This is a pre-print. The article is published at https://doi.org/10.1108/IntR-02-2017-0072

# Is Offensive Commenting Contagious Online? Examining Public vs. Interpersonal Swearing in Response to Donald Trump's YouTube Campaign Videos

K. Hazel Kwon
khkwon@asu.edu
Walter Cronkite School of Journalism,
Arizona State University
Phoenix, Arizona, USA

Anatoliy Gruzd
gruzd@ryerson.ca
Ted Rogers School of Management,
Ryerson University
Toronto, Ontario, Canada

Abstract

- Purpose: The current study explores the spillover effects of offensive commenting in online community from the lens of emotional and behavioral contagion. Specifically, it examines the contagion of swearing –a linguistic mannerism that conveys high arousal emotion –based upon two mechanisms of contagion: mimicry and social interaction effect.
- Design/methodology/approach: The study performs a series of mixed-effect logistic regressions to investigate the contagious potential of offensive comments collected from YouTube in response to Donald Trump's 2016 presidential campaign videos posted between January and April 2016.
- Findings: The study examines non-random incidences of two types of swearing online: public and interpersonal. Findings suggest that a first-level (a.k.a. parent) comment's public swearing tends to trigger chains of interpersonal swearing in the second-level (a.k.a. child) comments. Meanwhile, among the child-comments, a sequentially preceding comment's swearing is contagious to the following comment only across the same swearing type. Based on the findings, the study concludes that offensive comments are contagious and have impact on shaping the community-wide linguistic norms of online user interactions.
- Originality/value: The study discusses the ways in which an individual's display of offensiveness may influence and shape discursive cultures on the Internet. This study delves into the mechanisms of text-based contagion by differentiating between mimicry effect and social interaction effect. While online emotional contagion research to this date has focused on the difference between positive and negative valence, Internet research that specifically look at the contagious potential of offensive expressions remain sparse.

Keywords: Verbal Aggression, Offensive Comments, Emotional Contagion, Swearing and Profanity, Linguistic Mimicry, YouTube

**Is Offensive Commenting Contagious Online? Examining Public vs. Interpersonal Swearing in Response to Donald Trump's YouTube Campaign Videos**

Social interactions on the Internet have increasingly become emotional. Although emotional expressions may be viewed as matters of 'free speech' in various user interaction contexts, the exchange of blatant verbal aggressions often provoke anger and hostility among discussants (Kramarae and Kramer, 1995). Excessive emotional expressions can be problematic and undesirable because emotion carries power in meaning, and is easily contagious even by a slight inkling of someone else's feelings (Barsade, 2002).

Previous studies have explained offensive commenting on the Internet as an individual behavior driven by a psychological process such as deindividuation and disinhibition, often promoted by user anonymity (Cho and Kwon, 2015; Claessens et al., 2003; McKenna and Bargh, 2000). Less emphasized, however, is the fact that offensiveness can become a community-wide phenomenon through the process of "emotional contagion," defined as "the tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person's and, consequently, to converge emotionally" (Hatfield et al., 1993, p.96). A central mechanism of emotional contagion is "behavioral synchrony", an instantaneous behavioral copying that subsequently leads to emotional convergence (Hatfield et al., 1993, p. 97). An exposure to, and simultaneous mimicking of nonverbal behavioral cues are understood as common precursors for emotional contagion in traditional offline settings.

By contrast, in digitally mediated communication, the presence and immediate copying of a nonverbal signal is often absent because user interactions are predominantly text-based (e.g. discussion boards, microblogging, and online news commenting communities). Accordingly, Internet researchers have recently enquired whether or not emotions are nonetheless contagious in contexts limited to textual interactions. Several studies have shown that emotions can spread via text-based social interactions, most notably by copying linguistic styles (Hancock et al., 2008; Kramer et al., 2014). In other words, synchrony occurs in the form of “language matching” (Gonzales et al., 2010, p.3).

The current study advances the emotional contagion literature by examining the spillover effect of offensive comments in public online communities (i.e., on YouTube). For the purposes of this study, one particular act of emotional expression is investigated: swearing. Swearing is an explicit way to display a high-arousal emotion (Kwon and Cho, 2017). In face-to-face interpersonal interactions, the use of swear words may sometimes contribute to the atmosphere of informality (Cavazza and Guidetti, 2014). However, in online communities where interaction mostly occurs among strangers or in an anonymous public setting, swearing is most likely linked to emotional disinhibition that accompanies highly active negative emotionality such as anger, frustration, and/or hostility (Ivory and Kaestle, 2013; Kwon and Cho, 2017). Based on the assumption that swearing is a linguistic mannerism that conveys anger and verbal aggression to a varied degree, this study investigates whether swearing is contagious through user text-based interactions.

This study attempts to advance the literature in two ways. First, by examining the spillover effect of swearing, the study discusses the ways in which an individual’s display of offensiveness may influence and shape discursive cultures on the Internet. To date, most of online emotional contagion research has focused on the difference between positive and negative valence (Hancock et al., 2008; Kramer et al., 2014), neglecting the lower dimensions of emotionality. Offensive commenting conveys anger, a sub-category of negative emotion that fall in line with recent concerns over the rise of digital incivility. Studies that specifically look at the contagious potential of offensiveness in online contexts remain sparse. Second, a majority of emotional contagion research fail to differentiate between the effect of “simple exposures [to emotional cues]” and the effect of “experiencing an interaction” on the likelihood of contagion (Kramer et al., 2014, p. 8788). This study argues that text-based emotional contagion occurs not only by instantaneous exposure to an emotive linguistic marker but also through comment-based social interactions. Such nuanced effects are highlighted by separating and distinguishing the *exposure to interpersonal swearing* from the *exposure to public swearing.*

The study examines YouTube user comments posted on the official election campaign channel of newly elected President of the United States, Donald Trump. Akin to other social media platforms that are shaped by user comments and expressions (Hassan and Casalo Arino, 2016), YouTube is known for active user participation and content virality created by it (Chiang and Hsiao, 2015; Kahn and Vong, 2014; Oh et al., 2017). Simultaneously however, YouTube is known to contain a nontrivial portion of users’ anger outbursts. Previous studies have revealed concerns over the platform, and have focused on issues of user interactions, trolling and flaming (Halpern and Gibbs, 2013; Moor et al., 2010). Trump’s channel was selected due to the controversy surrounding his candidacy—inducing polemics from supporters and detractors alike at the time this study was conducted on Spring 2016.

## Background

### Online Emotional Contagion

The majority of emotional contagion research (in face-to-face contexts) posit that nonverbal behavioral cues convey greater emotionality than linguistic cues (Hatfield et al., 1993b). In recent times, however, a handful of Internet-based research has shed necessary light on the neglected role of textual messages in signaling emotional states (Berger and Milkman, 2012; Kramer et al., 2014; Stieglitz and Dang-Xuan, 2013). Other studies have suggested that the online public's emotional commentaries are contagious enough to facilitate participatory democracy, often assisting mobilization of sympathizers for social movements (Papacharissi, 2015). However, too much activation of negative valence emotions can deteriorate democracy by inciting biases, polarization and hate speech in online communities (Kramarae and Kramer 1995; Herring et al., 2002). In this regard, the outbursts of emotion that have become increasingly prevalent in today's digital culture are worth greater scholarly attention. This is especially the case when considering that emotions exchanged through text and online messages are contagious.

Studies on the role of emotion in group dynamics and its contagious potential have highlighted two dimensions of emotions. First, studies have examined whether the valence of emotion – positive and negative –produces disproportionate effects on the contagion process. For example, Orford's (1986) ground-breaking study found a negativity bias, highlighting that exposure to negative emotion escalates the chain of negative social interactions. Research on the effects of valence, however, have been mixed as Barsade's study (2002) and Small and Verrochi (2009) found strong evidence of contagion for both positive and negative emotion. In the online context, the mixed results seem even more common. For example, some studies found either no valence difference (Steiglitz and Dong-Xuan, 2013) or a positivity bias in online viral diffusion (Berger and Milkman, 2012; Gruzd et al., 2011; Gruzd, 2013); while in their study of Internet advertising videos containing depression prevention messages, Tseng and Huang (2016) found a direct link between both positive and negative emotion of the narrator and the audiences' intention to adopt health risk-reducing behaviors. Moreover, Lee et al.'s study (2013) showed that message senders' emotional valence (signalled by a profile avatar) had only a moderating effect on the product review based-purchase intention.

