Constructing artistic discourse: Amateur reviews of amateur movies in a large new media community

Jonathan Warren
School of Library and Information Science
Indiana University, Bloomington
jowarren@indiana.edu

2008-05-01

Introduction

Web 2.0 phenomena have been called the "read/write" Web (Berners-Lee, 2005), where users' contributing of comments and content can provide equivalent, and sometimes greater, value to a website than content merely posted in static pages by the site's owner. This is especially true for e-commerce websites, where users' product reviews and ratings add collaborative filtering capabilities that the site's proprietors could not fabricate on their own. This collaborative activity constitutes a "virtual community" (Hagel & Armstrong, 1997), where anyone with access to that site, regardless of their physical location or demographics, can socially engage with the site's user-base and content.

New media portals (e.g., Youtube) provide users with the capability to post original amateur multimedia content for a virtual community to view, including short animations and videos that resemble traditional films, as well as media with interactive features, such as simple video games. Although site proprietors exist, and can play editorial and administrative roles, they are generally much more peripheral to the process of creating and reviewing content on the site, compared with the community's members. Like e-commerce sites, such portals usually provide relevance feedback and collaborative filtering mechanisms, such as user comments, ratings, and message boards. They also often provide personal profile pages for individual users, as one finds on many social networking and blog sites (e.g., MySpace, Facebook). The primary difference between new media portals and blog portals is that users of the former produce multimedia artifacts on which other users comment, whereas bloggers produce texts.

Literature exists on the typical characteristics of blogs and on user reviews of professionally produced products (e.g., consumer technology, Hollywood films, fine art, etc.). However, no literature currently exists on either the interactions had between *amateur authors and reviewers* on new media sites or the themes present in comments posted to textual blogs. Therefore, the current study presents a thematic content analysis (Rosson, 1999) of amateur movie reviews on one of the oldest and largest new media portals on the Web: Newgrounds.com (2008). Additionally, it relates the results of that analysis to previous studies by our research group of Newgrounds' social network structure, the stylistic features and cultural references present in users' movies, and the demographic data present in users' profiles. This portal was chosen primarily because, unlike Youtube, it is a highly insular community, with users focused on developing both their skills with Adobe Flash animation as well as their reputations in this particular community. This should make the social relation and communication structures of the community all the more salient. Newgrounds' history is also considerably longer than Youtube's (8 years, compared with 3 years), meaning that this is one of the few places where long-term trends may be viewable for this new social phenomenon.

Literature review

This paper supplements recent work by our research group on this community. Paolillo, Warren, and Kunz (2007, 2008) explored the relationship between the social network of author-author "favorite movie" relations on the site and dimensions of variation in the stylistic features of movies

produced by those same authors. They found a seven-core network, with each core being associated with a prominent social clique. Also, by correlating the stylistic features of movies with social network positions, two types of genres became evident, where "genre" refers to typified rhetorical artifacts or performances with socially recognized form and communicative purpose (Crowston & Williams, 2000; Miller, 1994; Yates & Orlikowski, 1992). The first type of genre involved groups of authors that use the features of their movies to act on their social positions, called "crews" by users of the site (e.g., Clock Crew, Lock Legion, Glock Group, Star Syndicate, Kitty Krew, etc.). Crew movies typically represent the faces of users' avatars with clock, lock, etc. icons. Also, very much like a street gang, each crew is known for its own idiomatic slang words and movie features. The second type of genre was production-oriented. Semi-professional or otherwise famous authors often use the site to demonstrate their Flash skills or humor. Adam Phillips (2008) is an example of a professional (formerly Disney) animator who often presents original material on Newgrounds. Similarly, "animutation", a Flash genre created by Internet personality Neil Cicierega (2008), features dense collages of unusual pop-culture images that are animated and set to international pop-music, often with humorous mondegreens (mistaking a foreign language for English, due to homophony) of the songs' lyrics written in subtitles (Kendall, 2007).

Three forthcoming studies (Warren, FC) sought to address limitations with Paolillo et al. (2007, 2008), primarily the absence of a time dimension in the author-author network and movie feature analyses. How have the social network and movies changed over time? Also, the "favorite" network relation does not characterize well the relationships between authors. What are the qualities of these social relations? Finally, self-reported user demographic profiles were available on the site, but not used. Who are these people, and how do their personal characteristics relate to their behavior in this virtual community? To answer these questions, a snowball sample of 2,032,534 movie reviews was collected from the site,[1] with each review containing the following information: date, reviewer, author, score, review title, and review message. An exhaustive sample of 1.14 million user profiles was also collected and combined with the review data, yielding a subset of 853,341 review observations containing age and gender information for both reviewers and authors. Finally, 1000 random user locations were geocoded and plotted for a visual representation of users' physical locations.

Those studies found a 19-core social network, with each core gaining and losing members over time. They also found that cultural references and movie styles are tightly interwoven, and that changes in movie features parallel the network patterns and user-clique memberships over time. The result is a more complex and evolving version of the earlier studies. Demographically, most users report being from the USA, Europe, or Australia, with a few in South and Central America, the Middle East, and Southeast Asia. Most users are 18-22 year old males. Older (than 22) males and younger (than 18) females can hold privileged social positions in the community, often by using sexualized images of young women to get the attention of the young male majority. Teenage males who come to the site often post high volumes of reviews initially, and are the primary reviewers of the high volumes of lower-quality animations that constantly stream onto the site. This causes many of them to burn-out from reviewing over time. Teenage females, on the other hand, usually start reviewing slowly and increase their participation over time. Older males primarily review the most experienced, semi-professional authors, indicating connoisseur behavior.

Though they paint a richer picture of the community than did Paolillo et al. (2007, 2008), these three studies did not examine the nature of "reviewing", being perhaps the core social activity in this

1 Starting from a random seed of 10,000 movie pages, all reviews were collected for that movie. Then, all of the movie pages by any of those same authors were visited, their reviews were collected, and any new collaborators' movie pages were added to the crawler's frontier. This resulted in a snowball sample of 292,881 unique movie pages by 147,926 unique authors.

context. For a macro-scale quantitative study of this community, this suggests the need for a content analysis of themes present in review messages, so that those themes can be related to the social network, movie feature, and demographic results of the previous studies. Therefore, the following *research questions* arise:

1. What kinds of reviewers and review themes are usually associated with different review scores?
2. With what purposes have the different social positions reviewed each other over time?
3. What aspects of the stylistic and cultural content of movies are praised, criticized, or recommended over time, and by whom?

