



Community Norms



Mehreen Masood



“Broad, often unstated rules for what is acceptable on a specific social medium.”

Social Norms



Activity: In a group discuss social norms you think exist on these platforms:

Instagram

Tiktok

Facebook

How do you think they influence the way you use these platforms?



Each group picks a post from one platform, and their task is to "remix" it to fit another platform's norms without losing the original essence of the message

01

Ignore the Affordances; It's the Social Norms: How Millennials and Gen-Z Think About Where to Make a Post on Social Media

LEE TABER, SONIA DOMINGUEZ, STEVE WHITTAKER



How do these young adults, consisting of both Millennials and Gen-Z, who are using multiple media decide exactly where to post?



Multiple Factors

Goals

- Different motivations and motives
- Self presentation: Idealized and Authentic
- Poly-media swinging

Features

- Immutable, objective properties of the technology that cannot be changed without transforming the medium

Affordances

- What users perceive as possible
- Constructed from a set of features
- Anonymity, Content Persistence, Visibility etc

Social Norms

- Implicit group rules for what is acceptable on a social medium
- Helps regulate behavior and manage user expectations



Social Norms

- Implicit sanctions for non-compliance
- **TikTok:** Expressing a broad range of emotions, including negative ones, is acceptable, reflecting platform-specific norms of authenticity
- **Facebook:** Norms encourage more positive portrayals, where showing negative emotions might be subtly sanctioned or lead to negative feedback or being ignored
- Social norms are not static; they evolve and can vary significantly between different media platforms

Affordances and Features

Twitter

- Text-based
- Tweets go out to large groups of followers and can easily be retweeted to others

Snapchat

- Image and video-based
- Key Feature is ephemerality
- Posts generally go to a smaller audience

Tiktok

- Short form (15-60 sec) videos
- “For You” page, where users can view a stream of algorithmically selected content based on their viewing preferences

Instagram

- Image-based,
- Focused on creating albums of pictures or videos that go to a large group of people



Past research generally investigates the impact of a single factor on a single social medium rather than how multiple factors might influence medium posting choice.

Features, Affordances, Norms, and Goals (FANG) model

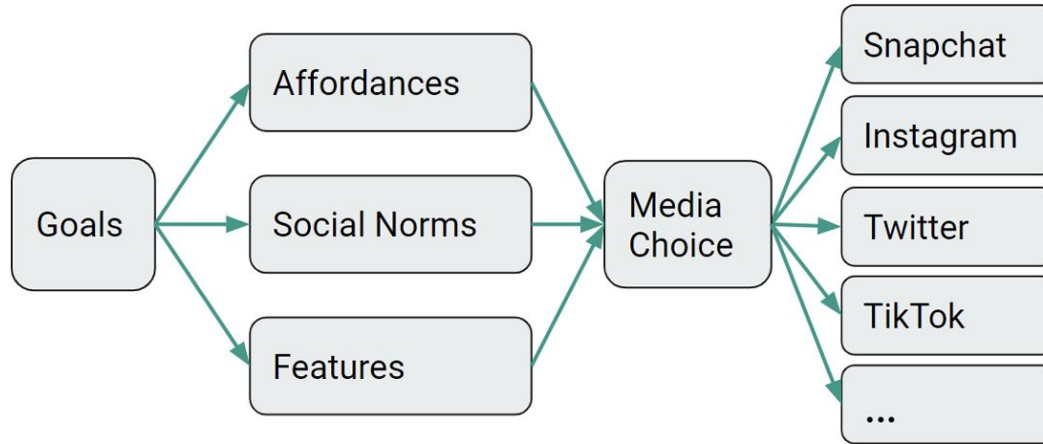


Fig.1. FANG model of Social Media Choice. Users start from a Goal, consider the Affordances, Social Norms, and Features needed, then choose the best social medium for a post.



Assesses the **FANG** model by assessing how Millennials and Gen-Z people who use multiple social media make **posting choices** between them



Research Questions

1. Does the FANG model reflect how participants actually choose between social media?
2. What factors influence participants' posting choices on social media?
3. Are any factors more important than others?
4. Are participants consistent in their media choices?

Methods

Interviews

- 19 undergraduate/graduate students
- Use at least 3 of the following media: Tiktok, Twitter, Instagram, and Snapchat
- Presented multiple scenarios, asking participants to make a posting choice
- Flowchart-building task for their mental model of social media choice

Authentic scenarios reflected Goals where the protagonist desired a genuine connection with others

Idealized scenarios reflected Goals where the protagonist wanted to express their best self

Do you think scenario-based interviews accurately capture real-life decision-making processes for choosing social media platforms? Why or why not?