Another important dimension of interest is the level of arousal in emotion, also known as "emotional energy" (Barsade, 2002) or "emotional activation" (Berger and Milkman, 2012). Studies have consistently found a positive effect of emotional arousal on the contagion process in online Internet cultures. For example, an analysis of retweeting on the Twitter platform (Stieglitz and Dang-Xuan, 2013) revealed that emotional intensity in tweets was associated with greater retweeting outcomes. Berger and Milkman (2012) also showed that emotional activation has a causal effect on the willingness to share online content.

**Interpersonal vs. Public Swearing**

Among different ways to express emotions, swearing is of particular interest in this paper. Swearing is an act of uttering aggressive languages –or "taboo" words –which is often deterred by "social convention" (Jay, 2009, p. 153). The high arousal of emotion is a defining characteristic of swearing (Jay, 2009; Kwon and Cho, 2017), and thus studying the pragmatics of swearing in the context of online social interactions begs scholarly understanding on the role of aggressive emotional expressions in defining and carving out an ambience of online discussion culture.

This study distinguishes two types of swearing that can occur in an online public setting. First, *interpersonal swearing* refers to a designative use of taboo-words, targeting specific individuals in the process of social interactions. Interpersonal swearing can trigger reciprocal flaming and trolling among anonymous users, as multiple studies have found negative effects of uncivil social interactions online (Alonzo and Aiken, 2004; Cho and Kwon, 2015; Coyne et al., 2001).

The second type of swearing is *public swearing*, distinguished from interpersonal swearing due to no target-specificity. Verbal aggression is not intended to be a direct interpersonal attack. Instead, public swearing functions to accentuate –in an aggressive manner – a speaker's feelings toward an entity, issue, or event beyond the involved discussants. While an immediate interpersonal attack is less obvious, public swearing is nonetheless a form of emotional outbursts, characterized as potentially agonistic and uncivil.

## Two Mechanisms for Swearing Contagion

Swearing as an emotional outburst may be contagious akin to other forms of emotional contagion. Note, however, that swearing in text-based social interactions is *both emotional and behavioral*: it displays activated emotion while it is also an act of verbal aggression. Two theoretical lenses are useful to explain both mechanisms of emotional and behavioral contagion: mimicry and social contagion theory.

**Mimicry.** Most of emotional contagion research is centered on mimicry theory. Mimicry is an interpersonal synchronization of emotion through imitating emotional cues of others (Chartrand and van Baaren, 2009). While mimicry can occur in both conscious and unconscious manners, most emotional contagion research has highlighted the automatic, unconscious imitation as a key precursor of contagion (Hatfield et al., 1993b). Nonverbal mimicry is an imitation of gestures, postures, and facial motions (Lakin et al., 2003). The majority of mimicry studies have been conducted in offline settings and focus on kinetics and facial expressions, and find that mimicry of nonverbal movements transcends the emotional states between communication partners. A recent study examined the mimicry via voice-to-voice communication (Rueff-Lopes et al., 2015). While testing mimicry in the context of voice communication is novel, the emphasis on nonverbal cues such as voice pitch and tones remains consistent with previous mimicry research.

Online text-based interactions do not accompany physical signals that are prevalent in offline settings, or even vocal signals inherent in voice-to-voice communication. Nonetheless, it is possible for users to mimic other users' writing mannerism and linguistic styles (Gonzales et al., 2010). For example, communication accommodation theory suggests that the convergence of conversation styles is frequently observed in interpersonal relations, which helps reduce social distance between communicators and facilitate social approval within the conversation community (Giles and Coupland, 1991). Welbers and de Nooy (2014) tested this theory using Internet forums, and found evidence of linguistic convergence among discussants. Studies have used the linguistic style matching technique (LSM) to examine the textual mimicry via digital social networks (Gonzales et al., 2010; Niederhoffer and Pennebaker, 2002; Welbers and de Nooy, 2014).

While swear words are one of the widely used linguistic cues for emotional expressions in online discussions, part of reason swearing contagion has not been examined from the lens of mimicry theory could be due to its anti-normative functionality. Most mimicry studies to date have focused on the prosocial functions of mimicry (i.e., imitation occurring as an instinctive attempt to blend into the immediate social context) and communication convergence (i.e., imitating others' communication style and mannerism reinforces social identity and facilitates a sense of cohesion and rapport) (Chartrand and Baaren, 2009). Other goals and motives that could drive mimicking behaviors such as competition or antagonism remain understudied. Although swearing can occur in an effort to blend oneself into a group that he or she identifies with (Lee, 2007), it may also occur purely to antagonize or compete with other discussants. Indeed, the mimicry of swearwords can be explained through motives of confrontation as opposed to social blending.

**Social contagion theory.** Whereas mimicry theory focuses on the instantaneous convergence of emotional signals, social contagion theory offers insights on the effects of social interaction on behavioral contagion. Social contagion literature explains social connections as the conduits of beliefs, attitudes, information, and behaviors. For example, Fowler and Christakis (2008) propose the three-degrees-of-separation rule of social contagion: contagion occurs not only through the direct contacts but also through indirect connectivity up to three degrees of separation (e.g., a friend of ‘a friend of my friend’ may affect my happiness, propensity to be obese, etc.).

Nevertheless, most of the robust findings from web-based studies have focused on the first degree of separation, that is, the influence of the directly connected others. For example, Suri and Watts (2011) conducted web experiments to understand contagion of cooperative behaviors, finding that the donating behavior of a directly connected neighbor positively influenced the focal actor’s decision to donate. However, no clear evidence was found regarding multi-degree contagion. Other studies have similarly highlighted the direct exposure effect in online networks. For example, Kwon et al. (2014) showed that the exposure to online friends’ behaviors influence the likelihood of focal actor’s engaging in similar behaviors on Facebook. Large-scale online field experiments on Facebook also suggest that exposure to the decisions of online friends influence an individual’s ad-clicking behavior (Bakshy et al., 2012), and voting intention (Bond et al., 2012).

While the aforementioned studies are mainly interested in the direct exposure effect, Tsvetkova and Macy's recent study (2014) is noteworthy in that they focus on the effects of more complex social interactions on behavioral contagion. Specifically, they (2014) investigated different types of social interactions, including direct reciprocity (i.e., A helps B, then B helps A), generalized reciprocity (i.e., A helps B, then B helps C), and vicarious experience (A helps B, and C observed this interaction and helps D), concluding that different social interaction mechanisms influence different dynamics of behavioral contagion. Tsvetkova and Macy (2015) also tested these social interaction effects on antisocial behavior contagion. While dissecting specific patterns of social interaction is beyond the scope of this study, the aforementioned research substantiates the need to differentiate between mimicry and social interaction effects in order to better understand the contagion of offensive comments.

## Hypotheses

Both the mimicry effects studied in emotional contagion literature and the direct exposure effects discussed in social contagion literature point to the same rule for contagion: contagion occurs through imitation, after "exposure" to certain information. At the same time, reciprocal interactions and other higher-order network effects described by social contagion literature emphasize the importance of social interactions for an individual's behavioral choice: contagion occurs by adopting others' behaviors after "experiencing" social interactions (Kramer et al., 2014). The distinction between the simple exposure and social interaction effect allows for hypothesizing different mechanisms relevant to public and interpersonal swearing on YouTube.

### Public Swearing as Exposure Effect

Public swearing has no specific targeted attack, and thus does not anticipate any reciprocal social interactions. Therefore, if a user reads someone else's public swearing, it is most likely to be a simple exposure to the expressed emotion. That is, the contagion effect of public swearing, if observed, may be understood as an outcome of exposure and subsequent verbal mimicry.

This logic allows for two different hypotheses for understanding causes of public swearing contagion in online discussion context. First, an online discussion thread, in particular on YouTube, always has a first-level comment (a.k.a., “parent” comment). The discussion thread begins when sub-comments, or the second-level comments are posted under the parent-comment (a.k.a., “child” comments). This nested structure infers that a child-comment is made after exposure to a parent-comment. Therefore, if a parent-comment has public swearing, a child-comment should be exposed to it, and then mimic the swearing behavior if a contagion occurs.

H1: Public swearing of a parent-comment increases the likelihood of a child-comment’s public swearing.

Second, if the discussion thread becomes long enough, the default setting of discussion threads on YouTube will make only the parent-comment and a couple of the most recent child-comments visible. The rest of the child-comments will be hidden unless a user clicks the option that shows all the replies. This hidden structure makes it likely that a user will be exposed to not only the parent-comment but also to the immediate prior in the sequence of child-comments.

In other words, the preceding child-comment’s public swearing could also have an exposure effect, such that the following child-comment mimics the practice of public swearing.

H2: Public swearing of a preceding child-comment increases the likelihood of the following child-comment’s public swearing.

Figure 1 exemplifies the structure of YouTube discussion thread, and public and interpersonal swearing.

