of comment mode on evaluative stance, categories from Martin and White (2005) associated with attitude were adapted and coded for the occurrence and polarity of each. Participation metrics were also calculated, and the results for all measures are compared by thread, gender, and mode.

The findings reveal that video comments were made mostly by males, and male comments are more negative, consistent with previous research on technology adoption (e.g., Gefen & Straub, 1997) and gender and politeness in CMC (e.g., Herring, 1994). Mode differences are also evident throughout the data, and provide support for the claims of early CMC theorists associating richer media with more social, interpersonal communication and text with more judgment and negativity. The latter findings lead us to revisit claims in the literature that have previously been largely discredited, including technological determinism itself, and to suggest an alternative account that reconciles the VoiceThread findings with those of previous research (e.g., Walther, 1996) regarding the positive sociability of textual CMC.

Background

VoiceThread.com was founded in Boca Raton, Florida by Benji Papell and Steve Muth and launched in March 2007.² Thus far, the platform has mainly attracted the attention of educators, who view it as an engaging learning and discussion tool (e.g., Weir, 2008). However, Millard (2010) analyzed user comments from 50 randomly-selected public VoiceThreads and found that the comments showed little evidence of interaction or collaboration; rather, they responded to the content featured in the multimedia slideshow. Other than this, little scholarship has been conducted on VoiceThreads that we were able to locate.

VoiceThread is of broader theoretical and practical interest, nevertheless, in that it offers users a choice of communication modes on a single platform. This has not been possible before on the same scale outside experimental contexts, although mode choice across separate communication platforms has been theorized and researched. Much of the research is based on early theoretical claims involving media “richness,” or the number and nature of communication channels a medium provides. Social presence theory (Short, Williams, & Christie, 1976) posits that participants will perceive greater social presence via richer media such as video than via telephone or written communication. The degree of social presence is equated to the degree of awareness of the other person in a communication interaction. Information richness theory posits that richer media such as

face-to-face (FTF) communication are better suited for tasks involving nuanced social communication, while leaner media such as text are best suited for simple, routine tasks such as arranging meetings (Daft & Lengel, 1984). Similar to social presence theory, Daft and Lengel posit a one-dimensional continuum, with FTF as the “richest” and numerical data via computer as the “leanest” medium. Both of these theories predated CMC, but they have been influential in CMC research. CMC modes can be situated along the continuum as in Figure 1.

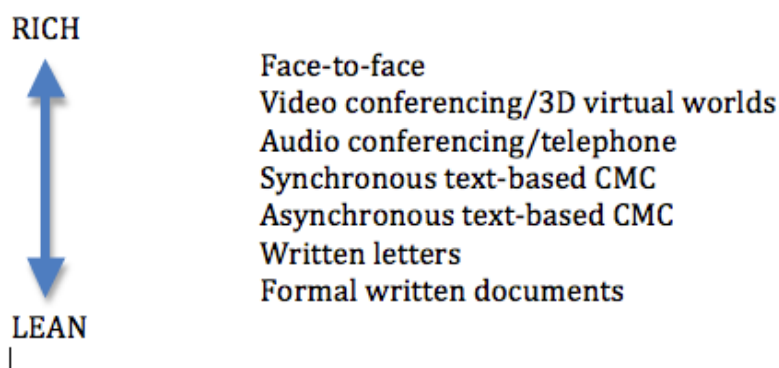


Figure 1. CMC modes in the continuum of media richness

In an empirical study of mode choice, Murray (1988) conducted an ethnographic study over a nine-month period of mode and medium choice by IBM employees. She found that FTF, telephone, paper, and computer (including both synchronous and asynchronous CMC) media were used for different purposes, as determined by the availability of prospective interactants, time management (which favored the use of CMC), and the sensitivity of the message (which tended to favor FTF conversation). Baym et al. (2004) had undergraduate students keep a diary and also conducted a survey regarding their medium use. One of the study’s findings was that “[p]eople were more likely to use FTF conversations and telephone calls in more intimate relationships” (p. 314), as predicted by media richness theory. Kim et al. (2007) conducted a social network analysis, via a web survey, of communication relationships in Korea through FTF, email, instant messaging, mobile phone, and SMS text messaging. They found that, whereas “face-to-face (FtF) seems to be a universal medium without significant differences across respondents, “mobile phones tend to be used in reinforcing strong social ties, and text-based CMC media tend to be used in expanding relationships with weak ties” (n.p.).

The research described above mainly focused on mode choice in relation to communicative purpose and/or nature of the interactants’ relationship. Other research has studied the social and psychological effects on behavior of different modes. An early and influential study by Kiesler et al. (1984) found mode differences in disinhibition (as measured by swearing, insults, name-calling, and hostile behavior), decision shifts, and equality of participation, all of which occurred more often in textual CMC (both synchronous and asynchronous) than in FTF conditions. The first two findings led the authors to predict that textual CMC would be ill suited for social interaction, consistent with the predictions of the social presence and media richness theories. Kiesler et al. explain their findings in terms of the paucity of social cues, such as facial expressions and prosody, in text-only CMC, which results in a process of ‘depersonalization’ whereby

messages senders feel detached from (experience less social presence with) their addressees.

The explosion of naturally occurring CMC on the Internet soon thereafter gave the lie to these predictions as regards the limitations of textual CMC, however—or so it seemed. As early as 1987, Rice and Love found a high incidence of socio-emotional content in the messages exchanged among medical professionals on a CompuServe discussion forum. Online communities arose in which participants reported feeling a sense of belonging and social presence (Rheingold, 1993); reports of online romances soon followed (e.g., Cooper & Sportolari, 1997). At the same time, consistent with Kiesler et al.'s findings of disinhibition in CMC, a high incidence of hostile verbal behavior, or 'flaming', was observed in some online forums (e.g., Kim & Raja, 1990). To account for these behaviors, Walther (1996) theorized that textual, asynchronous CMC was 'hyperpersonal' in that message recipients tended to overgeneralize from the limited social cues in computer-mediated messages, with the result that both positive and negative perceptions of the sender were exaggerated. Such perceptions fuel romantic feelings as well as feelings of hostility.

An alternative approach that accounts for the inconsistencies between studies showing that text-only CMC can be socially oriented and findings of studies such as that of Kiesler et al. (1984) is the Social Information Processing theory (Walther, 2002), which holds that close social relationships may be formed via leaner media, but that more time is required. Most previous experimental studies provided subjects with the same amount of time in each mode; to compensate for this, Walther and his colleagues allowed subjects in text-only CMC conditions up to four times as long to complete tasks (Tidwell & Walther, 2002). Under such conditions, few differences in social perceptions or socioemotional tone arose across modes, although when some communicators used one mode and others used a richer mode, the leaner mode was rated less satisfying (e.g., Walther & Bazarova, 2008). These results are consistent with the Electronic Propinquity Theory proposed by Korzenny (1978), which posits that "users will adopt the widest bandwidth communication medium available to them," but that "the fewer one's choices of media, the more closeness one may experience even through the lowest of bandwidths" (Walther, in press).

In contrast to explanations that invoke the influence of communication technology on mode choice and user behavior, Herring (1993, 1994, 1995, 2003) invoked a social factor, gender socialization, to explain why men and women participated in CMC in different ways. Herring noticed that 'flaming' was much more common in messages posted to online forums by males than by females, which should not have been the case if technology alone predisposed users towards expressing hostility. Men also participated more in most public forums and tended to post longer messages. Similarly, studies of technology adoption in organizations have argued for the necessity of including gender as a factor in models of technology acceptance. These studies found that women perceive email as higher in social presence than men do (Gefen & Straub, 1997), adopt new technologies for social reasons (such as usage by their peers) more, and are more concerned with ease of use, whereas men are motivated to adopt mainly in terms of the affordances of the technology (its perceived usefulness; Venkatesh & Morris, 2000). Despite being motivated to adopt, women tend to be slower to do so, consistent with previous observations that men tend to feel more at ease with computers (e.g., Frankel, 1990).

Research Questions and Hypotheses

The general research question that informs this study is: How do users comment in a given mode when multiple modes are available on a single Internet platform? Specifically:

- RQ1. What differences, if any, are there in sociability and negativity across the three commenting modes available on VoiceThread.com?
- RQ2. Are there gender differences in sociability and negativity, independent of mode choice?
- RQ3. Are there gender differences in mode choice and amount of participation?

