How Background Images Impact Online Incivility

JINKYUNG PARK, Rutgers University, USA
VIVEK K SINGH, Rutgers University, USA and Massachusetts Institute of Technology, USA

Despite the potential of online spaces for the democratic development of public discourse, concerns over aggressive and uncivil interactions in those spaces are rising. Online incivility is the term to define the features of a discussion that convey an unnecessarily disrespectful tone toward the participants or the topic. Concerning the adverse impact of online incivility on users, we sought to explore ways to reduce online incivility to promote secure and trustworthy online debate culture. Meanwhile, the web is becoming increasingly multimodal and images are known to be an effective way of improving emotions. The broaden-and-build theory of positive emotions suggests that positive emotions generated by positive images can impact incivility levels. Hence, with the goal to reduce online incivility, we conducted a one factorial between-subject online experiment with three conditions (N = 105). We compared the three conditions (with a positive background image in color, a positive background image in grayscale, and no background image) to identify an efficient way of reducing online incivility. The data gathered from surveys and participants' online comments were qualitatively and quantitatively analyzed to answer the research questions. The results showed that not only positive backgrounds in color but also positive backgrounds in grayscale may be effective in reducing online incivility. The results will pave way for designing more civil discussion platforms in online settings.

CCS Concepts: • **Security and privacy** → *Social aspects of security and privacy*; • **Human-centered computing** → **Empirical studies in HCI**; *User studies.*

Additional Key Words and Phrases: Online Incivility, Design Intervention, Positive Emotions, Online Experiment

ACM Reference Format:
Jinkyung Park and Vivek K Singh. 2022. How Background Images Impact Online Incivility. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 444 (November 2022), 23 pages. https://doi.org/10.1145/3555545

1 INTRODUCTION

The growth of digital media has offered a chance for the public to share their opinions on a variety of topics from everyday concerns to political viewpoints. Scholars have suggested that online discussions provide a promise to the democratic development of debate culture [4, 56, 60]. Anonymity and limited social context cues of the online environment allow participants to focus on ideas and opinions by lessening their fear of repercussions for expressing their opinion [24]. At the same time, concerns have been raised about online discussions due to the frequent presence of uncivil behaviors [11, 24, 61]. As per a recent study, 80% of respondents reported having experienced uncivil behavior at one time or another; 63% pointed to the Internet and social media as the cause of online incivility [65]. Online incivility is concerning as it may weaken political trust [51], polarize

Authors' addresses: Jinkyung Park, jinkyung.park@rutgers.edu, Rutgers University, 4 Huntington Street, New Brunswick, New Jersey, USA, 80901; Vivek K Singh, v.singh@rutgers.edu, Rutgers University, 4 Huntington Street, New Brunswick, New Jersey, USA, 80901 and Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, USA, 02139.

2573-0142/2022/11-ART444 $15.00
https://doi.org/10.1145/3555545

views [1, 33], and radicalize individuals [44]. Furthermore, in some cases, online incivility may lead to offline violence such as physical harassment, hate crime, and threats [8, 65].

Concerns over uncivil behavior are exemplified in the user comments on discussion forums that allow a high level of anonymity [60]. As a response to uncivil comments, many online news outlets began requiring users to register or verify some form of identity which connects to their real names or geolocation before commenting on their sites [31]. Some news outlets disallowed comments or abandoned the forums altogether on controversial topics. Some of the news topics – crime, religion, immigration, disaster, celebrity, and social issues – which tend to elicit a large number of uncivil comments are more likely to appear without comment forums than other topics [62]. However, disabling the comment forum faces the paradox that an effort in reducing incivility may silence the public, prevent public discourse, and challenge democracy.

In this study, we focus on the ways to reduce online incivility while preserving the anonymity of the online discussion. More specifically, we implement a computer-based design intervention to reduce uncivil comments in anonymous online discussion forums. Design interventions apply design concepts to provoke real-world action and intervene in human behavior, routine, and experiences [69]. As computers are integrated into individuals' everyday lives, computing systems have become more persuasive by design. Computers, mobile phones, and other systems with interfaces are now used to persuade users in changing attitudes and behaviors through contextual information, advice, and guidance [46]. Prior work has suggested that computer images can be part of a persuasive process that affects individuals' decision-making [40]. Recently, images are increasingly used to affect various aspects of people's lives such as lifestyle advertising [64], learning [10], and fake news propaganda [68]. It is undeniable that visuality has come into prominence by affecting individuals' perceptions and beliefs.

One of the ways to apply visuality on the web is the use of background images. The main background images of the website are known to have a positive visual impact on how the website is evaluated by the users [14, 78]. In the early days of the web, background images were applied on the web but slow internet speeds adversely impacted the user experience. However, with faster internet connections, there is a growing trend toward adopting background images for enhancing user experience [55, 78]. Social media platforms are now using images extensively as part of the digital environment to expand visual communications. For example, Google and Twitter recently added options that allow online users to add and customize background images [32, 45]. Moreover, as the web is increasingly becoming multimodal (e.g., metaverse) [54], these options are now seen as "intriguing options" [32] which add significantly to social media's offerings to improve engagement. Multiple online discussion forums (e.g., reddit.com/r/Appalachia/) are already using background images to support user experience on the discussion forums.

However, the impact of positive background images in online forums on the civility of the ensuing discussion is understudied. This study aims to systematically explore this question using a one factorial between-subject online experiment with three conditions (N = 105). We compared the three conditions (with positive background image in color, positive background image in grayscale, and no background image) to identify an efficient way of reducing online incivility. The results showed that not only positive background images in color but also positive background images in grayscale may be effective at fostering less out-group aversion, more humane discussions, and reducing online incivility. The results of the current study will pave way for leveraging design decisions to increase positive emotions and civil discourse in online settings.

2 RELATED WORK AND RESEARCH QUESTIONS

2.1 Online incivility

The definitions of incivility vary widely because incivility is a multifaceted construct that can be described by various online behaviors such as flaming or trolling [71]; more importantly, this is because incivility is in the eyes of the beholder [27]. In a broad sense, incivility can be operationalized as the set of behaviors that threaten democracy, deny people their freedoms, and stereotype social groups [56]. This definition builds on the concept of "face" from the politeness literature [6], a social construct to describe individuals' attempts to put forward a desired image of self in interaction. According to Papacharissi [56], civility is a positive collective face; deference to the social and democratic identity of an individual. On the contrary, incivility can be a negative collective face; disrespect for the collective traditions of democracy.

A more widely accepted definition is the features of discussion that convey an unnecessarily disrespectful tone toward the discussion forum, its participants, or the topic [11]. Focusing on incivility in online newspaper comments, this definition encompassed components of incivility that are likely to be present in online forums. In the current study, we adopted a definition suggested by [11] because the key components of the definitions (e.g., incivility in online discussion forums) are consistent with those employed in the current study.

Previous literature has shown that online incivility may have an adverse influence on online users' cognitions [1, 26, 30], emotions [4, 24, 44], and behaviors [24, 59, 81]. Accordingly, various studies have tried to identify strategies to reduce uncivil comments. One example is to apply social modeling, the idea that the level of incivility can be reduced by being exposed to civil conversation. For example, [25, 26] showed that participants who were exposed to civil cues tended to be more civil in their responses, stay on-topic, and offer additional perspectives in their comments. In a similar vein, [50] found that civil commenting behavior was modeled, yet uncivil behavior was not. At the same time, many social media platforms have implemented moderation strategies to discourage uncivil comments. These strategies include moderating comments through voting where users up-vote or down-vote each submission, reminding community rules to users, and allowing users to report or block abusive accounts so that posts of those accounts do not appear on the user's timeline or newsfeed [9, 37, 48].

The root of uncivil online behavior has been addressed in a few psychological studies and the computer-mediated communication (CMC) domain. One of the key reasons for online incivility is the anonymity of the Internet [7]. Specifically, anonymity has long been studied as the antecedent to yield deindividuation, described by reduced self-regulation and enhanced social identity. The construct of "deindividuation" describes as a state in which inner restraints are lost when individuals act in groups and do not see themselves as individuals [19]. When others cannot identify or single them out (unidentifiable) and no one can evaluate them, individuals feel that there is no need to concern for social evaluation. In this way, the deindividuated internal state allows individuals' uninhibited behavior to be released.

Building on the concept of deindividuation, social identity theory [74], and the self-categorization theory [77], the social identity model of deindividuation effects (SIDE) suggests that when individuals are anonymous during CMC, they tend to behave at the level of social group identity rather than personal identity and hence, are more likely to conform to the group-based norm [43]. For example, individuals may intentionally express their social identity to consolidate their identity or seek acceptance as an ingroup member (e.g., "we" talk). At the same time, they may state resistance against the outgroup to enhance the standing of their ingroup (e.g., "us" versus "them"). As a consequence, when political identity is salient, membership in a political group has important implications for online incivility [24].