The following is a brief survey several literatures related to this research area, in order to identify potentially relevant thematic variables in amateur movie reviews by amateur reviewers. No literature is known to exist exactly in this area.

Though plentiful, the machine learning literature on text and opinion mining, topic detection and tracking, and sentiment analysis – where automated means are used to classify content into pre-defined classes – is not particularly relevant to the research questions above. This is because the tricky task of interpreting which classes best represent the social semantics in a particular context is usually taken for granted by such studies, so that attention can be given to creating algorithms that can best learn those rules/associations and sort content according to them. Exemplary studies include Lee (2007), who assumes that user needs and product attributes have been pre-specified, then creates a matrix-based approach for finding rules/dimensions of associations between needs and attributes. Ghose (2006) similarly uses the network distance between each word in the review from the "good" and "bad" entries in the Wordnet ontology for classifying buyers' reviews of sellers. Dave et al. (2003) uses Amazon.com's product categories. Turney & Littman (2003) describe "semantic orientation" as being either positive or negative. Though all of these studies are also working on user reviews, they are clearly focusing on a different problem.

Not all work on classifying product reviews has been done by machine learning researchers. Gebauer et al. (2006, 2007), a business researcher, have several working papers in which product reviews for four new mobile information-communication technologies (ICTs) are coded for 45 inductively-derived variables related to the fit of that technology to users' task, context, and performance requirements. The first study uses exploratory factor analysis to reduce the 45 variables' covariations to five factors, and uses multiple regression to test whether those latent variables predict overall scores given by users to the technologies. The second study combines these steps using a structural equation model. These papers' objects of study (mobile technology) are different than digital new media artifacts, however their methods are generally applicable.

Next most relevant are studies that code the structural features (page layouts, content blocks, etc.) of websites. Such studies include Metzer (2005) and Toms & Taves (2004), which seek indicators of reputability and credibility in websites. Reese et al. (2007) and Herring et al. (2004) analyzed the structural qualities of weblogs, and Singer (2005) studied the features of political blogs. This literature, also, is not directly relevant to features of new media content, which are not guaranteed to obey the properties of websites. Website credibility, the topic most often studied by this literature, is not interesting in the same way for new media, because usernames and authors' previous work are usually well-indexed by a third-party portal (e.g., Newgrounds, deviantART, etc.), and the cost/risk of viewing fake movies is low for consumers of free and brief digital media. Although, there do exist identity-related reputability issues, such as authors taking on different usernames and profiles in order to become known for a different style of work. On Newgrounds, for example, the iconic StrawberryClock user, described in the results section, later changed his username to Coolboyman and produced movies

unrelated to the Clock Crew. Closer reading and more non-random investigation of multiple sources of evidence than can be done in the current analysis are required to uncover multiple identities.[2] Perhaps the closest that structural feature studies have come to studying new media reviews is Herring et al. (2004, 2006), which counted the number of occurrences of comments on blogs, though they did not include any interpretive theme analysis of comments' contents.

Closer to digital multimedia, a number of thematic content analyses have studied sex and violence in Hollywood films and their previews. Brennan, Jr. (2002) wrote a masters thesis coding violence in films using a grounded theory approach (Strauss & Corbin, 1998). Oliver & Kalyanaraman (2002) studied sex and violence in film previews, also using an inductive approach, in order to see to what quantity of these themes potentially young viewers could be exposed. Sapolsky & Molitor (1996, 1993) studied the link between sex and violence in gore/slasher films, specifically what kinds of sex and violence occur concurrently in films. Though not studying user reviews or amateur media, the facts that grounded theory is often used to define variables for measuring contemporary multimedia content, and that there is a great deal of sex and violence in Newgrounds movies, are worthy of taking note.

Finally, considerable breadth and history of literature exists in psychology on reviews of fine art by both experts and amateurs. Hekkert & van Wieringen (1996) provide a lengthy review of personality factors that may affect a reviewer's perceptions and preferences about a work of art, including extraversion and neuroticism (Eysenck, 1941, 1972), open-mindedness (Child, 1965), conservatism (Wilson et al., 1973), locus of control (Juhasz & Paxon, 1978), field-dependence and sensation-seeking (Tobacyk et al., 1981), and self-reference (Alexander & Marks, 1983). However, they point out that empirical research into these factors has suggested that they only account for a small amount of preference variance. Though, as they say, social aspects of aesthetic preferences are often neglected by psychologists – including social class (Bourdieu, 1989), cultural background (Heinrichs & Cupchik, 1985), and general educational level (Frances, 1976) – such factors also have been shown to be less determining of aesthetic preferences than are *art-related education and experience*. Adults without formal art training are usually guided by semantic features incorporated from their daily lives, things to which they can relate (Parsons, 1987; Schmidt et al., 1989; Gotshalk, 1962). Visual color preferences are the next most likely to affect aesthetic judgment (Valentine, 1962). However, when confronted by abstract art, inexperienced adults' responses approach those of experts (Schmidt et al., 1989), who are inclined to classify artwork irrespective of content, according to formal properties (Cupchik, 1992; Parsons, 1987), composition (Peel, 1945), style typicality (Hekkert & van Wieringen, 1990), originality, and technique (Getzels & Csikszentmihalyi, 1976).

Since 1996, this literature has primarily been picked up by Human Computer Interaction (HCI) researchers, such as the highly cited review by Lavie & Tractinsky (2004) on assessing aesthetic judgments of websites, and by empirical psychological studies of emotional reactions to art (e.g., Silva, 2005). As the current research questions are more social than psychological exploring public expressions of approval and disapproval for artworks that were often composed by amateurs in/for a collaborative community setting he distinctions noted by Hekkert & van Wieringen (1996) between expert and non-expert reviewing seem most interesting for studying reviews of new media on the Internet. In what contexts does “expertness” (analysis of formal properties) and “non-expertness” (content/semantic orientation) arise in a community of amateurs? What of offline artistic discourse survives in amateur reviews of amateur movies by teens and young adults online?

2 This connection between StrawberryClock and Coolboyman comes from reading the Newgrounds forums, Coolboyman's profile on DeviantART, and the Coolboyman entries in Urban Dictionary.