Based on the research surveyed above, we posited the following hypotheses:

- H1: “Richer” modes such as audio and video will convey more social communication and “leaner” text more impersonal, contentious communication.
- H2. Females will be more social and less contentious than males, independent of mode choice.
- H3. Males will participate more and comment more often in video and audio, which are more novel modes than text, than females will.

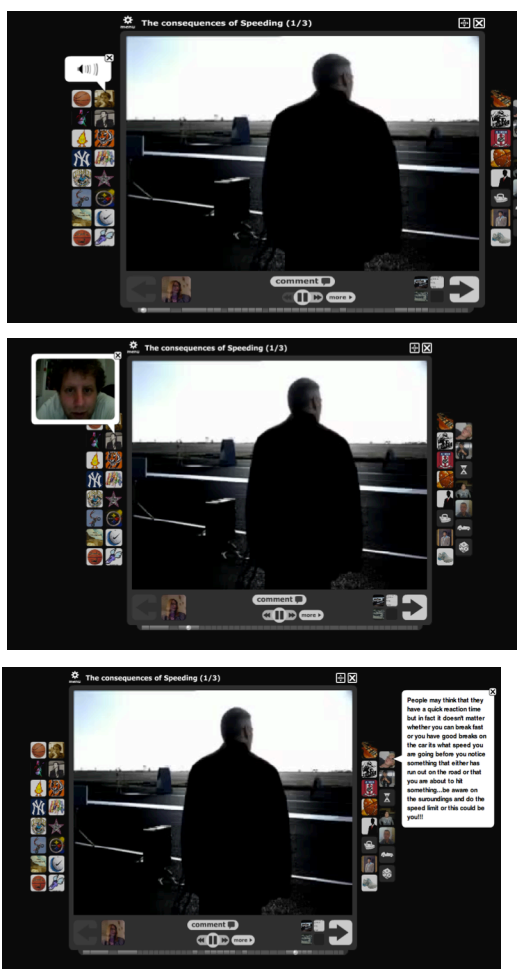


Figure 2. An audio comment (top left), a video comment (bottom left), and a text comment (right) in the Speeding thread

Methodology

Data

The data for this study are all comments posted in three public VoiceThreads. The threads are prompt-triggered discussions set in academic contexts, consistent with the most common use of the VoiceThread platform. Because we had little indication of what to expect given the paucity of previous research on VoiceThread, we sampled for diversity. The three threads vary in topic and the age of the participants: The first was produced by an elementary school class evaluating art in science fiction, the second was produced by a high school driver’s education class on the topic of the deleterious effects of speeding, and the third is a discussion among professional educators in response to the question: What does the network mean to you? The Sci Fi thread had 155 comments, the Speeding thread had 97 comments, and the Network thread had 111 comments at the time of our data collection in March 2011. The total corpus consists of 363 comments and 22,069 words.

Figure 2 shows how an audio, a video, and a text comment, respectively, appear during playback of the Speeding thread.³ Comments can be accessed either by clicking on the user profile pictures on the right and left sides of the central slide, or by

clicking on segments of the playback bar at the bottom of the interface. The playback bar replays the comments sequentially in the order in which they were posted.

Analytical Methods

For the purposes of analysis, all audio and video comments were transcribed by the authors and entered into an Excel spreadsheet along with the text comments. Each comment was first coded for the independent variables thread, mode, and gender of participant. Gender was included as an independent variable in this study, because previous research has found gender differences in participation in CMC (e.g., Herring, 2003), including in online learning environments (e.g., Barrett & Lally, 1999). The comments were then coded for the three dependent variables of participation, attitude, and metadiscourse, as described below.

Participation

The participation analysis consisted of a straightforward counting of the number of messages and the number of words, as well as a calculation of the number of words per message. The results of these measures are reported using descriptive statistics.

Attitude

To assess the effect of comment mode on judgment and negativity, we adapted categories from Martin and White (2005) associated with attitude. These are listed in Table 1 below (definitions are from Martin & White; examples are from our data).

<i>Category</i>	<i>Definition and Examples</i>
Affect	“Emotions; reacting to behavior.” (e.g., happy, sad)
Judgment	“Ethics; evaluating behavior.” (e.g., dangerous, wrong, should/shouldn’t)
Appreciation	“Aesthetics; evaluating text/process, natural phenomena.”
Appreciation:reaction	<ul style="list-style-type: none"> • to things: “do they please us, do they catch our attention?” (e.g., interesting, like, dislike)
Appreciation:composition	<ul style="list-style-type: none"> • the thing’s “balance and complexity.” Did it hang together? Is it well crafted?
Appreciation:valuation	<ul style="list-style-type: none"> • “How innovative, timely, authentic, etc.” Was it worthwhile? (e.g., useful, cool, great, weird)

Table 1. Martin and White’s attitude categories

We also found instances in our data that seemed to express attitude, but in a formulaic way, and that did not seem to fit into Martin and White’s categories. Thus we modified the coding rubric to include the three additional categories listed in Table 2 below, for a total of eight attitude coding categories.

<i>Category</i>	<i>Examples</i>
Formulaic affect	'I'd love to hear your thoughts,' 'I hate to break it to you'
Formulaic judgment	'You have to wonder,' 'Speed kills'
Formulaic appreciation	'Thanks for starting this thread,' 'take care'

Table 2. Additional attitude categories

In addition, we coded each expression of attitude for polarity: positive, negative, or neutral. The unit of analysis could be a word, phrase, utterance, or chunk of a message that expressed a particular attitude, and more than one attitude could be coded per message, including the same one multiple times.

Both authors participated in the coding. Attitude and polarity codes were assigned independently by each author for a portion of the data, code assignments were compared, and disagreements were resolved through discussion. This process was iterated until all messages had been jointly coded with 100% agreement (for Speeding and SciFi) or until better than 80% agreement was reached for both attitude and polarity assignment (for Network); in the latter case, the second author coded the remaining messages.

Metadiscourse

All comments were analyzed for the frequency of metadiscourse—language forms that indicate a degree of social awareness and an orientation on the part of the commenter to the addressees. Hyland (2005, p. 37) defines metadiscourse as “self-reflective expressions used to negotiate interactional meanings, assisting the writer (or speaker) to ... engage with readers.” This analysis was conducted to measure the amount of social presence in the comments in each mode.

Hyland groups metadiscourse into two broad types: interactive and interactional, each of which has multiple categories, as summarized in Table 3.

<i>Interactive Metadiscourse</i>	<i>Interactional Metadiscourse</i>
Code glosses	Attitude markers
Endophoric markers	Boosters
Evidentials	Self-mention
Frame markers:	Engagement markers
• Sequencing	Hedges
• Label stages	
• Announce goals	
Shift topic	
Transition markers	

Table 3. Hyland's metadiscourse categories

Interactive metadiscourse mainly involves reference to other parts of the discourse, while interactional metadiscourse involves the interaction of the writer or speaker with the audience.

In an appendix to his book, Hyland (2005) provides an extensive list of metadiscourse

terms in English, based on his studies of academic writing. We modified this list to exclude punctuation and non-alphabetic symbols (on the grounds that these are not possible in speech); formal terms and conventions mainly found in written text (such as ‘in chapter X’); terms with variable elements (such as date, sequences of numbers) that are difficult to search for; and common terms that do not mainly function as metadiscourse (such as ‘and’ and ‘go’). We also manually reviewed the messages in our sample and added terms that occurred there but were not already in Hyland’s list; these included references to the multimedia slide show at the center of the VoiceThread (e.g., ‘(in) the/this video/chart/diagram/picture/image/slide’) and spoken discourse phenomena such as contracted hedges (e.g., ‘kinda,’ ‘sorta’) and discourse markers (e.g., ‘oh’, ‘yeah’, ‘y’know’, ‘like’). Finally, we lemmatized terms with a common root (such as ‘seem/seems/seemed’ into seem*) to facilitate their retrieval from the corpus. The end result was a list of 390 lemmatized terms that we imported into a freely available concordancing program, CasualConc, which was used to sort and count the frequencies of each. The results returned by CasualConc for each term were manually filtered by the authors to exclude instances that did not function as metadiscourse in the contexts in which they occurred.

Figure 3 shows part of the spreadsheet in which term frequencies were recorded.