2.2 The Broaden-and-build Theory of Positive Emotions

The broaden-and-build theory of positive emotions provides a framework for understanding how positive background images may contribute to reducing online incivility. The broaden-and-build theory of positive emotions captures the unique effects of positive emotions on individuals' affection, cognition, and behavior. According to the theory, positive emotions (e.g., joy, gratitude, serenity, interest, hope, etc) can change an individual's cognition and behavior by (1) broadening people's momentary thought-action repertoires and (2) building enduring personal resources, ranging from physical and intellectual resources to social and psychological resources [20, 21].

The current study focuses particularly on the "broadening effect" of positive emotions characterized by broadened cognition such as openness and flexibility to new situations and ideas [21]. On an individual level, positive emotions may enhance thoughts that are flexible and inclusive [34], open to information [18], forward-looking and high-level [57], and openness to a variety of behavioral options [38]. For example, under the influence of positive emotions, individuals take in more of the contextual surrounding whereas negative emotions narrow individuals' field of view [63]. More importantly, positive emotions may improve self-regulation [76].

A series of experimental studies have shown that generated positive emotion resulted in an improvement in self-regulation [76], and people who are in a high self-control state had fewer intentions to behave aggressively [13]. That is, enhanced self-regulation may be a mediating process in the relationship between positive emotions and aggressive behavior. If this is the case, enhanced self-regulation – inner restrains against aggression – would work in countering online incivility, one form of aggressive behavior, derived from the anonymous nature of CMC.

The broaden-and-build theory of positive emotions also suggests that the broadening effect of positive emotions extends to the social level. On a social level, positive emotions may expand individuals' circle of trust [17], underlie the creation of a variety of bonds and interdependence opportunities [12], and form inclusive social categories and common in-group identities in a way that people are more likely to see "them" as "us" [15, 16, 35]. For example, individuals experiencing positive emotions had greater perspective-taking and compassion for a person from dissimilar cultural backgrounds [52].

In summary, deindividuation and SIDE theories suggest that the anonymous nature of CMC can contribute to uncivil online conversation due to individuals' lowered self-regulation and enhanced conformity to the group-based norm. Meanwhile, the broaden-and-build theory of positive emotions provides a framework for designing an online intervention to counter uncivil online conversation. Positive emotions may contribute to reducing online incivility by (1) enhancing self-regulation and hence decreasing human aggression, and (2) lowering the boundaries between in-group (us) versus out-group (them).

Existing literature that explored the ways to reduce online incivility involves noticeable interventions such as removing the anonymity of the platform [31, 59], showing civil example comments [25, 26, 50], adding reminders [48], or moderating potentially harmful comments [37]. However, removing anonymity can curtail participation from certain groups, including the most vulnerable, and explicit moderation or reminder actions may impact user experience with the platform. In fact, multiple recent studies have identified issues with moderation including the labor intensiveness, opaqueness, and the (perceived) unfairness of the process [36, 73]. Hence, there is a need to explore other ways to reduce online incivility.

Taking inspiration from Weiser and Brown's vision of "calm" design i.e., where the interaction between the technology and its user is designed to occur in the user's periphery rather than constantly at the center of attention [80], we attempted to empirically investigate the effectiveness of the exposure to positive background images as a way to reduce online incivility. To do so, we

implemented an online experiment with three conditions: positive image background in color, positive image background in grayscale, and no background (control). As positive images are widely used to generate positive emotions in the positive emotions literature [22], we embedded positive background images in the discussion forum to create positive conditions. The research questions in this study are:

RQ1: *Does exposure to positive background images help reduce online incivility?*

Given that incivility can be measured in multiple ways including as an *in-situ* experience of the participants in online forums and in terms of the number of incivility instances as observed by a third party, we consider two related sub-questions.

RQ1a: Does exposure to positive background images help reduce online users' perceived online incivility?

RQ1b: Does exposure to positive background images help reduce uncivil expression present in the discussion forum?

RQ2: *How do the user comments in positive (color), positive (grayscale), and control conditions differ?*

In this study, we consider multiple facets of comments including the direction of incivility (the target or aim of uncivil expression), the emotions conveyed, and the topics conveyed, yielding the following sub-questions.

RQ2a: How do the user comments in positive (color), positive (grayscale), and control conditions differ in terms of the direction of incivility expressed?

RQ2b: How do the user comments in positive (color), positive (grayscale), and control conditions differ in terms of positive/negative emotions manifested?

RQ2c: How do the user comments in positive (color), positive (grayscale), and control conditions differ in terms of the topics discussed?

3 METHODS

3.1 Experiment Settings

Three experimental conditions created in this work are (1) positive-color condition with a positive background image in color, (2) positive-grayscale condition with the same positive background image in grayscale, and (3) control condition with no background image. According to previous literature, compared to the color screen display, the grayscale screen display has a calming effect (e.g., reducing anxiety) for online users, though it maintains the same communication functionalities [28, 29]. Following the literature, we created the positive-color and the positive-grayscale conditions to test whether a positive background image in grayscale is calmer and less obtrusive for users than a positive background image in color but has a similar effect on reducing online incivility. Note that the use of a negative background image for a control condition was not considered due to the potential for harm.

A blog site (WordPress.com) was used to create all three experiment settings. Each condition was an entirely separate website so that the participants could not see or engage with comments from other groups. The site consisted of a blog post and a comment thread below the post where an anonymous discussion happened. The anonymous environment was manipulated by assigning a unique user ID to every participant to log in to the blog site. The positive images were embedded as the background of the blog site but the center part of the blog site (where the content of the blog post and the comments are displayed) had a plain white background so that the participants could read the post and the comments easily (see Figure 1).

The blog post was an excerpt of an existing news article from "The Washington Post" describing U.S. immigration. The title of the article was "Texas governor vows to build Mexico border wall and increase arrests of migrants." U.S. immigration was chosen as a topic of the blog post as it is likely to stimulate a robust response in the commenting forums of online newspapers [4, 61, 62]. Also, three comments taken from the Reddit forum were displayed as example comments in all three conditions. The blog posts and the example comments in all experimental conditions were identical; the only difference among the three experimental conditions was the background image.

Following the previous positive emotions literature [23], a positive image was used to generate positive emotions. In this study, 10 positive images in color were chosen from the standardized affective image dataset, Open Affective Standardized Image Set (OASIS) developed by Kurdi et al. [41]. Then we grayscaled the same 10 positive images chosen from the OASIS to create 10 positive images in grayscale. With the 10 positive images in color and the same 10 positive images in grayscale, an online survey was conducted to assess the valence (the extent to which an emotion is positive or negative) and the arousal (intensity of the emotion, high vs low) [41] of the images. A positive image in color and the same positive image in grayscale to be used in the experiment were chosen based on the average values of the valence scores and the arousal scores (see Appendix A.1 for the details on the process of selecting the images). Figure 1 shows the screenshots of the experiment settings that were used in this study.

3.2 Procedure

This study consisted of a three-stage between-subjects experimental design. The three stages included the 1) pre-task survey, 2) intervention, and 3) post-task survey. For the pre-task survey stage, participants answered the survey questionnaire asking about their demography, personality traits (e.g., Big Five personality traits, conflict avoidance), political ideology, and opinions regarding three different political issues (immigration, abortion, and gun control). After they filled out the pre-task survey, participants were randomly assigned to one of the three experimental conditions. The participants received an email with the invitation link, a unique user ID, and the password to access the blog site where the main intervention took place. The main intervention lasted for three days. On the first day, participants read a blog post describing the U.S. immigration policy and left one comment toward the blog post. On the second and the third day, they left at least two reply comments (at least one comment a day) to other participants' comments. Participants were able to see other participants' comments when reading the blog post on the first day and the following two days. After the participants left at least three comments for three days, they filled out the post-survey. The post-task survey asked participants to rate how controversial the discussion topic was and how uncivil the comments were. It also asked three open-ended questions regarding whether they noticed the background image, and whether the background image caused emotional and behavioral changes.