Methodology

Following many of the earlier studies, and because it is unclear whether psychological fine art theories are appropriate, a thematic content analysis was conducted following a grounded theory approach (Strauss & Corbin, 1998). 39 variables were inductively defined in seven categories, namely: generic, technical, technique, composition, impactfulness, community, and other. A codebook is provided in Appendix 1. With the exception of the "other" category, each variable was coded in the senses of praise, criticism, and recommendation; for example, one can praise, criticize, or recommend a soundtrack of/for a movie. All variables were coded as binary (presence/absence), because the small frequencies of themes in these brief (1-324 word) reviews is not of particular interest. Compared with the literatures reviewed above, these variables represent both the theoretical interests in praise and criticism of the machine learning researchers, sex and violence of the film researchers, and the fine art distinction between technique and content ("composition") of the psychologists. Additional variables created by the researcher relate to the virtual community context (i.e., "technical" difficulties and "community" clique references). Variables in the "impactfulness" and "other" categories are perhaps the closest to psychology in this study (e.g., humor and self-reference). However, only the public/social manifestations of those variables are presented here; no attempt was made to infer the psychological states of reviewers (e.g., inferring anger from using curse words).

The *generic* category refers to vague expressions of approval or disapproval. Generic recommendations (e.g., "could be better") were coded as generic criticism. The only *technical* variable refers to technical errors and difficulties experienced by the reviewer, such as program crashes or unexpected movie behavior. *Technique* variables are the most in evidence, including: imitating camera perspectives, fight scene choreography, use of text, animation type (vector-based, frame-by-frame, stop-motion, etc.), images used, drawing capability, props and colors used, pacing and event timing, interactive features (e.g., game controls), samples of professional music, and other sound effects. The *composition* variables refer to the strategies used in organizing the movie, such as its premise, plot/characters, use of sex and violence, originality, and length. *Impactfulness* refers to how the movie affects the reviewer, regarding aesthetics/style, humor, and difficulty (to play or comprehend). The *community* variables regard social or genre-based cliques, such as the "crews" mentioned earlier, references to the Newgrounds community overall, to "stick" movies, collaborations, "tweens" (pre-teenage users), and several authors whose work is considered socially memetic or exemplary (e.g., the "Foamy" character by user illwillpress, and claymations by user Knox). "Stick" movies feature a protagonist, drawn as a stick figure, who moves through a series of scenes, often involving military infiltration or espionage, killing anyone they encounter, often with complex visual effects imitating the Matrix films. Finally, an *other* category contains variables that did not necessarily involve either praise, criticism, or recommendations. These include condolences, apologies, helpfulness, thanks, challenges, self-centeredness in the form of advertisements as well as more general pridefulness and conceit, requests for responses, and identifications ("Hey, I know you!").

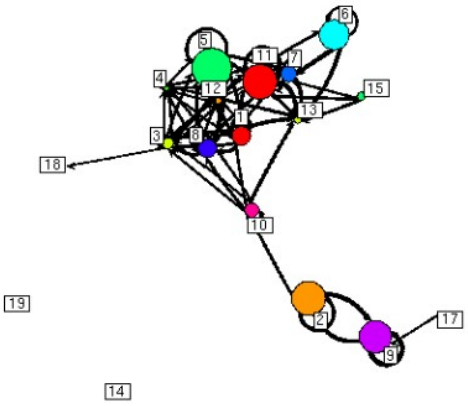


Figure 1: SNA

Two samples were taken from the previous studies' large datasets. The first sample was intended for comparison with the social network analysis from Warren et al. (FC; see sociogram in Figure 1). Of the 19 clusters in that network, in yearly time periods, 2% of reviews from each cluster at each time were randomly sampled, with reviewer and author network position assignments being added to the coding sheet. This resulted in n=180 reviews, a purely pragmatic

sample size. The second sample was intended for comparison with the principal components analysis of movie stylistic features and cultural references from Warren et al. (FC). That study sampled 890 movies from the seven social network positions of Paolillo et al. (2007, 2008) and coded them for 3016 stylistic and cultural features, also using a grounded theory approach. The current study randomly sampled n=200 of those 890 movies, combining the 3016 features from the movies with the 39 variable codings of the review texts in a single coding sheet. For answering the first research question (how review themes and demographics predict scores), all n=380 (180+200) of the review codings were combined in a single matrix, the age and gender demographic information for each reviewer and author were added to the matrix, and the social network position assignments and movie features were removed. No reviews were duplicated in either of the samples or combined sheets. Unfortunately, it is not feasible to combine the network cluster and movie feature analyses in a single statistical model, because the two were drawn from different networks with different clusterings, namely Paolillo et al.'s (2007, 2008) 7-cluster favorite-author network ending in 2006, and Warren et al.'s (FC) 19-cluster review-based network ending in 2007.

The coding was done by one coder, the author, who has a great deal of experience with this community. A second coder who had no previous experience with the community re-coded 10% (n=38) of the reviews using the codebook provided in Appendix 1, and achieved a Krippendorf's alpha value of __.[3] Due to the high insularity and idiosyncrasy of the community, successful coders must often additionally conduct basic definitional searches of the site and Internet in order to understand terminology used in the review texts. These could not be provided in the codebook. For example, one must know that Pube Muppet is a series depicting Kermit the Frog as an aggressive homosexual, many episodes of which were created by Coolboyman.

For the first research question, a general linear model was constructed, with score (0-10) as a continuous dependent variable, and review themes, review date (accurate to the day), and user demographics (reviewer and author age and gender) as independent variables. Interactions between every independent variable and time, measured in days from the first review in the sample, were also included. No link function was used, and the descriptive statistics showed Gaussian distributional assumptions to be appropriate. This model does not account for autocorrelations in observations by the same person over time, however this is not of particular concern, since observations for the same reviewer are rare in the sample. The resulting conclusions are at a community-wide level of abstraction.

For the second and third research questions, two matrices were created, one containing review themes and clusters in the columns, and the other containing review themes and movie features in the columns. Each matrix was column-wise centered about the mean and z-score scaled. Each of these matrices was sub-divided into the same 6 yearly time periods, which were each subjected to column-wise principal components analyses (PCA) with varimax rotations. A single narrative interpretation across the two sets of analyses is provided in the results section. PCA was used instead of exploratory factor analysis (FA; cf. Gebauer et al, 2006), because FA is the same as PCA with an added error coefficient on each measured variable. FA is most statistically appropriate when human subjects provide measurements with potentially valuable variations (e.g., motor coordination speeds), which should be preserved in the model. PCA is most appropriate as a dimensionality reduction tool, where, as is the case here, the measurement of many research variables on insentient artifacts by a trained analyst, having minimal and well-controlled variation, are summarized in a statistically principled way.