1	Item	Metadiscourse category	Metadiscourse type	Search term(s)	Network F auc	Network F tex	Network F vid
357	undisputedly	Boosters	Interactional	undisputed*	0	0	0
358	undoubtedly	Boosters	Interactional	undoubtedly*	0	0	0
359	unexpected unexpectedly	Attitude Markers	Interactional	unexpected*	0	0	0
360	unfortunate unfortunately	Attitude Markers	Interactional	unfortunate*	0	0	0
361	unlikely	Hedges	Interactional	unlikely	0	0	0
362	unusual unusually	Attitude Markers	Interactional	unusual*	0	0	0
363	us (incl)	Self-mention	Interactional	us	3	1	0
364	use	Engagement markers	Interactional	use*	1	0	0
365	usual	Attitude Markers	Interactional	usual*	0	0	0
366	usually	Hedges	Interactional	usually	1	0	0
367	very	Boosters	Interactional	very	14	0	2
368	want to	Frame markers: Announce goals	Interactive	want* to	4	1	2
369	we (inclusive)	Self-mention	Interactional	we	9	0	2
370	well	Frame markers: Shift topic	Interactive	well	2	0	0
371	What you’re about to see/watch	Frame markers: Announce goals	Interactive	What you’re about to *	0	0	0
372	whereas	Transition Markers	Interactional	whereas	0	0	0
373	while	Transition Markers	Interactional	while	0	1	1
374	whole	Boosters	Interactional	whole	2	0	1
375	wish to	Frame markers: Announce goals	Interactive	wish* to	0	0	0
376	with regard to	Frame markers: Shift topic	Interactive	with regard to	0	0	0
377	with that	Frame markers: Label stages	Interactive	with that	0	0	0
378	without doubt	Hedges	Interactional	without doubt	0	0	0
379	would like to	Frame markers: Announce goals	Interactive	would like to	0	0	0
380	would wouldn’t	Hedges	Interactional	would*	11	1	2
381	x above	Endophoric markers	Interactive	above	0	0	0
382	x before	Endophoric markers	Interactive	before	1	0	0
383	x below	Endophoric markers	Interactive	below	0	0	0
384	x earlier	Endophoric markers	Interactive	earlier	0	0	0
385	x later	Endophoric markers	Interactive	later	0	0	0
386	yeah	Discourse markers	Interactional	yeah	2	0	0
387	yet	Transition Markers	Interactional	yet	0	0	0
388	you know yknow y’kn y’know	Discourse markers	Interactional	you know	2	0	1
389	you know yknow y’kn y’know	Discourse markers	Interactional	yknow y’kn y’know	3	0	1
390	you your yours	Engagement markers	Interactional	you your yours	49	9	6

Figure 3. Spreadsheet in which metadiscourse frequencies were recorded

Statistical Analysis

After the participation, attitude, and metadiscourse data were collected, counts for each category within each variable were aggregated to reflect counts in various subcorpora, first for each combination of thread, gender, and mode, and then at higher levels of aggregation (thread, gender, mode, and secondary combinations of thread-gender, thread-mode, and

mode-gender combinations). In order to allow for direct comparisons of values from subcorpora of different sizes, results were normalized: attitude counts, by the number of messages within the subcorpus, metadiscourse counts by total word count in the subcorpus, scaled to counts per 1000 words.

After compiling descriptive statistics, we further assessed attitude and metadiscourse data differences through the use of Chi square tests to determine statistical significance. Chi square tests were employed due to the nature of the data collected as unranked counts of exclusive categories, rather than scalar measurements.

Results

Participation

Table 4 summarizes the participation results for the corpus overall. Percentages add up to 100% vertically for each independent variable.

	<i>Words (%)</i>	<i>Valid msgs (%)</i>	<i>Avg words/msg</i>
All	22,069	320 ⁴	69.0
<i>Gender</i>			
Male	12,702 (57.6%)	177 (55.3%)	71.8
Female	8824 (40.0%)	108 (33.75%)	81.7
Unknown	543 (2.5%)	35 (10.9%)	15.5
<i>Mode</i>			
Video	2972 (13.4%)	15 (4.7%)	198.1
Audio	14,801 (67.1%)	102 (31.9%)	145.1
Text	4296 (19.5%)	203 (63.4%)	21.2
<i>Thread</i>			
Speeding	2865 (13.0%)	87 (27.2%)	32.9
Network	16,853 (76.4%)	109 (34.1%)	154.6
SciFi	2351 (10.6%)	124 (38.7%)	19.0

Table 4. Participation results, overall corpus

Overall, males contributed more messages (comments) and more words than females. The majority of comments were via text, followed by audio; relatively few video comments were posted (the SciFi thread contained none at all). Conversely, video comments contained the most words, followed by audio comments; text comments were much shorter. There was also considerable variability in participation across threads: SciFi, the thread with elementary school children, had the most messages but the fewest words (most comments were via text), while Network, the thread with academic professionals, had the most words and longest messages (most comments were via audio).

Table 4 represents fairly well the patterns within each thread, with one exception: Females did not post longer messages than males in any of the threads. The overall result that female messages were longer is caused by the fact that a majority of females in the total sample are in the Network thread, which also has the longest messages. The distribution of participation by gender is shown in Table 5.

	<i>Words (%)</i>	<i>Valid msgs (%)</i>	<i>Avg words/msg</i>
<i>Speeding</i>			
Male	1538 (54%)	42 (48.3%)	36.6
Female	1154 (40.0%)	34 (39.0%)	33.9
Unknown	173 (6.0%)	11 (12.6%)	15.7
<i>Network</i>			
Male	9464 (56.2%)	53 (48.6%)	178.6
Female	7273 (43.2%)	48 (44.0%)	151.5
Unknown	116 (0.7%)	8 (7.3%)	14.5
<i>SciFi</i>			
Male	1700 (72.3%)	82 (66.1%)	20.7
Female	397 (16.9%)	26 (21.0%)	15.3
Unknown	254 (10.8%)	16 (12.9%)	15.9

Table 5. Participation by gender

Because comments posted by participants of unknown gender are a minority in each thread, and because they tend to be shorter than comments by gender-identifiable participants and contain less information, making them harder to interpret, they were excluded from further analysis.

Attitude

Figure 4 presents the attitude results for the corpus overall. Positive expressions of attitude are the most frequent throughout the corpus (totaling 488, or 70% of all attitude expressions), with negative being second-most frequent at 146, or 21%. Neutral comments constitute the remaining 10%, with 68 occurrences. Apprec:valuation acts were used most often, followed by apprec:reaction acts, then judgment, then affect. Appreciation: composition and the formulaic acts were used least often.

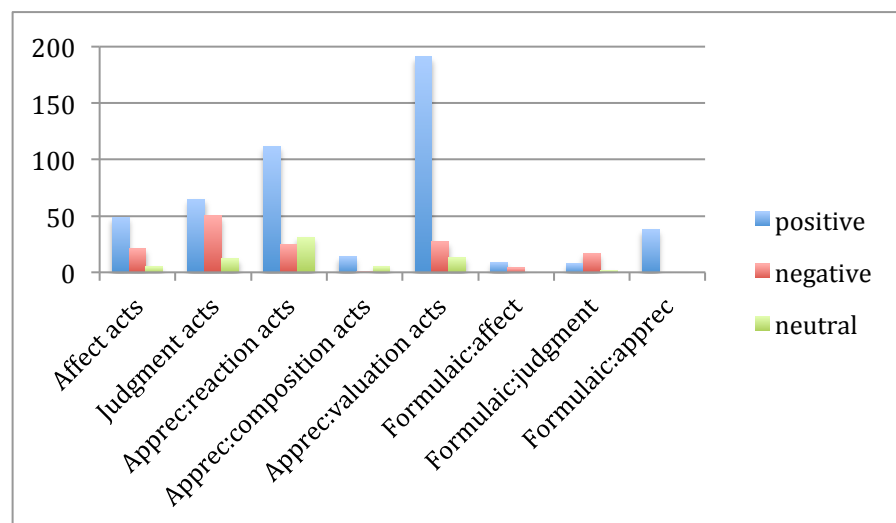


Figure 4. Attitude results, overall corpus

The profiles for attitude results vary considerably by thread, however, as shown in Figure 5. The Network thread, in which participants mostly praise the virtues of the Internet, is overwhelmingly positive and valuative, while comments in the Speeding thread, which concerns the dangers of driving over the speed limit, have mostly negative judgment. The SciFi thread, in which participants commented on art that was sometimes violent and sometimes beautiful, has both positive and negative comments and contains more reactions.