3.3 Participants

A total of 120 participants were recruited via the *Prolific* crowdsourcing platform. The participants (1) were over 18; (2) were comfortable with written English; (3) had access to a desktop with an active Internet connection; (4) had a valid email address; (5) resided in the U.S; (6) had engaged with an online forum in the last three months. Participants were compensated $20 for completing the pre-survey, intervention, and post-survey. Three participants who most frequently left comments that encouraged further discussion for other participants received an extra compensation of $50, $30, and $20. This evaluation was undertaken by the lead author. We informed the potential of this additional reward when we recruited participants to motivate vibrant conversations during the experiment. An electronic consent form was provided to each participant in the first stage of the

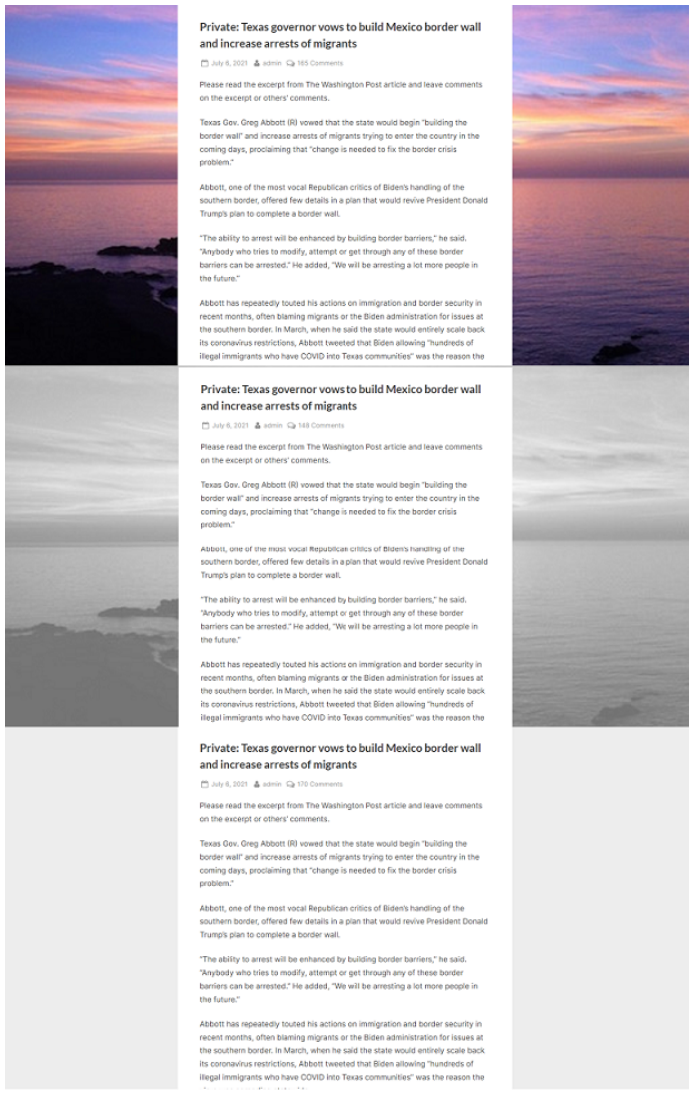

Fig. 1. Screenshots of Positive-color (top), Positive-grayscale (middle), and Control (bottom) Conditions

study and consent was required for participation. The participants were also informed that they can withdraw their participation at any time. The study protocol was approved by the institutional review board at the authors' home institution.

105 of the 120 participants completed all three stages of the experiment (85% completion rate). Among the 105 participants, 34 participants were assigned to the positive-color and positive-grayscale conditions (respectively), and 37 participants were assigned to the control condition. The participants' age ranged from 18-24 (19%), 25-34 (39%), 35-44 (25.7%), 45-54 (11.4%), and 55 and older (4.8%); 52.4% of them were female, 46.7% were male, and 1% were non-binary; 70.5% were White, 13.3% were African American, 9.5% were Asian or Pacific Islander, 3.8% were Hispanic, and 2.9% were other; 48.6% had a Bachelor's degree, 23.8% had some college education, 18.1% had a

Master's degree, 6.7% were high school graduate, 1.9% had a Doctorate degree, and 1% had some high school education.

3.4 Measurement

3.4.1 Perceived Incivility. The perceived online incivility was measured by four items created by [39]. The four items asked participants to rate the degree to which the comments were "rude," "disrespectful," "necessary," and "civil" on seven-point Likert-type scales from (1) strongly rude to (7) strongly polite, (1) strongly disrespectful to (7) strongly respectful, (1) strongly necessary to (2) strongly unnecessary, and (1) strongly civil to (2) strongly uncivil. The responses to "rude" and "respectful" were reverse coded so that the higher scores contribute to higher perceived online incivility. The item "necessary" was excluded as it had low reliability and low factor loading compared to the other three items. A principal component factor analysis confirmed that the three items loaded on a single factor (eigenvalue = 2.41; 80.38% variance; all factor loading > .70). The perceived incivility scores were averaged over the three items for all the cases as there was no missing data. The reliability was good (α = .88) and the scores ranged from 1 to 6 (M = 3.0, SD = .99).

3.4.2 Incivility Score. The incivility score measured the frequencies of uncivil expressions that are manifested in each experimental condition. Following previous literature [11, 58], six types of incivility (name-calling, aspersion, lying, vulgarity, pejorative for speech, others) were used to assess the number of uncivil expressions in all comments collected from the three experimental groups (N = 457). The definitions and the examples of the six categories of incivility are shown in Table 1. Note that all the example comments in the current study were taken from the user comments collected from the three experimental conditions. Each comment was checked for the presence of uncivil expressions (i.e., word or phrase that could be assigned to one of the six categories of incivility). The total number of uncivil expressions in each comment was counted to generate a total incivility score for the comment. The total incivility score for the comments ranged from 0 to 9 (M = .81, SD = 1.25). The average values of incivility for all comments within an experimental condition were used for comparison (results follow).

An outside coder was recruited to establish intercoder reliability and to code the comments. The researcher (first author) and the coder had multiple training sessions during Fall 2021. After the training sessions, 10% of all comments (n = 45) were coded to establish intercoder reliability between the researcher and the coder. The intercoder reliability scores for the total incivility score were 0.88 (percent agreement) and 0.72 (Krippendorff's alpha). Both scores met the criteria suggested by the existing literature [42, 53] for a good agreement.

3.4.3 Direction of Incivility. The direction of incivility was evaluated by the first author and the same outside coder. Following the previous literature [56, 60, 72], three categories were used to assess the direction of incivility. The three categories of direction included: interpersonal, other-directed, and impersonal. If incivility was directed to someone present within the discussion forum, it was coded as interpersonal. If incivility was directed to someone outside of the discussion forum, it was coded as other-directed. Finally, if incivility was directed at something other than a person (e.g., idea, policy, plan, etc.), it was coded as impersonal. The definitions and examples of the direction of incivility are shown in Table 2.

The intercoder reliability for the direction of incivility was established between the researcher and the outside coder using 10% of the comments (n = 45, 0.91 percent agreement and 0.82 Krippendorff's alpha).

Table 1. Definition and Examples of Types of Incivility

Type	Definition	Example
Name-calling	Mean-spirited or disparaging words directed at a person or group of people	"Damn these immigrants who come here legally"
Aspersion	Mean-spirited or disparaging words directed at an idea, plan, policy, or behavior	"Abbott's blaming of immigrants for the issues in Texas is quite honestly disgusting and blatantly xenophobic"
Lying	Stating or implying that an idea, plan, policy, or public figure was disingenuous	"Abbot is just spreading lies"
Vulgarity	Using profanity of language that would not be considered proper in professional discourse	"something is kinda shitty about making Texas seem like such a horrible place to live"
Pejorative for speech	Disparaging remark about the way in which a person communicates	"How he's whined and moaned about the mandates"
Others	All comments that may be deemed uncivil, but do not fall into any of the previous categories of incivility	"WHY DO THEY NEED TO WORK FOR FREE?"

Table 2. Definition and Examples of Direction of Incivility

Direction	Definition	Example
Interpersonal	Incivility was directed to someone present within the discussion forum	(1) "Check yourself before you wreck yourself!" (2)"blaaaaaaaah go to hell dude"
Other-directed	Incivility was directed to someone outside of the discussion forum	(1) "Damn these immigrants who come here illegally" (2) "Governor Abbott is Self righteous dumbass"
Impersonal	Incivility was directed at something other than a person (e.g., idea, policy, plan, etc.)	(1) "the idea of a wall being the end-all-be-all to stop illegal immigration is ridiculous" (2) "there hasn't been a damn thing done"

3.4.4 Sentiment. The positive and the negative sentiment scores for each comment in all three conditions were generated using the sentiment detection classifier "SentiStrength." The SentiStrength is a lexicon-based classifier that also uses non-lexical linguistic information to detect positive and negative sentiment in informal English text [75]. It adds the total values of positive/negative sentiment scores of the texts within each comment. The sentiment classifier with the two scales (positive and negative) was used in this study because even short texts can contain both positivity and negativity and the goal was to detect the sentiment expressed in the comments rather than their overall polarity. The positive sentiment scores ranged from 2 to 405 (M = 104.73 , SD = 65.81), while the negative sentiment scores ranged from -2 to -419 (M = -108.42 , SD = 68.00).

3.4.5 Topics Discussed. Topics discussed in the three experimental conditions were identified using the latent Dirichlet allocation (LDA) model. The LDA is an unsupervised algorithm that identifies latent thematic patterns of word occurrence using the distribution of words in a collection of text documents [3]. We applied the LDA topic model because it combines an inductive approach with quantitative computations, hence, making it suitable for exploratory analysis of text data [47].

Table 3. Unadjusted and Covariate-Adjusted Descriptive Statistics for Perceived Incivility

Group	Unadjusted		Adjusted 1 (demography)		Adjusted 2 (personality)		Adjusted 3 (ideology/opinion)	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Positive (color) (n = 34)	2.74	.16	2.81	.16	2.76	.17	2.72	.16
Positive (grayscale) (n = 34)	2.81	.14	2.82	.16	2.96	.17	2.85	.17
Control (n = 37)	3.41	.18	3.39	.15	3.37	.16	3.39	.16

Before running the LDA model, the texts (457 comments collected from the three experimental groups) were pre-processed (e.g., removing the stop words and infrequent words, lemmatizing, tokenizing, etc.). In addition, various words that refer to the same phenomenon (e.g., COVID, Corona, Corona virus to refer to COVID-19) were replaced with one unified word (e.g., virus). After the pre-processing, a document term matrix was created where each row is an intervention group, each column is a term, and the number in each cell shows how often each term appears in each intervention group. After creating the document-term matrix, the number of topics the LDA model should classify (K) was specified. The number of topics was decided by looking at the perplexity score (how surprised the model is at seeing the data, smaller the better) [3] and the coherence score (how often the topic words for each topic appear together in a document, closer to zero is better) [49]. The two scores confirmed that the LDA model with the five topics (K = 5) was a good fit.