3 The re-coding by a second coder could not be completed in time for this draft's deadline.

Results

Variables' distributions

Figures 2-5 show the variables' distributions for each time-period, sorted according to praise, criticism, recommendation, and other, respectively. The curving lines come from linearly interpolating between data values for each period using a cubic function; there are no observations between periods. Time period 1 runs from April 2000 to April 2001, and so on up to time period 5. Due to the dual samples described in the methods section, period 6 includes both 2006 and 2007's observations.

In Figure 2, generic praise is seen to have consistently been the most frequent (blue line) and took a dramatic upturn in 2005, with praise of humor (black) the next most frequent, followed by praise of plots and premises. Figure 3 shows criticisms. Generic criticism (blue) again trumps the other themes, having been briefly overtaken in 2004 by plot and plot and premise criticisms (yellow and red). Criticisms of sound (light blue) have been on the rise recently, and 2005 again seems to have been a year of generic and non-substantive reviews. Figure 4, recommendation variables, has lower frequencies, making the lesser-used variables' oscillations more apparent. Recommendations for more sex and violence (red) appear to be in a long-term oscillation, being initially high, declining over several years, and peaking again in 2005. Plot and drawing recommendations were most common in 2003; introduction and conclusion recommendations in 2004; plot, interactivity, and animation recommendations 2005; and text, plot, sex and violence, and length recommendations in 2006-7. Finally, figure 5, the "other" variables, show somewhat larger frequencies, with abstract self-centered review themes (green) being most prominent overall, and with response requests (blue), condolences (red), and thanks (dark brown) being next most common.

All of these descriptive results are skewed by the frequency distribution of reviews overall on the site, which follows a normal distribution, with a peak in 2005 and with the right third of the curve in the future. This suggests that reviewing on Newgrounds has somewhat died down since its peak in 2005. This growth and decay pattern would be interesting to correlate with the popularity of Flash elsewhere, though this is outside the scope of this study. 2005 was also the founding year of Youtube, suggesting a possible migration of people from Newgrounds to there.

Figure 2: Praise variables

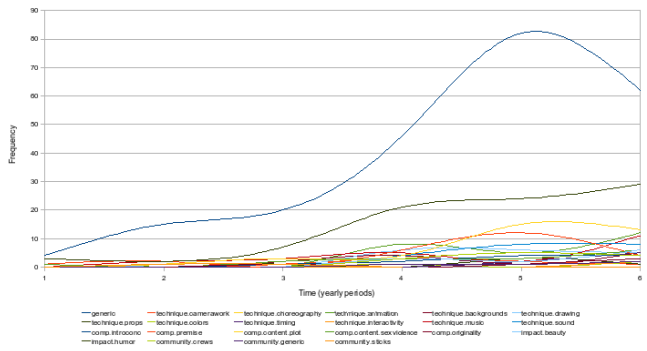


Figure 3: Criticism variables

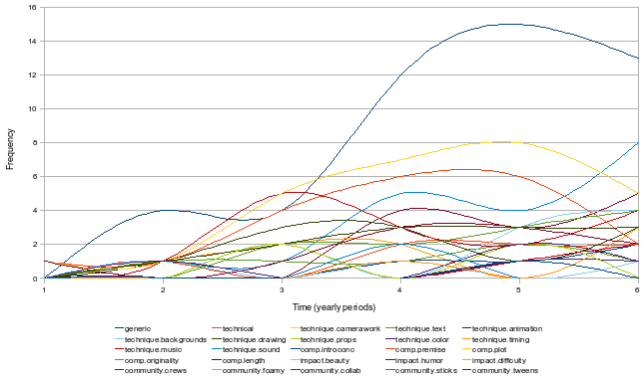


Figure 4: Recommendation variables

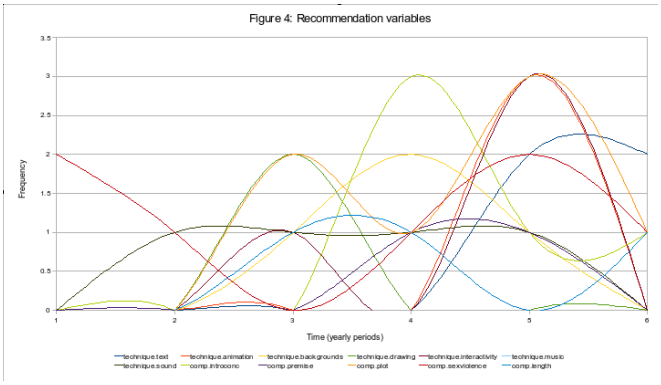
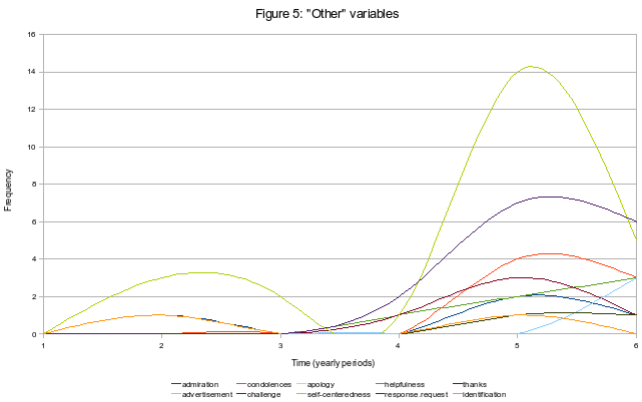


Figure 5: "Other" variables



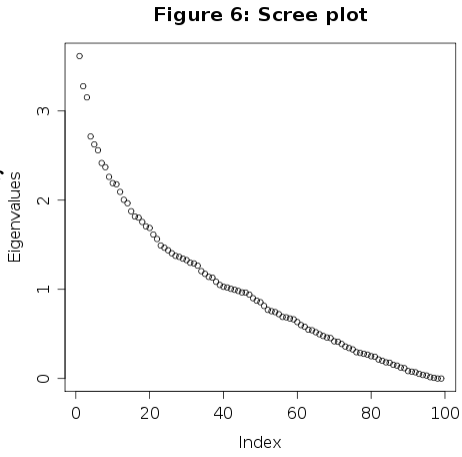
Research question 1: predicting scores

The best-fitting and most significant general linear model is summarized in Table 1. These results describe the ability of review themes to affect scores for the whole community over its entire history. Scores average from 7.1-7.7, with no significant change over time. Generic praise and praise of images can boost a score to near 10 (perfect). Older authors are penalized, but this has lessened over time; younger authors are preferred. No significant gender effects are in evidence here, though they are evident in the forthcoming demographic study (Warren, FC). Generic & premise criticisms lower scores the most. Generic criticism is having less impact over time (though the effect is of marginal significance, $p < 0.08$), and the effect of premise criticism has not changed over time. Oddly, technical criticisms positively affect scores, and this has decreased slightly over time. Criticisms of music and humor have had a worsened effect over time.