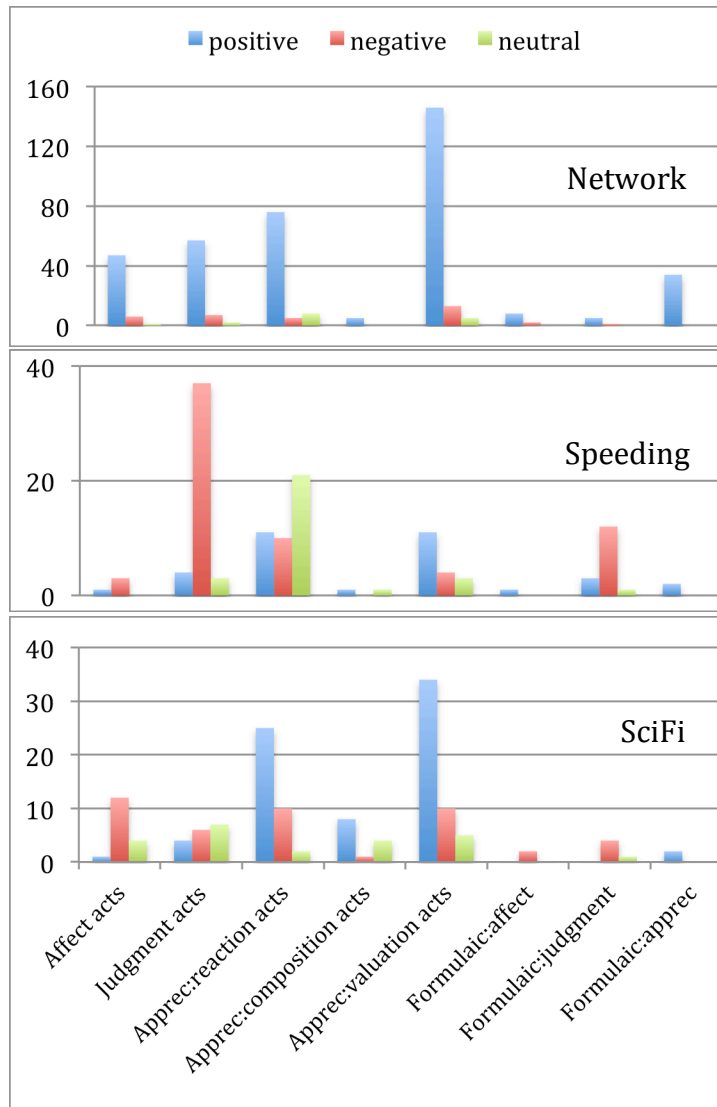


Figure 5. Attitude results for (from top to bottom) Network, Speeding, and SciFi thread

Figure 6 shows that there are more similarities than differences in the overall attitude profiles of males and females. However, females are more positive, especially as regards judgment acts, which they appear to use more than males do. Females also express more formulaic appreciation (which is all positive) and males express more formulaic judgment (which is mostly negative).

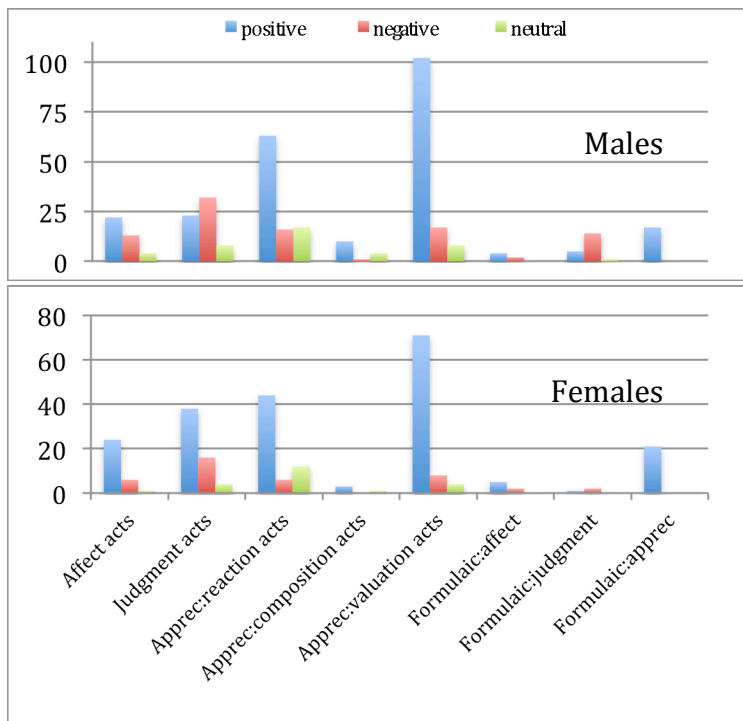


Figure 6. Attitude results for (from top to bottom) males and females

In contrast, the mode results in Figure 7 reveal a striking pattern: Video and audio pattern together in favoring positivity and valuation acts, in contrast to text, which is more negative and has proportionately more judgment, formulaic judgment, and appreciation:reaction acts, all of which tend to be more negative than the corresponding acts in the other two modes.

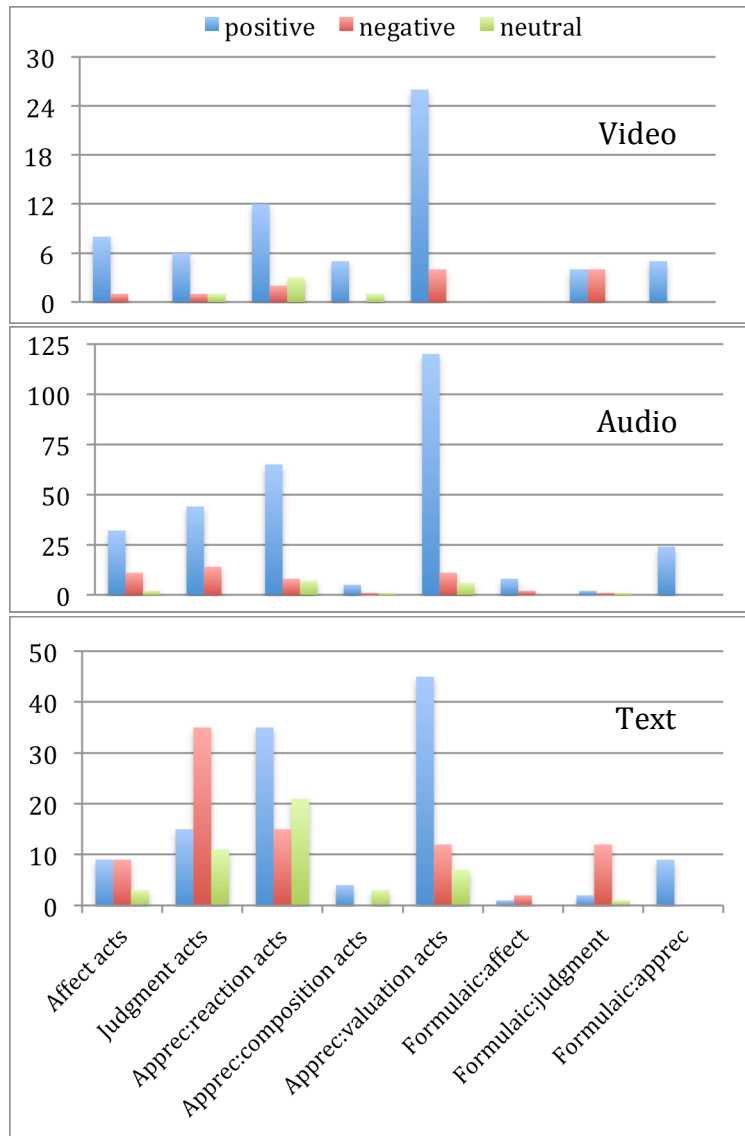


Figure 7. Attitude results for (from top to bottom) video, audio, and text

Chi square tests

Chi square statistical tests were conducted for the trends and patterns suggested by the descriptive statistics presented above. Table 6 presents the results for those patterns that were found to be significant. Chi square tests were run in relation to proportions derived from higher categories than those being tested. Specific polarities within a category were tested against overall proportions of a category. Overall counts of a specific category were tested against proportions of total attitude counts. Total attitude counts were tested against proportions of total message counts.