[Figure 1 Here]

### Interpersonal Swearing as Social Interaction Effect

Contrary to public swearing, interpersonal swearing attacks a specific user and anticipates a negative reaction from the targeted user or others within the community. Different interaction patterns may be conceived to induce interpersonal swearing contagion, for example direct reciprocity (A swears to B, and B responds to A by swearing back), collective attack (A swears to B, and C joins A by swearing to B as well), and chain swearing (A swears to B, and B swears to C). While the underlying motivation associated with each of these interaction patterns may indeed differ, a shared commonality is that swearing becomes spiral, through *sequences of social interactions*. It is highly unlikely that a parent-comment will initiate interpersonal swearing in online discussions where discussants hardly know each other, hence we hypothesize the contagion effect of interpersonal swearing only in terms of the child-comment effect.

H3: Interpersonal swearing of a preceding child-comment increases the likelihood of the following child-comment's interpersonal swearing.

Furthermore, it is possible that interpersonal swearing could create a culture of generalized swearing. That is, swearing may become normative behavior whereby the attacked user, or the user who observed others' interpersonal swearing may in turn engage in outburst swearing towards not only a specific person but also an unspecified audience. Such community-wide swearing, if any, may suggest the potential for swearing to diffuse as an epidemic practice among online participants.

H4: Interpersonal swearing of a preceding child-comment increases the likelihood of the following child-comment's public swearing.

## Research Design

### Data Collection

YouTube was chosen as an empirical site, wherein the frequent presence of profanity makes its comment data ideal for conducting reliable statistical modeling of swearing contagion. The publicly accessible comments data were collected from 38 videos posted to the official channel of Donald Trump ("Donald J. Trump for President") between January 18, 2016 and April 29, 2016, using the API tool developed by Digital Methods Initiative at the University of Amsterdam. Among the initial 38 videos, three videos blocked user commenting, resulting in null data. In sum, the dataset included the total of 23,925 comments from 35 videos. Among them, 13,852 comments constituted 2,075 discussion threads, each of which contained one parent-comment and at least one child-comment. While the unit of analysis was the child-comments (*N*=11,777), the analysis plan accounted for the multilevel structure (each child-comment nested under a parent-comment, which in turn is nested under its corresponding video). Also collected are the metadata associated with each video (e.g., when it was uploaded, the number of likes and dislikes, the date and time when each comment was posted, and the total reply counts for parent-comments). There were a few of non-English comments, mostly in Spanish. These comments were automatically translated into English using Google Translate and Google Spreadsheet.

**Swearing Dictionary**

To automatically detect swearing occurrences, this study developed a dictionary of swear words. The dictionary was developed based on the two primary sources: (a) public lists of English swear words shared freely on websites such as www.noswearing.com; and (b) a custom-built dictionary of swear words and abbreviations (e.g., smfh, stfu, wtf, wth) derived from the manual reviews of over 60,000 Twitter messages, developed as a part of one of the authors' ongoing project. The inter-coder reliability of the Twitter-derived swear words achieved 92.04% agreement, with kappa alpha = .87.

After combining swear words from both sources, the research team manually reviewed the resulting list and removed any ambiguous words to avoid false positives such as ‘killer’, ‘gay’, etc. In total, the dictionary consisted of 437 words (including derived forms) (see Appendix). The resulting dictionary was used to compute the occurrences of swear words in each comment.

**Variables**

**Swearing in parent-comment.** Public swearing was operationalized as an occurrence of swear word without any call-out of specific user name. Interpersonal swearing was defined as the occurrence of swear words along with the call-out of specific user name in the same message. The call-out of a specific user was expressed in the forms of either a direct response to the target user (i.e., by starting a comment with ‘+*username*’) or a hyperlink to the target user’s profile.

As expected, none of the parent-comments included a specific interpersonal marker, and thus all swearing comments were considered to be public swearing. The total number of swear words was counted within each parent comment, assuming that the more swear words the higher activation of emotion. Presented below are exemplary comments with varied number of swear words included (original texts).

*“You <u>fucking</u> dictator! <u>Fuck</u> you! You don’t know what it’s like to live without a house and without freedom <u>motherfucker</u>! make America great again? Brainwashing people into voting for you! This is the new <u>fucking</u> Adolfo hitler <u>motherfuckers</u>!”* (5 swear words)

*“At least Hillary doesn’t discriminate people like that nazi <u>fuck</u> Trump. You see how your boy Trump made fun of a disabled reporter a while back some guy. He hates women as well but*

*your too blind to see that. I hope you enjoy voting for that cold hearted celebrity as our president''* (1 swear word)

**Swearing in the preceding comment.** First, in line with the parent-comments, the total number of swear words in each child-comment was counted to be added as a predictor for modeling purposes. Second, a categorical variable – 'types of swearing' –was created, with 0 = no swearing, 1 = interpersonal swearing, 2 = public swearing. Public and interpersonal swearing of a child-comment were defined in the same manner to parent-comments. That is, a comment is public swearing if it has a swear word without an interpersonal marker; a comment is interpersonal swearing if the occurrence of swear words accompanies the call-out of specific user name in the same message. Then, the child-comment that appears right before a focal child-comment in the chronologically ordered thread was defined to be the preceding comment of the focal child.

**Dependent variable**. Dependent variables pertain to the types of a focal child-comment. Specifically, three binary dependent variables are concerned: (a) an occurrence of any swearing, (b) an occurrence of public swearing, and (c) an occurrence of interpersonal swearing in the focal child-comment.

**Comment-level control variables.** Four factors were considered as comment-level control variables. (1) It is possible that an occurrence of swear words be a byproduct of the length of message. Accordingly, the total words used in a focal child-comment was counted to measure the message length effect. (2) The temporal effect was controlled by addressing time lag between the time of video upload and of the focal comment's posting time. (3) Popularity of a thread could influence the way in which child-comments interact with one another. Popularity of a thread was measured by the total number of replies, that is child-comments. (4) Given that swearing is an emotional expression, the exposure to different types of emotional markers could affect the likelihood of swearing. Therefore, the number of uppercased words in the parent- and preceding child-comment were controlled, assuming that uppercased words could convey some activation of emotion. Mindful of abbreviations of media and other organizational names (uppercased names like NBC, CNN, FBI) only the words with at least four consecutive uppercases were counted in the sample.

**Video-level control variables.** Video characteristics may also affect the likelihood of swearing. Two factors were considered. (1) If most people dislike a video, its comments may include frequent swearing revealing an overall dissatisfaction or disagreement with the video, although swearing in some cases can also be a form of agreement. To account for disliking of a video, the proportion of dislike votes out of the sum of likes and dislikes was taken into consideration. (2) Similarly, comments in response to polarizing videos may contain frequent swearing. The polarizing tendency of a video was represented by Simpson's diversity index (D) of like and dislike votes, with "0" indicating no polarization at all, and "0.5" indicating the complete split between likes and dislikes (Eveland and Hively, 2009)[1].

## Results

### Manipulation Check

To confirm whether swearing is an exemplar of offensive linguistic markers, two coders evaluated the level of verbal aggression and anger in a randomly selected sample of 500 comments. The modified Buss and Perry's items (1992) were used to create a codebook comprised of 6 anger and 6 verbal aggression items (5-point Likert scale).[2] Researchers computed composite scores of anger and verbal aggression for each coder, then performed reliability analysis based on the intra-class correlation coefficients (ICC). The verbal aggression scale resulted in the ICC of .72 (single measure) and .84 (average measure, equivalent to Cronbach's alpha); the anger scale resulted in the ICC of .62 and .74.

The anger and verbal aggression scores were averaged between the two coders. *T*-tests were used to examine the difference between swearing comments and non-swearing comments. The results showed that verbal aggression was significantly greater in swearing comments ($M = 2.35$) than non-swearing comments ($M=1.64$), $t = 10.82$, $p < .001$. Likewise, anger was significantly higher in swearing comments ($M = 2.54$) than in non-swearing ($M = 1.67$), $t = 16.67$, $p < .001$ (Figure 2).

[Figure 2 Here]

**Multilevel Logistic Regression**

On average, a child-comment was about 34 words long, and was posted about 14 days after the initial video upload. About one fourth (25.2%) of child-comments contained swearing to some extent, mostly interpersonal swearing (17.8%). Among the preceding comments, 10.5% included public swearing, with 15.8% considered interpersonal swearing. On average, both the preceding child-comments and parent-comments had 0.4 swear words per message; On average, the proportion of dislike votes out of the total votes made to a video was 54.18%, and the average Simpson's D score was 0.38, indicating some level of polarization. Table 1 summarizes descriptive statistics.