Variable	**Parameter estimate**	**Standard Error**	**t value**	**Pr(>|t|), significance** (0.001 ***, 0.001 **, 0.01 *, 0.05 .)
Intercept	7.428	0.3145	23.621	< 2e-16, ***
Generic praise	1.383	0.2743	5.044	0.000000754, ***
Image praise	1.714	0.7586	2.260	0.024477, *
Author age	-0.03918	0.01790	-2.189	0.029296, *
Generic criticism	-5.231	1.205	-4.342	0.0000188, ***
Premise criticism	-2.289	0.5495	-4.165	0.0000398, ***
Technical criticism	5.400	2.561	2.108	0.035756, *
Author age * time	0.00002706	0.00001107	2.444	0.015033, *
Generic crit. * time	0.001401	0.0007998	1.752	0.080666, .
Technical crit. * time	-0.003625	0.001709	-2.121	0.034709, *
Music crit. * time	-0.001480	0.0004880	-3.032	0.002621, **
Humor crit. * time	-0.001520	0.0004375	-3.475	0.000579, ***

Table 1: General linear model

Research questions 2 & 3: review themes alongside movie features and network position



The following is a narrative of the two sets of principal components analyses, regarding how review themes co-vary over time with both movie features and the social network positions of reviewers and authors. Examination of the eigenvalues and scree plots of each analysis indicates that most had 1-3 dimensions accounting for most of that period's variation, however most also had long tails, indicating many more dimensions of variation that could be interpreted. Figure 6 shows a typical scree plot illustrating the effect. For brevity, only the top three principal components will be discussed in the following interpretation. The following interpretation was also facilitated by a qualitative analysis of the members of each network position (cf. Figure 1). For example, the authors in network position 13 are the most famous in the community for producing semi-professional movies.

In the beginning of the portal, a congratulatory environment existed between the original Clocks

and the semi-professional authors, with both groups and their followers non-expertly praising each others' plots, premises, and use of sex and violence. However, dissidence developed among the follower groups, who began criticizing the short lengths of movies by the popular authors.

In 2002, the general praise among popular groups continued, but the technique-related issues of how to use props and text in movies emerged in both the popular and follower groups. During this time, Clock themes and non-expert praise for movies with action, sex, and violence were dominant in the movies' features. Among the semi-professional crowd, praises for humor were also supplemented with recommendations for better music and sound effects. The stick figure group became known for having especially good cinematic music. Movies featuring Japanimation ("anime") themes and popular culture figureheads similarly began experiencing criticisms wishing for better sound effects and drawing at this stage.

In 2003, the old Clocks and semi-professional authors were praised by the new Clocks for the improved lengths, difficulty, and drawing of their movies and games. The new Clocks and sticks also gained increased attention for their innovations in sex and violence put to music, though the sticks were recognized as having better sound effects and images, and the early new Clock movies were judged to be too short and poorly drawn. In the movie features during this time, generic praise for horror movies and violent video games abounded; movies about terrorism by Arabs came under attack for being unoriginal, not funny, and poorly drawn; and appreciation grew for drama, role-playing, and dancing movies that had good plots, props, and music.

In 2004, the new Clocks tried to distinguish themselves from the sticks in terms of sound effects, music, animation, and drawing. The semi-professional group suggested that the new Clocks plots, aesthetics, and timing could also be improved, but liked their premises and camera work. They also gave the same praises and criticisms to the old Clocks. The semi-professional group was criticized at this time by the followers and stick groups for having too short of movies with poor text, and was praised by the new Clocks for their premises, aesthetics, animation, and drawing styles. Follower and mixture groups in 2004 criticized the old Clocks for getting involved with the sticks and the Foamy character, and suggested that their sound and music could be better, but approved of their plots and images. In the movie features, new Clock movies were praised for their images and colors, but criticized for poor premises and humor. The old Clocks were praised for their humor, plots, sex and violence, sound effects, and drawing, and criticized for having poor premises, animation, and music. An interesting group of movies also emerged during this period that resisted the new and old Clock milieu with comforting domestic and nature themes.

In 2005, the new Clocks praised the aesthetics, humor, and sound of themselves and the semi-professional authors. The old Clocks and their followers recommended better animation, images, and interaction from the sticks, who responded by praising the drawing and camera work of the old Clocks, but criticizing their images, sound effects, and use of text. Follower groups at this time largely targetted their reviews towards the new Clocks, most often praising their animation and sound effects, and asking for better music and images. Movie features during 2005 followed review themes in two general ways. There was praise for premises, plots, humor, sex and violence, drawing, and sound related to mafia movies, movies about violence in nature, about dictators (who are highly correlated with the old Clocks), and about mummies and zombies. Alternatively there was generic praise for video games, teen action and comedy actors, fast food, and classic rock and roll.

Most recently, in 2006-2007, the Clocks have been interacting more with themselves, the periphery, and more central follower groups. The new Clocks praised their own colors and timing, and criticized their own humor, texts, and short movie lengths. The old Clocks praised the color and timing, and criticized the text and sound effects, of their follower groups. Groups on the periphery of the network criticized the old Clocks and their followers for almost all of the technique variables as

well as plot and humor. Movie features have recently found a balance between admiration for the new Clocks and comedians outside of Newgrounds, on the one hand, and praise for the firearms and obedience ("ownage") themes of the old Clocks, on the other. An interest also persists in more domestic, natural, and outdoors themes.