<i>Trend</i>	χ^2 value	significance ^a
<i>Overall</i>		
Positive > Negative	184.4858	****
Positive > Neutral	317.2662	****
Network > Speeding (Overall attitude)	101.7041	****
Network > Sci Fi (Overall attitude)	185.2269	****
Speeding > Sci Fi (Overall attitude)	4.8334	****
Females > Males (Overall attitude)	5.331	*
Video > Audio (Overall attitude)	13.929	***
Audio > Text (Overall Attitude)	175.851	****
Video > Text (Overall Attitude)	165.713	****
<i>Thread</i>		
Network > Speeding (Positive)	52.2693	****
Network > Sci Fi (Positive)	17.6825	****
Speeding > Network (Negative)	52.2693	****
Speeding > Sci Fi (Negative)	11.518	***
Network > Sci Fi (Appreciation:valuation)	18.1266	****
Speeding > Network (Judgment)	17.3875	****
Speeding > Sci Fi (Judgment)	14.4801	****
<i>Gender</i>		
Male > Female (Negative) ^b	11.11	***
Female > Male (Judgment, Positive)	5.2505	*
Female > Male (Formulaic:appreciation)	4.4041	*
Male > Female (Formulaic:judgment)	5.9287	*
<i>Mode</i>		
Video > Text (Positive) ^c	12.4295	***
Audio > Text (Positive) ^c	26.6178	****
Text > Video (Negative) ^c	8.3384	**
Text > Audio (Negative) ^c	30.8495	****
Audio > Text (Appreciation:valuation)	3.9999	*
Text > Video (Judgment) ^c	5.3915	*
Text > Audio (Judgment) ^c	6.0062	*
Text > Audio (Formulaic:judgment) ^d	12.8627	***
Text > Audio (Judgment, Negative) ^d	7.6414	**
Text > Audio (Apprec:reaction, Negative) ^d	3.9615	*
<i>Gender-Mode</i>		
Text:Male > Text:Female (Negative)	3.9361	*
Text:Female > Text:Male (Positive)	3.8598	*
Text:Male > Video:Male (Negative) ^c	8.8826	**
Text:Male > Audio:Male (Negative) ^c	22.4989	****
Video:Male > Text:Male (Positive) ^c	14.0265	***
Audio:Male>Text:Male (Positive) ^c	22.4102	****

^a * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

^b Female > Male (Positive) also nearly achieved significance ($p = .06$).

^c No significant differences were found between video and audio on these measures.

^d Insufficient video tokens were available to allow Chi square tests to be done for Text > Video.

Table 6. Chi square results for attitude analysis

Most of the trends suggested by the descriptive statistics turn out to be statistically significant. To summarize the main trends, expressions of attitude are mostly positive overall, but there is significant variation by thread: Network is overwhelmingly positive, while Speeding is mostly negative, and SciFi is mixed. Males are more negative than females, especially in judgment acts. Text is more negative than video and audio messages, and text also contains more judgment.

Metadiscourse

The results for the metadiscourse analysis for the overall sample are presented according to Hyland's categories in Figure 8. Self-mentions were used most frequently (1,423 occurrences, or 39%), followed by boosters (645, 18%), hedges (452, 12.5%), and engagement markers (438, 12%), especially 2nd person pronouns.

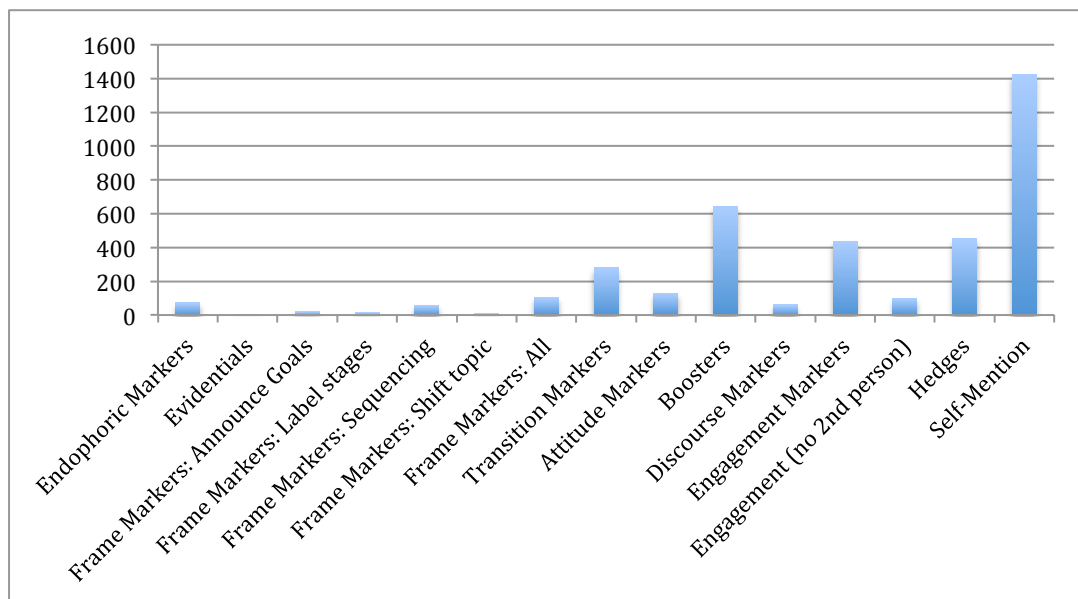


Figure 8. Metadiscourse results, overall corpus

Interactional metadiscourse was used much more than interactive metadiscourse. Because of the small numbers for the interactive categories, and because they are less 'social' than the interactional categories and therefore less relevant to the present study, we do not analyze them further here.

Figure 9 shows the distribution of the interactional metadiscourse terms by thread normalized per 1000 words. The Network thread, in which participants responded to the question "What does the network mean to me?", overwhelmingly favored self-mentions (e.g., 'Hi, this is [name] and I just wanted to say what the network means to me is...'), whereas the Speeding thread favored engagement markers (especially non-referential 'you', as in 'you shouldn't speed'), boosters (e.g., 'speeding is definitely not worth the consequences'), and attitude markers ('its amazing how 5km can make such a big difference'). The SciFi thread, in which participants evaluated and interpreted art works, slightly favored the use of hedges (e.g., 'possibly the person that made the drawing is

probably an atheist’) and discourse markers (e.g., ‘Ok, so, I kinda feel like there's um.. a more deeper aspect to this picture’).

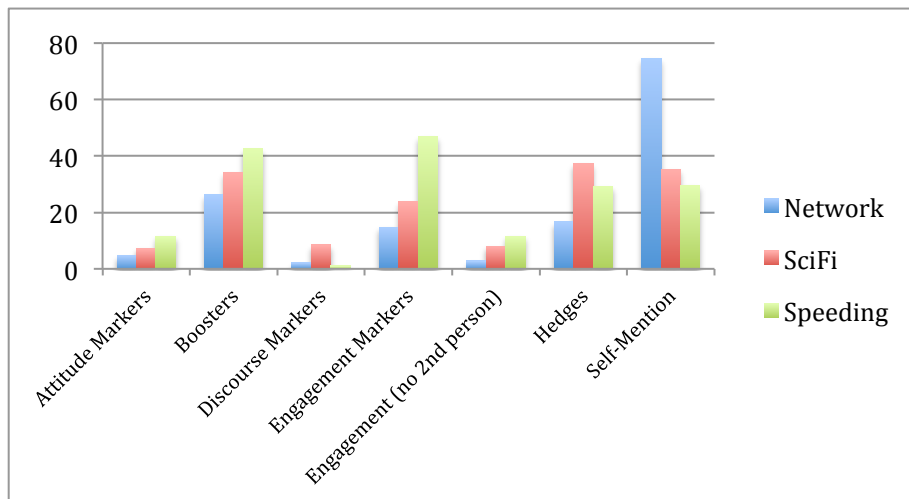


Figure 9. Interactional metadiscourse results by thread

In Figure 10, the interactional metadiscourse results are broken down by gender. Females use more boosters, and males use more hedges, 2nd person pronoun engagement markers, and more discourse markers; otherwise the genders use similar proportions of interactional metadiscourse.

