

Tied In on TikTok: Tie Strength and Emotional Dynamics in Algorithmic Communities

Charles Bickham^{1,2}, Minh Duc Chu^{1,2}, Arianna Yuan¹, Valerie Lookingbill³, Ehsan Mohammadi³, Stuart Murray⁴, Kristina Lerman⁵, Emilio Ferrara^{1,2}

¹University of Southern California, Los Angeles, California, United States

²Information Sciences Institute, University of Southern California, Marina Del Rey, California, United States

³University of South Carolina, Columbia, South Carolina, United States

⁴University of California, Los Angeles, Los Angeles, California, United States

⁵Indiana University Bloomington, Bloomington, Indiana, United States

cbickham@usc.edu, dchu@isi.edu, yuana@usc.edu, lookingv@email.sc.edu, ehhsa2@sc.edu, sbmurray@mednet.ucla.edu, lerman@isi.edu, emiliofe@usc.edu

Abstract

Whether genuine communities can form on algorithmically-driven short-form video platforms like TikTok remains an open question, given that user interactions are often brief, dispersed, and difficult to trace. Building on theories of tie strength and online community formation, we examine whether eating disorder (ED) discourse on TikTok exhibits behavioral and emotional signatures of strong ties, including more frequent, reciprocal, and affectively intense interactions. In this paper, we analyze 43,040 ED-related TikTok videos and over 560,000 comments, alongside a Non-ED comparison dataset. We find that at the user-pair level, greater interaction frequency is associated with increasingly positive emotional expression, a pattern that is amplified in ED-related conversations. This trend is also reflected linguistically, with pairs that interact more frequently exhibiting more of a positive tone. At the same time, how a relationship starts matters: pairs that begin with positive exchanges usually stay mostly positive as they continue interacting, while pairs that begin negatively may add some positive exchanges over time but rarely become mostly positive. To contextualize these dynamics, we classify ED videos into three content types (Pro-Recovery, Pro-ED, and ED Experiences) and find that each exhibits distinct emotional interaction patterns. These findings suggest that dense, emotionally structured relationships can emerge within ED discourse on TikTok. More broadly, our work provides one of the first empirical demonstrations of how community-like relational dynamics form and persist on algorithmically driven short-form video platforms.

Dataset —

<https://github.com/cbickham3232/Tied-In-on-TikTok>

Introduction

Communities play a central role in social life, providing individuals with emotional support, shared identity, and access to information. Classic sociological theory conceptualizes communities as networks of *strong social ties*—close, emotionally meaningful, and reciprocal relationships that bind individuals into cohesive groups (Granovetter 1973). These

ties sustain trust and belonging, forming the foundation of social support systems that help individuals navigate personal challenges and make sense of the world. Studies of large-scale communication networks provide empirical support for theories of tie strength: individuals who frequently and persistently interact with each other also share mutual contacts, signifying their embeddedness within dense communities (Onnela et al. 2007).

As social life has increasingly moved online, digital platforms have become critical spaces for community formation and support. Online communities can promote meaningful interaction, information sharing, and solidarity, particularly among vulnerable or stigmatized populations. For individuals facing mental health challenges (Wang et al. 2017; Juarascio, Shoaib, and Timko 2010; Wang et al. 2018) or belonging to marginalized groups (Clark 2015; Jackson et al. 2018; Lubbers 2022), social media can provide both a refuge and a risk. Eating disorder (ED) communities exemplify this duality: exposure to content glorifying eating disorders as a desirable aesthetic (i.e., pro-ED content) can disseminate harmful narratives, normalize disordered behaviors, and reinforce illness identities (Bardone-Cone and Cass 2007; Harper, Sperry, and Thompson 2008; Mulveen and Hepworth 2006; Custers 2015); conversely, pro-recovery spaces allow users to share coping strategies, affirm progress, and offer mutual encouragement (Au and Cosh 2022; Juarascio, Shoaib, and Timko 2010).

Defining community on social media platforms, however, can be challenging. Social media platforms such as X (formerly known as Twitter) and Instagram rely on explicit social graphs, including follows and retweets, through which communities manifest as dense clusters of interpersonal connections (Conover et al. 2011; Wang et al. 2018; Lerman et al. 2025). These structures produce persistent, relationally anchored communities that reflect ideological, cultural, or emotional bonds. In contrast, TikTok’s social architecture is algorithmically curated rather than network-based. The platform’s “For You Page” (FYP) delivers personalized content through algorithmic recommendation rather than social ties. As a result, scholars argue that TikTok cultivates “interest clusters”—ephemeral, algorithmically-linked collectives

based on shared exposure to content rather than social relationships (Lee et al. 2022; Gerbaudo 2024; Gombar 2025). This leaves an open question whether genuine online communities can exist in an environment where users are connected primarily by algorithmic coincidence rather than intentional interaction.

This question is especially important in eating disorder related discourse, where repeated interactions can meaningfully influence emotional support, recovery, or risk. TikTok has become a major venue for ED content (Bickham et al. 2024; Herrick, Hallward, and Duncan 2021; McCashin and Murphy 2023; Griffiths et al. 2024), especially among adolescents and emerging adults who face elevated risk for disordered eating behaviors (Pruccoli et al. 2022; Sauerwein 2025; Wokke 2025; Sjöström et al. 2024). Algorithmic personalization further concentrates ED related content among vulnerable users (Griffiths et al. 2024), creating repeated opportunities for interaction that may shape both emotional tone and social dynamics over time. Whether these interactions remain superficial or develop into relationally meaningful exchanges has important implications for understanding social support, emotional contagion, and risk within algorithmically mediated spaces.

Contributions of this work. In this paper, we provide an empirical examination of whether and how relational structure emerges on TikTok at the level of user dyads. Consistent with literature (Cheng et al. 2024), we operationalize tie strength through repeated interaction between pairs of users and study how emotional expression evolves as interactions accumulate. Focusing on ED related discourse and a Non-ED dataset as baseline for comparison, we characterize how algorithmically mediated interaction gives rise to stable relational patterns in the absence of explicit social networks. Using a zero-shot prompting framework, we systematically categorize eating disorder-related content into distinct themes (Lookingbill, Mohammadi, and Cai 2023), including content promoting recovery vs eating disorders.

Guided by this perspective, we address the following research questions:

- **RQ1:** How does increasing tie strength relate to emotional expression in user-user interactions on TikTok, and does this relationship differ in ED-related and Non-ED content?
- **RQ2:** How is tie strength associated with positivity in ED-related user-pair interactions, and how does the initial interaction tone condition this relationship?
- **RQ3:** Do recovery-oriented, pro-ED, and experience-sharing discussions exhibit different patterns of emotionally intense interaction as ties strengthen, and what do these differences reveal about the formation of strong ties within distinct ED-related communities?

Using 43,040 ED-related TikTok videos and over 560,000 comments, we show that emotional expression of interactions (comments) at the user-pair level becomes more positive as interaction frequency increases, with substantially more pronounced patterns in ED-related discourse than in Non-ED discourse. In other words, repeated exchanges are associated with a more positive tone. At the same time, early

interactions play a defining role in these dynamics: relationships that begin positively tend to sustain positivity, whereas those that begin negatively may incorporate more positive exchanges over time but largely remain negative overall. These patterns are sensitive to the context of discourse, with different types of ED content exhibiting distinct emotional trajectories as ties strengthen.

Our findings challenge prevailing views of TikTok as a platform of purely transient engagement and demonstrate how community-like relational dynamics can emerge under algorithmic recommendation. More broadly, this work highlights how online communities continue to fulfill enduring social functions of support, identity, and belonging, even as the mechanisms organizing interaction shift from social networks to algorithmic curation.