Discussion

Judging from the variables' distributions, praise most often regards non-expert themes, and this has been true throughout the history of the community. Non-expert themes also appear to dominate critical reviews, though more technique-oriented expert themes have arisen recently regarding sound. More expert artistic appraisals are seen in recommendations, with non-expert views being replaced somewhat by technique criticisms since 2005. As in Hekkert & van Wieringen (1996), the more personal, and possibly psychological, "other" variables appear to be an undercurrent, though not insignificant in size. The deepest conclusion available from this study is that such comments are primarily self-serving and secondarily empathetic. More study by psychologists in this area could be fruitful.

Regarding the first research question, over the history of the site, the review themes that have had the greatest effect on review scores have been non-expert in nature: praise and criticism in general, about humor, movie premises, and technical difficulties. There is also a small bias in favor of younger authors, which has been lessening with time. Additionally, more expert, technique-related themes are in evidence to a lesser degree, especially regarding authors' use of images and music.

For research questions two and three, the social network and movie feature analyses suggest that expert themes in review messages have emerged slowly from non-expert themes. The first technique-related themes were from less popular groups generally complaining about movies being too short. Over the following year, the core groups of the early Clock Crew and semi-professional authors developed internal discourses about using props, text, images, sound effects, and music samples in their movies. The rise of the stick figure genre, with its accompanying Matrix-inspired soundtrack, further encouraged praise from the general community regarding music selection in movies. This focus on music extended into 2003, when the rise of the new Clocks brought requests from community members for better drawing techniques in order both to resist the old Clocks' characteristically simple style and to imitate the semi-professional authors. In trying to distinguish their style from the sticks, who are known for advanced animation techniques, the new Clocks' self-reflective, and semi-professional authors', reviews in 2004 incorporate animation, timing, and camera angles into their existent discourse on sound, drawing, and images. The missing piece of artistic critique on the site at this point seems to have been interactivity, extending movies with player controls and video game features. The old Clocks asked this of the sticks in 2005, as did general followers and mixture groups of the new Clocks. It is interesting that the old Clocks were not asked for more interactivity, presumably because it is not consistent with their traditional, simple, and controlling public image. The most recent innovation in expert artistic discourse in the community has regarded color. After 2006, the new Clocks began praising their own use of color, the old Clocks praised their follower groups' use of color, and the peripheral groups criticized the old Clocks' use of color. This has happened alongside a parallel trend towards less dark and serious features of movies, and more focus on nature, activities, and domestic themes.

Unexpected results include the importance of technical criticisms in the general linear model. Why should having a bug in one's movie or game cause a reviewer's score to increase sharply? It intuitively seems that bugs should frustrate reviewers. The reason might be that reviewers who go to the trouble of exploring and commenting on bugs do so because they otherwise enjoyed the movie, and therefore want to help and praise the author overall. The "technical criticism" effect in the model (see

Table 1) has relatively low significance ($p < 0.03$), and has a large standard error (2.6), so future results could clarify this.

Implications of these findings for the literature are considerable. In this community of amateur artists, within the first year, those with the most interest in producing movies found others of similar experience and skill, and formed discursive collaborations, initially just for fun, based on their everyday teenage senses of humor and personal interests (e.g., video games). Many other smaller groups of less-committed amateurs and spectators formed around those first groups, and their criticisms started the formation of more technical/expert artistic discourses in, or in reference to, the main groups. The exact development of those discourses depends on the social positions and artistic choices/styles that certain groups have chosen to occupy and make their own. However, eventually, the discourse matured, such that the appropriate groups had become, or been made, aware of techniques known to be relevant for their chosen genre, and a common artistic vocabulary emerged. The elements of that vocabulary may have come from outside the community (e.g., the stick genre imitating the music and camera work of the Matrix films) or from inside (e.g., clock face images for avatars). Also, since social identity here is tied to one's reputation for producing a certain genre of movies, doing something new can require remaking one's profile, user image, and career/history of movie-making (e.g., Coolboyman). This seems very much like "expert" schools of art and thought one finds in the academie. Also like academia, this community has allowed one group, the semi-professional group, to remain aloof from the social-positioning and genre-enforcing environment, so long as its members produce virtuosic work. However, just as in the non-virtual world, as for every other author in the community, their talent must be appreciated by the community, in order for it to be approved (not "blammed") for public posting on the site (i.e., publication). One can see this in the works of Adam Phillips (2008), which possess both the artistic refinement of recent Disney animated films as well as the harshness and violence ("Bitey of Brackenwood") expected by a western male teenage audience.

Conclusion

In summary, this paper sought to understand the nature of amateur reviews of amateur new media on the Internet by studying one of the largest, oldest, and most successful new media virtual communities. This topic was prompted by previous social network, genre, and demographic analyses of this interesting new Web 2.0 phenomenon. Review of the genre and aesthetics literatures suggested a difference between experiential non-expert judgments of art and technique-related expert judgments. A thematic content analysis of n=380 movie reviews using a grounded theory approach yielded 39 variables, which were categorized and related to the previous literature. Modeling these observations in terms of review scores, social network position, and movie features found that initially non-expert judgments formed through social interaction over time into a stable technical vocabulary within the community, which resembles many of the peer-reviewed institutions of the offline world.

Implications of these findings include confirmation that, even in the relative absence of formal institutions of art, motivated amateurs will seek out inspiration, from within themselves and the community or outside of it, and will construct an ever-more-refined and stable discursive tradition and vocabulary over time. For designers and administrators of artistic virtual communities such as these, this means that providing an interesting creative activity, a basic means of personal identity and history management, and a minimal and well-archived discussion and review mechanism may be key components. Members of such communities find a supportive and challenging creative environment, however only a few prodigies can escape the peer pressure to conform to social norms of (sexual, violent, etc.) movie content and artistic form, and their work is still judged generally by those norms.

There are two notable limitations to this study. First, two different networks from previous studies were used to sample the reviews. This meant that the reviews vs. network and reviews vs.

movie features analyses could not be combined into a single quantitative model. In this case, this was done because re-viewing and re-coding the features of a sufficiently large sample of movies based on the 19-node network was prohibitively time-consuming. Second, despite centering and scaling, the very large number of movie feature variables effectively drowned out the review data, such that only the most frequent (and generic) review themes clearly correlated with dimensions of movie features. Fortunately, the number of network positions was comparable to the number of review variables, and the number of observations was sufficient, so the evolution of an artistic vocabulary in the community was still observable in great detail. More conservatively choosing movie features for the analysis should improve the results.