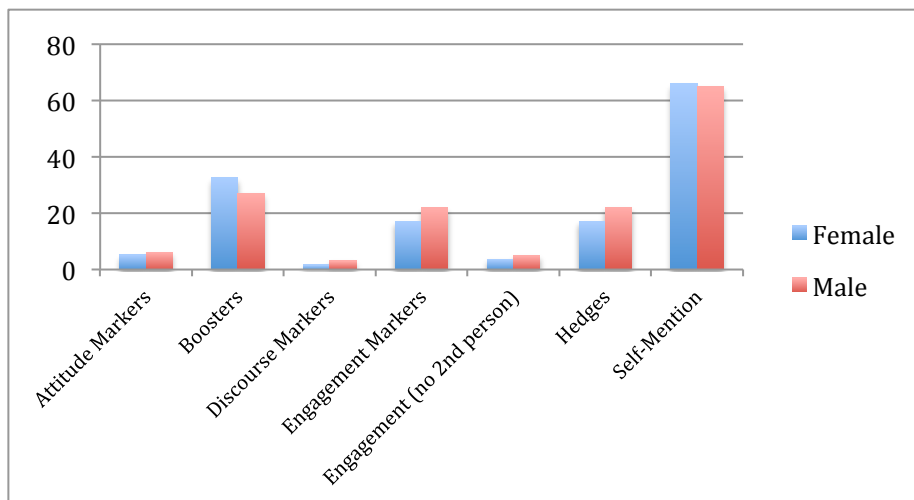


Figure 10. Interactional metadiscourse results by gender

In contrast, numerous differences emerge in interactional metadiscourse use by mode. Video and audio comments contain many more self-mentions, while text comments contain more engagement markers (including, but not limited to, 2nd person pronouns), hedges, and attitude markers. See Figure 11.

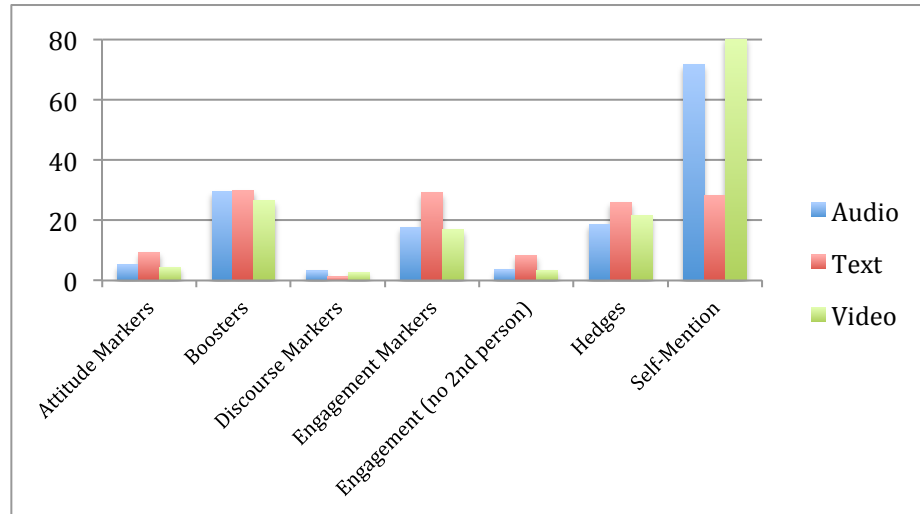


Figure 11. Interactional metadiscourse results by mode

Chi square tests

Chi square statistical tests were conducted for the metadiscourse trends and patterns suggested by the descriptive statistics presented above. The Chi square tests were run in relation to proportions derived from total counts of words in the tested subsets. Table 7 shows the results for those patterns that were found to be significant.

<i>Trend</i>	χ^2 value	significance ^a
<i>Overall</i>		
Interactional > Interactive ^b	2000.546	****
Speeding > Network (Overall metadiscourse)	8.0095	**
Video > Text (Overall metadiscourse) ^c	15.8028	****
Audio > Text (Overall metadiscourse) ^c	16.5534	****
<i>Thread</i>		
Network > Speeding (Self-mentions)	72.0725	****
Network > Sci Fi (Self-mentions)	45.1856	****
Sci Fi > Network (Hedges)	46.8777	****
Speeding > Network (Hedges)	21.5244	****
Speeding > Network (Boosters)	22.8357	****
Sci Fi > Network (Boosters)	4.6058	*
Speeding > Network (Engagement)	130.2452	****
Sci Fi > Network (Engagement)	10.8868	***
Speeding > Sci Fi (Engagement)	18.6528	****
Speeding > Network (Attitude Markers)	19.6604	****
Sci Fi > Network (Discourse Markers)	27.7765	****
<i>Gender</i>		
Female > Male (Boosters)	5.1122	*
Male > Female (Hedges)	6.6942	**
Male > Female (Engagement)	5.6535	*
Male > Female (Discourse Markers)	4.2774	*
<i>Mode</i>		
Video > Text (Self-mentions) ^c	96.7819	****

Audio > Text (Self-mentions) ^c	100.5152	****
Audio > Text (Discourse Markers)	4.9392	*
Text > Audio (Engagement) ^c	21.1007	****
Text > Video (Engagement) ^c	10.9684	***
Text > Audio (Hedges) ^d	9.0758	**
Text > Video (Attitude Markers) ^c	5.8628	*
Text > Audio (Attitude Markers) ^c	9.2009	**
<i>Gender-Mode</i>		
Video:M > Video:F (Hedges)	4.8521	*
Audio:M > Audio:F (Engagement)	8.6958	**
Audio:F > Audio:M (Boosters)	8.8934	**

^a * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

^b *There is no significant difference in the ratio of interactional:interactive for any of the variables below. Results for 'overall' apply equally to 'interactional'.*

^c *No significant differences were found between video and audio on these measures.*

^d *Text > Video (Hedges) is not significant, nor is Video > Audio (Hedges).*

Table 7. Chi square results for metadiscourse analysis

To summarize the main findings, the corpus contains more interactional than interactive metadiscourse overall. Speeding contains the most metadiscourse, followed by SciFi, then Network, and each thread favors different terms, reflecting the topic of the thread. Males use significantly more hedges (especially in video), engagement markers (especially in audio),⁵ and discourse markers, while females use more boosters (especially in audio). Video and audio comments contain significantly more metadiscourse overall than text, especially self-mentions, and audio contains more discourse markers. In contrast, text contains significantly more engagement markers, hedges, and attitude markers than video and audio.

Discussion

Research Questions and Hypotheses Revisited

Our first research question asked: What differences, if any, are there in sociability and negativity across the three commenting modes available on VoiceThread.com? We hypothesized that “richer” modes such as audio and video would convey more social communication and “leaner” text more impersonal, contentious communication. The first part of this hypothesis was partially supported. Significantly more metadiscourse terms were found in the audio and video comments than in the text comments overall. The difference was mainly accounted for by the much greater frequency of self-mentions in video and audio, however. Text comments actually had more engagement markers, including ‘you’, more hedges, and more attitude markers. Thus text does not appear to be impersonal. If anything, these results make text appear more other-aware, in contrast to audio and video, which seem to make commenters more self-conscious.

The second part of Hypothesis 1 was also partially supported. The greater frequency of attitude metadiscourse in text is consistent with the results of the attitude content analysis, which found more expressions of judgment in text. As hypothesized, text comments were also significantly more negative than audio and video comments overall,

although male commenters alone were responsible for this result. This gender difference is consistent with previous findings that flaming and hostility are more frequent in text comments by males than by females in discussion forums, newsgroups, and chat rooms (e.g., Herring, 1994, 1995, 2003).

Our second research question asked whether there were gender differences, independent of mode choice. We hypothesized that females would be more social and less contentious than males. The results showed that females were significantly less negative than males in all modes, in support of the second part of this hypothesis. However, few gender differences were found for sociability, as measured by the use of metadiscourse terms. If anything, males used slightly more metadiscourse terms overall, relative to the number of words in their comments. Moreover, females' greater use of boosters and males' greater use of hedges is the opposite of what has been observed in spoken discourse (cf. Coates, 1993) and Internet discussion forums (cf. Herring, 1995; Herring, Johnson, & DiBenedetto, 1998). Thus the first part of the hypothesis was not supported. A possible explanation for the lack of predicted gender differences in sociability is that the task/topic was the same for both genders within each thread, and almost all comments were on-task. Herring and Paolillo (2006) found that gender differences in word frequencies disappeared in blog posts when the general topic of the blog (personal reflections vs. commentary on events external to the writer) was held constant. Similarly, topic may exercise a greater influence than gender in determining metadiscourse term frequency.

Finally, the third research question asked whether there were gender differences in mode choice and amount of participation, hypothesizing that males would participate more and comment more often in video and audio than females would. Males posted more and longer messages in each thread, consistent with the first part of this hypothesis. Males also chose to comment using video much more often than females did, although there was no gender difference in choice of audio.

The hypotheses and results of the study are summarized in table 8.

