Mind Your Social Media Manners: Pseudonymity, Imaginary Audience, and Incivility on Facebook vs. YouTube

GI WOONG YUN
University of Nevada, Reno, USA

SASHA ALLGAYER
State University of New York at Geneseo, USA

SUNG-YEON PARK
University of Nevada, Reno, USA

A quantitative social media content analysis was conducted to examine the role of anonymity, imagined audience and impression management, and incivility in the context of a controversial Eurovision Song Contest win. User comments to posts featuring the artist's winning performance from Facebook and YouTube were culled and analyzed. Although comments about the singer's performance were predominantly positive on both platforms, non-performance-related comments were far more negative on Facebook than on YouTube. Though valence of the replies targeting other social media users and the Eurovision organization ranged from neutral to negative, as predicted by notions of imagined audience and impression management, YouTube comments were more negative than Facebook comments toward other users. YouTube comments also contained more profanity than Facebook comments. In terms of interacting with other users, YouTube comments more frequently targeted other users and used more profane language in doing so. Overall, the results suggest that anonymity may not necessarily promote negative commenting behaviors; however, the desire for impression management triggered by the nature of imaginary audience could influence social media user interactions.

Keywords: anonymity, Eurovision Song Contest, pseudonymity, imaginary audience, social media, YouTube, perceived reach

With more than 2 billion Facebook users (Hutchinson, 2019) and 2 billion YouTube user activities a month (Iqbal, 2020), social media have become an integral part of social interaction (e.g., Chen, 2017; Gil de Zúñiga, Jung, & Valenzuela, 2012). One factor that was found to influence posting behaviors of social
media users is identity exposure (e.g., Yun & Park, 2011). More specifically, the mechanisms provided to users that allow posting with anonymity or pseudonymity can disinhibit people from the social courtesy of filtering or softening communication messages (e.g., Lee, 2007). Perceived like-mindedness of others in the immediate online communication environment was also found to influence whether people engage in social interactions through posting and, if they do, what to say in the posts (e.g., Lee, 2007; Yun & Park, 2011; Yun, Park, Lee, & Flynn, 2018). Given that users can choose their preferred social media platforms with a wide array of user identification mechanisms, and social media platforms set up their networks differently, the relationships between anonymity and network openness afforded by different social media platforms and the nature of social interactions on the platforms deserve more scrutiny (e.g., Halpern & Gibbs, 2013; Treem & Leonardi, 2013).

To best examine the role of anonymity and relationship among participants of a particular social media discussion in shaping social media manners of people from diverse cultural and national backgrounds, a highly controversial topic that invites emotionally charged user comments was chosen for this study: 2014 Eurovision Song Contest (hereafter referred to as Eurovision) winner, Conchita Wurst. Conchita Wurst is the female embodiment (thus referred to as she/her, the female pronouns throughout this article) created by Thomas Neuwirth, a biological male. Although three other drag queens performed in Eurovision’s history, Conchita Wurst became the first drag queen performer to win the contest, and none of the other acts matched Conchita Wurst in terms of political and cultural controversies, some quite prejudiced and homophobic, surrounding her. For example, a Russian politician had asked the Russian Ministry of Culture to ban Conchita Wurst from performing in Russia (Blistein, 2014). A spokesperson for a Polish political party stated that “Conchita Wurst is a symbol . . . of Europe I don’t want. My Europe is based on Christian values” (Gawęda, 2014, para. 3). After Conchita’s win, the Turkish Foreign Affairs Committee Chairman said, “Gut, dass wir da nicht mehr mitmachen” (It’s good that we don’t participate anymore; Güsten, 2014, para. 3).

Following such polarizing comments, this current research adopted a quantitative content analysis method to examine public discourse about Conchita Wurst on two of the most popular social media platforms—YouTube and Facebook (Perrin & Anderson, 2019)—which differ in terms of anonymity and in relationships among participants. Whereas YouTube provides almost complete pseudonymity and is not built on preexisting offline social networks, Facebook offers lesser anonymity because it was originally built for users’ existing network of friends, such as Harvard University students (Kaplan, 2003). In addition, participation in social interactions on a given Facebook page, with the exception of open pages, typically requires admission to the network. These distinct characteristics of YouTube and Facebook present an opportunity to observe the differences in social interaction based on the level of anonymity and perceived audiences of the social media posts and comments in a setting that closely resembles a field experiment.

---

1 They are Ketil Stokkan, of Norway, in 1986; Sestre, of Slovenia, in 2002; and Verka Serduchka, of Ukraine, in 2007 (Toronidis, 2014). There are also non-gender-binary singers who won the contest: Dana International, a trans singer from Israel, in 1998, and Marija Šerifović, a self-identified lesbian from Serbia, in 2007.