Future research could confirm these results on other new media communities as well as communities of weblogs. Comparing the vocabularies that emerge in different virtual communities could also be worthwhile. Also interesting would be to combine the relational and attributional datasets in a single statistical model (e.g., Snijders et al., 2006), to estimate the probabilities of individuals' group memberships as they change over time (rather than clustering based on the whole network), and to model the network as a time-series rather than in terms of several long periods.

(6590 words up to this point)

Appendix 1

Codebook

Syntax

1. Variable category
    1. Variable name
        1. Examples of how a variable's value might manifest

Praise, criticism, and recommendation

1. **Generic**
    1. **Generic**
        1. Praise – awesome, great, good, cool, insane/crazy (depends on context)
        2. Criticism – sucks, crap, terrible, awful, worst ever
        3. Recommendation – (coded as criticism) could be better, keep working on it
2. **Technical**
    1. **Errors & difficulties**
        1. bugs, crashes, unexpected behavior, slow computers, didn't play
3. **Technique**
    1. **Camera work, perspective**
        1. 3D-like effects, first-person perspective, other perspectives taken by the camera/viewer
    2. **Fight scenes**
        1. quality of fight scene choreography, rendering, planning
    3. **Text**
        1. labels, menu text, titles, subtitles, credits, links, etc. used in the movie; not to be confused with “interactivity” below
    4. **Animation**
        1. “animation”, vector-based movement/transforms, frame-by-frame, stop-motion, live-action/video
    5. **Images**
        1. images used as backgrounds, in collages, as icons
    6. **Drawing, art**
        1. hand-drawing capability of the artist, as vectors or raster
    7. **Props**
        1. objects referenced in the movie: cars, guns, bombs, buildings
    8. **Colors**
        1. references to specific colors: blue, red, green
    9. **Timing**
        1. pacing of the story/game: speed of the plot, scenes, scrolling speed, responsiveness of game controls
    10. **Interactivity**
        1. presence of interactive functions: menu buttons, on-screen/keyboard/mouse controls
    11. **Music**
        1. music/songs that have been professionally recorded and are sampled in the movie
    12. **Sound**
        1. sound effects, foley art, voice-overs by the author, lip-synching
4. **Composition, organization, strategies**
    1. **Introduction & conclusion**

    1. references to introduction/beginning or conclusion/ending sections of the movie
2. **Premise**
    1. the point of, or motivation/reason for, the movie; common criticisms are that movies are random or ill-conceived
    2. references to "series" of movies by that author were coded as premise
3. **Plot, narrative, characters**
    1. specific characters or scenes in the movie: why did you use Goku (from the Dragon Ball Z anime series) here?
4. **Sex & violence**
    1. asking for more or fewer sexual or violent acts in the movie
5. **Originality**
    1. creativity; absence of plagiarism; originality of plot, premise, technique
6. **Length**
    1. Is the movie too long or short?

5. **Impactfulness**
    1. **Aesthetics, style**
        1. Does the reviewer approve of the aesthetic sense of the author? Beautiful, ugly, great style
    2. **Humor**
        1. Is the movie funny?
    3. **Difficulty**
        1. Usually for games, too hard or easy; also difficulty to comprehend/follow the plot or premise
6. **Community**
    1. **Crews**
        1. Clock Crew, Lock Legion, Glock Group, Star Syndicate, Gost, Kitty Krew
    2. **Newgrounds overall**
        1. references to the Newgrounds community at large, "reviewers" in general
    3. **Sticks**
        1. sticks, stickdeath
    4. **Foamy**
        1. foamy, rant, squirrel, illwillpress, germaine
    5. **Collaborations between authors**
        1. collabs, work with, find someone to do your sound for you
    6. **Knox**
        1. the knox author, claymation
    7. **Tweens**
        1. tweens, younger-than-teenage users

Other

1. **Condolences, sympathy**
    1. rest in peace, I feel your pain, I wish you all the best
2. **Apology**
    1. sorry, apologize, please forgive me
3. **Helpfulness**
    1. hope this helps, giving general/career advice
4. **Thanks**

    1. thanks, thanx, thank you, appreciate
5. **Challenge**
    1. to a rematch, duel, to make another equally-good movie
6. **Self-centeredness: advertisement**
    1. service-oriented self-centeredness: you should collaborate with me, I could fix that for you
7. **Self-centeredness: abstract**
    1. gloating, bragging, favoritism, self-praise, possessiveness, jealousy, I want
8. **Response request**
    1. asking a non-rhetorical question: how did you do that, please let me know if you received my submission (for collaboration managers)
9. **Identification**
    1. I know you or your work from somewhere

References

Alexander, B., & Marks, L. E. (1983). Aesthetic preference and resemblance of viewer's personality to paintings. *Bulletin of the Psychonomic Society*, 21, 384-386.

Berners-Lee, T. (2005). Berners-Lee on the read/write web. Retrieved April 28, 2008, from http://news.bbc.co.uk/2/hi/technology/4132752.stm

Bourdieu, P. (1989). *Distinction: A social critique of the judgement of taste* (R. Nice, Trans.). London: Routledge. (Original work published 1979)

Brennan Jr, R. D. (2002). *A multidisciplinary approach to the effects of violence in motion pictures*. Georgetown University.

Child, I. L. (1965). Personality correlates of esthetic judgment in college students. *Journal of Personality*, 33, 477-511.

Cicierega, N. (2008). Personal homepage. Retrieved April 30, 2008, from http://www.eviltrailmix.com/

Crowston, K. & Williams, M. (2000). Reproduced and emergent genres of communication on the world wide web. *The Information Society, 16*, 201-215.

Cupchik, G. C., & Gebotys, R.J. (1988). The search for meaning in art: Interpretive styles and judgments of quality. *Visual Arts Research*, 14, 38-50.

Dave, K., Lawrence, S., & Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. *Proceedings of the 12th international conference on World Wide Web*, 519-528.

Eysenck, H. J. (1941). Personality factors and preference judgments. *Nature*, 148, 776.

Eysenck, H. J. (1972). Personal preferences, aesthetic sensitivity and personality in trained and untrained subjects. *Journal of Personality*, 40, 544-557.

Frances, R. (1976). Comparative effects of six collative variables on interest and preference in adults of different educational levels. *Journal of Personality and Social Psychology*, 33, 62-79.