Hypothesis	Supported?
Video and audio more social	YES overall and for self-mentions; NO for engagement, hedges, attitude markers
Text more judgmental, negative	YES
Females more social	NO
Females less contentious	YES
Males participate more	YES
Males use audio and video more	YES for video; NO for audio

Table 8. Summary of results

Implications for CMC Theory

Overall, our hypotheses predicting modal differences in discourse are supported by the VoiceThread data. These hypotheses were based on early CMC theories—the social presence theory (e.g., Short, Williams, & Christie, 1976), the media richness theory (e.g., Daft & Lengel, 1984), and lack of social cues theory (e.g., Kiesler et al., 1984; Sproull & Kiesler, 1992)—which in turn were grounded in the notion that technical properties of communication media determine the nature of the communication that takes place through

those media. The notion of technological determinism has been much criticized in subsequent literature, especially given the evidence that naturally occurring text communication on the Internet, in email, discussion forums, chatrooms, instant messaging, blogs, etc., can be highly social, even “hyperpersonal” (Walther, 1996). How can we reconcile the findings of the present study, which are also based on naturally occurring communication on the Internet, with these findings of sociable text-only CMC?

Walther and Bazarova (2008) invoke electronic propinquity theory (Korzenny, 1978) to explain why, under experimental conditions, subjects express greater satisfaction and experience more “propinquity” (perceived closeness) with their interlocutor(s) in a given mode of communication when no other modes are present in the environment to compare it with. The tenet that “the fewer one's choices of media, the more closeness one may experience even through the lowest of bandwidths” (Walther, in press) is not directly relevant to the present study, since more than one mode was available to VoiceThread users, and perceptions of closeness were not addressed in this study. Nonetheless, electronic propinquity theory helps explain the apparent inconsistency between our results and those of other studies of natural computer-mediated discourse, if one considers that Internet users for the most part previously had only textual CMC available. With only text to work with, users adapted their practices to enable rich communication. VoiceThread, in contrast, offers a choice of multiple modes, and in that context of choice, differentiation of behavior occurs across modes, consistent, for the most part, with technologically deterministic theories and the findings of multi-mode experimental studies. Thus the VoiceThread results provide hitherto unavailable evidence that the technological properties of a mode shape the nature of the communication when multiple modes are available to choose from in a natural (i.e., non-experimental) context.

The alternative explanation—that the VoiceThread participants selected a given mode because they wished to communicate in a certain way; that purpose shaped their choice, rather than the choice shaping their communication—seems less likely, for several reasons. First, topic was a constant within each VoiceThread; thus we may assume that the purpose for communicating was similar for each participant. The on-topicness of all but a few random comments supports this assumption. Second, almost all the video comments included some reflection on the fact that the commenter was ‘trying out’ the video mode for the sake of experimentation, because it was novel. That almost all video commenters were male is consistent with the findings of Gefen and Straub (1997) and Venkatesh and Morris (2000) that males tend to adopt new technology earlier than females do. Finally, the majority of commenters used text, but not because they wanted to express negative judgment; many text comments were positive. Participants chose text presumably because text commenting was most familiar to them; it was the default choice. Video commenting was selected least often, most likely because it was unfamiliar and perhaps intimidating to users of the platform. This finding violates Korzenny’s (1978) tenet that users will adopt the widest bandwidth communication medium available to them.

Thus the VoiceThread data partially support several previous theories, but they do not fully support any one of them. They do not support the prediction of the social presence, media richness, and electronic propinquity theories that richer media will be preferred for social communication. Although the comments in video and audio contained more linguistic expressions of sociability—interactional metadiscourse—than the text comments, considerations of novelty and familiarity appear to have influenced mode

choice to a greater extent than the affordances of the modes themselves. The VoiceThread results do provide compelling evidence, however, that once a mode is selected, its technical properties influence the communicative behavior that takes place in that mode. In particular, as Kiesler et al. (1984) reported long ago based on experimental results, text communication tends to be more contentious. This has also often been observed in discussion forums and chatrooms: At the same time that textual communicators are more socially “hyperpersonal,” they tend to be anti-socially disinhibited (Walther, 1996).

However, we do not assume that this is because text is depersonalized or that the message sender lacks awareness of the addressee; in our data, text messages contained more other-awareness markers than audio or video. Rather, we subscribe to the view that asynchronous textual CMC provides a buffer of temporal and spatial distance between the sender and the receiver that allows senders to express themselves freely without having to deal with the receiver’s immediate responses. In this respect, textual CMC can be superior to richer media, including FTF communication, for certain sensitive tasks, such as firing an employee or professing one’s love (e.g., Sproull & Kiesler, 1992; Walther, 1996). At the same time as it retains these medium benefits, we propose that text commenting has become “enriched” through decades of use as the primary mode of CMC, such that it is now an unmarked mode capable of fulfilling a wide range of functions. This proposition is supported by Sindoni (in press), who recently surveyed users about when and why they choose to use written text during synchronous videochats. The respondents indicated that they used text for precision, secrecy, and, most tellingly, intimacy or to express face-threatening feelings—a varied set of uses, including those considered by the media richness theory to occupy both ends of the “richness” continuum.

Thus rather than positing FTF communication as the “universal medium” (Kim et al., 2007), this account casts text as the default mode in CMC contexts. It is in the *marked* modes, asynchronous audio and video commenting, that the technical properties of the modes appear to exercise greater influence, presumably because their novelty leads users to notice them more. On VoiceThread.com, the immediate reflection audio and video provide of the speakers as they record their messages appears to increase the speakers’ self-consciousness. Similarly, Sindoni (2011) observes that the videochat users in her study displayed self-consciousness, including arranging their hair as if their video image were a mirror, when communicating via video. As Sindoni (in press) writes:

Observing oneself during a conversation produces a series of psychological effects influencing the verbal and non verbal characteristics of the online exchange. [... C]asting a sidelong glance at oneself during a conversation may change, if not determine, the way one speaks, gesticulates, smiles and so on. (p. 10)

While this is especially true of video, audio commenting also has a defamiliarizing effect in that it requires the commenter to speak out loud to a computer, which can feel strange, as anyone who has tried to rehearse a presentation without an immediate audience can attest. The resulting self-consciousness can affect behavior and language use.

Conclusions

This study asked how different commenting modes are used on a multimodal commenting platform, VoiceThread.com. The results suggest that text commenting is the default on the

platform, as it is for CMC on the Internet in general. However, independent of the reasons for which users choose to comment in one mode or another, the selection of a mode predisposes them to communicate in certain ways. In particular, multiple modes being available and all other factors being equal, the technical properties of textual CMC appear to encourage openness, brevity, and precision, whereas video and audio increase users' self-consciousness and self-focus.⁶

Other factors are not always equal, of course. Previous experience and degree of comfort with a technology may affect mode choice. Social factors such as gender may also condition choice; for example, peer influence appears to influence women more than men, and men tend to be early adopters of new technology. Moreover, the greater social presence of audio and video may be disadvantageous for users in contexts where they do not wish their gender to be revealed.

Situational factors may condition mode effects, as well. The lack of availability of other modes to choose from can lead an otherwise lean channel of information to become enriched, as proposed by the electronic propinquity theory. Over time, users' adaptations of the mode to serve a range of purposes can effectively widen the scope of the mode itself, as we suggest has taken place for textual CMC. Conversely, some of the self-consciousness observed in the present study in video and audio comments may be attributable to the novelty of the voice-based asynchronous commenting modes, which renders their technological properties salient to users.

Limitations of the Study and Directions for Future Research

The small, heterogeneous sample analyzed is a limitation of the present study. The three sample VoiceThreads proved to be significantly different in their expressions of attitude, polarity, and metadiscourse, as well as in message length. Future study of VoiceThread could scale up while focusing on a single demographic and discussion type, such as pedagogical discussion among adults, to obtain more robust and consistent findings. Further, the commenting modes on VoiceThread proved not to be equally easy to use. More invalid messages were posted in video and audio than in text; this could have biased mode use away from video and towards text. This asymmetry should be taken into account in interpreting the results of this study. Finally, our analysis did not directly address reasons for mode choice, satisfaction with modes, or perceived social presence in different modes, as did previous experimental research; rather we focused on discourse behavior in each mode. As a consequence, our findings are not directly comparable with those of most previous studies of cross-mode differences, although we believe that legitimate parallels can be inferred, as we have attempted to do here.