2 See Facebook Help for access details (https://www.facebook.com/help/).
Continuing the tradition of social networks research in the 1960s (see boyd, 2008; Simmel & Wolf, 1964; Wellman & Berkowitz, 1988), social media research has grown exponentially in part owing to the growth in the field of big data (Manovich, 2012) and developments in human network research frameworks (e.g., Barabasi, 2003; Christakis & Fowler, 2009). Readily available social media posts and user comments data also contribute to the burgeoning social media research because data can be extracted from social media sites without interaction with participants (Zimmer, 2010). Indeed, social media offer unique opportunities for studying unfiltered content created by Internet users and some researchers (e.g., Rogers, 2009) argue that the content deserves closer examination, not for its implications for off-line interactions, but for its own significance as a major social phenomenon.

Of precise interest related to the case of Conchita Wurst and this research are the different social media environments with different degrees of identity protection mechanisms and the posted comments based on the premise of perceive reach (Yun et al., 2018). In fact, the environment could modulate the willingness to post homophobic and hateful comments against Conchita and the LGBTQ groups in general, as has been observed in the political and/or traditional media scenes (Yun & Park, 2011). Likewise, certain identity protection mechanisms of social media platforms may facilitate culturally tolerant comments that express diversity and freedom of artistic and self-expression (e.g., Proudfoot, Wilson, Valacich, & Byrd, 2018; Vitak, Blasiola, Patil, & Litt, 2015). After all, members of the wider European (and worldwide) audience have varying attitudes, beliefs, and values regarding gender performance, sexuality, and online behavior. Thus, the aim of this study is to investigate whether platforms with self-identifying information such as Facebook foster politeness in their conversation with their imagined audience compared with more anonymous social media such as YouTube.

To identify comparable user comments about Conchita Wurst on YouTube and Facebook, we examined the official Eurovision Song Contest pages on the two social media platforms. Because both featured the official video of her winning performance, user comments from two identical posts, one on YouTube \( (n = 1,069) \) and the other on Facebook \( (n = 1,050) \), were harvested from the social media pages and subsequently analyzed by four coders.

**Anonymity and Deindividuation in Social Media**

In terms of identity exposure, early Internet users were not particularly concerned about it because of a relatively small number of people online and strong camaraderie among them at the time (Timberg, 2015). With the explosive expansion of the Internet user base, however, rude or otherwise disorderly behaviors came into the focus of communication researchers, and the social identity model of deindividuation effects was first offered as a theoretical framework to explain the emerging undesirable communication behaviors (Lee & Nass, 2002). When applied to the context of social media use, social media deindividuation process, or depersonalization process, facilitates uninhibited communication. With no mandate of revealing true identity, users’ behaviors can substantially change from the measured off-line behavioral patterns to the unruly and uninhibited, revealing the aggressive side of their perspectives or personality (Halpern &

---

3 This article uses the term “uninhibited” instead of “disinhibited” to emphasize the unruly nature of the social media comments (see Reid, 1999, for more discussion on the distinction).
Gibbs, 2013; Lee, 2007). In fact, some users even express their views aggressively using profane and/or bigoted words in their posts or online comments. Behavioral psychologists posit that such uninhibited behavior ranges from insensitivity toward tragic situations (e.g., bystander effects) to participation in violent riots (for more on this topic, see Halpern & Gibbs, 2013; Lee, 2006, 2007).

YouTube represents a pseudonymous environment where users are not obligated to reveal their true identity. They can create an ID not anchored to any aspect of their off-line identity and use it as their pseudonym without restrictions, quite similar to having complete anonymity. YouTube users with Google accounts could link their YouTube activities to their Google IDs (Madden et al., 2013), but they could also stay anonymous on YouTube if they chose. In addition, YouTube users are more likely to be aware of the high heterogeneity of people within the loosely defined network who drop by, watch publicly available videos, and leave comments (Lange, 2007). This perceived anonymity of one’s own identity and unknown others as the audience of their posts and comments could potentially prompt more users to exhibit uninhibited communication behaviors. As a result, YouTube comments may have more aggressive, offensive, hurtful, and damaging words than comments on less anonymous social media platforms.

In contrast, Facebook is regarded as an example of nonanonymous social media (Correa, Silva, Mondal, Benevenuto, & Gummadi, 2015). The social media platform’s identification policy is stricter, and the audience of one’s Facebook posts and comments are more likely to be friends or acquaintances rather than total strangers. The fundamental basis of the Facebook network is “connected friends” who communicate in a mediated environment. This “friend” status allows users to peek at virtually all of their friends’ posts and comments, unless their friends’ individual privacy controls are set otherwise. Even when users decide to use complete pseudonyms on “pages,” a feature of Facebook that is open to all Facebook users, their real-life friends typically know who they are because of the shared network component (Stokes, 2019). The awareness that their comments could be seen by real friends on Facebook even when the friends are not on the Facebook page could lead Facebook users to be inhibited when commenting on open pages (Halpern & Gibbs, 2013). Furthermore, such comments can be pushed to friends who have opted in to the feature, resulting in self-monitoring (see Jung & Rader, 2016).