[Table 1 Here]

**Baseline model.** The data structure was hierarchical: child-comments nested in a parent-comment, and parent-comments nested in a video. Accordingly, mixed effect modeling was employed, specifically multilevel logistic regressions[3], to take the video-level and parent-comment level random effects into account. The Likelihood Ratio (LR) tests confirmed that the random effects were significant, suggesting non-independence due to the hierarchical data structure (Table 2).

[Table 2 Here]

For a baseline model, the research team examined whether swearing in a parent-comment and a preceding child-comment increased the chance of the focal child-comment's swearing (whether interpersonal or public). The model results suggested that, when the whole population was considered, swearing comments had the odds of 0.13 times lower than the non-swearing comments. That is, non-swearing comments were 7.69 times higher to occur than swearing comments.

As expected, message length had a significant effect on swearing occurrences ($b$ = .008, odds ratio = 1.008, $z$ =16.00, $p$ < .001). While a one-unit increase effect was small, note that the unit of length being each word. For example, the odds of swearing in a 40-word long comment were 32% greater than the odds of swearing in a 10-word long comment. Also, the popularity of a thread, measured by the total number of replies, also increased the likelihood of the focal child-comment's swearing ($b$ = .006, odds ratio = 1.006, $z$ =4.67, $p$ < .001). For example, a child-comment nested in a thread replied by 100 child-comments showed 57% higher chance of swearing than the one nested in a thread with only five child-comments. Posting time also showed a significant effect, albeit weak ($b$ = .003, odds ratio = 1.003, $z$ =2.382, $p$ < .05). For example, a comment posted a month later had a 9% greater chance of containing swear words than a comment on the day of video upload.

As seen in the baseline model, swearing in both a parent- and preceding child-comment increased the likelihood of the following child-comment's swearing. The odds of focal child-comment's swearing increased by 15.6% for a one swear word contained in a parent-comment;

increased by 31.2% for two swear words contained in a parent-comment; increased by 46.8% for three swear words, and so on ($b$ = .145, odds ratio = 1.156, $z$ =4.913, $p$ < .001). In the same vein, the odds of focal child-comment's swearing increased by 9.8% for a one swear word contained in a preceding child-comment; increased by 19.6% for two swear words in a preceding child-comment; increased by 29.4% for three swear words, and so on ($b$ = .094, odds ratio = 1.098, $z$ =4.089, $p$ < .001).

**Public vs. interpersonal swearing models**. To address the hypotheses, additional models were designed by (1) separating two outcome variables (focal child-comment's public and interpersonal swearing) and by (2) adding "swearing type" of the preceding comment as another categorical predictor (public =1, interpersonal swearing =2). As seen in the baseline model, the mixed-effect modeling resulted in significant random effects, indicating non-independence due to the nested data structure (Table 3).

[Table 3 Here]

Results suggested as follows. First, the number of swear words in a parent-comment, showed a positive effect on the likelihood of public swearing of a child-comment ($b$ = .111, odds ratio = 1.117, $z$ =3.296, $p$ < .001). While this result confirmed H1, the effect of parent-comments, all of which were public swearing, were equally significant in terms of the likelihood of interpersonal swearing of a child-comment ($b$ = .111, odds ratio = 1.118, $z$ =3.379, $p$ < .001). These significant results confirmed that a parent-comment's *public* swearing increased *both public and interpersonal swearing* of a child-comment.

Second, when the preceding comment's swearing type was taken into account, the number of swear words in the preceding comment was no longer significant. Instead, the results indicated that the contagion effect of the preceding comment was valid only for the *same kind* of swearing. Specifically, *public* swearing of the preceding comment increased by 61.8% of the likelihood of the focal comment's *public* swearing ($b$ = .481, odds ratio = 1.618, $z$ =3.928, $p$ < .001); whereas *interpersonal* swearing of the preceding comment increased by 22.8% of the likelihood of the *interpersonal* swearing of the focal comment ($b$ = .206, odds ratio = 1.228, $z$ =2.383, $p$ < .05). In other words, H2 and H3 were confirmed, but not H4.

Third, control variables showed somewhat different effects between public and interpersonal swearing of focal child-comments. Although video polarization levels did not affect the likelihood of public swearing, the dislike proportion showed a significant effect: the odds of public swearing increased by 0.9% per one-percent increase in the video's dislikes proportion ($b$ = .009, odds ratio = 1.009, $z$ =2.306, $p$ < .05). For example, a video with 50% dislikes proportion would show a 1.36 times higher chance of having a comment with public swearing than a video with 10% dislike proportion.

On the other hand, none of video-level variables affected the likelihood of interpersonal swearing. Instead, interpersonal swearing was influenced by the posting time. For example, a comment posted a month later would have 15% higher chances of interpersonal swearing than a comment posted on the day of video upload ($b$ = .005, odds ratio = 1.005, $z$ =2.903, $p$ < .05). Interestingly, a thread's popularity influenced the chance of interpersonal swearing in child-comments, with a 1% increase of swearing per reply added to the thread ($b$ = .01, odds ratio = 1.01, $z$ =6.186, $p$ < .001). These temporal and thread popularity effects suggest that interpersonal swearing could indeed be a product of social interactions.

Message length effect was significant for both public and interpersonal swearing, however in an opposite direction to each other. That is, the longer the message the *more* likely interpersonal swearing ($b$ = .011, odds ratio = 1.011, $z$ =19.363, $p$ < .001). Conversely, the longer the message was, the *less* likely public swearing was included ($b$ = - .007, odds ratio = .991, $z$ = -5.552, $p$ < .01). This result is possibly due to the fact that interpersonal swearing often occurs in a contextualized social interaction, whereas public swearing is more instantaneous and shorter than interpersonal swearing, and thus lacks contextual information.

Figure 3 presents a visualization of the predicted probability of public and interpersonal swearing of child-comments. The graphs show that, in general, interpersonal swearing has higher predicted probability, and the contagion effect increases by the intensity of swearing in a parent-comment. Interpersonal swearing, however, does not show much difference across the types of preceding child-comments. Surprisingly, results indicate the effects of preceding comment's interpersonal swearing on the focal child-comment's interpersonal swearing to be quite small. On the contrary, the swearing types of preceding comments show disproportionate effects on public swearing occurrences in child-comments. *Public* swearing of the preceding child-comment has a fairly high contagion effect on *public* swearing of the focal child-comment.

[Figure 3]

## Discussion and Conclusion

Aggressive emotional exchanges have become increasingly common in contemporary digital culture. When the Internet's culture of self-expression meets with polemical topics like controversial political issues / politicians, belligerent commentaries that threaten mutual respect seem to be, unfortunately, one of the byproducts. It is especially concerning if an individual's offensive comment creates chain reactions such that it affects and transforms the implicit norms that surround community-wide discussions.

In line with recent attention to text-based contagion of emotions, this study demonstrated the ways in which offensive emotional displays become contagious in textual online social interactions on YouTube. This study examined swearing as an explicit speech act that provokes anger and verbal aggression. The function of swearing as a high-arousal emotional marker may be especially prominent in text-based interactions where other nonverbal cues are largely absent.

This study was based on two theories of emotional and behavioral contagion: mimicry and social contagion theory. Mimicry theory suggests that being exposed to an emotional cue is a sufficient trigger for an imitative pattern to emerge. Based on this logic, the study proposed public swearing contagion be the "exposure" mechanism for contagion of offensive comments. This study used social contagion literatures (social interaction dynamics in behavioral adoption) to examine interpersonal swearing as the "social interaction" mechanism for contagion. Moreover, two sources of contagion were identified, 1) a parent-comment and 2) a sequentially preceding child-comment. The results are in line with previous research on online emotional contagion and thus add one more evidence of negative emotional contagion (Kramer et al., 2014).

One interesting finding is that, despite each swearing thread initiated with a parent's *public* swearing, the parent's public swearing was prone to catalyzing chains of *interpersonal* swearing as well as reiterating public swearing. The predicted probability for a focal child-comment's interpersonal swearing was indeed greater than that for public swearing. This result suggests that simple exposure to another's aggressive speech online has a spillover effect such that subsequent users may adopt the swearing as a linguistic style and reuse it in a dyadic social interaction setting.