Discussion



Results

- Goals:
 - Goals drive platform choice
 - Break Goals into sub-elements and see which sub-elements a particular medium best supports
 - Examples: Different goals for different audiences to achieve different gratifications
 - Revisited When No Matching Social Norm
- Social Norms:
 - Important and also drove media choice
 - Aware of a variety of Social Norms on different media
 - Violating them could lead to consequences

“I see a lot of people posting their feelings on Twitter. Versus Instagram because [...] it’s more of a picture place. If I wanted to see written posts about something I’d go to Twitter. Snapchat is the same thing, I kind of go to Snapchat for pictures and happy experiences.” (P13).



Results

- Affordances:
 - Less Important Than Goals or Social Norms
 - Described Affordances as parasitic on the Social Norms they support
 - P5: Snapchat doesn't possess certain Affordances related to Privacy and Anonymity due to Social Norms around its usage
- Features:
 - Type of content found in different media
 - Post the type of content that Social Norms dictate
 - Rather than choice being based Features, it is driven by how others typically use that medium (Norms)

These observations suggest the current FANG model needs revision, with the role of Affordances and Features being downgraded.

Revised FANG Model

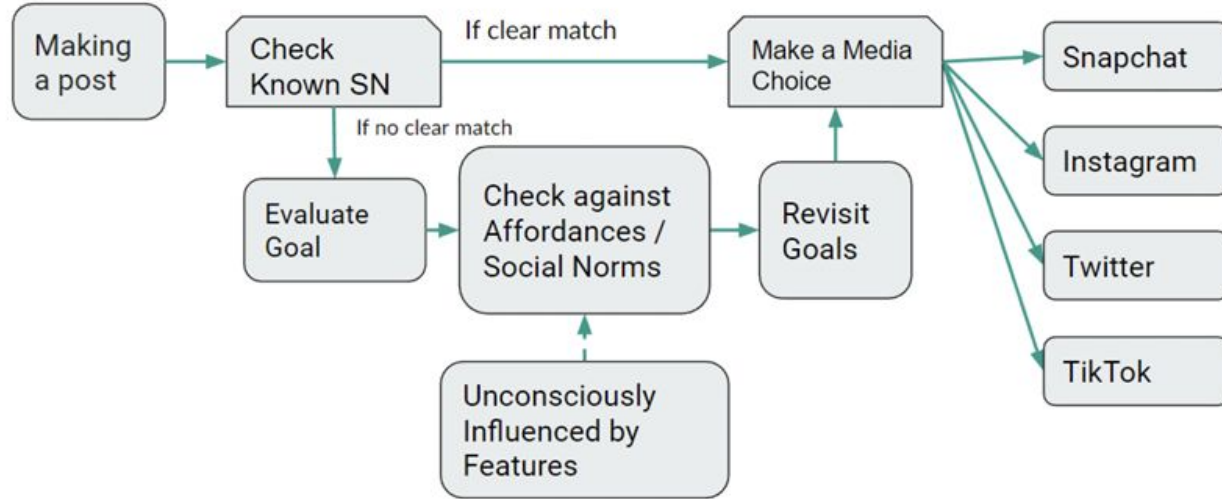


Fig. 3. Revised FANG Model: Millennial and Gen-Z young adults start by evaluating if their intended post matches an existing Social Norm. If there is an exact match, they post it on that social medium. If not, they evaluate their Goal, check Social Norms and Affordances (influenced by Features), revisit their Goal based on their thoughts, then choose a medium.

What are your thoughts on the revised model?
How can this model be used to inform the design
of technologies?

Discussion

02

Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions

Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, Jure Leskovec



Trolls are born and not made: those engaging in trolling behavior have unique personality traits and motivations.

Prior Work



Other research suggests that people can be influenced by their environment to act aggressively.

So this paper seeks to answer:



- Is trolling caused by particularly antisocial individuals or by ordinary people?
- Is trolling behavior innate, or is it situational?
- What are the conditions that affect a person's likelihood of engaging in such behavior?
- If people can be influenced to troll, can trolling spread from person to person in a community?

Trigger Mechanisms

Mood

- Negative mood and external stressors link to aggression and impaired social interactions.
- Trolling varies with time, peaking with negative mood at week's start and night.

H1: Negative mood increases a user's likelihood of trolling.

Discussion Context

- Online behavior tends to mimic preceding comments, influenced by inferred social norms.
- The topic of discussion can significantly impact the prevalence of abusive comments.

H2: The discussion context (e.g., prior troll posts by other users) affects a user's likelihood of trolling.



Influence and antisocial behavior

- People are influenced by their surroundings
- Can lead to breakdown of communities
- Negative entities are more contagious than positive ones
- Bad impressions are quicker to form and more resistant to disconfirmation

H3: Trolling behavior can spread from user to user.



What are your thoughts on the choice of these two mechanisms as focal points for the study? Do you believe these factors sufficiently cover the complexity of trolling behavior, or are there other dimensions that you feel the study could have explored?