In sum, social media pseudonymity is less secure on Facebook than on YouTube. Facebook users typically have no perceptions of anonymity on the platform as they often do on YouTube. Indeed, 92% of teen Facebook users use their real name (Madden et al., 2013). Consequently, deindividuation and uninhibited communication behaviors are less likely to happen on Facebook than on YouTube.

**Eurovision Song Contest and the Bearded Lady**

Eurovision began in 1956 as a technological experiment in live television because of its attempt to connect multiple nations through an international network (“In a Nutshell,” 2017). The concept behind the contest was that nations throughout the European continent would send a singer or group to perform on live television, with a winner decided by a professional jury. Audience voting was introduced decades later.

---

4 Since the retirement of Google’s social media service, Google Plus, in 2019 (Wakabayashi, 2018), it became harder to trace user’s identity from Google ID.
The original purpose of the contest was to promote peace and unity in the war-torn continent (Engelhart, 2014), and the stakes of winning include hosting duties the following year.

Though it began as a small music contest that only included participants from Belgium, France, Germany, Italy, the Netherlands, Switzerland, and the United Kingdom, it has since grown into a massive spectacle, with as many as 43 countries taking part and stretching the continental boundaries to include Armenia, Australia, Azerbaijan, Georgia, Israel, Morocco, and Turkey. It has run uninterrupted annually and attracted an estimated 204 million viewers in 2016 (“Eurovision Song Contest,” 2016), with the critical acclaim that it “has become a modern classic, strongly embedded into Europe’s collective mind” (Govaerts, 2014, para. 5).

Winners of the contests typically enjoy fame and recognition throughout the continent and the world. Conchita Wurst, the 2014 winner, however, also accrued an unusually high amount of controversies (Engelhart, 2014). Beyond being a queer performer, she is also deemed, by some critics, to challenge the hegemonic notions of masculinity and femininity by sporting a clearly noticeable facial beard, hence earning the nickname “the Bearded Lady.” Some regarded it as mere personal taste, but others were seriously offended to the extent that Hungarian conservative weekly Heti Válasz featured Conchita on a bull with the title “The Rape of Europa: The Gay Lobby Won Song Contest” (Dunin-Wąsowicz, 2014).

Indeed, her win was highly politicized. On the one hand, some politicians expressed outrage, such as Vladimir Zhirinovsky, of Russia, who proclaimed “no limit to our outrage” and “there are no more men or women in Europe, just it” (Davies, 2014, para. 3). On the other hand, the European Union invited Conchita to become a spokesperson for LGBTQ rights (Riegert, 2014). It is important to note that anti-LGBTQ laws were introduced and passed in Russia shortly before the 2014 Eurovision, and the laws and public discourse about her painted an image “of a backward, homophobic Eastern Europe and a progressive LGBT-friendly Western Europe” (Ulbricht, Sircar, & Slootmaeckers, 2015, p. 156). As such, political discourse about Conchita was mostly constrained to three themes: Conchita as normal, perverse, or normal and/or perverse (Weber, 2016).

When applied to user comments attached to Conchita Wurst’s winning performance video on the official Eurovision Song Contest Facebook page and YouTube channel, the anonymity-based deindividuation would predict more negative comments on YouTube than on Facebook. Because Conchita Wurst’s fame or notoriety was driven not only by her music but also by her gender identity, the valence of the comments was also examined in the two aspects:

\[ H1: \text{YouTube comments will be more negative than Facebook comments in general and specifically regarding (a) Conchita Wurst’s song, performance, or talent, as well as (b) her gender identity.} \]

\section*{Imaginary Audience and Impression Management}

The notion of imaginary audience can be helpful in examining the difference between the perceived audiences of one’s social media comments (i.e., friends’ comments vs. strangers’ comments). Researchers demonstrated that social media users imagine their audience when they make posts or comments (Ranzini
& Hoek, 2017). Imagining the audiences for their posts and comments, they deploy impression-management strategies (Ranzini & Hoek, 2017; Vitak et al., 2015), whether the strategy is content based (what to post/comment on and how to post/comment) or network based (tailoring or limiting access to their posts; see more about this discussion in Ranzini & Hoek, 2017). This role of imaginary audience in impression management on social media has been studied in various contexts, such as projection of personality (Back et al., 2010; Hall & Pennington, 2013), adolescent development (Cingel & Krcmar, 2014), selfie posting (Zheng, Ni, & Luo, 2019), privacy management (Proudfoot et al., 2018; Vitak et al., 2015), and self-monitoring (Turnley & Bolino, 2001).

It is likely that people will deploy impression management strategies to maximize their satisfaction obtained through their use of the social media comment features. In expressing negative emotions, social media users may be particularly strategic: They could strategically target other commenters in their replies, diffuse their anger or frustration by targeting a third party rather than other commenters, or direct the negative emotion to the subject in the original post.