Another more convincing explanation of this phenomenon could be sought out from Balance theory (Cartwright and Harary, 1956). The most straightforward rule of balanced triadic network is "my friend's enemies are my enemies." Hence, while public swearing may not attack a specific message recipient per se, it can target third-party individuals, events, or objects with which the recipient maintains a strong affinity with. In this case, emotional aggression toward the third-party could hurt the recipient user's social identity, who may in turn reciprocate his or her hurt feeling by attacking the initial commenter. For example, if public swearing occurred against Trump, a supporter for Trump might feel offended and obliged to swear back by targeting the initial commenter. If Balance theory is the mechanism of public swearing contagion, public swearing should be understood within the complexity of social network dynamics. While the current study cannot address whether or not social network dynamics intervene in the process of 'public-to-interpersonal' swearing spillover, further research in this area is recommended.

Another interesting finding is that swearing contagion from a preceding child-comment was effective only for the *same* kind of swearing. Meaning, public swearing in a preceding comment was contagious only for the public swearing of the focal comment, and interpersonal swearing was contagious only for the interpersonal swearing. These findings are consistent with the proposed hypotheses, highlighting that different contagion mechanisms are in effect. Specifically, public swearing could spread through instant convergence of linguistic styles, whereas interpersonal swearing could be a product of more contextualized social interactions. The contagion effect of public-to-public swearing (in terms of preceding comments) was especially large, falling in line with previous research on linguistic convergence (e.g., Gonzales et al., 2009; Niederhoffer and Pennbaker, 2002; Welbers and de Nooy, 2014).

Although one should be cautious about equating contagion of swearing *behaviors* to actual *emotional* convergence, the results of this study demonstrate that swearing comments do indeed contain higher anger than non-swearing comments. An individual act of swearing may propagate from one comment to another comment, echoing some of the existing concerns about negative chain reactions of incivility in online discussions (Moor et al., 2010). Swearing is a verbal marker of highly activated emotionality as well as a speech habit, the spread of which could potentially shape hostile discussion environments online. Interestingly, the majority of emotional contagion literature has predominantly focused on prosocial and harmonious function of mimicry (Chatrand and Van Baaren, 2009), paying little attention to different goals and motives linked to competition or enmity. The gap between the existing theory and the phenomena of online swearing contagion and other hostile emotional and behavioral contagion calls for further theoretical elaboration.

One limitation of this study is that the analysis could not delve into the effects of different social interaction patterns. Examination of different interaction patterns such as direct reciprocity, collective swearing, and swearing chains could have enriched understandings of the underlying motivations that induce contagion of offensive comments. Also, the current findings are based on one particular political campaign (Donald Trump) on a particular social media platform (YouTube). The nonsignificant effects of video attributes could be due to this rather narrow topic selection. Future work would benefit from a comparative element, whereby the results between different political candidates or across different social media platforms are contrasted.

From a practical perspective, the findings of this study suggest an important role of initial comments in setting the tone for the subsequent online discussions. When there is a need to moderate an online community for the sake of maintaining respectful discussions and promotion of civility, it is recommended that community managers to pay special attention to the parent-posts and implement intervention efforts during the initial phase of discussions as needed.

## Notes

[1] Simpson's D = 1 - $\sum Pi^2$, where $Pi$ is the proportion of like and dislike votes.

[2] Verbal Aggression: The commenter tells readers openly that he or she disagree with someone/ disagrees with others/ is annoyed by others and telling them what he or she thinks of them/ cannot help getting into argument/ is argumentative/ is verbally attacking someone. Anger: The commenter flares up quickly/ is frustrated and lets his or her irritation show/ is an even-tempered (inverse)/ is a hothead/ is angry/ has trouble controlling his or her temper.

[3]. The program R was used. See "R data analysis examples: Mixed effects logistic regression," UCLA: Statistical Consulting Group (http://www.ats.ucla.edu/stat/r/dae/melogit.htm).

**Acknowledgement**

Added upon acceptance

**References**

Alonzo, M. and Aiken, M. (2004), "Flaming in electronic communication", *Decision Support Systems*, Vol. 36 No. 3, pp. 205–213.

Bakshy, E., Eckles, D., Yan, R. and Rosenn, I. (2012), "Social influence in social advertising: Evidence from field experiments", *Proceedings of the 13th ACM Conference on Electronic Commerce*, ACM, New York, NY, USA, pp. 146–161.

Barsade, S.G. (2002), "The ripple effect: Emotional contagion and its influence on group behavior", *Administrative Science Quarterly*, Vol. 47 No. 4, pp. 644–675.

Berger, J. and Milkman, K.L. (2012), "What makes online content viral?", *Journal of Marketing Research*, Vol. 49 No. 2, pp. 192–205.

Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D.I., Marlow, C., Settle, J.E. and Fowler, J.H. (2012), "A 61-million-person experiment in social influence and political mobilization", *Nature*, Vol. 489 No. 7415, pp. 295–298.

Buss, A.H. and Perry, M. (1992), “The aggression questionnaire”, *Journal of Personality and Social Psychology*, Vol. 63 No. 3, pp. 452–459.

Cartwright, D. and Harary, F. (1956), “Structural balance: a generalization of Heider’s theory”, *Psychological Review*, Vol. 63 No. 5, pp. 277–293.

Cavazza, N. and Guidetti, M. (2014), “Swearing in political discourse: Why vulgarity works”, *Journal of Language and Social Psychology*, Vol. 33 No. 5, pp. 537–547.

Chartrand, T.L. and van Baaren, R. (2009), “Human mimicry”, *Advances in Experimental Social Psychology*, Vo. 41, pp. 219-274.

Chiang, H.-S. and Hsiao, K.-H. (2015), “YouTube stickiness: the needs, personal, and environmental perspective”, *Internet Research,* Vol. 25 No. 1, pp.85-106.

Cho, D. and Kwon, K.H. (2015), “The impacts of identity verification and disclosure of social cues on flaming in online user comments”, *Computers in Human Behavior*, Vol. 51, Part A, pp. 363–372.

Claessens, J., Díaz, C., Goemans, C., Dumortier, J., Preneel,B., and Vandewalle, J. (2003), "Revocable anonymous access to the Internet?", *Internet Research,* Vol. 13 No. 4, pp.242-258.

Coyne, S.M., Stockdale, L.A., Nelson, D.A. and Fraser, A. (2011), “Profanity in media associated with attitudes and behavior regarding profanity use and aggression”, *Pediatrics*, Vol. 128 No. 5, pp. 867–872.

Eveland, W.P. and Hively, M.H. (2009), “Political discussion frequency, network size, and ‘heterogeneity’ of discussion as predictors of political knowledge and participation”, *Journal of Communication*, Vol. 59 No. 2, pp. 205–224.

Fowler, J.H. and Christakis, N.A. (2008), “Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham heart study”, *BMJ*, Vol. 337, p. a2338.

Giles, H. and Coupland, N. (1991), *Language: Contexts and Consequences*, Vol. xvi, Thomson Brooks/Cole Publishing Co, Belmont, CA, US.

Gonzales, A.L., Hancock, J.T. and Pennebaker, J.W. (2010), “Language style matching as a predictor of social dynamics in small groups”, *Communication Research*, Vol. 37 No.1, pp. 3–19.

Gruzd, A. (2013), “Emotions in the Twitterverse and implications for user interface design”, *AIS Transactions on Human-Computer Interaction*, Vol. 5 No. 1, pp. 42–56.

Gruzd, A., Doiron, S. and Mai, P. (2011), “Is happiness contagious online? A case of Twitter and the 2010 winter olympics”, *2011 44th Hawaii International Conference on System Sciences*, presented at the 2011 44th Hawaii International Conference on System Sciences, pp. 1–9.

Halpern, D. and Gibbs, J. (2013), “Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression”, *Computers in Human Behavior*, Vol. 29 No. 3, pp. 1159–1168.

Hassan, M., and Casaló Ariño, L.V. (2016), “Consumer devotion to a different height: How consumers are defending the brand within Facebook brand communities”, *Internet Research*, Vol. 26 No. 4, pp.963-981.

Hancock, J.T., Gee, K., Ciaccio, K. and Lin, J.M.-H. (2008), “I’m sad you’re sad: Emotional contagion in CMC”, *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, ACM, New York, NY, USA, pp. 295–298.

Hatfield, E., Cacioppo, J.T. and Rapson, R.L. (1993), “Emotional contagion”, *Current Directions in Psychological Science*, Vol. 2 No. 3, pp. 96–100.

Herring, S., Job-Sluder, K., Scheckler, R. and Barab, S. (2002), “Searching for safety online: Managing ‘trolling’ in a feminist forum”, *The Information Society*, Vol. 18 No. 5, pp. 371–384.

Ivory, A.H. and Kaestle, C.E. (2013), “The effects of profanity in violent video games on players’ hostile expectations, aggressive thoughts and feelings, and other responses”, *Journal of Broadcasting & Electronic Media*, Vol. 57 No. 2, pp. 224–241.