Related Works

Social Ties & Online Communities Prior work in sociology and network science has examined how communities emerge through strong social ties, typically defined by frequent and reciprocal interaction and structural embeddedness within networks (Granovetter 1973; Onnela et al. 2007). These concepts have been operationalized across a range of offline and online settings to study cohesion, information flow, and social support.

On platforms like X and Instagram, online communities are commonly identified through explicit social graphs, such as follow relationships or retweet networks. In these settings, dense clusters of interaction often correspond to shared interests, identities, or ideological alignment. For example, political communication on Twitter has been shown to organize into polarized echo chambers (Conover et al. 2011). Similar network-based approaches have been applied to eating disorder related discourse, where retweet networks reveal tightly connected subcommunities, including both pro-ED groups that circulate potentially harmful content and pro-recovery groups oriented toward support and treatment (Wang et al. 2018; Lerman et al. 2025). In contrast, TikTok operates with limited reliance on explicit social connections, instead emphasizing algorithmic content recommendation. Prior research suggests that this form of personalization reshapes patterns of exposure and identity expression (Lee et al. 2022). As a result, TikTok communities are often described as “interest clusters,” formed through shared algorithmic exposure and trend participation rather than persistent interpersonal ties (Gerbaudo 2024; Gombar 2025).

Social Media & Emotions Emotions are a central part of both daily interactions and engagement on social media. Research shows that people frequently mirror the emotions they encounter online (Goldenberg and Gross 2020). Emotional states such as happiness (Fowler and Christakis 2008), loneliness (Cacioppo, Fowler, and Christakis 2009), and depression (Rosenquist, Fowler, and Christakis 2011; De Choudhury et al. 2013) have been shown to spread across social networks, shaping how individuals express themselves. Repeated exposure to a particular emotional tone increases the likelihood of responding in a similar way (Ferrara and Yang 2015), and even something as simple as view-

ing a smiling selfie can raise viewers' smile intensity by 15% (Sasaki et al. 2021).

These dynamics are also evident in specific online communities. For example, participation in depression-related forums can intensify emotional experiences, where positive interactions promote hope while negative ones strengthen feelings of hopelessness and fear (Tang, Yu, and Yao 2021). Exposure to toxic content similarly provokes strong affective reactions, with more severe toxic replies linked to greater expressions of anger and sadness (Aleksandric et al. 2024). Emotional expression also influences patterns of engagement on social platforms. In cancer-related Twitter discussions, joy, sadness, and hope tended to receive more likes, while joy and anger were more likely to be amplified through retweets (Wang and Wei 2020).

Social Media & Mental Health Communities Prior research has examined how social media enables both supportive and harmful community dynamics around mental health. Frequent and sustained engagement has been linked to stronger online participation (Chhikara and Chakraborty 2024) and social connectedness among individuals with mental illness (Brusilovskiy et al. 2016; Gowen et al. 2012). In the context of EDs, scholars have analyzed how online communities form, interact, and evolve across platforms. Early work on X and Instagram revealed that pro-eating disorder networks display strong homophily and insular communication patterns, with recovery-oriented users forming smaller, less connected subgroups (Wang et al. 2017, 2018). Other studies conceptualize the growth of pro-ana discourse as an online radicalization process that reinforces harmful norms through repeated exposure (Lerman et al. 2023). On Instagram, pro-ana and thinpiration content has become increasingly mainstreamed, embedding ideals of self-discipline and restriction within broader youth culture (Ging and Garvey 2018), while recovery spaces offer peer validation and representation but face risks of unmoderated harmful content (Au and Cosh 2022). Network analyses of ED discourse also highlight how influential users amplify unhealthy behaviors and emerging trends such as “#meanspo,” reflecting the shifting language of these communities (Nova et al. 2022). Finally, large-scale computational studies show that mental illness severity in pro-ED posts has increased over time, underscoring the urgency of understanding how community structures and repeated engagement shape online mental health discourse (Chancellor et al. 2016).

Methodology

Data

For this study, we utilize our publicly-available TikTok eating disorder dataset (Bickham et al. 2025). This collection contains 43,040 TikTok videos accompanied by 388,332 top-level comments and 188,379 replies, spanning the period from January 2019 to June 2024. To provide a comparison group, a Non-ED dataset was collected using the same keywords employed in collecting the EDTok corpus but applying a NOT operator to exclude eating disorder-related terms. The resulting Non-ED dataset consists of 37,530 videos, 203,709 original comments, and 102,902 replies.

Data for the Non-ED dataset were retrieved via the TikTok Research API. Each video entry contains its own ID, the text content, the date of the comment and, where relevant, a reply ID. This structure makes it possible to reconstruct conversation threads and examine patterns of interaction.

In addition, we incorporate an independently annotated dataset of 200 TikTok videos from (Lookingbill, Mohammedi, and Cai 2023), which we use as ground truth for evaluation. Each entry in this dataset contained the original mp4 video file, the unique video ID, and human-provided annotations labeling the video into one of four eating disorder-related content types. This dataset provides an external benchmark that allows us to evaluate classification performance and to calibrate our zero-shot prompt before scaling to the full dataset.

Measuring Emotions

To analyze emotional expression in the comments, we employed two emotion classification models: Demultiplexer (Demux) (Chochlakis et al. 2023) and RoBERTa-GoEmotions.¹ Demux is a multi-label emotion recognition model derived from SpanEmo (Alhuzali and Ananiadou 2021). It encodes two inputs—the emotion categories and the text sample—to generate contextual embeddings that yield probability scores across emotions. In contrast, RoBERTa-GoEmotions is a transformer-based classifier trained on the GoEmotions dataset (Demszky et al. 2020), which contains 28 emotion classes. For comparability, we aligned its predicted categories with those of Demux and grouped each emotion by polarity (positive, neutral, or negative) to ensure consistent aggregation across models (see Table 4 in the Appendix).

Both models were applied to the full set of comments, targeting emotions such as anger, anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise, trust, and none. To assess accuracy, we manually annotated a random sample of 500 comments. On this benchmark, Demux achieved F1 scores of 92.26% for comments, while RoBERTa-GoEmotions performed considerably worse with a score of 33.62% (see Table 1). Agreement between the two models was also limited, with Cohen's Kappa of 0.21 for comments and overall agreement rates of 27.80%. Given its substantially stronger alignment with human annotations, Demux was selected as the primary model for emotion detection in this study.

Linguistic Cues

We applied the 2022 release of the Linguistic Inquiry and Word Count (LIWC) tool (Boyd et al. 2022) to examine psychological and linguistic characteristics of the comments. LIWC is a widely used software package for text analysis that identifies the proportion of words in a document that fall into theoretically and empirically validated categories. These categories span both structural aspects of language (e.g., pronouns, articles, and function words) as well as psychological dimensions such as affective expression, cogni-

¹https://huggingface.co/SamLowe/roberta-base-go_emotions

	Demux			RoBERTa-Go		
	P	R	F1	P	R	F1
Results	92.77	92.20	92.26	47.33	27.59	33.62

Table 1: Precision, recall, and F1-score for Demux and RoBERTa-GoEmotions on the annotated sample of 500 comments.

tive processes, and social interaction. Access to LIWC requires a license, which we obtained for this research.

Our analysis concentrated on the Positive Tone and Negative Tone categories, which capture sentiment rather than discrete emotions. These dimensions include not only words directly linked to emotional states (e.g., happy, joy, sad, angry) but also terms associated with emotional contexts (e.g., birthday, beautiful, kill, funeral). For example, in the transcript “I am very happy,” one out of four words is coded as Positive Tone, resulting in a score of 25. LIWC was selected for this task because it provides a continuous lexical measure of sentiment that can be aggregated across multiple interactions, a property that has been shown to be effective for estimating aggregate sentiment in large collections of short social media texts. This property is particularly useful for relationship-level analysis, where tone must be summarized across many short comments exchanged between the same pair of users. Across our dataset, the mean Positive Tone score was 11.36 and the mean Negative Tone score was 2.01, both of which exceed the averages reported in prior Twitter-based analyses (6.05 and 1.85, respectively) (Boyd et al. 2022). These differences provide a comparative benchmark for interpreting the prominence of sentiment expression in TikTok eating disorder-related content.