Because of her highly controversial presentation and subsequent politicization by others, it is expected that negativity will be present in the user comments to Conchita Wurst’s video, regardless of the platform. At the same time, the extent of negativity is predicted to be different because the imaginary audience of a user comment on YouTube consists of unidentifiable viewers of the winning performance video, whereas the imaginary audience of a user comment on Facebook consists of the Eurovision Facebook page viewers plus the commenter’s friends. Accordingly, YouTube is expected to have a higher proportion of negative comments to other users on the platform than on Facebook. However, Facebook is expected to feature a higher percentage of negative comments about Eurovision itself than YouTube is because the commenters may be less comfortable offending friends and more comfortable affronting Eurovision or a third party. Though Facebook users are more likely to target the subject in the original post or a third party in negative comments, YouTube users may not mind expressing negative emotions toward other users because their imaginary audience are more likely to be strangers.

In addition, fewer bigoted comments about Conchita’s identity, whether homophobic, misogynistic, or xenophobic, are expected to be observed on Facebook than on YouTube because Facebook users would imagine their friends among potential viewers of their comments and thus self-censor their politically incorrect opinions. Hence, the following hypotheses are generated about impression-management strategies of Facebook and YouTube users:

\[ H2: \quad \text{YouTube comments about Conchita Wurst’s Eurovision win will be (a) more negative toward other users and (b) less negative toward Eurovision than Facebook comments.} \]

Whereas negative comments can be either beneficial or detrimental to the social media discourse depending on their actual substance, using profane language certainly degrades the conversation. Although classical studies of computer-mediated communication projected mostly positive aspects of the anonymous life in cyberspace (e.g., Turkle, 1995), and anonymous online communication still facilitates plenty of positive experiences, one of the major concerns raised about the anonymous nature of online communication is the use of inappropriate language. It is necessary to keep social media free of incivility for them to function
as open spaces of free-flowing ideas (Gil de Zúñiga et al., 2012; Halpern & Gibbs, 2013; Oz, Zheng, & Chen, 2018; Papacharissi, 2004).

Here, the impression-management strategy prompted by the imaginary audience allows us to predict that YouTube facilitates more incivility that Facebook does. Specifically, YouTube users’ comments contain profane language more frequently than Facebook users’ comments because the imaginary audience of Facebook users includes friends whereas the imaginary audience of YouTube users barely involves friends. Though this study recognizes that the concept of social media incivility encompasses various forms of negative comments (Chen, 2017), it focuses on the use of profane language—the rather simple yet unequivocal type of social media incivility. Such an operationalization provides a clarify of measuring the concept when the comments to Conchita Wurst’s winning performance video were written in more than 60 different languages that also connoted various cultural nuances. In addition, this conservative approach probably reduced the chances of Type I error.

**H3:**YouTube comments about Conchita Wurst’s Eurovision win will contain more profane language than Facebook comments.

**Reciprocity of Social Media Comments**

Reciprocity also becomes a key concept in understanding human interactions on social media because the very existence of social media is premised on the assumption that the rights and responsibilities for relationship maintenance are dispersed throughout the network among participants (Lewis, 2015). To date, reciprocity in social media has been discussed in the contexts of technology ethics (Vallor, 2012), online interface design (Pelaprat & Brown, 2012), and journalism (Lewis, 2015). Still, a clear definition of reciprocity in social media is yet to emerge and empirical investigation of reciprocity in social media in general is scarce. In one study, reciprocity of communication was computationally examined by analyzing the direction of messages exchanged among a group of Twitter users, and the authors identified a long list of network features (e.g., the indegree/outdegree ratio, incoming/outgoing message ratio, and others) predicting reciprocity among Twitter users (Cheng, Romero, Meeder, & Kleinberg, 2011).

Because user comments on Facebook are more likely to be exposed to existing relationships, it is expected that they will subsequently trigger more diverse motivations to respond to the comments than user comments on other more anonymous social media platforms. As a result, user comments on Facebook are more likely to receive replies from other users. On the other hand, because YouTube comments are less likely to be exposed to existing relationships, they are thus less likely to trigger a reciprocity motivation. H4 examines a difference between Facebook and YouTube in the level of user comment reciprocity.

**H4:** Facebook comments about Conchita Wurst’s Eurovision win will include replies to other user comments more than YouTube comments.