4 RESULTS

4.1 Perceived Incivility

Multiple one-way Analysis of Covariance (ANCOVA) tests were conducted to compare the effectiveness of positive background images in reducing perceived online incivility. The covariates included participants' demography (model 1), personality (model 2), and political ideology & issue opinion (model 3). There was a significant difference in perceived online incivility between the experimental groups when controlling for demography [$F(2, 91) = 3.40$, $p = .038$, $\eta^2 = .09$], personality [$F(2, 99) = 4.99$, $p = .009$, $\eta^2 = .07$], and political ideology & issue opinion [$F(2, 97) = 4.55$, $p = .013$, $\eta^2 = .09$]. The values of partial eta-squared indicated the effect size of between medium and large (Cohen, 2013; Miles & Shevlin, 2001). The estimated means (Table 3) showed that the perceived incivility was the lowest in the positive-color condition and the highest in the control condition for all three models. None of the demographic characteristics, personality traits, political ideology, and issue opinion had a significant effect on the perceived incivility score when controlling for the effect of the intervention group. Post-hoc pairwise t-tests (with Bonferroni correction) showed that the mean difference was statistically significant between the positive-color and control conditions (p = .009), and between the positive-grayscale and control conditions (p = .025). However, the difference was *not* significant between the positive-color and positive-grayscale conditions (p = .939).

There were also three sets of open-ended questions asking (1) whether they recognized the existence of the background image and if they did, what was the background image (2) whether the participants experienced emotional changes during the experiment and if they did, why such changes occurred to them, and (3) whether the participants experienced behavioral change and if they did, why such changes occurred to them.

The first set of questions explored whether the participants recognized the background image while they are reading and writing the comments on a discussion forum. For the positive-color condition, 20 out of 34 participants reported that they recognized the background image, while 9

out of 34 participants from the positive-grayscale condition reported that they did. This difference across the two conditions was statistically significant based on a Chi-square test (χ^2 (1, N = 68) = 7.28, p = .007). That is, the participants in the positive-grayscale condition were less likely to recognize the existence of the background image compared to the participants in the positive-color condition. Among those who recognized the background image, 19 out of 20 from the positive-color condition, and 8 out of 9 participants from the positive-grayscale condition described the background images correctly. 1 out of 37 participants from the control condition reported that they recognized the background image when there was no background image at all (the participant recognized a plain background as a gray-colored background).

The second set of open-ended questions asked whether the background image caused any emotional changes while they are reading and writing the comments on a discussion forum. A total of 4 participants from the positive-color condition reported that they experienced some emotional changes. They also mentioned reasons why they experienced emotional changes. Some of the reasons include:

"*It sets a peaceful tone, so I'm sure it affected us all on a subconscious level (user 1, positive group)*"

"*The background was peaceful ... the background would cause me to feel peaceful and remember that I was sharing my thoughts with another person, not an enemy (user 5, positive group)*"

Two participants from the positive-grayscale condition reported that they experienced some emotional changes. The reasons for emotional changes included:

"The background was relaxing and elicited a serene feeling (user 1, positive-grayscale group)"

"I think the background gave a sense of calmness, peace, and relaxation (user 2, positive-grayscale group)"

None of the participants from the control condition reported that they experience any emotional changes.

The third set of open-ended questions asked whether the background image caused any behavioral changes while they were reading and writing the comments on a discussion forum. For behavioral changes, 5 participants from the positive-color condition reported that they experienced some behavioral changes. Some of the reasons for behavioral changes included:

"I think with the peaceful background, people were subconsciously leaving more civil comments ... I believe the softer colors prevented people from acting unruly and rude (user 3, positive group)"

"I think maybe there was a slight softening and calming effect on people reading/commenting, which may have encouraged a more civil discussion (user 4, positive group)"

"The positive mood of the image encouraged me to be polite (user 5, positive group)"

At the same time, 2 participants from the positive-grayscale condition reported that they experienced some behavioral changes. The reasons for behavioral changes included:

"The backdrop created a more relaxed atmosphere (user 1, positive-grayscale group)"

"It helped to calm me, and made me think more about what I was typing as a comment (user 2, positive-grayscale group)"

None of the participants from the control condition reported that they experience any behavioral changes.

The thematic content analysis (independent qualitative descriptive approach for identifying, analyzing, and reporting patterns or themes within data [5]) was conducted to identify common themes across the participants' answers. Overall, the most frequently mentioned reasons for the emotional or behavioral changes for the positive-color group were "peaceful" and "calming," followed by "soft color." For the positive-grayscale group, the most commonly mentioned reasons for the emotional or behavioral changes were "relaxing," followed by "calming."

Table 4. Unadjusted and Covariate-Adjusted Descriptive Statistics for Incivility Score

Group	Unadjusted		Adjusted	
	Mean	SE	Mean	SE
Positive (color) (n = 157)	0.49	0.06	0.50	0.10
Positive (grayscale) (n = 137)	0.73	0.08	0.75	0.10
Control (n = 163)	1.18	0.13	1.16	0.10

Table 5. Direction of Incivility and Experiment Groups (%)

	Positive (color)	Positive (grayscale)	Control
Interpersonal	6.49	7	13.02
Other-directed	67.53	71	77.08
Impersonal	25.97	22	9.9
Total	100	100	100

4.2 Incivility Score

An ANCOVA test was conducted to compare the mean values of incivility scores for the comments in the three experimental conditions (with the word count per comment as a covariate). A significant difference was found in terms of the mean incivility scores across the three experimental conditions, F (2, 453) = 11.97, $p < .001$, $\eta^2 = .08$. Post-hoc pairwise t-tests (with Bonferroni correction) confirmed that the mean difference was statistically significant between the positive-color and control conditions ($p < .001$), and between the positive-grayscale and control conditions ($p = .011$). This means that both the positive image in color and the positive image in grayscale were effective in reducing the number of uncivil expressions. When comparing the mean scores for the positive-color and positive-grayscale conditions, the mean incivility score was lower for the positive-color condition than it was for the positive-grayscale condition. However, the difference was not statistically significant. Table 4 shows the mean values of incivility per comment across the three groups.

4.3 Direction of Incivility

The most common direction of incivility was "other-directed" in all three experimental conditions (positive-color condition = 67.53%, positive-grayscale condition = 71%, control condition = 77.08%). For the positive-color and the positive-grayscale conditions, the second most common direction was "impersonal" and the least common was "interpersonal." On the other hand, for the control condition, the second common direction of incivility was "interpersonal" and the least common was "impersonal" direction (see Table 5). This difference across the three conditions was found to be statistically significant based on a Chi-square test (χ^2 (4, N = 253) = 15.46, p = 0.0038). That is, the comments in the positive-color and the positive-grayscale conditions were more likely to be impersonal and less likely to target other participants than those in the control condition.

4.4 Sentiment

An ANCOVA analysis was conducted to compare the mean values of positive and negative sentiment for comments in the three experimental conditions with the word count per comment as the covariate. For the positive sentiment, a significant difference was found among the three

Table 6. Unadjusted and Covariate-Adjusted Descriptive Statistics for Positive Sentiment Score

Group	Unadjusted		Adjusted	
	Mean	SE	Mean	SE
Positive (color)	94.81	4.8	97.97	3.57
Positive (grayscale)	98.29	5.5	103.19	3.82
Control	119.69	5.5	112.53	3.51

Table 7. Unadjusted and Covariate-Adjusted Descriptive Statistics for Negative Sentiment Score

Group	Unadjusted		Adjusted	
	Mean	SE	Mean	SE
Positive (color)	-98.64	4.9	-101.90	3.71
Positive (grayscale)	-101.85	5.7	-106.89	3.98
Control	-123.36	5.7	-115.98	3.65

experimental groups, F (2, 453) = 4.31, p = .014 (see Table 6 for means). Post-hoc pairwise t-tests (with Bonferroni correction) showed that the comments in the control condition manifested more positive sentiment compared to the comments in the positive-color condition (p = .002). In addition, the comments in the positive-grayscale condition manifested more positive sentiment compared to those in the positive-color condition; however, the difference was not statistically significant (p = .891).

For the negative sentiment, a significant difference was found among the three experimental groups, F (2, 453) = 3.74, (p = .024) (see Table 7 for means). Pair-wise t-tests (with Bonferroni correction) confirmed that the comments in the control condition manifested more negative sentiment compared to the comments in the positive-color condition (p = .003). In addition, the comments in the positive-grayscale condition manifested more negative sentiment compared to those in the positive-color condition; however, the difference was not statistically significant (p = .913).