Discussion



Contributions

- Present an experiment that shows that both negative mood and discussion context increases the likelihood of trolling
- Validate these findings with a large-scale analysis of a large online discussion community
- Use these insights to develop a predictive model that suggests that trolling may be more situational than innate

Methods

Field Experiment

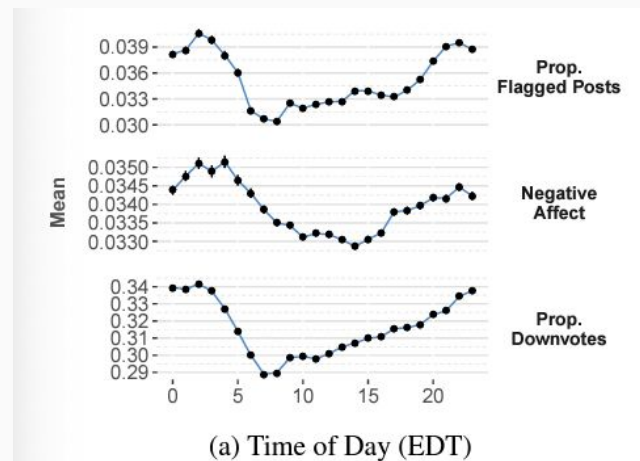
- Triggered Bad Mood in Participants
- Mood and Discussion Context
- Replicate a typical online discussion of a news article

Observational Analysis

- Studied mood through correlates: temporality
- Analysis of over 16 million posts on CNN.com

Results

- Confirmed both H1 and H2, that negative mood and the discussion context (i.e., prior troll posts) increase a user's likelihood of trolling
- Patterns of trolling behavior correspond predictably with mood
- Bad mood may also persist across discussions, but its effect diminishes with time
- Trolling in a past discussion, and participating in a discussion where trolling occurred, affects whether a user trolls in the future discussion





A MODEL OF HOW TROLLING SPREADS

Incorporates:

- **Mood Impact:** Indirectly modeled via seasonality (time of day/week) and recent posting history (time since last post, previous post flagged status).
- **Discussion Context Impact:** Modeled through previous posts' status (flagged/not, same user's posts), and discussion topic (e.g., politics).
- **Innate Trolling Tendency:** Utilizes User ID for base propensity and overall history of flagged posts.

Prediction Task: Determining the likelihood of a post being flagged using logistic regression, based on discussion context and user behavior. Random sampling (N=116,026) balanced flagged vs. non-flagged posts.



A MODEL OF HOW TROLLING SPREADS

Findings:

- **Context is Key:** Discussion context features were most predictive of trolling (AUC=0.74).
- **Mood's Role:** Indirect measures of mood were less predictive due to inability to measure mood directly.
- **Individual Propensity:** Users' history of trolling moderately predictive but does not surpass discussion context and recent history in explaining trolling behavior.

Conclusion: Trolling behavior is more situational than innate, significantly influenced by the immediate discussion environment and possibly by mood carried over from previous interactions.



Negative mood can persist and transmit trolling **norms** and **behavior** across multiple discussions

Negative **norms** can be reinforced and that downvoted users go on to downvote others



What are your thoughts about this paper? How do you think these findings translate to platforms other than CNN.com?

Discussion



How might the design and norms of different platforms (e.g., Twitter, Reddit, Facebook) influence the prevalence or form of trolling behaviors? Can you think of examples where platform design either promote or limits trolling?

Discussion

03

Norms Matter: Contrasting Social Support Around Behavior Change in Online Weight Loss Communities

Stevie Chancellor, Andrea Hu, Munmun De Choudhury



Healthy Weight Loss

Subversive Behavior
Change

surface similarities but are
motivated by radically different
intentions manifesting as
distinct behaviors



Situates norms as contextual and embedded in community language

Norms decide what behavior is appropriate as well as identifies outsiders to communities



Understand norms and language of support in weight loss OHCs to tease apart differences in how they encourage healthy and subversive behavior change



Methods

- Two Reddit communities compared: r/loseit (promotes healthy weight loss) vs. r/proED (associated with disordered eating behaviors)
- Computational linguistic analysis to identify content and context differences in community support
- Goal: Differentiate between healthy and subversive behavior change support.



Results

- Support in both communities use similar words with similar frequencies despite different norms and intentions
- Contextual analysis reveals r/loseit focuses on healthy, sustainable weight loss, whereas r/proED often veils subversive behavior change
- Content and context combined can predict community alignment with 78% accuracy, highlighting the nuanced influence of community norms



r/loseit: "You got the stretch marks when you tore your dermis...subcutaneous fat you were putting on." Highlights biological and physiological discussions on weight loss.

r/proED: "I looked in the mirror this morning and noticed my squishy fat wings...were gone!" Shows self-deprecation and unhealthy fixation on body image.



Implications for Design and Moderation

- Provided a computational linguistic approach to examine norms in social support for behavior change in different online health communities (OHCs)
- By identifying linguistic indicators of harmful norms, community moderators can more effectively target and mitigate subversive behavior change
- Understanding the language and norms of health-promoting communities can inform the development of more effective online interventions that leverage the positive aspects of social support for behavior change