Whereas reciprocity is often discussed in the context of prosocial behaviors such as kindness and altruism (Gouldner, 1960), there is nothing inherently altruistic about the term itself. It could also be extended to explain and regulate antisocial behaviors in social media. In other words, the traditional concept
of reciprocity would predict that the replies to other user comments would be responsive to others and thus socially desirable. Though replying to other user comments with profane language is being responsive, it is not in a way that would be appreciated by other participants of the online discussion and contribute to a healthy communication environment. The concepts of imaginary audience and impression management suggest that user replies to comments on Facebook would be less likely to contain profane language than user replies to comments on YouTube because Facebook users’ imaginary audience are more likely to include their friends, and thus they may choose their language more carefully on Facebook. The role of impression management in negative reciprocity is examined by H5.

\[ \text{H5: Facebook comments about Conchita Wurst’s Eurovision win will include fewer replies with profane language to other user comments than YouTube comments.} \]

**Method**

The official Eurovision YouTube channel and Facebook page both featured Conchita’s winning performance video and allowed user comments on the video post. The two sites had comparable numbers of subscribers and likes: The YouTube channel had 1,287,833 subscribers and the Facebook page had 1,313,698 likes (likes on Facebook pages are equivalent to subscriptions on YouTube channels). In comparison, the official Eurovision Twitter account had 328,000 followers and the official Instagram account had 231,000 followers at the time of data collection. Given these user statistics, the YouTube channel and Facebook page were determined to be two most popular official social media platforms regarding Eurovision, and thus were chosen as data collection sites for this research.

**Data Gathering and Translating**

On YouTube, Conchita’s performance of the winning song—“Rise Like a Phoenix” (Wurst, 2014)—from the semifinal was selected as the target post from which the user comments were collected because it had more viewers, comments, and likes/dislikes than the performance for the grand finale. The two performances were identical, otherwise. The video was posted on May 8, 2014, the day of the performance. As of November 4, 2016, the date when the comments were harvested, the video had attracted more than 25,000,000 views, more than 126,000 likes, nearly 75,000 dislikes, and nearly 34,000 comments. A total of 1,069 comments were extracted from the YouTube video for analysis by using NVIVO plug-in, Ncapture. The data set contained comments from May 20, 2015, to October 4, 2015.\(^5\)

Facebook comments were collected from the post that declared Conchita as the winner on the official Eurovision Facebook page. It had a cover photo with a clear headshot of Conchita and the winning performance video posted. The page included the words “Congratulations to Austria for Winning the Eurovision Song Contest Copenhagen 2014” written on the left side of the photo. The photo was posted on May 10, 2014 (the day Conchita won), and had received more than 65,000 likes, more than 7,000 likes, and more than 7,000 dislikes.

\(^5\) Instead of random sample selection from the comment universe, we decided to use the fixed time period because of our interest in the reply thread of comments. Randomized selection could have made the research more complicated to execute and less intuitive because we may not be able to follow along with threads.
comments, and nearly 7 thousand shares as of April 22, 2016. The comments were collected for analysis using the Facebook app Netvizz. A total of 1,050 Facebook comments from the similar data collection period were selected for the analysis.

Because of the global attraction to Eurovision, the sample comments were written in 62 different languages or their subvariations. To capture nuances of the comments in many different languages, human coders were employed. Four coders who together understood eight different languages (i.e., Bosnian, Croatian, English, French, Korean, Portuguese, Serbian, and Spanish) were hired to translate some of the comments on their own. Comments in the languages outside of the four coders’ comfort zone were sent to native speakers of those languages to be translated. The unit of analysis was individual comments, and non-English comments were translated to English before they were coded. Translators outside of the four coders did not partake in the coding process itself, but they were asked to generate a direct/literal translation of each comment first and then add any notes about its meaning whenever the meaning of words may have been lost in translation. Translations were cross-checked by another translator for accuracy.

Table 1 shows the language composition of Facebook and YouTube comments and lists the 15 most frequently employed languages in the comments.

<table>
<thead>
<tr>
<th>Language</th>
<th>YouTube</th>
<th>Facebook</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>857</td>
<td>597</td>
<td>1,454</td>
</tr>
<tr>
<td>Greek</td>
<td>2</td>
<td>133</td>
<td>135</td>
</tr>
<tr>
<td>Russian</td>
<td>70</td>
<td>19</td>
<td>89</td>
</tr>
<tr>
<td>German</td>
<td>23</td>
<td>45</td>
<td>68</td>
</tr>
<tr>
<td>Emoticons only</td>
<td>0</td>
<td>47</td>
<td>47</td>
</tr>
<tr>
<td>Polish</td>
<td>37</td>
<td>9</td>
<td>46</td>
</tr>
<tr>
<td>French</td>
<td>20</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>Spanish</td>
<td>10</td>
<td>16</td>
<td>26</td>
</tr>
<tr>
<td>Hungarian</td>
<td>4</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>0</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Portuguese</td>
<td>10</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>Swedish</td>
<td>4</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Turkish</td>
<td>5</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Georgian</td>
<td>1</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Azeri</td>
<td>1</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

We determined that 1,000 comments from each platform would capture a generalizable amount of information regarding the variables. The data collection period was set to capture comparable numbers of YouTube and Facebook comments in the proximity of 1,000 each. Based on these considerations, the comments collected were posted between May 20, 2015, and October 4, 2015.
English was the dominant language, followed by Greek, Russian, and German. Diversity of the Eurovision audience resulted in various combinations of languages appearing in single comments, such as English/French, English/Russian, English/Dutch, and others. In one extreme case, six languages—English, Greek, French, Italian, German, and Japanese—were used in one comment.