Overall, the ANCOVA tests confirmed that comments in the control condition manifested more positive and negative tones compared to the positive-grayscale and positive-color conditions. In other words, the comments in the positive-grayscale and positive-color conditions were less emotional compared to the comments in the control condition. The difference in the emotional tone was statistically significant between the positive-color and the control condition, whereas it was not between the positive-color and the positive-grayscale condition.

4.5 Topics Discussed

The five most common topics discussed in the three experimental groups and the top words in each topic were generated by the LDA topic model (see Appendix for the top words and their beta values generated by the LDA topic model). The beta values i.e., per-topic-per-words probabilities, for the top words were also calculated. The beta represents the probability that the term would be generated for the topic [67]. That is, the higher the beta, the more often the word would appear within the topic. We reviewed the top words and read through the entire comments from the three conditions to deductively name a title describing the substantive content of the five topics generated by the LDA model.

Table 8. Top Five Topics and Example Quotes

Topic	Title	Example Quotes
Topic 1	Costs related to the immigration issues	"People worry about all of the taxpayer money that is going to be needed to make this happen," "Why would anyone in their right mind still be in favor of such a blatant waste of taxpayers' money?"
Topic 2	Human aspects of the immigration issues	"Give them a place to live and a way to make a better life," "Some people need to come here for a better life for themselves and their families"
Topic 3	Building of the border wall in Texas	"They have been saying they will build the border wall for years now, and it has never happened," "I would also like to know a cost-benefit analysis of building a border wall, having people guard the border wall, making arrests with drones and helicopters, etc."
Topic 4	Immigrants and the spread of COVID-19	"It seems pretty hypocritical to blame illegal immigrants for spreading COVID, when Texas lifted state restrictions before it was considered safe to do so nationwide," "I think the fact that they want to try and blame the spread of covid on immigrants is just crazy"
Topic 5	Other priorities in Texas (e.g., power grid)	"I feel the priorities at this stage in time should be focused on more important things,""Instead of focusing on issues that Texans clearly have been affected by (did we already forget about the power grid coming to a standstill this past winter?)"

The five topics were assigned the titles: "Costs related to the immigration issues," "Humanity aspects of the immigration issues," "Building of the border wall in Texas," "Immigrants and the spread of COVID-19," and "Other priorities in Texas." Table 8 shows the top five topics generated by the LDA topic modeling and some example quotes that represent each topic.

Next, a gamma value was calculated to see how common the five topics were in each experimental group. Gamma is the estimated proportion of terms in the document that may be generated by each topic [67]. The higher the gamma, the more the group is likely to talk about the topic. For example, a gamma value of 0.17 for Topic 1 in the positive-color condition means that 17% of the words in the user comments from the positive-color condition were generated from Topic 1. The gamma values of each topic in the three intervention groups are present in Figure 2.

Based on a Chi-square test, the proportion of topics discussed differed significantly across the three intervention groups, χ^2 (8, N = 457) = 53.69, p <0.001. In the positive-color condition, the most commonly discussed topic was defending the illegal immigrants from being blamed for the spread of the COVID-19 in the U.S (topic 4), followed by the building of the border wall in Texas (topic 3). In the positive-grayscale condition, the most discussed topic was the building of the border wall in Texas (topic 3). The second most discussed topic was the human aspects of the immigration issues such as the value of life and family (topic 2). Meanwhile, participants in the control condition discussed the most regarding costs of the immigration issue (e.g., U.S tax money) (topic 1), followed by other priorities in Texas (topic 5).

5 DISCUSSION

In the present study, we focused on the impact of positive background images on online incivility. RQ1a asked, "Does exposure to positive background images help reduce online users' perceived

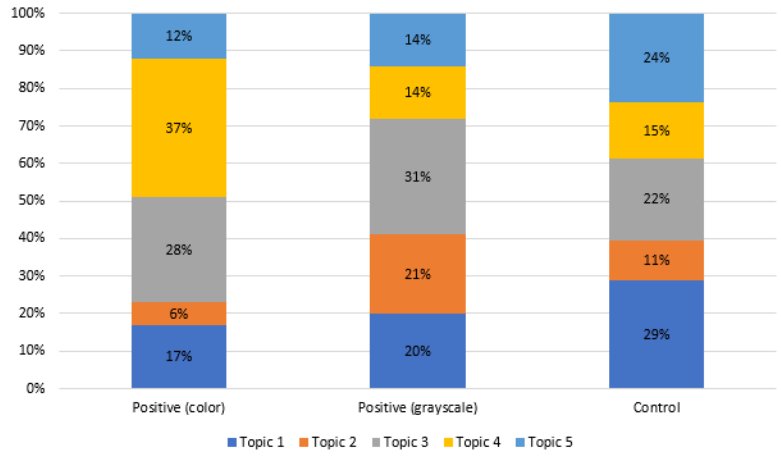

Fig. 2. The Estimated Proportion of Topics Discussed

online incivility?" The results of the study showed that the positive images in color and the positive images in grayscale had a significant effect on reducing online users' perceived incivility. The results indicated that regardless of its color, having a positive background image can help reduce online incivility compared to a discussion forum that has no background image at all. When comparing the positive background image in color and the positive background image in grayscale, the positive background image in grayscale was less likely to be recognized by the participants. Participants' input suggests that the soft colors of the positive image in color may have generated peacefulness in the users. Meanwhile, the gray tone of the positive image in grayscale may have generated a relaxing and calming atmosphere for the users.

An interesting finding from the open-ended post-survey was that some participants from the positive-color condition explicitly mentioned that the *positive mood* that they felt from the positive image encouraged them to be *polite* and that the softer colors (of the image) *prevented people from acting unruly and rude*. According to previous literature, positive emotions could enhance individuals' self-regulation (inner restrains against aggression) and people who are in high self-control are less likely to behave aggressively. Participants' answers from the study suggest that positive emotions indeed improve online users' self-regulation and hence, are effective in countering online incivility. Some participants mentioned that the background was "relaxing" and elicited a "serene" feeling. In fact, "serenity" is one of the ten representative positive emotions in the broaden and build theory of positive emotions [22]. Hence, the results add empirical evidence to existing positive emotions literature. At the same time, they motivate further research that examines self-regulation as the mediating process between positive emotions and reduced online incivility, and teases out the relative efficacy of different kinds of positive emotions (e.g., serenity vs. joy) for reducing incivility.

Another interesting quote from the participants' answers was that the positive background caused them to "remember that they were sharing their thoughts with another person, not an enemy." The participant's answer may indicate that positive emotions indeed broaden social group concepts and break down the sense of "us versus them" [15, 23]. Given that the participants interacted in an anonymous experiment setting, the findings from the study may imply that positive emotions could mitigate the uncivil and partisan conversations, particularly on anonymous online platforms. Future research may examine the effectiveness of positive emotions in reducing online incivility,

particularly focusing on an online environment where social identities are salient and partisan issues are discussed.

RQ1b asked, "Does exposure to positive background images help reduce uncivil expression present in the discussion forum?" The results of the study showed that both the positive background image in color and the positive background image in grayscale were effective in reducing the frequencies of uncivil expressions in the discussion forums. The findings suggest that embedding positive background images in grayscale may be almost as effective as embedding positive background images in color in reducing online incivility.

As mentioned by some of the participants in the study, the positive background in grayscale may elicit "a sense of calmness, peace, and relaxation (user 2, positive-grayscale group)" in online users. Given that the positive background image in grayscale was less likely to be recognized by the participants in this study, embedding positive background images in grayscale could be *calm* yet *effective* strategy to reduce uncivil conversations on online discussion forums.

For the direction of incivility (RQ2a), the majority of the uncivil expressions in all three experimental conditions were directed toward third persons outside the discussion forum (e.g., Texas governor) and their policies. The noticeable difference was the proportion of *interpersonal* uncivil expressions in the three conditions. The comments in the control condition were almost twice as likely to be directed to other participants in the discussion forum than the comments in the positive-color or positive-grayscale conditions. The trend was the opposite in terms of *impersonal* (not directed at a particular person) incivility. The control condition had less than half of the impersonal incivility ratio as the positive-color and positive-grayscale conditions did. That is, uncivil expressions in the positive-color and positive-grayscale conditions were more often used as a way to articulate a controversial *idea* or an *issue* while uncivil expressions in the control condition were used to offend other *individuals* discussing the same topic.

Positive emotions literature states that positive emotions may broaden social cognition in the form of enhanced attention to others and reduced distinction between self and others. Especially, just as in the current study setting, when a close relationship does not yet exist, generated positive emotions may underlie the creation of bonds and interdependence opportunities [23]. Positive emotions generated in the current study may have broadened the social cognition of the participants, hence, they became less likely to leave uncivil comments directed toward other participants in the positive-color and the positive-grayscale conditions. Further research can compare the effectiveness of positive emotions to reduce different types and directions of incivility in the anonymous discussion forums.