Jay, T. (2009), “The utility and ubiquity of taboo words”, *Perspectives on Psychological Science*, Vol. 4 No. 2, pp. 153–161.

Khan, G.F. and Vong, S. (2014), “Virality over YouTube: an empirical analysis”, *Internet Research*, Vol. 24 No. 5, pp.629-647.

Kramarae, C. and Kramer, J. (1995), “Legal snarls for women in cyberspace”, *Internet Research,* Vol. 5 No. 2, pp.14-24.

Kramer, A.D.I., Guillory, J.E. and Hancock, J.T. (2014), “Experimental evidence of massive-scale emotional contagion through social networks”, *Proceedings of the National Academy of Sciences*, Vol. 111 No. 24, pp. 8788–8790.

Kwon, K.H. and Cho, D. (2017), “Swearing effects on citizen-to-citizen commenting online: A large-scale exploration of political versus nonpolitical online news sites”, *Social Science Computer Review*, Vol. 35 No. 1, pp. 84–102.

Kwon, K.H., Stefanone, M.A. and Barnett, G.A. (2014), “Social network influence on online behavioral choices: Exploring group formation on social network sites”, *American Behavioral Scientist*, Vol. 58 No. 10, pp. 1345–1360.

Lakin, J.L., Jefferis, V.E., Cheng, C.M. and Chartrand, T.L. (2003), "The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry", *Journal of Nonverbal Behavior*, Vol. 27 No. 3, pp. 145–162.

Lee, E.-J. (2007), "Deindividuation effects on group polarization in computer-mediated communication: The role of group identification, public-self-awareness, and perceived argument quality", *Journal of Communication*, Vol. 57 No. 2, pp. 385–403.

Lee, M., Kim, M. and Peng, W. (2013), "Consumer reviews: reviewer avatar facial expression and review valence", *Internet Research*, Vol. 23 No. 2, pp. 116–132.

McKenna, K.Y.A. and Bargh, J.A. (2000), "Plan 9 from cyberspace: The implications of the Internet for personality and social psychology", *Personality and Social Psychology Review*, Vol. 4 No. 1, pp. 57–75.

Moor, P.J., Heuvelman, A. and Verleur, R. (2010), "Flaming on YouTube", *Computers in Human Behavior*, Vol. 26 No. 6, pp. 1536–1546.

Niederhoffer, K.G. and Pennebaker, J.W. (2002), "Linguistic style matching in social interaction", *Journal of Language and Social Psychology*, Vol. 21 No. 4, pp. 337–360.

Oh, S., Baek, H., and Ahn, J. (2017), "Predictive value of video-sharing behavior: Sharing of movie trailers and box-office revenue", *Internet Research*, Vol. 27 No. 3, doi: 10.1108/IntR-01-2016-0005

Papacharissi, Z. (2015), *Affective Publics: Sentiment, Technology, and Politics*, Oxford University Press.

Rueff-Lopes, R., Navarro, J., Caetano, A. and Silva, A.J. (2015), "A markov chain analysis of emotional exchange in voice-to-voice communication: Testing for the mimicry hypothesis of emotional contagion”, *Human Communication Research*, Vol. 41 No. 3, pp. 412–434.

Small, D.A. and Verrochi, N.M. (2009), “The face of need: Facial emotion expression on charity advertisements”, *Journal of Marketing Research*, Vol. 46 No. 6, pp. 777–787.

Suri, S. and Watts, D.J. (2011), “Cooperation and contagion in web-based, networked public goods experiments”, *PLOS ONE*, Vol. 6 No. 3, p. e16836.

Tseng, C.-H. and Huang, T.-L. (2016), “Internet advertising video facilitating health communication: Narrative and emotional perspectives”, *Internet Research*, Vol. 26 No. 1, pp. 236–264.

Tsvetkova, M. and Macy, M.W. (2014), “The social contagion of generosity”, *PLOS ONE*, Vol. 9 No. 2, p. e87275.

Tsvetkova, M. and Macy, M.W. (2015), “The social contagion of antisocial behavior”, *Sociological Science*, Vol. 2, pp. 36–49.

Welbers, K. and de Nooy, W. (2014), “Stylistic accommodation on an Internet forum as bonding: Do posters adapt to the style of their peers?”, *American Behavioral Scientist*, Vol. 58 No. 10, pp. 1361–1375.

Table 1. Descriptive Statistics (*N* = 11,777)

| | *M* | *sd* | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. % of Dislikes (Video) | 54.18 | 23.81 | $-.26^{*}$ | $.33^{**}$ | $.06^{**}$ | $.09^{**}$ | $.05^{**}$ | $.05^{**}$ | $.03^{*}$ | .00 | $.05^{**}$ | $.06^{**}$ | .01 | .0 | $.14^{**}$ |
| 2. Polarization (Video) | .38 | .07 | | $-.08^{**}$ | $-.08^{**}$ | $-.04^{**}$ | $-.02^{*}$ | .00 | $-.03^{*}$ | -.01 | $-.03^{*}$ | -.01 | $-.02^{+}$ | $.02^{+}$ | $-.20^{**}$ |
| 3. Thread Popularity | 34.13 | 42.64 | | | $-.03^{**}$ | $.05^{**}$ | $.05^{**}$ | $-.07^{**}$ | $.13^{**}$ | -.01 | $.08^{**}$ | .00 | $.09^{**}$ | $.08^{**}$ | $-.05^{*}$ |
| 4. Parent SWC | .40 | .95 | | | | $.16^{**}$ | $.26^{**}$ | $.23^{**}$ | $.04^{**}$ | $.05^{**}$ | $.08^{**}$ | $.06^{**}$ | $.05^{**}$ | $.03^{*}$ | $.04^{*}$ |
| 5. Parent Upper | .95 | 3.29 | | | | | $.08^{**}$ | $.03^{*}$ | $.02^{*}$ | $.21^{**}$ | $.03^{**}$ | $.03^{*}$ | $.02^{+}$ | $.02^{+}$ | $.10^{*}$ |
| 6. Preceding SWC | .43 | 1.00 | | | | | | $.40^{**}$ | $.53^{**}$ | $.13^{**}$ | $.11^{**}$ | $.04^{**}$ | $.09^{**}$ | $.04^{**}$ | $.05^{*}$ |
| 7. Preceding PSW | .11 | .31 | | | | | | | $-.15^{**}$ | $.04^{**}$ | $.04^{**}$ | $.08^{**}$ | .00 | $-.05^{**}$ | -.02 |
| 8. Preceding ISW | .16 | .36 | | | | | | | | $.06^{**}$ | $.10^{**}$ | -.01 | $.12^{*}$ | $.09^{**}$ | $.05^{*}$ |
| 9. Preceding Upper | .49 | 2.83 | | | | | | | | | .01 | .00 | .00 | $.02^{+}$ | $.03^{*}$ |
| 10. DV: ASW | .25 | .43 | | | | | | | | | | $.49^{**}$ | $.80^{**}$ | $.17^{**}$ | $.05^{*}$ |
| 11. DV: PSW | .07 | .26 | | | | | | | | | | | $-.13^{**}$ | $-.05^{**}$ | .00 |
| 12. DV: ISW | .18 | .38 | | | | | | | | | | | | $.23^{**}$ | $.05^{*}$ |
| 13. Message Length | 34.04 | 52.34 | | | | | | | | | | | | | .00 |
| 14. Time Lag | 14.33 | 21.44 | | | | | | | | | | | | | |

Note. $^{**}p < .001$, $^{*}p < .01$, $^{+}p < .05$; Upper = uppercased words, SWC = Swearing count, PSW=Public swearing; ISW = Interpersonal swearing; ASW = Any type of swearing; DV = Dependent variable.

Table 2. Baseline Model: Contagion Effects on a Child-Comment Swearing (*N* = 11,777)

| | Coefficient | | | 95% C.I. | | |
|---|---|---|---|---|---|---|
| | Est | SE | Odds Ratio | LL | UL | z-value |
| % of Dislikes (Video) | .003 | .003 | 1.003 | .998 | 1.009 | 1.176 |
| Polarization (Video) | .008 | .762 | 1.008 | .240 | 4.234 | .010 |
| Thread Popularity ** | .006 | .001 | 1.006 | 1.004 | 1.009 | 4.670 |
| Parent SWC ** | .145 | .030 | 1.156 | 1.001 | 1.335 | 4.913 |
| Parent Upper | -.001 | .010 | .999 | .980 | 1.018 | -.110 |
| Preceding SWC ** | .094 | .023 | 1.098 | .989 | 1.220 | 4.089 |
| Preceding Upper | -.012 | .010 | .988 | .970 | 1.006 | -1.214 |
| Time lag + | .003 | .001 | 1.003 | 1.001 | 1.006 | 2.383 |
| Message length ** | .008 | .001 | 1.008 | 1.007 | 1.009 | 16.000 |
| (Intercept) | -2.043 | .31 | .13 | .072 | .233 | -6.588 |
| Random effect (intercept) | | | | | | |
| Video-level | .234 | .061 | | | | |
| Thread-level | .622 | .048 | | | | |
| LR test $\chi^2(2) = 182.67$, $p < .001$ | | | | | | |
| Log-likelihood =6296.003, Wald $\chi^2(9)=347.75$, $p < .001$ | | | | | | |

Note. $^{**}p<.001$, $^{+}p<.05$; Generalized linear mixed model fit by maximum likelihood (Laplace approximation); Upper = uppercased words, SWC = Swearing count, PSW=Public swearing; ISW = Interpersonal swearing.