**Coding Variables**

A coding scheme was created based on the theoretical framework and expanded after a preliminary reading of the comments. Based on the reading, concepts such as gender identity, comment to Eurovision rather than toward the singer, profane language, and replies to others were operationalized before the actual coding began.

**Valence of Comments About Conchita Wurst's Song/Performance/Talent**

This variable was coded on a 5-point Likert type scale, ranging from 1 (very negative) to 5 (very positive). Each comment was first classified into one of these three categories: negative, neutral, positive. Afterward, positive or negative comments were further assessed for the intensity of the valence.

**Comments About Conchita Wurst's Gender Identity**

Comments were coded as 0 (“not about gender identity”) or 1 (“about gender identity”).

**Valence of Comments About Conchita Wurst's Gender Identity**

This variable was similarly measured in a 5-point scale (1 = very negative; 5 = very positive). Examples of very negative comments included “This is end of Europe. This is all!!! This is man!!! No this is woman!!! No this is IT”; “Smart tactic sending a faggot so all the sick people can vote for this bastard whatever it is!”; “Why didn’t you join the freak show or run with the circus you freak!” These negative comments often used gender pronouns (he/she/it) and/or references to sex organs.

**Valence of Comments About Conchita Wurst’s Other Aspects**

Although the vast majority of the comments were focused on Conchita Wurst’s performance and gender identity, a substantial number of the comments were also addressing other aspects of the singer, such as appearance and nationality. Hence, the comments were also analyzed in terms of how positive or negative they were about Conchita Wurst on the 5-point Likert-type scale, even though no specific hypothesis was generated regarding this variable.

**Overall Valence**

This measure was created by averaging the valence regarding Conchita Wurst’s performance, gender identity, and the other aspects.
Comments Replying to Specific Others

Comments were coded as 0 ("not targeting specific others") or 1 ("targeting specific others"). To be coded as targeting specific others (rather than Conchita Wurst or unspecified other social media users), the comment had to address other user(s) by their username(s) on the social media platform—for example, "XXX (username of a previous comment) is so narrow-minded."

Valence of Comments Replying to Specific Others

Comments targeting specific others were assessed for their valence toward the others on the same 5-point scale.

Comments Replying to Unspecified Others

Comments were coded as 0 ("not targeting unspecified others") or 1 ("targeting unspecified others"). To be coded as targeting unspecified others (rather than Conchita Wurst or specific other social media users), the comment should have not included a specific username and yet referred to a theme(s) of preceding comments (e.g., "It is great to see that many others agree with me on how great Conchita is!").

Valence of Comments Replying to Unspecified Others

Comments targeting unspecified others were assessed for their valence toward the unspecified others on the same 5-point scale.

Comments Targeting Eurovision

Comments were coded as 0 ("not targeting Eurovision") or 1 ("targeting Eurovision"). To be coded as targeting Eurovision (rather than Conchita Wurst or other social media users), the comment had to be specific to the Eurovision Song Contest (e.g., "Now every country needs to send more LGBT people to the contest in order to win.").

Valence of Comments Targeting Eurovision

Comments targeting Eurovision were assessed for their valence toward the international popular music competition on the same 5-point scale.

Comments Containing Profane Language

Comments were coded as 0 ("not containing profane language") or 1 ("containing profane language").

Frequency of Profane Language Use

The number of curse words used in a comment (damn, f*ck, shit, and other foreign-language equivalents) was counted.
Intercoder Reliability Test and Coding

After five conferences, each lasting about one and a half hours, among the four coders, a random sample of 150 comments was coded to test intercoder reliability. All of the variables registered high reliability indicated by the Krippendorf’s alpha scores⁷ (Lombard, Snyder-Duch, & Bracken, 2002): valence of the comments about song/performance/talent = .98; valence of the comments about gender identity = .97; valence of the comments on nationality = .96; valence of the comments about other singers = .98; valence of the comments replying to others = .99; valence of the comments targeting Eurovision = .98; frequency of profane language use = .94.

Because the 150 comments were already analyzed for the intercoder reliability testing, the remaining 1,969 comments were divided among the four coders. Each coder analyzed approximately 500 comments; one half were YouTube comments and the other half were Facebook comments.

Results

H1 hypothesized that the overall valence of the comments would be more negative on YouTube than on Facebook. The result did not support the hypothesis. Rather, the opposite was observed. The overall mean valence of YouTube comments was more positive ($M = 1.69, SD = .98$) than Facebook comments ($M = .56, SD = .34$), and the difference was statistically significant ($p < .001$; see Table 2 for the tabulated results).