The overall trend of a positive and negative sentiment of user comments (RQ2b) across the three experimental groups was that both positive and negative emotions were stronger in the control condition compared to those in the positive-color and the positive-grayscale conditions. In other words, the user comments in the positive-color condition and the positive-grayscale condition were less emotional than those in the control condition.

The trend of the negative sentiment scores across the three groups was in line with the results of perceived incivility and incivility score. Given that the incivility scores were highest in the control condition and the tone of uncivil comments is negative, it was intuitive that the comments in the control condition were more negative in tone compared to those in the positive-grayscale and the positive-color conditions.

The results were counter-intuitive for the positive sentiment scores; comments in the control condition were more positive in tone compared to those in the positive-grayscale and the positive-color conditions. According to positive affect literature, people in a positive mood may avoid thinking about negative information to preserve positive affect (mood maintenance), while people in a negative mood may seek positive information or behavioral opportunities (e.g., using positive

sentiment words) to improve their mood (mood repair) [2]. This might have been the case for the current study that participants in the two positive conditions reacted conservatively to the negatively valenced information (e.g., other participants' uncivil comments) to preserve their positive mood, hence, they were more likely to leave dispassionate comments. On the other hand, the control condition might have witnessed an interplay of strongly emotional comments, where some participants used a strong negative tone and others (a potentially smaller group) were actively trying to "repair" the discussion by bringing up positive aspects.

In addition, in the positive-color and the positive-grayscale conditions, people may have felt fewer "intergroup emotions," i.e., general emotional feelings (both positive and negative emotions) toward the ingroup [70], compared to those in the control condition due to generated positive emotions, thus, their comments manifested a less general affective tone. Given that some of the user comments from all three conditions reflected intergroup emotions such as group-based pride (e.g., *the great American*) and collective guilt (e.g., *as American, I feel sorry for what we have done to other countries*), the degree to which participants felt intergroup emotions may have affected the emotional tone of user comments across three experimental groups. Understanding such emotion dynamics under different experimental conditions is an interesting area for future research.

Finally, the most discussed topics across the three experimental groups (RQ2c) were different. The comments in the positive-color and the positive-grayscale conditions discussed more human-centered approaches to the immigration issue, while the comments in the control condition discussed more economic-centered approaches to the immigration issue. That is, the comments in the positive-color and the positive-grayscale conditions showed more compassion toward the immigrants than those in the control condition. The results are in line with the positive emotions literature suggesting that positive emotions can enhance attention to others, reduce the distinction between self and others, and underlie the creation of bonds and interdependence opportunities [23]. As interpersonal incivility was less likely to happen in the positive conditions, participants in a positive mood may have lower hostility toward the illegal immigrants and, thus, may have seen the immigration issues from a humanitarian perspective.

Taken together, the results suggest that exposure to both positive images in color and positive images in grayscale can help reduce online incivility and at the same time yield more positive discussion, fewer personal attacks, and more humane discussion for the topics at hand.

Like most research, the current study also has some limitations. The controlled setting of the experiment may have biased the participants' behavior. For instance, participants may not leave uncivil comments as they would in real life because they know that they are part of the experiment and that their behavior may be monitored by a researcher. The fact that the researcher will be reading the comments may put social pressure on the participants. In this case, a controlled setting can be an external factor that confounds the results. Another limitation of the current study could be sampling bias. The participants in the current study were recruited from a crowdsourcing platform. Therefore, there is a self-selection effect as the participants in the study must have chosen to sign up for a certain crowdsourcing platform and to participate in the study. In addition, we consciously recruited the participants who had engaged in an online forum in the last three months. Given that we focused on reducing incivility in a specific online setting (e.g., online discussion forums), we tried to recruit those who are likely to engage in the same online setting. Thus, we recognize that the results from the current study may not be the same for other populations. At the same time, this allows us to quantify the efficacy of the proposed approach with a group most likely to engage with online forums. Finally, this study may have exposed participants to an uncomfortable experience. Reading others' uncivil comments could have made them feel anger or fear. Further, the background images could have also influenced their mood. Having said that, the issue of online

incivility is an important one, and experiments like the current study are necessary to identify a path forward toward more civil discourse in online settings.

Despite the limitations, the present study has important implications. This study extends research in CSCW by adding theoretical and empirical evidence to support the application of positive emotions to counter online incivility. Methodologically, the mixed methods applied in this research can provide examples of a holistic approach to measuring online incivility using randomized controlled experiments, user surveys, and content analysis of logged user comments. An improved understanding of positive emotions and online incivility can motivate further research on similar topics in different contexts. For example, further research can test different kinds of emotional cues or similar kinds of positive emotions in different settings to observe which interventions are the most effective in reducing online incivility.

Practically, this study provides design implications to online community designers, developers, and moderators. Following the rapid growth of social media, individuals have increasingly been exposed to aggressive online communication such as online incivility. In 2020, 75% of the victims of online abuse reported that their most recent experience was on social media; 79% agreed that social media companies are doing only a "fair" or "poor" job at addressing online harassment on their platform [8]. Hence, identifying ways to counter risks such as online incivility is vital for social media companies such as Facebook, Twitter, and Reddit. Particularly, the background setting used in this study is an easy intervention already supported by some social media sites (see the Appalachia subreddit at https://www.reddit.com/r/Appalachia/). The utility of a "calm" approach to reducing incivility and the efficacy of even a grayscaled positive background image to reduce online incivility are useful takeaways from this work and can help expand the repertoire of design tools available to designers to reduce incivility.

The results from the current study can motivate social media designers and developers to consider and test other creative ways of exposing users to positive emotions, especially when the discussion forums focus on controversial topics. For instance, the designers can utilize positive imagery in the side- banners, loading screens, create (or only allow) positive advertisements, or create interactive elements that support the generation of positive emotions. With the growing adoption of 360-degree/ immersive content, the development of "calm" technologies and design decisions on what is shown in the peripheries will be just as important as what is shown in the center region. Such approaches may also be explored to identify ways to counter other forms of online harm such as cyberbullying, sexual harassment, and more.

Finally, in every intervention, the positives and the negatives should be considered and the interventions should be applied in practice only if the positives outweigh the negatives. Our analysis of the literature yielded no solutions that were free of side effects. In this experiment, we found that positive background images were associated with more dispassionate comments. This can be an issue if maintaining/promoting passionate discussion is a priority. At the same time, there can be many scenarios, especially when dealing with sensitive topics (e.g., immigration, gun rights), where the social media system designers may want to reduce incivility and have a less passionate discussion. We acknowledge that a translation from a controlled setting to a large-scale social media deployment must be done with care and this may require additional *translational* research. For instance, the current study was conducted in strict compliance with the institution's research ethics guidelines. Before conducting the experiment, the purpose of the study and the entire procedure of the experiment was clearly communicated to the participants, and written consent was collected from all participants. Also, we consciously chose not to include a condition that may generate a negative mood to minimize potential harm to the participants. Ensuring such ethics review, transparency, and permission at a social media scale would require additional important work. Evaluation of long-term impact should be another important component of this

translation work. Yet, there is a need to explore the design space and identify the potential positive and negative effects of different moderation approaches (e.g., explicit reminders, voting, etc.) before the important translation work can be undertaken. This study contributes to the exploration of that design space.

6 CONCLUSION

Making online space secure continues to be one of the critical societal challenges. The main purpose of the current study was to understand the effectiveness of exposure to positive background images in reducing online incivility. Building upon the broaden-and-build theory of positive emotions, the study designed and implemented an online experiment with 105 participants with three different settings: positive background image in color, positive background image in grayscale, and no background image embedded in online discussion forums. The results showed that not only positive background images in color but also positive background images in grayscale may be effective at fostering less out-group aversion, more humane discussions, and reducing online incivility. Such a "calm" approach to reducing incivility can be a useful tool for designers to reduce incivility in future online spaces. With the growth in immersive technologies, we expect the importance of design decisions on the peripheral and background content to only increase. This study marks a vital first step toward utilizing the peripheral space to increase positive emotions and civil discussions in online settings.

ACKNOWLEDGMENTS

This work was supported by Rutgers University School of Communication and Information.

REFERENCES

[1] Ashley A Anderson, Dominique Brossard, Dietram A Scheufele, Michael A Xenos, and Peter Ladwig. 2014. The "nasty effect:" Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication* 19, 3 (2014), 373–387.

[2] Lisa G Aspinwall. 1998. Rethinking the role of positive affect in self-regulation. *Motivation and emotion* 22, 1 (1998), 1–32.

[3] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research* 3 (2003), 993–1022.

[4] Porismita Borah. 2014. Does it matter where you read the news story? Interaction of incivility and news frames in the political blogosphere. *Communication Research* 41, 6 (2014), 809–827.

[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[6] Penelope Brown, Stephen C Levinson, and Stephen C Levinson. 1987. *Politeness: Some universals in language usage*. Vol. 4. Cambridge university press.