Content Classification

(Lookingbill, Mohammadi, and Cai 2023) developed a taxonomy to categorize eating disorder-related content on TikTok, distinguishing harmful material that promotes eating disorders (pro-ED), videos sharing personal experiences with eating disorders, and content emphasizing social support and recovery (pro-recovery). This taxonomy enables a systematic assessment of eating disorder-related discourse on the platform.

The taxonomy was developed through thematic analysis of 200 TikTok videos related to eating disorders, which the authors coded to identify recurring patterns. From this process, they distilled 23 subthemes into four overarching categories:

- **Pro-ED Content:** Encouraging eating disorders by promoting them as a lifestyle choice, seeking support to maintain disordered behaviors, or urging others to adopt them.
- **Pro Recovery Content:** Sharing recovery stories by offering personal experiences, advice, encouragement, and challenges faced during the recovery process.
- **ED Experiences:** Sharing personal experiences with eating disorders to raise awareness or educate others, includ-

Category	Accuracy	Precision	Recall	F1
Recovery Content	78.26%	48.65%	78.26%	60.00%
Pro-ED Content	82.69%	93.48%	82.69%	87.76%
ED Experiences	63.22%	74.32%	63.22%	68.32%
Social Support	15.38%	50.00%	15.38%	23.53%

Table 2: Ground truth classification results.

Category	Precision	Recall	F1-Score	# Videos
Recovery Content	80.18%	95.70%	87.25%	22,425
Pro-ED Content	66.67%	100.00%	80.00%	1,247
ED Experiences	94.20%	78.31%	85.53%	11,987
Social Support	75.00%	37.50%	50.00%	2,472

Table 3: Classification results on the EDTok dataset.

ing how they developed and the physical and emotional symptoms they face.

- **Social Support:** Creators give, seek, or share experiences of receiving social support both online and offline.

We first applied Google Gemini² in a zero-shot setting (see Figure 7 in the Appendix) to the existing annotated dataset of 200 videos from (Lookingbill, Mohammadi, and Cai 2023). Classification performance on this dataset is reported in Table 2. Based on this evaluation, we applied the same zero-shot prompt to label the full ED dataset. To further validate labeling performance on the ED data, two authors independently constructed a ground truth set of 200 videos. Classification results on this manually annotated subset are shown in Table 3. Given the relatively low classification performance across both evaluation datasets, the Social Support category was excluded from subsequent analyses.

Results

RQ1: Is Tie Strength Related to Stronger Emotional Expression in TikTok Interactions?

We begin by operationalizing *tie strength* at the level of user dyads as the total number of comments exchanged between two users across all videos in the dataset. This measure captures persistence of interaction. The distribution of user-user exchanges in both datasets is highly skewed, with most pairs interacting only once but a small subset engaging in many repeated interactions (see Figure 6 in the Appendix). Notably, the ED dataset exhibits a heavier tail than the Non-ED dataset, indicating a greater concentration of dyads with sustained, repeated interactions, with some user pairs interacting hundreds of times.

To examine how emotional expression varies across these levels of relational intensity, we group user pairs into discrete interaction bins corresponding to increasing tie strength: 1, 2–5, 6–10, 11–19, and 20 or more exchanged

²<https://gemini.google.com/>

comments. For each bin, we compute the aggregate emotional tone across all comments exchanged within those relationships.

We analyze the Non-ED dataset to establish a baseline pattern. As shown in Figure 1a, user-pairs that interact more frequently exhibit a higher share of positive emotional interactions and a corresponding decline in negative interactions. This trend indicates that repeated interactions between the same users are increasingly characterized by positivity. This association is supported by a Chi Square Test of Independence between interaction bins and emotional tone categories. Examination of the standardized residuals (Figure 2a) shows that user-pairs with more than one interaction are overrepresented in the positive tone category, with the strongest effect observed in the 20+ bin (residual = 10.2, $p < 0.05$). These same user-pairs are underrepresented in the negative tone category (residual = -5.6, $p < 0.05$), indicating that increased interaction frequency is associated with a shift away from negative expression in Non ED discourse.

Next, we apply the same analysis to the ED dataset. Figure 1b shows a similar pattern, but with substantially stronger trend. User-pairs in ED discourse display a sharper increase in positive emotional tone and a more pronounced decline in negative tone with interaction frequency. The Chi Square Test of Independence further confirms this pattern. Standardized residuals for ED related interactions (Figure 2b) reveal strong overrepresentation of positive tone among relationship pairs in the 20+ interaction bin (residual = 62.2, $p < 0.05$), alongside substantial underrepresentation of negative tone (residual = -41.0, $p < 0.05$).

Summary. Across both datasets, stronger tie-strength is associated with systematically more positive and less negative emotional expression at the relationship level. This association is present in Non-ED discourse but is significantly amplified in ED related interactions, particularly among frequently interacting user-pairs.

RQ2: Linguistic Tone and Initial Interaction Effects in ED Relationships

We next examine how the emotional tone and interaction trajectories vary across user pair relationships for ED TikTok content. To quantify emotional tone at the relationship level, we use LIWC-derived sentiment scores rather than the Demux classifier, since LIWC provides a continuous lexical signal that can be aggregated across many short interactions between the same pair of users. We compute a *tone score* for each user pair (i, j) . Let \mathcal{C}_{ij} be the set of all comments exchanged between i and j in either direction ($i \rightarrow j$ or $j \rightarrow i$), and let $N_{ij} = |\mathcal{C}_{ij}|$. For each comment $c \in \mathcal{C}_{ij}$, let p_c and n_c denote the positive-tone and negative-tone values from LIWC. We define:

$$\bar{p}_{ij} = \frac{1}{N_{ij}} \sum_{c \in \mathcal{C}_{ij}} p_c, \quad \bar{n}_{ij} = \frac{1}{N_{ij}} \sum_{c \in \mathcal{C}_{ij}} n_c$$

$$\text{ToneScore}_{ij} = \bar{p}_{ij} - \bar{n}_{ij} = \frac{1}{N_{ij}} \sum_{c \in \mathcal{C}_{ij}} (p_c - n_c)$$

A positive score indicates that the pair’s exchanges are, on average, more positive than negative, while a negative score indicates the opposite. This formulation takes into account *all* comments between each user pair, ensuring that interaction frequency is directly reflected in the score. As shown in Figure 3, tone scores generally shift upward and display less variance as the number of interactions increases. Importantly, once pairs surpass three reciprocal interactions (see Figure 8 in the Appendix), their tone scores are consistently positive. This suggests that repeated and reciprocated engagement between the same users reinforces positivity at the linguistic level.

To further contextualize these patterns, we examine how interaction patterns differ depending on whether a relationship begins with a positive or negative exchange in ED-related content. We restrict this analysis to user pairs with at least two interactions and classify each relationship as either positive-start or negative-start based on the tone of its initial exchange. For each group, we analyze interaction composition across increasing levels of tie strength.

For negative-start relationships relatively few pairs transition to positive majority, defined as cases where more than half of all interactions between two users are positive in tone. Across interaction bins, only a minority of negative-start relationships reach this threshold, even among pairs with high interaction frequency (see Figure 4). This indicates that while negative-start relationships may incorporate positive exchanges over time, these interactions often coexist with sustained negativity rather than replacing it. In contrast, positive-start relationships are far more likely to maintain a positive majority as interactions accumulate. Even at lower interaction frequencies, most positive-start relationships remain above the majority-positive threshold, reflecting a tendency for early positivity to persist as the dominant emotional mode of interaction.