H1a predicted that YouTube user comments about Conchita Wurst’s song, performance, or talent would be more negative than Facebook user comments. The result indicated that the valence of YouTube comments ($M = 3.80, SD = 1.60$) about the song, performance, or talent was more positive compared with that of Facebook comments ($M = 3.05, SD = 1.49$). The difference was statistically significant ($p < .001$). The result failed to support H1a, and instead the opposite was the case.

Hypothesis 1b predicted that YouTube comments about Conchita Wurst’s gender identity would be more negative than Facebook comments. However, on average, the valence expressed in the comments was marginally more negative in Facebook comments ($M = 2.28, SD = 1.21$) than in YouTube comments ($M = 2.47, SD = 1.57; p < .10$). The result similarly failed to support H1b, and instead the opposite was the case.

---

⁷ Because the measurements included nominal and ratio variables, the Krippendorff’s alpha method was used (see Lombard et al., 2002).
Table 2. Comparison Between YouTube and Facebook User Comments About Conchita Wurst’s Eurovision Video.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>YouTube</th>
<th>Facebook</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1 Valence: performance</td>
<td>3.80 (1.60)</td>
<td>3.05 (1.49)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Valence: gender identity</td>
<td>2.47 (1.57)</td>
<td>2.28 (1.21)</td>
<td>&lt;.10</td>
</tr>
<tr>
<td>Valence: overall</td>
<td>1.69 (.98)</td>
<td>.56 (.34)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>H2 Valence: reply to others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unspecified</td>
<td>2.29 (1.04)</td>
<td>2.73 (1.05)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Specified</td>
<td>1.10 (.36)</td>
<td>1.70 (.82)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Valence: Eurovision</td>
<td>2.50 (1.62)</td>
<td>2.07 (1.02)</td>
<td>ns</td>
</tr>
<tr>
<td>Number of comments: targeting Eurovision</td>
<td>24</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>H3 Number of profane words (per comment)</td>
<td>1.29 (.68)</td>
<td>1.14 (.38)</td>
<td>&lt;.10</td>
</tr>
<tr>
<td>Number of comments with profanity</td>
<td>118</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>H4 Number of comments: reply to others</td>
<td>378</td>
<td>329</td>
<td></td>
</tr>
<tr>
<td>H5 Number of comments: reply to others with profanity</td>
<td>50</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses. N = 2,118

Hypothesis 2a predicted more negative replies to others’ comments on YouTube than on Facebook. Two measures were used to test this hypothesis: valence of replies to specific other users and valence of replies to unspecified other users. As H2a posited, more negativity toward both specific and unspecified other users was observed on YouTube (specific others M = 1.10, SD = .36; unspecified others M = 2.29, SD = 1.04) than on Facebook (specific others M = 1.70, SD = .82; unspecified others M = 2.73, SD = 1.05). Both differences were highly statistically significant (p < .001). Thus, H2a was supported.

According to H2b, YouTube comments targeting Eurovision were expected to be less negative than Facebook comments. However, there was no statistical difference between the two (YouTube: M = 2.50, SD = 1.62; Facebook: M = 2.07, SD = 1.02). Yet the sheer number of comments about Eurovision was higher on Facebook (n = 89) than on YouTube (n = 24).

Hypothesis 3 compared YouTube and Facebook comments in terms of the use of profane language. As predicted, there were more YouTube comments (n = 118, 11%) containing some profane language than Facebook comments (n = 93, 8.8%). When only the comments that contained some profane language were compared, there was only a marginal difference between YouTube comments (M = 1.29) and Facebook comments (M = 1.14) in the number of profane words contained in them (p < .10). Although the mean difference of profane words used in comments was only marginally significant, the difference in the number of comments containing profane language was large enough (25 = 118 – 93) to support H3.
Hypothesis 4 predicted that Facebook comments would include replies to other comments more frequently than YouTube comments. However, the percentages of reply comments were comparable between YouTube ($n = 378, 35\%$) and Facebook ($n = 329, 31\%$). Hypothesis 4 was not supported.

When referring and replying to other users’ comments, H5 predicted that fewer Facebook reply comments would contain profane language than YouTube reply comments. Indeed, Facebook had 15 replies with profane language, in comparison with 50 YouTube replies containing profane language. The result supported H5.

**Discussion**

Given the role of social media in contemporary society that mixes entertainment and politics, this study produced valuable insights. Particularly, it brought some order to the messy state of social media commenting by applying several theoretical constructs such as anonymity, imaginary audience and impression management, and positive and negative reciprocity.

First, the valence of the comments regarding Conchita’s song, performance, or talent was more positive than neutral on both YouTube and Facebook. Although extremely negative comments were observed on both YouTube and Facebook, they were on the minority side and shadowed by positive comments. This is somewhat surprising given the bad reputation social media have when it comes to norm-violating content. The finding also cautiously projects that the comment filtering approaches taken by *The New York Times* and other media outlets might be a worthy investment. When users are tasked to only make comments that are relevant to the topic at hand, the social media user comment function might be able to facilitate a more focused and civil exchange of ideas.