Finally, the fact that VoiceThread is an emergent technology and most users are inexperienced with the platform means that the snapshot we have presented here is likely ephemeral. Our analysis leads us to predict that communication especially in the novel commenting modes, asynchronous audio and video, will evolve as users become more familiar with their use. Such familiarity will likely lead to more strategic usage that takes advantage of the inherent strengths of each mode but that is also susceptible to influence from social and situational factors present in different contexts of use. It will be interesting to track this evolution in future research.

The results of this study have theoretical and methodological implications for research on other interactive multimodal platforms that offer users a choice of communication modes, such as YouTube and World of Warcraft, as well as sites that combine text and audio/video more generally. Some theoretical implications have been explored in this study as regards media richness. Multimodal platforms, especially those that combine entertainment content with CMC, also raise methodological challenges: the need to devise parallel transcription and visualization displays for textual and non-textual communication; the need to devise new analytical methods drawn from disciplines outside communication and linguistics (e.g., visual semiotics); and, finally, the need to conceptualize approaches that capture the interaction among multiple semiotic systems. This study has demonstrated one bottom-up approach to addressing this last challenge, by treating modes of communication as building blocks of a larger whole and seeking to identify the effects each contributes.

Notes

- ¹ The term 'mode' is used here to refer to a CMC option within a single platform, as distinct from the term 'medium,' which refers to, e.g., CMC in contrast to face-to-face communication or mediated communication hosted on a non-Internet platform (see also Murray, 1988).
- ² Retrieved August 18, 2011 from <http://www.appappeal.com/app/voicethread/>
- ³ The contents of the three comments are as follows. Audio: "Hey ya' guys, welcome to .. VoiceThread. I want you to watch this video clip, and give your feedback, either through a video recording, or a voice recording. It's a short clip on speeding, and I'm interested to hear your thoughts." Video: "Hey hey! It's not actually Steve Perry, it's just me ... checking this thing out. Uh, this video clip ... is pretty neat, I think ... I especially enjoy his Australian accent, and I know it's Australian bec- for a few reasons, uh .. the biggest one of them obviously being that the car's steering wheel ... is on the right hand side of the car! So. Another cool thing that I ... liked about this video, other than the information in it, was the special effects they used, he- he's sorta walked out in front of the car, uh which .. I wouldn't think anybody would normally do, but he did it! And, uh, Steve Perry - come back for us all, please!" Text: "People may think that they have a quick reaction time but in fact it doesn't matter whether you can break fast or you have good breaks on the car its what speed you are going before you notice something that either has run out on the road or that you are about to hit something...be aware on the surroundings [sic] and do the speed limit or this could be you!!!"
- ⁴ The full corpus contains 363 comments, of which 43 were invalid because they were empty or contained nothing but noise. Invalid comments appeared to be the result of problems using the commenting technology. The breakdown by mode of invalid comments is: video (29%), audio (17%), text (7%). The thread with the youngest participants, SciFi, had the greatest proportion of invalid comments (20%), compared with the teen Speeding thread (8.4%) and the adult Network thread (1.8%).
- ⁵ Contributing to this result is the fact that 11 out of 29 of the female audio comments to the Network thread, which contained the most and longest audio comments, appeared to

be scripted in advance, resulting in speech that sounded more written-like than spoken-like. Engagement markers are common in unplanned speech.

- ⁶ Differences can also be observed across modes that are due to differences between speaking and writing. For example, audio and (especially) video comments, because they typically are not planned in advance, contain more false starts, pause fillers (um, hm, etc.), and phatic communication (such as greetings and closings) than does text, and voice comments tend to be more rambling and less concise, as can be observed in the examples in note 3. (But cf. note 5.)

References

- Barrett, E., & Lally, V. (1999). Gender differences in an on-line learning environment. *Journal of Computer Assisted Learning*, 15, 48-60.
- Baym, N., Zhang, Y., & Lin, M. (2004). Social interactions across media: Interpersonal communication on the Internet, telephone and face-to-face. *New Media & Society*, 6(3), 299-318.
- Coates, J. (2003). *Women, men, and language*, 2nd ed. London: Longman.
- Cooper, A., & Sportolari, L. (1997). Romance in cyberspace: Understanding online attraction. *Journal of Sex Education and Therapy*, 22, 7-14.
- Daft, R., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organizational design. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (pp. 191-233). Greenwich, CT: JAI Press.
- Frankel, K. A. (1990). Women and computing. *Communications of the ACM*, 33(11), 34-45.
- Gefen, D., & Straub, D. W. (1997). Gender differences in the perception and use of e-mail: An extension to the Technology Acceptance Model. *MIS Quarterly*, 21(4), 389-400.
- Herring, S. C. (1993). Gender and democracy in computer-mediated communication. *Electronic Journal of Communication*, 3 (2). Retrieved August 30, 2011 from <http://ella.slis.indiana.edu/~herring/ejc.txt>
- Herring, S. C. (1994). Politeness in computer culture: Why women thank and men flame. In: *Cultural Performances: Proceedings of the Third Berkeley Women and Language Conference* (pp. 278-294). Berkeley Women and Language Group.
- Herring, S. C. (1995). Men's language on the Internet. *Nordlyd*, 23, 1-20.
- Herring, S. C. (2003). Gender and power in online communication. In J. Holmes & M. Meyerhoff (Eds.), *The handbook of language and gender* (pp. 202-228). Oxford: Blackwell.
- Herring, S. C., Johnson, D., & DiBenedetto, T. (1998). Participation in electronic discourse in a "feminist" field. In J. Coates (Ed.), *Language and gender: A reader* (pp. 197-210). Oxford: Blackwell.
- Herring, S. C., & Paolillo, J. C. (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10(4), 439-459.

- Hyland, K. (2005). *Metadiscourse. Exploring interaction in writing*. Oxford: Continuum.
- Kiesler, S., Siegel, J. & McGuire, T. W. (1984) Social psychological aspects of computer-mediated communication. *American Psychologist*, 39, 1123-1134.
- Kim, H., Kim, G. J., Park, H. W., & Rice, R. E. (2007). Configurations of relationships in different media: FtF, email, instant messenger, mobile phone, and SMS. *Journal of Computer-Mediated Communication*, 12(4), article 3. <http://jcmc.indiana.edu/vol12/issue4/kim.html>
- Kim, M-S., & Raja, N. S. (1990). *Verbal aggression and self-disclosure on computer bulletin boards*. ERIC document (ED334620).
- Korzenny, F. (1978). A theory of electronic propinquity: Mediated communications in organizations. *Communication Research*, 5, 3-24.
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. New York: Palgrave Macmillan.
- Millard, M. (2010). Analysis of interaction in an asynchronous CMC environment. *Web Science Conf. 2010*, April 26-27, Raleigh, NC. Retrieved August 18, 2011 from http://journal.webscience.org/391/2/websci10_submission_106.pdf
- Murray, D. E. (1988). The context of oral and written language: A framework for mode and media switching. *Language in Society*, 17, 351-373.
- Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier*. Reading, MA: Addison Wesley.
- Rice, R. E., & Love, G. (1987). Electronic emotion: Socioemotional content in a computer-mediated communication network. *Communication Research*, 14(1), 85-108.
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. London: Wiley.
- Sindoni, M. G. (2011). *"Mode-switching": Speech and writing in videochats*. Paper presented at GURT 2011: Discourse 2.0: Language and New Media. Washington, DC, March 11.
- Sindoni, M. G. (in press, 2012). Through the looking glass: A social semiotic introduction to the study of videochats. In G. Kress, E. Adami, & R. Facchinetti (Eds.), *The challenge of multimodal analysis to notions of text* [special issue of *Text & Talk*].
- Sproull, L., & Kiesler, S. (1992). *Connections: New ways of working in the networked organization*. Cambridge, MA: MIT Press.
- Venkatesh, V., & Morris, M. G. (2000). Why do not men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior, *MIS Quarterly*, 24(1), 115-139.
- Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 19, 52-90.
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23, 1-43.

- Walther, J. B. (in press). Visual cues in computer-mediated communication: Sometimes less is more. In A. Kappas & N. Krämer (Eds.), *Face-to-face communication over the Internet: Issues, research, challenges*. Cambridge, UK: Cambridge University Press.
- Walther, J. B., & Bazarova, N. N. (2008). Validation and application of electronic propinquity theory to computer-mediated communication in groups. *Communication Research*, 35(5), 622-645.
- Weir, L. (2008, April 16). *Edutopia: Voicethread Extends the Classroom with Interactive Multimedia Albums*. Retrieved November 10, 2009 from: <http://www.edutopia.org/Voicethread-interactive-multimedia-albums>