Summary. Across user pair relationships, repeated interaction is associated with increasingly positive tone, indicating that tie strengthening reinforces affective alignment over time. However, early emotional conditions matter: relationships that begin positively tend to sustain positivity as interactions accumulate, while those that begin negatively rarely transition to a predominantly positive state despite incorporating some positive exchanges.

RQ3: Differences in Relational Emotional Dynamics Across ED Discussion Spaces

Lastly, we examine whether the emotional expression of frequent interactions differs across distinct ED-related discussion spaces. Following (Lookingbill, Mohammadi, and Cai 2023), we categorize ED TikTok videos into three content types: Pro-Recovery, Pro-ED, and ED Experiences. For each category, we analyze how emotional expression evolves within *user-user relationships* as tie strength increases. Figure 9 in the Appendix shows the distribution of interaction frequencies across content types, highlighting substantial variation in the prevalence and depth of repeated dyadic interactions.

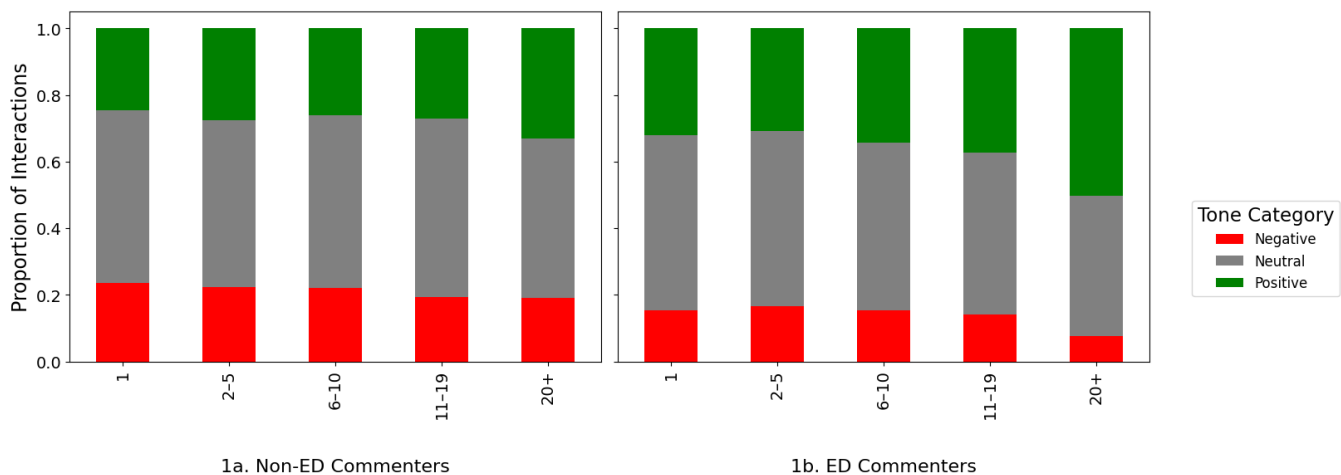


Figure 1: Distribution of emotional tone across interaction bins for Non-ED (1a) and ED (1b) commenters. Each bar represents the proportion of negative (red), neutral (gray), and positive (green) comments within a given interaction frequency range. Compared to the Non-ED baseline, ED commenters show a sharper decline in negativity and a stronger amplification of positivity as interaction frequency increases, particularly in the 20+ bin.

Pro-Recovery Content Comments on Pro-Recovery videos often express affirmation, pride, reassurance, and mutual recognition of progress, including references to treatment milestones, fear food challenges, and personal recovery experiences. As shown in Figure 5a, Pro-Recovery content exhibits a strong association between tie strength and emotional tone. Among user pairs with more than 20 interactions, positive emotion is substantially overrepresented (residual = +25.4, $p < 0.05$), while negative emotion is strongly underrepresented (residual = -25.0, $p < 0.05$). This emotional pattern shows that as relationships deepen in Pro-Recovery spaces, they also become stronger emotionally, consistent with formation of strong social ties.

Pro-ED In contrast, interactions around Pro-ED content exhibit a markedly different emotional profile. Comments between recurring user pairs primarily involve transactional or informational exchanges, such as calorie counts, weight references, or body comparisons—rather than affective engagement. The overall tone is less emotional compared to other types of ED content. As shown in Figure 5b, standardized residuals remain close to zero across all interaction bins. Even among the most frequently interacting pairs, emotional expression does not systematically intensify in either a positive or negative direction (negative residual = -2.3, $p < 0.05$). These patterns suggest that repeated interaction in Pro-ED spaces does not translate into strong ties.

ED Experiences ED Experiences content consists primarily of personal disclosures related to ongoing struggles, relapses, or lived experiences with eating disorders. Comments often include expressions of empathy and shared vulnerability, alongside disclosures of distress or ambivalence. As shown in Figure 5c, positivity is modestly overrepresented for user pairs who had one comment (residual = +4.7, $p < 0.05$), indicating that early engagement is often supportive. However, this pattern attenuates with increased interac-

tion frequency. These patterns suggest that while ED Experience content elicits supportive responses early in interaction histories, sustained engagement corresponds to more complex and emotionally mixed relational dynamics.

Summary The relationship between tie strength and emotional expression varies substantially across ED content types, reflecting differences in discourse norms and interactional goals. Pro-Recovery interactions show strong emotional reinforcement, with increased tie strength associated with amplified positivity, whereas Pro-ED content exhibit stable, largely neutral or balanced emotional patterns regardless of interaction frequency. ED Experiences occupy an intermediate position, where early interactions are often supportive but sustained engagement leads to more emotionally mixed dynamics rather than convergence toward a single dominant tone.

Discussion & Conclusions

This study provides one of the first systematic examinations of how online communities function on an algorithmically curated platform such as TikTok, where explicit community boundaries and relational structures are often opaque. We observe that repeated interactions are associated with increasingly positive emotional expression on TikTok, with this pattern especially pronounced in eating disorder related discourse. We further show that this relationship is not uniform across eating disorder content, but varies systematically when ED-related videos are categorized into Pro-Recovery, Pro-ED, and ED Experiences, with each category exhibiting distinct emotional interaction patterns as ties strengthen. These findings extend strong-tie theory to algorithmically mediated environments, where relational structure emerges without explicit social graphs.

Our results have several implications for the study and governance of online communities, especially in vulnera-



Figure 2: Standardized residuals from Chi-Square Tests of Independence for Non-ED commenters (2a) and ED commenters (2b). Positive values (red) indicate overrepresentation of tone categories relative to expectation, while negative values (blue) indicate underrepresentation. ED commenters show a stronger amplification of positivity and sharper underrepresentation of negativity at higher interaction bins (20+), compared to the more moderate shifts observed in Non-ED commenters.

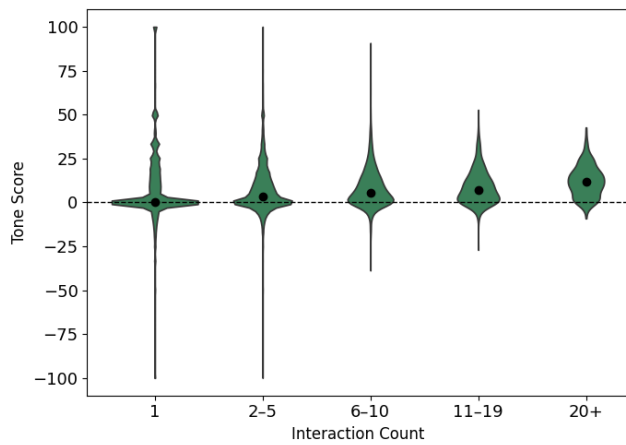


Figure 3: Tone score distributions across user pair relationships by interaction bin for ED TikTok content. Violin plots show the distribution of relationship-level tone scores within each bin, with black dots marking medians and the dashed line indicating neutrality. Distributions shift toward higher values as interaction frequency increases, reflecting more positive and more stable emotional expression.