The importance of encouraging users to stick to the relevant topic in their comments is only accentuated by the noticeable decline in the tone of user comments regarding Conchita Wurst’s gender identity. On both YouTube and Facebook, user comments became more neutral than positive on average. Further decline in the overall valence encompassing all comments—including the comments about Conchita Wurst’ performance, gender identity, appearance, nationality, etc.—only reinforces the importance of enforcing the comment policy.

Contrary to our predictions, between the two, Facebook user comments were more negative than YouTube comments. Two potential interpretations are offered here. First, many social media users simply did not employ face-saving impression-management strategies in their commenting activities even when their highly negative posts could be viewed by everyone in their social network. Second, the negative commenting behavior might have been bolstered rather than deterred by the desire for impression management. If a Facebook user considers that their friends disapprove of Conchita Wurst’s queer identity and wearing a beard, the negative comments about the nonperformance aspects of the singer might be norm conforming rather than norm violating and thus beneficial rather than detrimental to impression management. Although nonanonymous social media are often assumed to be better at keeping deviant behaviors in check, that may not necessarily be the case.
The results concerning whom the social media users targeted in their comments and how uncivil they were in the targeted comments—as indicated by the use of profane language—were in general supportive of the predictions based on the notion of imagined audience and impression management. Whereas YouTube users referred to other people on the platform more than Facebook users did, Facebook users targeted the nonhuman actor, Eurovision, more than YouTube users. Valence of the reply comments or comments targeting Eurovision was from neutral to negative, suggesting that the purpose of the targeted comments was to criticize other users or Eurovision rather than to provide support. Deploying an impression-management strategy, it seems that Facebook users were more comfortable criticizing the organization than criticizing fellow peers on Facebook. On the contrary, YouTube users were more confrontational toward other users than Facebook users were. Similarly, the higher number of YouTube comments containing profane language compared with Facebook comments supported the hypothesis that Facebook’s nonanonymous system would discourage its users from exposing their incivility to their imaginary audiences of friends on the social network.

Overall, hostile exchanges were not common on either platform. Between the two, the level of interaction among users was higher on YouTube than on Facebook. Moreover, YouTube comments replying to others were more likely to include profane language. These results suggest that negative reciprocity might be mediated by the nature of imaginary audience and the desire to save face when the imaginary audience includes friends on a social network.

Other social media platforms that facilitate an identity confirmation process and evoke a high-stake imaginary audience appear to have lower hostile exchanges among their users. For example, LinkedIn users’ identities on the platform are strongly tied to their professional identities and their imaginary audiences are the people whose help they would like to enlist to advance their professional career. Anecdotal evidence (Leetaru, 2019) points that LinkedIn has less profanity in the exchanges of their users. As social media platforms adopt new policies, such as hiding “like” counts on Instagram (Paul, 2019), that could affect user interactions, we will have more data to test this theoretical framework.

As part of the study design, we deliberately selected one of the most controversial performers in contemporary European pop culture—Conchita Wurst—to observe the role of platform affordances in user comments to the original post as well as the interactions among multinational users. As demonstrated by numerous examples (Zittrain, 2017), online discourses have a real impact on our off-line world, and thus opinions online matter. People take online comments seriously and use them to shape their opinions (e.g., Yun & Park, 2011). In the context of LGBTQ rights, online spaces are vulnerable to prejudiced attacks, and if left unaddressed, the incivility could infect more areas of the cyberspace and normalize such behaviors. At the same time, social media that only allow real identities, such as the Chinese national ID system, could also hinder free flow of information and exchange of ideas. More studies in this area are needed to better understand the mechanisms of online discourse. In the end, we need to balance the need to protect free speech and the need to protect people from being harassed or even threatened online.

Limitations and Suggestions

One clear limitation of this study is in the research design. This study analyzed existing user comments on two different social media platforms rather than prospectively assigning people to the two
platforms and examining their commenting behaviors. Because of the distinct affordance differences between YouTube and Facebook, social media users might selectively choose their preferred social media platform at any given point. It is also possible that the two different social media platforms attract different types of people and thus the findings could be attributed to the users rather than to the platforms, as some researchers cautioned about social media comparison research (Oz et al., 2018). If that is the case, there might be confounding factors that better explain the differences found in this study. Further studies using an experimental design will clarify this question.

Nonetheless, this study analyzed a large volume of text data and has high external validity. In the age of big data, computerized sentiment analysis combined with network analysis can yield important insights about many different social media platforms. The categorization framework used in this research can help researchers develop future studies aided by powerful computing and analysis algorithms such as machine learning. In addition, the high external validity we achieved by using two popular social media platforms should not be overlooked. Once joining a platform, users are constrained to its affordances. Because this research focused on behaviors (i.e., commenting), an analysis of the internal validity of the measurements could provide further insights on users’ social media comment manners.

References


In a nutshell. (2017, November 7). Retrieved from https://eurovision.tv/history/in-a-nutshell