[7] Pew Research Center. 2017. Online Harassment 2017. Retrieved Jan. 12, 2022 from https://www.pewresearch.org/internet/2017/07/11/online-harassment-2017/

[8] Pew Research Center. 2021. The State of Online Harassment. Retrieved Jan. 12, 2022 from https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/

[9] Eshwar Chandrasekharan, Chaitrali Gandhi, Matthew Wortley Mustelier, and Eric Gilbert. 2019. Crossmod: A cross-community learning-based system to assist reddit moderators. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–30.

[10] Sungwon Chung and Jongpil Cheon. 2020. Emotional design of multimedia learning using background images with motivational cues. *Journal of Computer Assisted Learning* 36, 6 (2020), 922–932.

[11] Kevin Coe, Kate Kenski, and Stephen A Rains. 2014. Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication* 64, 4 (2014), 658–679.

[12] Michael A Cohn and Barbara L Fredrickson. 2006. Beyond the moment, beyond the self: Shared ground between selective investment theory and the broaden-and-build theory of positive emotions. *Psychological Inquiry* 17, 1 (2006), 39–44.

[13] C Nathan DeWall, Roy F Baumeister, Tyler F Stillman, and Matthew T Gailliot. 2007. Violence restrained: Effects of self-regulation and its depletion on aggression. *Journal of Experimental social psychology* 43, 1 (2007), 62–76.

[14] Soussan Djamasbi, Marisa Siegel, and Tom Tullis. 2014. Can fixation on main images predict visual appeal of homepages?. In *2014 47th Hawaii International Conference on System Sciences*. IEEE, 371–375.

[15] John F Dovidio, Samuel L Gaertner, Alice M Isen, and Robert Lowrance. 1995. Group representations and intergroup bias: Positive affect, similarity, and group size. *Personality and Social Psychology Bulletin* 21, 8 (1995), 856–865.

[16] John F Dovidio, Alice M Isen, Paula Guerra, Samuel L Gaertner, and Mary Rust. 1998. Positive affect, cognition, and the reduction of intergroup bias. (1998).

[17] Jennifer R Dunn and Maurice E Schweitzer. 2005. Feeling and believing: the influence of emotion on trust. *Journal of personality and social psychology* 88, 5 (2005), 736.

[18] Carlos A Estrada, Alice M Isen, and Mark J Young. 1997. Positive affect facilitates integration of information and decreases anchoring in reasoning among physicians. *Organizational behavior and human decision processes* 72, 1 (1997), 117–135.

[19] Leon Festinger, Albert Pepitone, and Theodore M Newcomb. 1952. Some consequences of de-individuation in a group. *The Journal of Abnormal and Social Psychology* 47, 2, Suppl (1952), 382–389.

[20] Barbara L Fredrickson. 1998. What good are positive emotions? *Review of general psychology* 2, 3 (1998), 300–319.

[21] Barbara L Fredrickson. 2001. The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *American psychologist* 56, 3 (2001), 218.

[22] Barbara L Fredrickson. 2013. Positive emotions broaden and build. In *Advances in experimental social psychology*. Vol. 47. Elsevier, 1–53.

[23] Barbara L Fredrickson and Michael A Cohn. 2008. Positive Emotions. *Handbook of emotions* (2008), 777–796.

[24] Bryan T Gervais. 2015. Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics* 12, 2 (2015), 167–185.

[25] Soo-Hye Han and LeAnn M Brazeal. 2015. Playing nice: Modeling civility in online political discussions. *Communication Research Reports* 32, 1 (2015), 20–28.

[26] Soo-Hye Han, LeAnn M Brazeal, and Natalie Pennington. 2018. Is civility contagious? Examining the impact of modeling in online political discussions. *Social Media+ Society* 4, 3 (2018), 2056305118793404.

[27] Susan Herbst. 2010. *Rude democracy: Civility and incivility in American politics*. Temple University Press.

[28] Alex J Holte and F Richard Ferraro. 2020. True colors: Grayscale setting reduces screen time in college students. *The Social Science Journal* (2020), 1–17.

[29] Alex J Holte, Desiree T Giesen, and F Richard Ferraro. 2021. Color me calm: Grayscale phone setting reduces anxiety and problematic smartphone use. *Current Psychology* (2021), 1–13.

[30] Mark Hsueh, Kumar Yogeeswaran, and Sanna Malinen. 2015. "Leave your comment below": Can biased online comments influence our own prejudicial attitudes and behaviors? *Human communication research* 41, 4 (2015), 557–576.

[31] Matthew W Hughey and Jessie Daniels. 2013. Racist comments at online news sites: a methodological dilemma for discourse analysis. *Media, Culture & Society* 35, 3 (2013), 332–347.

[32] Andrew Hutchinson. 2021. Twitter Adds New Option to Add Background Video and Images in Fleets. Retrieved Jan. 12, 2022 from https://www.socialmediatoday.com/news/twitter-adds-new-option-to-add-background-video-and-images-in-fleets/598747/

[33] Hyunseo Hwang, Youngju Kim, and Catherine U Huh. 2014. Seeing is believing: Effects of uncivil online debate on political polarization and expectations of deliberation. *Journal of Broadcasting & Electronic Media* 58, 4 (2014), 621–633.

[34] Alice M Isen and Kimberly A Daubman. 1984. The influence of affect on categorization. *Journal of personality and social psychology* 47, 6 (1984), 1206.

[35] Alice M Isen, Paula M Niedenthal, and Nancy Cantor. 1992. An influence of positive affect on social categorization. *Motivation and Emotion* 16, 1 (1992), 65–78.

[36] Shagun Jhaver, Darren Scott Appling, Eric Gilbert, and Amy Bruckman. 2019. " Did You Suspect the Post Would be Removed?" Understanding User Reactions to Content Removals on Reddit. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–33.

[37] Shagun Jhaver, Sucheta Ghoshal, Amy Bruckman, and Eric Gilbert. 2018. Online harassment and content moderation: The case of blocklists. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 2 (2018), 1–33.

[38] Barbara E Kahn and Alice M Isen. 1993. The influence of positive affect on variety seeking among safe, enjoyable products. *Journal of Consumer Research* 20, 2 (1993), 257–270.

[39] Kate Kenski, Kevin Coe, and Stephen A Rains. 2020. Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research* 47, 6 (2020), 795–814.

[40] Wesley C King Jr, M Marie Dent, and Edward W Miles. 1991. The persuasive effect of graphics in computer-mediated communication. *Computers in Human Behavior* 7, 4 (1991), 269–279.

[41] Benedek Kurdi, Shayn Lozano, and Mahzarin R Banaji. 2017. Introducing the open affective standardized image set (OASIS). *Behavior research methods* 49, 2 (2017), 457–470.
[42] J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics* (1977), 159–174.
[43] Martin Lea and Russell Spears. 1991. Computer-mediated communication, de-individuation and group decision-making. *International journal of man-machine studies* 34, 2 (1991), 283–301.
[44] Barbara Lewandowska-Tomaszczyk. 2017. Incivility and confrontation in online conflict discourses. *Lodz Papers in Pragmatics* 13, 2 (2017), 347–367.
[45] Abner Li. 2021. Google Keep Now lets you set delightful note backgrounds on Android, iOS. Retrieved Jan. 12, 2022 from https://9to5google.com/2021/07/28/google-keep-background-notes-image/
[46] Dan Lockton, David Harrison, and Neville A Stanton. 2010. The Design with Intent Method: A design tool for influencing user behaviour. *Applied ergonomics* 41, 3 (2010), 382–392.
[47] Daniel Maier, Annie Waldherr, Peter Miltner, Gregor Wiedemann, Andreas Niekler, Alexa Keinert, Barbara Pfetsch, Gerhard Heyer, Ueli Reber, Thomas Häussler, et al. 2018. Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures* 12, 2-3 (2018), 93–118.
[48] J Nathan Matias. 2019. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences* 116, 20 (2019), 9785–9789.
[49] David Mimno, Hanna Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. Optimizing semantic coherence in topic models. In *Proceedings of the 2011 conference on empirical methods in natural language processing*. 262–272.
[50] Rocío Galarza Molina and Freddie J Jennings. 2018. The role of civility and metacommunication in Facebook discussions. *Communication studies* 69, 1 (2018), 42–66.
[51] Diana C Mutz and Byron Reeves. 2005. The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review* 99, 1 (2005), 1–15.
[52] Donna Webster Nelson. 2009. Feeling good and open-minded: The impact of positive affect on cross cultural empathic responding. *The Journal of Positive Psychology* 4, 1 (2009), 53–63.
[53] Kimberly A Neuendorf. 2017. *The content analysis guidebook*. sage.
[54] Casey Newton. 2021. Mark in the metaverse. Retrieved Jan. 12, 2022 from https://www.theverge.com/22588022/mark-zuckerberg-facebook-ceo-metaverse-interview
[55] nobackgroundimages.com. 2014. Pros and Cons of Applying Background Images For Your Web Pages. Retrieved Jan. 12, 2022 from https://www.nobackgroundimages.com/pros-and-cons-of-applying-background-images-for-your-web-pages/
[56] Zizi Papacharissi. 2004. Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New media & society* 6, 2 (2004), 259–283.
[57] Jin Seok Pyone and Alice M Isen. 2011. Positive affect, intertemporal choice, and levels of thinking: Increasing consumers' willingness to wait. *Journal of Marketing research* 48, 3 (2011), 532–543.
[58] Stephen A Rains, Kate Kenski, Kevin Coe, and Jake Harwood. 2017. Incivility and political identity on the Internet: Intergroup factors as predictors of incivility in discussions of news online. *Journal of Computer-Mediated Communication* 22, 4 (2017), 163–178.
[59] Leonie Rösner and Nicole C Krämer. 2016. Verbal venting in the social web: Effects of anonymity and group norms on aggressive language use in online comments. *Social Media+ Society* 2, 3 (2016), 2056305116664220.
[60] Ian Rowe. 2015. Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, communication & society* 18, 2 (2015), 121–138.
[61] Arthur D Santana. 2014. Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism practice* 8, 1 (2014), 18–33.
[62] Arthur D Santana. 2016. Controlling the conversation: The availability of commenting forums in online newspapers. *Journalism studies* 17, 2 (2016), 141–158.
[63] Taylor W Schmitz, Eve De Rosa, and Adam K Anderson. 2009. Opposing influences of affective state valence on visual cortical encoding. *Journal of Neuroscience* 29, 22 (2009), 7199–7207.
[64] Aslı Sezgin and Zaliha İnci Karabacak. 2019. Background of "Pinned" Images: Lifestyle Advertising in Social Media. In *The Role of Language and Symbols in Promotional Strategies and Marketing Schemes*. IGI Global, 267–284.
[65] Weber Shandwick. 2019. *Civility in America 2019: Solutions for Tomorrow*. Retrieved Jan. 12, 2022 from https://www.webershandwick.com/wp-content/uploads/2019/06/CivilityInAmerica2019SolutionsforTomorrow.pdf
[66] Eric Shouse. 2005. Feeling, emotion, affect. *M/c journal* 8, 6 (2005).
[67] Julia Silge and David Robinson. 2017. *Text mining with R: A tidy approach*. " O'Reilly Media, Inc.".
[68] Vivek K Singh, Isha Ghosh, and Darshan Sonagara. 2021. Detecting fake news stories via multimodal analysis. *Journal of the Association for Information Science and Technology* 72, 1 (2021), 3–17.