Table 3. Comparison of Public and Interpersonal Swearing Contagion Effect ($N$ = 11,777)

| | Focal Comment's Public Swearing | | | | | | Focal Comment's Interpersonal Swearing | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est | SE | Odds Ratio | 95% C.I. LL | 95% C.I. UL | z-value | Est | SE | Odds Ratio | 95% C.I. LL | 95% C.I. UL | z-value |
| (Intercept) | -2.763$^{**}$ | .443 | .063 | .026 | .150 | -6.244 | -2.680$^{**}$ | .321 | .069 | .037 | .129 | -8.353 |
| % of Dislikes (Video) | .009$^{+}$ | .004 | 1.009 | 1.001 | 1.016 | 2.306 | -.001 | .003 | .999 | .993 | 1.004 | -.488 |
| Polarization (Video) | -.695 | 1.019 | .499 | .068 | 3.676 | -.682 | .456 | .765 | 1.578 | .352 | 7.064 | .596 |
| Thread Popularity | -.001 | .001 | .999 | .996 | 1.001 | -1.071 | .010$^{**}$ | .002 | 1.010 | 1.007 | 1.013 | 6.186 |
| Parent SWC | .111$^{**}$ | .034 | 1.117 | 1.046 | 1.193 | 3.296 | .111$^{**}$ | .033 | 1.118 | 1.048 | 1.193 | 3.379 |
| Parent Upper | .012 | .011 | 1.012 | .992 | 1.034 | 1.163 | -.001 | .011 | .999 | .977 | 1.021 | -.122 |
| Preceding SWC | .005 | .045 | 1.005 | .920 | 1.098 | .113 | .045 | .032 | 1.046 | .983 | 1.113 | 1.426 |
| Preceding ISW | .025 | .126 | 1.026 | .801 | 1.313 | .201 | .206$^{+}$ | .086 | 1.228 | 1.037 | 1.454 | 2.383 |
| Preceding PSW | .481$^{**}$ | .122 | 1.618 | 1.273 | 2.057 | 3.928 | .096 | .101 | 1.101 | .904 | 1.341 | .953 |
| Preceding Upper | -.005 | .015 | .995 | .967 | 1.024 | -.324 | -.017 | .012 | .983 | .961 | 1.007 | -1.412 |
| Time lag | .001 | .002 | 1.001 | .996 | 1.005 | .325 | .005$^{*}$ | .002 | 1.005 | 1.002 | 1.008 | 2.903 |
| Message length | -.007$^{*}$ | .001 | .993 | .991 | .996 | -5.552 | .011$^{**}$ | .001 | 1.011 | 1.010 | 1.012 | 19.363 |
| Random effect (intercept) | | | | | | | | | | | | |
| Video | .284 | .078 | | | | | .231 | .068 | | | | |
| Parent comment | .312 | .088 | | | | | .646 | .551 | | | | |
| | LR test: $\chi^2(2) = 28.39^{**}$ | | | | | | LR test: $\chi^2(2) = 151.76^{**}$ | | | | | |
| | Log-likelihood = 3010.001, Wald $\chi^2(11)=93.27^{**}$ | | | | | | Log-likelihood = 5073.524, Wald $\chi^2(11)=476.76^{**}$ | | | | | |

Note. $^{**}p$ <.001, $^{*}p$ <.01, $^{+}p$ <.05; Upper = uppercased words, SWC = Swearing count, PSW=Public swearing; ISW = Interpersonal swearing.

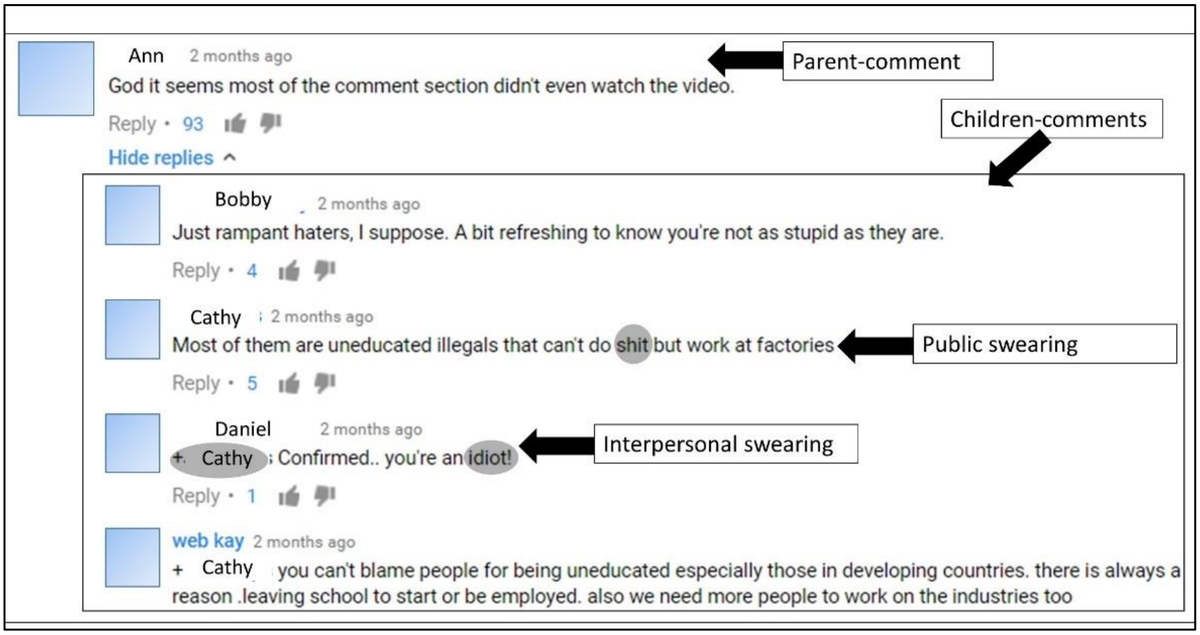


Figure 1. Parent-child comment structure and examples of public and interpersonal swearing (Names are aliases and photos were hidden for privacy).

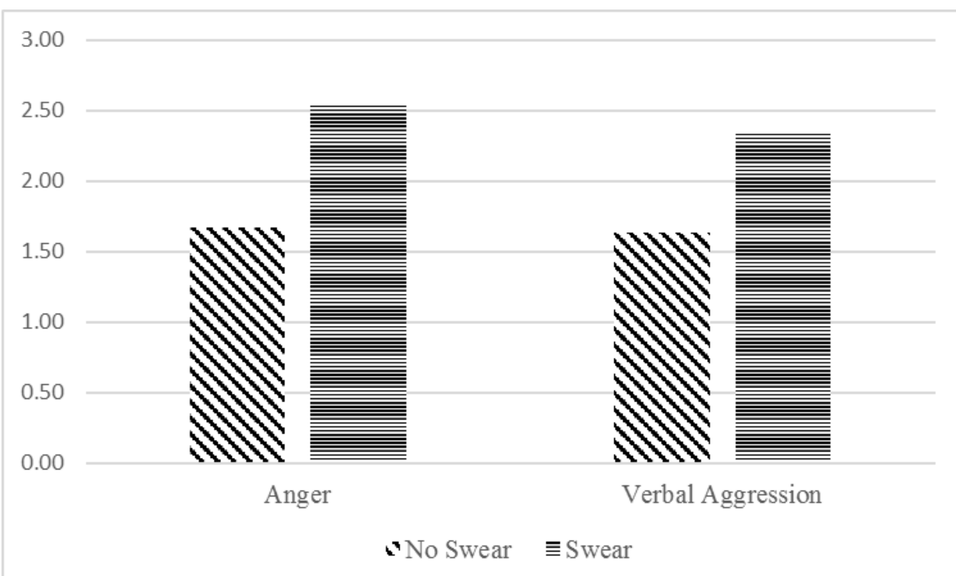


Figure 1. Difference in anger and verbal aggression between non-swearing and swearing comments on YouTube.