Gebauer, J. & Ginsburg, M. (2006). Exploring the black box of task-technology fit: The case of mobile information systems. Retrieved April 28, 2008, from http://www.business.uiuc.edu/Working_Papers/papers/06-0120.pdf

Gebauer, J., Tang, Y., & Baimai, C. (2007). User requirements of mobile technology- results from a content analysis of user reviews. Retrieved April 28, 2008, from http://www.business.uiuc.edu/Working_Papers/papers/07-0107.pdf

Getzels, J. W., & Csikszentmihalyi, M. (1976). *The creative vision: A longitudinal study of problem finding in art*. New York: Wiley-Interscience.

Ghose, A., Ipeirotis, P. G., & Sundararajan, A. (2006). *The dimensions of reputation in electronic markets*. Stern School of Business: New York University.

Gotshalk, D. W. (1962). *Art and the social order*. New York: Dover.

Hagel, J., & Armstrong, A. (1997). *Net gain: expanding markets through virtual communities.* Harvard Business School Press.

Heinrichs, R. W., & Cupchik, G. C. (1985). Individual differences as predictors of preference in visual art. *Journal of Personality*, 53, 502-515.

Hekkert, P., & Wieringen, P. C. W. V., van (1990). Experts' appraisal of cubist paintings. In L. Halasz (Ed.), *Proceedings the XIth International Congress on Aesthetics* (118-121). Budapest, Hungary: Institute for Psychology of the Hungarian Academy of Sciences.

Hekkert, P., & Wieringen, P. C. W. V., van (1996). Beauty in the eye of expert and nonexpert beholders: A study in the appraisal of art. *The American Journal of Psychology*, 109(3), 389-407.

Herring, S., Scheidt, L., Bonus, S., & Wright, E. (2004). Bridging the gap: A genre analysis of weblogs. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, 101-111.

Herring, S., and Scheidt, L. A., and Kouper, I., & Wright, E. (2006). A longitudinal content analysis of weblogs: 2003-2004. *Blogging, Citizenship, and the Future of Media*, 3-20.

Juhasz,J. B., & Paxon, L. (1978). Personality and preference for painting style. *Perceptual and Motor Skills*, 46, 347-349.

Kendall, L. (2007). Colin Mochrie vs. Jesus H. Christ: Messages about masculinities and fame in online video conversations. *40th Annual Hawaii International Conference on System Sciences (HICSS)*.

Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60(3), 269-298.

Lee, T. Y. (2007). Needs-based analysis of online customer reviews. P*roceedings of the ninth international conference on electronic commerce*, 311-318.

Metzger, M. (2005). Understanding how internet users make sense of credibility: A review of the state of our knowledge and recommendations for theory, policy, and practice. *Symposium on Internet Credibility and the User*.

Miller, C. (1994). Genre as social action. *Genre and the new rhetoric.*

Newgrounds. (2008). Newgrounds flash portal. Retrieved April 26, 2008, from http://www.newgrounds.com/portal/

Oliver, M. B., & Kalyanaraman, S. (2002). Appropriate for all viewing audiences? An examination of violent and sexual portrayals in movie previews featured on video rentals. *Journal of Broadcasting & Electronic Media*, 46(2), 283-299.

Paolillo, J. C., Warren, J., & Kunz, K. (2007). Social network and genre emergence in amateur Flash multimedia. *40th Annual Hawaii International Conference on System Sciences (HICSS)*.

Paolillo, J. C., Warren, J., & Kunz, K. (2008). Morphological genre analysis of amateur digital multimedia. In A. Mehler, S. Sharoff, G. Rehm, & M. Santini (Eds.), *Genres on the Web: corpus studies and computational models*. New York: Springer.

Parsons, M.J. (1987). *How we understand art: A cognitive developmental account of aesthetic experience*. Cambridge: Cambridge University Press.

Peel, E. A. (1945). On identifying aesthetic types. *British Journal of Psychology*, 35, 61-69.

Phillips, A. (2008). Stuff about me. Retrieved April 26, 2008, from http://www.brackenwood.net/blog/index.php/about

Reese, S. D., Rutigliano, L., Hyun, K., & Jeong, J. (2007). Mapping the blogosphere: Professional and citizen-based media in the global news arena. *Journalism*, 8(3), 235.

Rosson, M. (1999). I get by with a little help from my cyber-friends: Sharing stories of good and bad times on the Web. *Journal of Computer-Mediated Communication,* 4(4).

Sapolsky, B., & Molitor, F. (1996). Content trends in contemporary horror films. In J. Weaver (Ed.), *Horror Films: Current Research on Audience Preferences and Reactions* (33-48). Lawrence Erlbaum Associates.

Schmidt, J. A., McLaughlin, J. P., & Leighten, P. (1989). Novice strategies for understanding paintings. *Applied Cognitive Psychology*, 3, 65-72.

Silvia, P. J. (2005). Emotional responses to art: From collation and arousal to cognition and emotion. *Review of General Psychology*, 9(4), 342-357.

Singer, J. B. (2005). The political j-blogger: 'Normalizing' a new media form to fit old norms and practices. *Journalism*, 6(2), 173.

Snijders, T., Pattison, P., Robins, G., & Handcock, M. (2006). New specifications for exponential random graph models. *Sociological Methodology,* 36, 99-153.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Thousand Oaks, CA: SAGE Publications.

Tobacyk, J. J., Myers, H., & Bailey, L. (1981). Field-dependence, sensation-seeking and preference for paintings. *Journal of Personality Assessment*, 45, 270-277.

Toms, E. G., & Taves, A. R. (2004). Measuring user perceptions of Web site reputation. *Information Processing and Management*, 40(2), 291-317.

Turney, P. D., & Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 21(4), 315-346.

Valentine, C. W. (1962). *The experimental psychology of beauty*. London: Methuen.

Warren, J., Paolillo, J. C., & Kelley, K. (Forthcoming). Longitudinal network and genre emergence on a large new media Web portal in relation to user demographics.

Wilson, G., Ausman,J., & Mathews, T. (1973). Conservatism and art preference. *Journal of Personalityand Social Psychology*, 25, 286-288.

Yates, J., & Orlikowski, W. (1992). Genres of organizational communication: A structurational approach to studying communication and media. *The Academy of Management Review, 17*, 299-326.