[69] Carey Hill Smith. 2019. Design Interventions — (Prototyping User Experience 2/3). Retrieved Jan. 12, 2022 from https://medium.com/@careyhillsmith/design-interventions-76a8d1827ad7
[70] Eliot R Smith and Diane M Mackie. 2008. Intergroup emotions. *Handbook of emotions* (2008), 428–439.
[71] Jennifer Stromer-Galley and Alexis Wichowski. 2011. Political discussion online. *The handbook of Internet studies* 11 (2011), 168.
[72] Leona Yi-Fan Su, Michael A Xenos, Kathleen M Rose, Christopher Wirz, Dietram A Scheufele, and Dominique Brossard. 2018. Uncivil and personal? Comparing patterns of incivility in comments on the Facebook pages of news outlets. *New Media & Society* 20, 10 (2018), 3678–3699.
[73] Nicolas Suzor, Tess Van Geelen, and Sarah Myers West. 2018. Evaluating the legitimacy of platform governance: A review of research and a shared research agenda. *International Communication Gazette* 80, 4 (2018), 385–400.
[74] Henri Ed Tajfel. 1978. *Differentiation between social groups: Studies in the social psychology of intergroup relations.* Academic Press.
[75] Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. 2010. Sentiment strength detection in short informal text. *Journal of the American society for information science and technology* 61, 12 (2010), 2544–2558.
[76] Dianne M Tice, Roy F Baumeister, Dikla Shmueli, and Mark Muraven. 2007. Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of experimental social psychology* 43, 3 (2007), 379–384.
[77] John C Turner, Michael A Hogg, Penelope J Oakes, Stephen D Reicher, and Margaret S Wetherell. 1987. *Rediscovering the social group: A self-categorization theory.* Basil Blackwell.
[78] Wesley van der Hoop. [n. d.]. Short Run Website Evaluation: The Effect of Human Aspects in Images. *Carleton University, Masters thesis* ([n. d.]).
[79] Heather A Wadlinger and Derek M Isaacowitz. 2006. Positive mood broadens visual attention to positive stimuli. *Motivation and emotion* 30, 1 (2006), 87–99.
[80] Mark Weiser and John Seely Brown. 1996. Designing calm technology. *PowerGrid Journal* 1, 1 (1996), 75–85.
[81] Adam G Zimmerman and Gabriel J Ybarra. 2016. Online aggression: The influences of anonymity and social modeling. *Psychology of Popular Media Culture* 5, 2 (2016), 181.

A RESEARCH METHODS

A.1 Identifying background images based on emotions

Following the previous positive emotions literature [23], a positive image was used to generate positive emotions. This builds on the literature on affect and emotion. Emotion is the display of a sensation that has been checked against previous experiences and affection is a non-conscious experience of intensity [66]. Although they share a pleasant subjective feel, conceptually, positive affection and positive emotion differ in that emotions require evaluation of some stimulus while affection does not [23]. Empirically, the same techniques are used to generate affection and emotions, and affective standardized images such as the International Affective Picture System (IAPS) are widely used as positive stimuli to generate positive emotions in the previous literature [63, 79]. In this study, the 10 positive images were chosen from the standardized affective image dataset, Open Affective Standardized Image Set (OASIS) developed by [41]. The 10 positive-grayscale images were created by gray-scaling the same 10 positive images chosen from the OASIS.

We validated the valence (the extent to which an emotion is positive or negative) and the arousal (intensity of the emotion, high vs low) [41] of the chosen images to make sure that those images were perceived as positive before using them in the experiment. To do this, we shortlisted 10 colored images that had positive valence scores and shared similar arousal scores from the OASIS dataset. Next, to mimic the experimental setting, we added the article text with white background over the middle part of the image. This yielded images that were similar but not the same as the OASIS dataset. Then, we grayscaled the 10 color images to create another 10 positive images in grayscale. These two image groups (colored and grayscaled) were separately labeled for two aspects (valence and arousal) to avoid any prior-exposure effects. We recruited 4 separate groups of 10 online participants (N = 40) via Prolific to provide the above-mentioned ratings. This means that each image received 10 independent ratings for valence and 10 for arousal. Following the OASIS, we used a 7-point scale (from 0 = very positive to 7 = very negative) to assess the valence and a

Table 9. Descriptive Statistics (Mean, Standard Deviation) of the User Comments Across Different Conditions

	Positive (color) (n = 157)	Positive (grayscale) (n = 137)	Control (n = 163)
Word count per comment	87.57 (43.93)	85.92 (47.27)	97.58 (45.09)
Number of sentences per comment	5.52 (3.59)	4.45 (2.41)	5.69 (2.95)
Word per sentence	18.60 (9.04)	20.42 (7.80)	18.32 (5.88)

7-point scale (from 0 = very high to 7 = very low) to assess the arousal of the 10 candidate images [41].

B RESULTS

B.1 Descriptive statistics of the user comments

Table 9 shows the mean and the standard deviation values of word count per comment, number of sentences per comment, and word per sentence for the three intervention groups. There were no statistically significant differences in terms of the above-mentioned textual features across the three conditions.

B.2 Top Words and the beta values generated by LDA topic model

Figure 3 shows the top terms in five topics and their beta values (the probability that the term would be generated for the topic) that were generated by the LDA topic model. The higher the beta, the more often the term would appear within the topic.

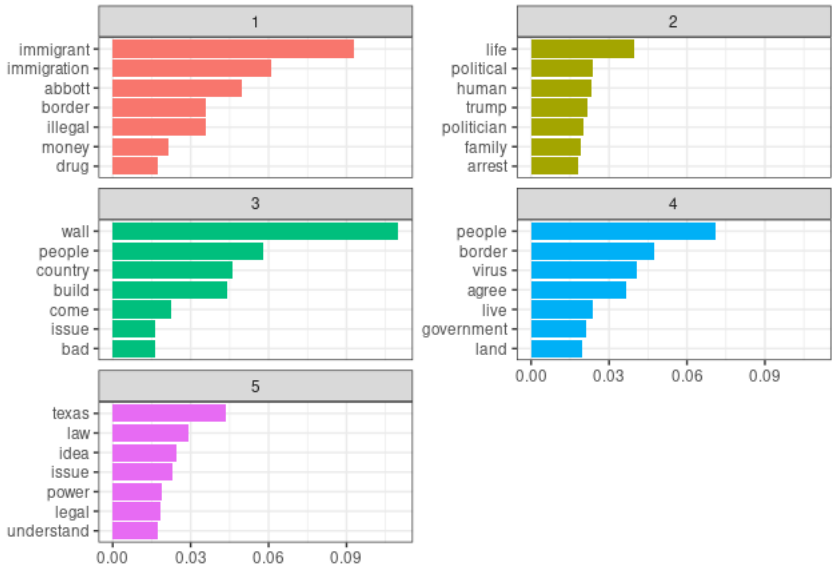

Fig. 3. Top Words and the Beta Values

Received January 2022; revised April 2022; accepted May 2022