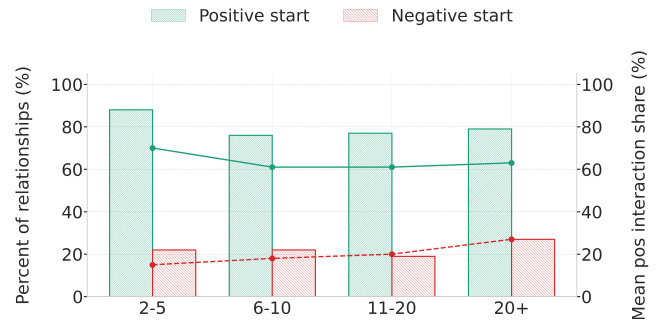


Figure 4: Relationship-level interaction patterns by initial emotional tone. Bars (left y-axis) show the percentage of relationships that are mostly positive at different interaction levels, while lines (right y-axis) show the average share of positive interactions.

ble discourse spaces. First, the tendency for emotional expression to become more positive with repeated interaction suggests that meaningful social structure can arise even in platforms dominated by algorithmic content recommendation. Prior research on online mental health communities similarly finds that sustained interaction and peer engagement can create supportive emotional environments and reinforce community norms (Naslund et al. 2016; Barak, Boniel-Nissim, and Suler 2008; Wright et al. 2013; Chao et al. 2025). This highlights the importance of analyzing interaction-level dynamics, rather than relying solely on net-

work representations, when studying community formation and emotional influence on short-form video platforms.

Second, the strong dependence of emotional trajectories on early interaction tone indicates that initial exchanges play a critical role in shaping long-term relational dynamics. This pattern aligns with relational development theory, which emphasizes that early interactions often establish expectations and norms that influence how relationships evolve over time (Altman and Taylor 1973). Evidence from online mental health communities similarly suggests that early supportive responses can have a “contagious” effect, where users who receive social support in their initial interactions are more likely to return and subsequently provide empathic support to others (Chen and Xu 2021). For platform designers and moderators, this suggests that early interventions in comment threads may be particularly impactful in sensitive contexts. Encouraging supportive responses early in interaction

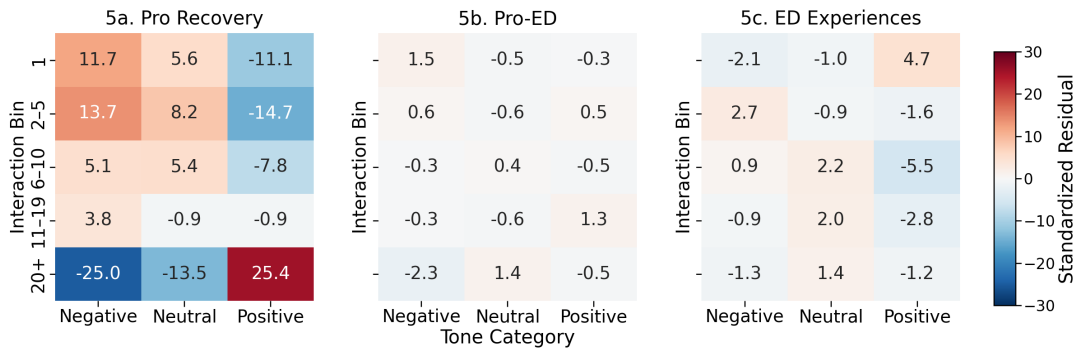


Figure 5: Standardized residuals from Chi-Square Tests of Independence for emotional tone (negative, neutral, positive) across interaction bins in ED-related TikTok content. Panels show differences by content type: (5a) Pro Recovery Content, (5b) Pro-ED, (5c) ED Experiences. Recovery content exhibits strong amplification of positivity and suppression of negativity at high interaction frequencies, while Pro-ED remain near neutral with little systematic change. ED Experiences show early supportive responses that diminish with more interactions.

histories could help stabilize more positive emotional climates, whereas unchecked negative exchanges may entrench persistently adverse relational patterns.

Third, the heterogeneity observed across eating disorder content types underscores the need for context-aware moderation and intervention strategies. Pro-Recovery spaces appear to amplify supportive affect through repeated engagement, while Pro-ED spaces do not exhibit comparable emotional reinforcement dynamics. The amplification of positive expression within Pro-Recovery spaces echoes literature on social support and buffering, which argues that sustained, reciprocal interpersonal exchanges can provide emotional and instrumental resources that reduce distress and promote resilience (Cohen and Wills 1985). Treating ED-related discourse as emotionally uniform risks obscuring these distinctions and may result in ineffective or even counterproductive interventions. Instead, moderation approaches should recognize that interactions within different types of ED content exhibit distinct emotional tones.

Building on these findings, future research could improve content classification by fine-tuning models directly on ED-related videos, allowing for more precise differentiation between nuanced and overlapping content categories. Beyond content, relationship-level analyses could be extended to examine directionality and reciprocity within user pairs, distinguishing mutual interactions from one-sided engagement patterns. Finally, integrating multimodal signals such as video features, on-screen text, or temporal posting behavior could provide a richer understanding of how emotional dynamics unfold across interaction modalities on short-form video platforms.

Limitations. Although our findings are promising, this study has several limitations. First, our analysis relied primarily on a single emotion classifier, which (despite being quite accurate) may not fully capture linguistic variation across different comments. While we partially addressed this by testing a second model and manually labeling 500 video descriptions and 500 comments, future research should expand annotation efforts and incorporate multiple

classifiers to reduce bias and improve robustness. Second, our relationship-level tone score relies on LIWC, a lexicon-based method that may be sensitive to vocabulary variation and informal language commonly used on social media. While lexicon approaches provide transparent and interpretable sentiment measures, they may miss contextual nuances such as sarcasm or platform-specific expressions. In this study, aggregating tone across multiple interactions within each user pair helps mitigate these effects, though future work could explore sentiment models trained directly on social media text. Third, the F1 scores for labeling ground truth video categories (Pro-Recovery, Pro-ED, ED Experiences, Social Support) were modest, reflecting the challenge of differentiating nuanced and overlapping content types. Fourth, most user pair relationships in our data involve only a single interaction, making repeated interactions concentrated among a more active subset of users. As a result, emotional patterns observed at higher interaction levels may reflect this subset rather than the broader population, despite our use of interaction binning and a Non-ED baseline for comparison. Fifth, algorithmic amplification, affective homophily, or selective engagement may also contribute to the observed emotional patterns and cannot be fully disentangled in this observational setting. Finally, this study is correlational and does not support causal inference. Despite these limitations, the consistency of observed patterns across multiple complementary analyses and across distinct datasets supports the reliability and robustness of our findings.

Ethical Statements. All data in this study was publicly available and collected in compliance with TikTok’s Research API Terms of Service. No identifiable user information, such as usernames or profile details, was included in the analysis. User identifiers were hashed to prevent re-identification while preserving dataset integrity. The comments underwent de-identification, with personally identifiable information (e.g., usernames, direct mentions, links) removed or replaced with generic placeholders. Additionally, all analyses were conducted on aggregated data to further minimize the risk of tracing individual user interactions.

These measures were taken to minimize potential risks and ensure ethical research practices. The authors declare no competing interests.

Acknowledgements

This project was in part supported by the NSF (Award Number 2331722). We confirm that all text in this paper was written by the authors. ChatGPT was used for grammar, spelling, and clarity improvements, while all content and intellectual contributions remain those of the human authors.