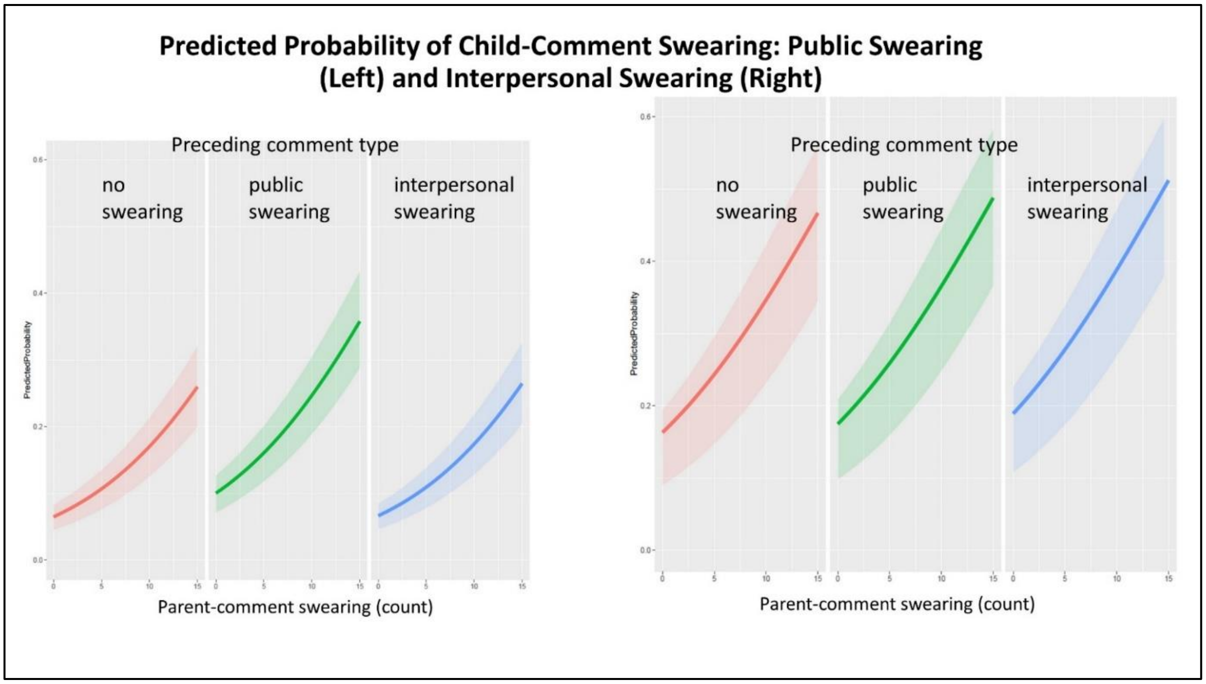


Figure 2. Predicted probability of a focal child comment's public and interpersonal swearing (X-axis is the number of swear words in a parent comment; Y-axis is the predicted probability of a focal child-comment; Each color represents the type of swearing of the preceding comment.

Appendix 1. The list of swear words (N = 437 words, including repetition between online search and Twitter)

| From Online Search | | | | | | From Twitter | |
|---|---|---|---|---|---|---|---|
| anus | chinc | dickbag | fuckhead | lesbo | shitbag | a-hole | motherfucker |
| arse | chink | dickbeaters | fuckhole | mcfagget | shitbagger | ass | motherfuckers |
| arsehole | choad | dickface | fuckin | mick | shitbrains | asshole | nigga |
| ass | chode | dickfuck | fucking | minge | shitbreath | assholes | niggas |
| assbag | clit | dickfucker | fucknut | mothafucka | shitcanned | bastard | nut |
| assbandit | clitface | dickhead | fucknutt | mothafuckin | shitcunt | bastards | nuts |
| assbanger | clitfuck | dickhole | fuckoff | motherfucker | shitdick | bitch | nutter |
| assbite | clusterfuck | dickjuice | fucks | motherfucking | shitface | bonehead | nutters |
| assclown | cock | dickmilk | fuckstick | muff | shitfaced | boo | pfff |
| asscock | cockass | dickmonger | fucktard | muffdiver | shithead | bullshit | pimp |
| asscracker | cockbite | dicks | fucktart | munging | shithole | bumped-up | pimping |
| asses | cockburger | dickslap | fuckup | negro | shithouse | butt | piss |
| assface | cockface | dicksucker | fuckwad | nigaboo | shitspitter | buttheads | pothead |
| assfuck | cockfucker | dicksucking | fuckwit | nigga | shitstain | cocksucker | prick |
| assfucker | cockhead | dicktickler | fuckwitt | nigger | shitter | coward | pricks |
| assgoblin | cockjockey | dickwad | fudgepacker | niggers | shittiest | cowardice | psycho |
| asshat | cockknoker | dickweasel | gayass | niglet | shitting | cowards | psychopath |
| asshead | cockmaster | dickweed | gaybob | nut sack | shitty | crap | psychopaths |
| asshole | cockmongler | dickwod | gaydo | nutsack | shiz | crapostan | psychos |
| asshopper | cockmongruel | dike | gayfuck | paki | shiznit | craze | pussies |
| assjacker | cockmonkey | dildo | gayfuckist | panooch | skank | craziness | pussy |
| asslick | cockmuncher | dipshit | gaylord | pecker | skeet | crazy | rat |
| asslicker | cocknose | doochbag | gaytard | peckerhead | skullfuck | creeps | scum |
| assmonkey | cocknugget | dookie | gaywad | penis | slut | cunt | scumbag |
| assmunch | cockshit | douche | goddamn | penisbanger | slutbag | cunts | shit |
| assmuncher | cocksmith | douche | goddamnit | penisfucker | smeg | damn | shits |
| assnigger | cocksmoke | douchebag | gooch | penispuffer | snatch | damnit | shitty |
| asspirate | cocksmoker | douchewaffle | gook | pissflaps | spic | damning | silly |
| assshit | cocksniffer | dumass | gringo | polesmoker | spick | darn | sleuths |
| assshole | cocksucker | dumb ass | guido | pollock | splooge | demon | smfh |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| asssucker | cockwaffle | dumbass | handjob | poon | spook | devil | stupid |
| asswad | coochie | dumbfuck | heeb | poonani | suckass | dick | stupidstan |
| asswipe | coochy | dumbshit | hell | poonany | tard | dipshits | sucker |
| axwound | coon | dumshit | ho | poontang | thundercunt | douchebag | thugs |
| bampot | cooter | dyke | hoe | porch monkey | tit | dumb | wierdo |
| bastard | cracker | fag | homo | porchmonkey | titfuck | dumbass | witch |
| beaner | cum | fagbag | homodumbshit | prick | tits | dumbasses | wtfu |
| bitch | cumbubble | fagfucker | honkey | punanny | tittyfuck | dumbest | wtf |
| bitchass | cumdumpster | faggit | humping | punta | twat | evil | wth |
| bitches | cumguzzler | faggot | jackass | pussies | twatlips | fag | |
| bitchtits | cumjockey | faggotcock | jagoff | pussy | twats | fool | |
| bitchy | cumslut | fagtard | jap | pussylicking | twatwaffle | fools | |
| blow job | cumtart | fatass | jerk off | puto | unclefucker | frak | |
| blowjob | cunnie | fellatio | jerkass | queef | va-j-j | freaking | |
| bollocks | cunnilingus | feltch | jigaboo | queer | vag | iffrig | |
| bollox | cunt | flamer | jizz | queerbait | vagina | libtard | |
| boner | cuntass | fuck | jungle bunny | queerhole | vajayjay | liar | |
| brotherfucker | cuntface | fuckass | junglebunny | renob | vjayjay | liars | |
| bullshit | cunthole | fuckbag | kike | rimjob | wank | loser | |
| bumblefuck | cuntlicker | fuckboy | kooch | ruski | wankjob | losers | |
| butt plug | cuntrag | fuckbrain | kootch | sand nigger | wetback | lunatic | |
| butt | cuntslut | fuckbutt | kraut | sandnigger | whore | lunatics | |
| buttfucka | dago | fuckbutter | kunt | schlong | whorebag | maniac | |
| buttfucker | damn | fucked | kyke | scrote | whoreface | maniacs | |
| camel toe | deggo | fucker | lameass | shit | wop | mofo | |
| carpetmuncher | dick | fuckersucker | lardass | shitass | | monsters | |
| chesticle | dick | fuckface | | | | mother-fucker | |