References

- Aleksandric, A.; Roy, S. S.; Pankaj, H.; Wilson, G. M.; and Nilizadeh, S. 2024. Users' Behavioral and Emotional Response to Toxicity in Twitter Conversations. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, 29–42.
- Alhuzali, H.; and Ananiadou, S. 2021. SpanEmo: Casting multi-label emotion classification as span-prediction. *arXiv preprint arXiv:2101.10038*.
- Altman, I.; and Taylor, D. A. 1973. *Social penetration: The development of interpersonal relationships*. Holt, Rinehart & Winston.
- Au, E. S.; and Cosh, S. M. 2022. Social media and eating disorder recovery: An exploration of Instagram recovery community users and their reasons for engagement. *Eating Behaviors*, 46: 101651.
- Barak, A.; Boniel-Nissim, M.; and Suler, J. 2008. Fostering empowerment in online support groups. *Computers in human behavior*, 24(5): 1867–1883.
- Bardone-Cone, A. M.; and Cass, K. M. 2007. What does viewing a pro-anorexia website do? An experimental examination of website exposure and moderating effects. *International Journal of Eating Disorders*, 40(6): 537–548.
- Bickham, C.; Kazemi-Nia, K.; Luceri, L.; Lerman, K.; and Ferrara, E. 2024. Hidden in plain sight: Exploring the intersections of mental health, eating disorders, and content moderation on tiktok. *arXiv preprint arXiv:2404.15457*.
- Bickham, C.; Ramirez-Gonzalez, B.; Chu, M. D.; Lerman, K.; and Ferrara, E. 2025. Edtok: A dataset for eating disorder content on tiktok. *ICWSM'25 Workshop on Data for the Wellbeing of Most Vulnerable*.
- Boyd, R. L.; Ashokkumar, A.; Seraj, S.; and Pennebaker, J. W. 2022. The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*, 10: 1–47.
- Brusilovskiy, E.; Townley, G.; Snethen, G.; and Salzer, M. S. 2016. Social media use, community participation and psychological well-being among individuals with serious mental illnesses. *Computers in Human Behavior*, 65: 232–240.
- Cacioppo, J. T.; Fowler, J. H.; and Christakis, N. A. 2009. Alone in the crowd: the structure and spread of loneliness in a large social network. *Journal of personality and social psychology*, 97(6): 977.
- Chancellor, S.; Lin, Z.; Goodman, E. L.; Zerwas, S.; and De Choudhury, M. 2016. Quantifying and predicting mental illness severity in online pro-eating disorder communities. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, 1171–1184.
- Chao, M.; Wang, Y.; Liu, J.; and Gu, L. 2025. Relationship between online social support and problematic internet use: a three-level meta-analysis. *BMC psychology*, 13(1): 745.
- Chen, Y.; and Xu, Y. 2021. Exploring the Effect of Social Support and Empathy on User Engagement in Online Mental Health Communities. *International Journal of Environmental Research and Public Health*, 18(13).
- Cheng, X.; Yang, C.; Zhao, Y.; Wang, Y.; Karimi, H.; and Derr, T. 2024. A comprehensive analysis of social tie strength: Definitions, prediction methods, and future directions. *arXiv e-prints*, arXiv–2410.
- Chhikara, G.; and Chakraborty, A. 2024. # LetsTalk: Understanding Social Media Usage of Self-Harm Users in India. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, 289–301.
- Chochlakis, G.; Mahajan, G.; Baruah, S.; Burghardt, K.; Lerman, K.; and Narayanan, S. 2023. Leveraging label correlations in a multi-label setting: A case study in emotion. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. IEEE.
- Clark, M. 2015. Black Twitter: Building connection through cultural conversation. *Hashtag publics: The power and politics of discursive networks*, 205–218.
- Cohen, S.; and Wills, T. A. 1985. Stress, social support, and the buffering hypothesis. *Psychological bulletin*, 98(2): 310.
- Conover, M.; Ratkiewicz, J.; Francisco, M.; Gonçalves, B.; Menczer, F.; and Flammini, A. 2011. Political polarization on twitter. In *Proceedings of the international aai conference on web and social media*, volume 5, 89–96.
- Custers, K. 2015. The urgent matter of online pro-eating disorder content and children: clinical practice. *European journal of pediatrics*, 174(4): 429–433.
- De Choudhury, M.; Gamon, M.; Counts, S.; and Horvitz, E. 2013. Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media*, volume 7, 128–137.
- Demszky, D.; Movshovitz-Attias, D.; Ko, J.; Cowen, A.; Nemade, G.; and Ravi, S. 2020. GoEmotions: A Dataset of Fine-Grained Emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4040–4054.
- Ferrara, E.; and Yang, Z. 2015. Measuring emotional contagion in social media. *PloS one*, 10(11): e0142390.
- Fowler, J. H.; and Christakis, N. A. 2008. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *Bmj*, 337.
- Gerbaudo, P. 2024. TikTok and the algorithmic transformation of social media publics: From social networks to social interest clusters. *New media & society*, 14614448241304106.

- Ging, D.; and Garvey, S. 2018. ‘Written in these scars are the stories I can’t explain’: A content analysis of pro-ana and thinspiration image sharing on Instagram. *New Media & Society*, 20(3): 1181–1200.
- Goldenberg, A.; and Gross, J. J. 2020. Digital emotion contagion. *Trends in cognitive sciences*, 24(4): 316–328.
- Gombar, M. 2025. Algorithmic echo chambers and user interaction in understanding the societal impact of personalized content. In *2025 MIPRO 48th ICT and Electronics Convention*, 1477–1483. IEEE.
- Gowen, K.; Deschaine, M.; Gruttadara, D.; and Markey, D. 2012. Young adults with mental health conditions and social networking websites: seeking tools to build community. *Psychiatric rehabilitation journal*, 35(3): 245.
- Granovetter, M. S. 1973. The strength of weak ties. *American journal of sociology*, 78(6): 1360–1380.
- Griffiths, S.; Harris, E. A.; Whitehead, G.; Angelopoulos, F.; Stone, B.; Grey, W.; and Dennis, S. 2024. Does TikTok contribute to eating disorders? A comparison of the TikTok algorithms belonging to individuals with eating disorders versus healthy controls. *Body image*, 51: 101807.
- Harper, K.; Sperry, S.; and Thompson, J. K. 2008. Viewership of pro-eating disorder websites: Association with body image and eating disturbances. *International Journal of Eating Disorders*, 41(1): 92–95.
- Herrick, S. S.; Hallward, L.; and Duncan, L. R. 2021. “This is just how I cope”: An inductive thematic analysis of eating disorder recovery content created and shared on TikTok using #EDrecovery. *International journal of eating disorders*, 54(4): 516–526.
- Jackson, S. J.; Clark, M. D.; Freelon, D.; and Lopez, L. 2018. How Black Twitter and other social media communities interact with mainstream news.
- Juarascio, A. S.; Shoaib, A.; and Timko, C. A. 2010. Pro-eating disorder communities on social networking sites: a content analysis. *Eating disorders*, 18(5): 393–407.
- Lee, A. Y.; Mieczkowski, H.; Ellison, N. B.; and Hancock, J. T. 2022. The algorithmic crystal: Conceptualizing the self through algorithmic personalization on TikTok. *Proceedings of the ACM on Human-computer Interaction*, 6(CSCW2): 1–22.
- Lerman, K.; Chu, M. D.; Bickham, C.; Luceri, L.; and Ferrara, E. 2025. Safe spaces or toxic places? Content moderation and social dynamics of online eating disorder communities. *EPJ Data Science*, 14(1): 55.
- Lerman, K.; Karnati, A.; Zhou, S.; Chen, S.; Kumar, S.; He, Z.; Yau, J.; and Horn, A. 2023. Radicalized by thinness: Using a model of radicalization to understand pro-anorexia communities on twitter. *arXiv preprint arXiv:2305.11316*.
- Lookingbill, V.; Mohammadi, E.; and Cai, Y. 2023. Assessment of accuracy, user engagement, and themes of eating disorder content in social media short videos. *JAMA Network Open*, 6(4): e238897–e238897.
- Lubbers, M. J. 2022. Social networks and the resilience of marginalized communities. In *A research agenda for social networks and social resilience*, 1–16. Edward Elgar Publishing.
- McCashin, D.; and Murphy, C. M. 2023. Using TikTok for public and youth mental health—A systematic review and content analysis. *Clinical Child Psychology and Psychiatry*, 28(1): 279–306.
- Mulveen, R.; and Hepworth, J. 2006. An interpretative phenomenological analysis of participation in a pro-anorexia internet site and its relationship with disordered eating. *Journal of health psychology*, 11(2): 283–296.
- Naslund, J. A.; Aschbrenner, K. A.; Marsch, L. A.; and Bartels, S. J. 2016. The future of mental health care: peer-to-peer support and social media. *Epidemiology and psychiatric sciences*, 25(2): 113–122.
- Nova, F. F.; Coupe, A.; Mynatt, E. D.; Guha, S.; and Pater, J. A. 2022. Cultivating the community: inferring influence within eating disorder networks on twitter. *Proceedings of the ACM on Human-Computer Interaction*, 6(GROUP): 1–33.
- Onnela, J.-P.; Saramäki, J.; Hyvönen, J.; Szabó, G.; Lazer, D.; Kaski, K.; Kertész, J.; and Barabási, A.-L. 2007. Structure and tie strengths in mobile communication networks. *Proceedings of the national academy of sciences*, 104(18): 7332–7336.
- Pruccoli, J.; De Rosa, M.; Chiasso, L.; Perrone, A.; and Parmeggiani, A. 2022. The use of TikTok among children and adolescents with Eating Disorders: experience in a third-level public Italian center during the SARS-CoV-2 pandemic. *Italian Journal of Pediatrics*, 48(1): 138.
- Rosenquist, J. N.; Fowler, J. H.; and Christakis, N. A. 2011. Social network determinants of depression. *Molecular psychiatry*, 16(3): 273–281.
- Sasaki, W.; Nishiyama, Y.; Okoshi, T.; and Nakazawa, J. 2021. Investigating the occurrence of selfie-based emotional contagion over social network. *Social Network Analysis and Mining*, 11(1): 8.
- Sauerwein, J. 2025. Scrolling Through Recovery: Youth-Centred Priorities for Supporting Eating Disorder Recovery in the TikTok Era.
- Sjöström, D. K.; de Mendonca Lindström, T.; Kapetanovic, S.; and Claesdotter-Knutsson, E. 2024. Helpful or not? A qualitative study on female adolescents’ experience of TikTok when recovering from anorexia nervosa. *International Journal of Eating Disorders*, 57(11): 2217–2227.
- Tang, J.; Yu, G.; and Yao, X. 2021. Emotional contagion in the online depression community. In *Healthcare*, volume 9, 1609. MDPI.
- Wang, J.; and Wei, L. 2020. Fear and hope, bitter and sweet: Emotion sharing of cancer community on twitter. *Social Media+ Society*, 6(1): 2056305119897319.
- Wang, T.; Brede, M.; Ianni, A.; and Mentzakis, E. 2017. Detecting and characterizing eating-disorder communities on social media. In *Proceedings of the Tenth ACM International conference on web search and data mining*, 91–100.
- Wang, T.; Brede, M.; Ianni, A.; and Mentzakis, E. 2018. Social interactions in online eating disorder communities: A network perspective. *PloS one*, 13(7): e0200800.

Wokke, L. 2025. The impact of TikTok's #WhatIEatInADay on the development of Eating Disorders and Body Image among Teenagers.

Wright, K. B.; Rosenberg, J.; Egbert, N.; Ploeger, N. A.; Bernard, D. R.; and King, S. 2013. Communication competence, social support, and depression among college students: A model of Facebook and face-to-face support network influence. *Journal of health communication*, 18(1): 41–57.

Paper Checklist

1. For most authors...
 - (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? **Yes**
 - (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? **Yes**
 - (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? **Yes, See Limitations Sections**
 - (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? **Yes, See Limitations Sections**
 - (e) Did you describe the limitations of your work? **Yes, See Limitations Sections**
 - (f) Did you discuss any potential negative societal impacts of your work? **Yes, see Ethics Statement**
 - (g) Did you discuss any potential misuse of your work? **Yes, see Ethics Statement**
 - (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? **Yes, see Ethics Statement**
 - (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? **Yes, and we have ensured the paper conforms to them.**
2. Additionally, if your study involves hypotheses testing...
 - (a) Did you clearly state the assumptions underlying all theoretical results? **Yes.**
 - (b) Have you provided justifications for all theoretical results? **Yes, see Discussion.**
 - (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? **Yes, see Limitations**
 - (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? **Yes, see Limitations.**
 - (e) Did you address potential biases or limitations in your theoretical framework? **Yes, see Limitations**
 - (f) Have you related your theoretical results to the existing literature in social science? **Yes, see Discussion**
- (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? **Yes, see Discussion**
3. Additionally, if you are including theoretical proofs...
 - (a) Did you state the full set of assumptions of all theoretical results? **Yes, all assumptions for the theoretical results are stated clearly**
 - (b) Did you include complete proofs of all theoretical results? **Yes, mathematical formulations are included where applicable**
4. Additionally, if you ran machine learning experiments...
 - (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? **NA**
 - (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? **NA**
 - (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? **NA**
 - (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? **NA**
 - (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? **Yes**
 - (f) Do you discuss what is “the cost” of misclassification and fault (in)tolerance? **Yes, see Limitations**
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity...**
 - (a) If your work uses existing assets, did you cite the creators? **Yes, all datasets are appropriately cited.**
 - (b) Did you mention the license of the assets? **NA**
 - (c) Did you include any new assets in the supplemental material or as a URL? **Yes, a link to the anonymized dataset is in the appendix**
 - (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? **Yes, the data used was curated from publicly available sources in accordance with platform terms of service.**
 - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? **Yes, See Ethics Statement**
 - (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR? **Yes, the dataset is publicly accessible, includes clear documentation, and supports reuse. We include additional information in the linked resources**
 - (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset ? **Yes, we include a datasheet in the supplemental material/link**
6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity...**
 - (a) Did you include the full text of instructions given to participants and screenshots? **NA**

- (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? NA
- (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? NA
- (d) Did you discuss how data is stored, shared, and de-identified? NA

Appendix

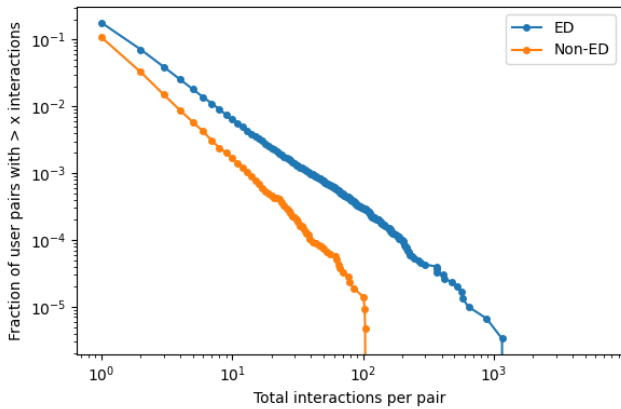


Figure 6: Log–log complementary cumulative distribution functions (CCDFs) of total interaction counts per user pair, aggregated by unique interaction counts. For a given interaction threshold x , the y-axis indicates the fraction of user pairs whose interaction frequency exceeds x . Both distributions are dominated by one-off interactions, as indicated by the sharp decline at low values, but ED-related pairs exhibit a consistently heavier tail, reflecting a higher prevalence of persistent, high-intensity dyadic interactions compared to non-ED pair

Table 4: Mapping of GoEmotions Labels to Demux Categories. Rows are shaded by grouped emotion polarity: **green** indicates **positive emotions**, **red** indicates **negative emotions**, and **gray** indicates **neutral emotions**.

GoEmotions Label	Mapped Demux Emotion
admiration	love
amusement	joy
anger	anger
annoyance	anger
approval	joy
caring	love
confusion	surprise
curiosity	anticipation
desire	love
disappointment	sadness
disapproval	sadness
disgust	disgust
embarrassment	sadness
excitement	joy
fear	fear
gratitude	love
grief	sadness
joy	joy
love	love
nervousness	anticipation
optimism	optimism
pride	joy
realization	surprise
relief	joy
remorse	sadness
sadness	sadness
surprise	surprise
neutral	none

Figure 7: Prompt card used for zero shot classification of TikTok videos into ED related categories and for consistent coder instructions.

This is an academic research project studying social media content patterns for health communication research. The classification is for educational and research purposes only.

Role. You are an expert researcher classifier analyzing social media content for academic research on health communication patterns.

Research taxonomy for content analysis.

Type 1: Pro ED Content
Definition: Encouraging the development or sustainment of eating disorders: creators promote eating disorders as a lifestyle, ask other users to help sustain their eating disorder, or actively encourage others to develop or sustain an eating disorder.
Subthemes: thinspo or meanspo, eating with an eating disorder, counting calories, dieting, food guilt, WIEIAD, losing weight to reach a goal, exercising to reach a goal.

Type 2: ED Experiences
Definition: Sharing physical and emotional experiences with eating disorders to raise awareness or educate others, such as onset, symptoms, or lived experience.
Subthemes: best practices for talking to someone with an ED, challenging misconceptions, revealing the unglamorous side, combating pro ED content, personifying EDs, sharing onset of ED, using humor.

Type 3: Recovery Content
Definition: Sharing narratives of recovery, including advice, encouragement, milestones, and challenges for those living with or recovering from an ED.
Subthemes: interacting with health care professionals, celebrating recovery, eating in recovery, recovery struggles.

Type 4: Social Support
Definition: Creators provide, seek, or describe receiving social support, online or offline.
Subthemes: providing support, receiving support, seeking support.

Research definitions.

- **Thinspo:** Thinspiration, content that glorifies extreme thinness to motivate weight loss.
- **Meanspo:** Mean inspiration, uses harsh or critical language to motivate weight loss.
- **WIEIAD:** What I eat in a day.

User prompt template. Carefully analyze this video considering all available information:

- Video file path: see video at {video_path}
- Caption: {caption}
- Audio transcript: {transcript}

When analyzing the video, pay special attention to:

1. Embedded or overlaid text in the frames
2. Visual content and imagery
3. Audio or spoken words
4. Overall context and messaging

Task. Classify the video into *exactly one* of the following categories: Pro ED Content, ED Experiences, Recovery Content, Social Support.
Output only the category name, no explanations.

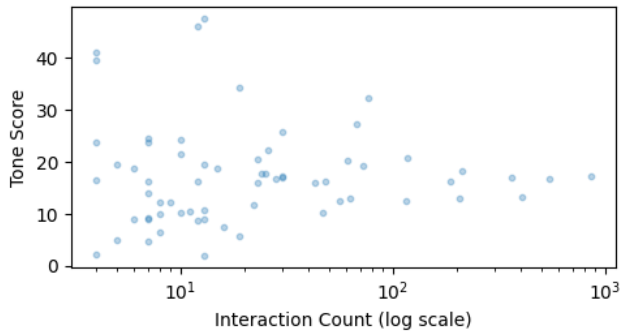


Figure 8: Tone scores for user pairs with at least three reciprocated interactions, plotted against interaction count (log scale). Results show that once reciprocal engagement is established, tone scores are consistently positive and remain stable as interactions increase, suggesting that repeated mutual exchanges create more supportive and affirmational communication.

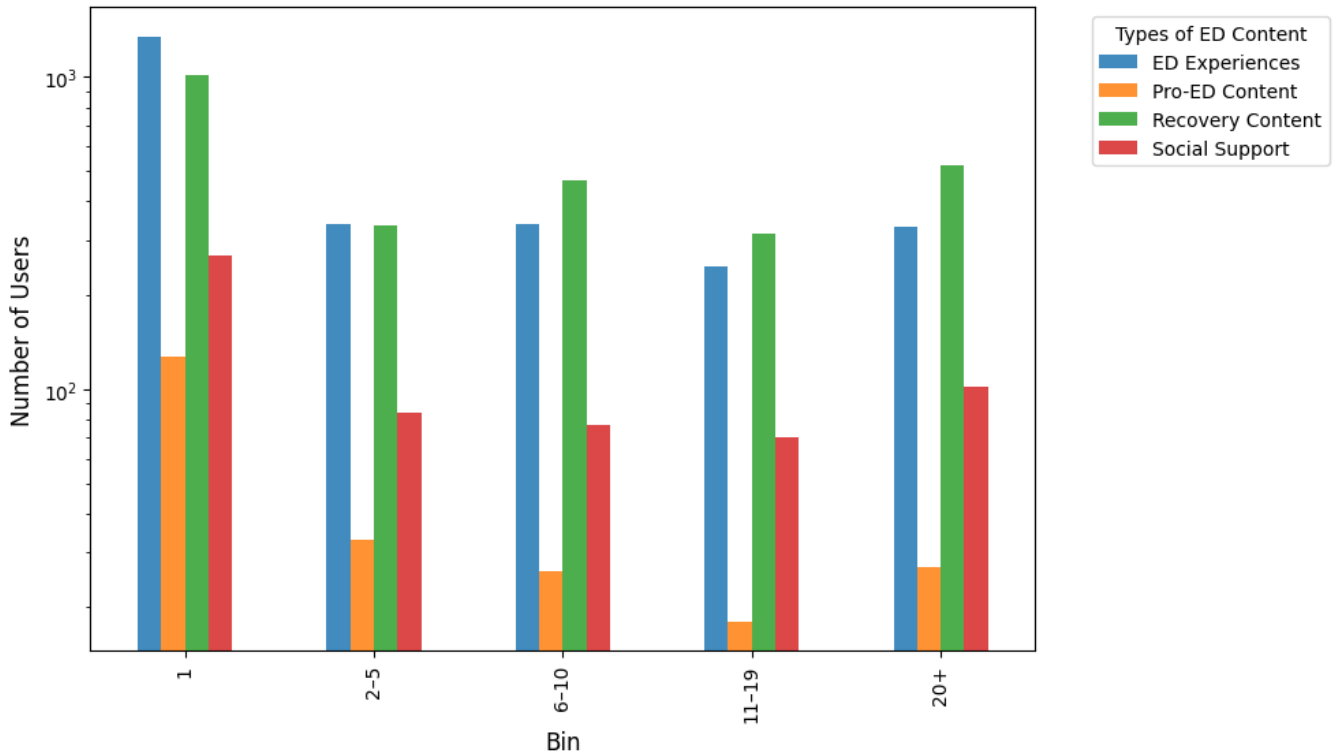


Figure 9: Distribution of user interaction frequencies across different types of ED-related TikTok content. Bars show the number of unique users in each interaction bin (1-3, 4-5, 6-10, 11-20, 20+ comments) for four content categories: ED Experiences, Pro-ED Content, Recovery Content, and Social Support.