Article

Not peer-reviewed version

From Connection to Anxiety: The Dual Effect of Social Media on Well-Being and Thematic Evolution—A BERTopic Analysis

Yu-Cheng Lin *

Posted Date: 9 April 2026

doi: 10.20944/preprints202601.0149.v2

Keywords: digital anxiety; BERTopic; Fear of Missing Out

Article

From Connection to Anxiety: The Dual Effect of Social Media on Well-Being and Thematic Evolution—A BERTopic Analysis

Yu-Cheng Lin

Ming Chuan University, Taipei; 12130340@me.mcu.edu.tw

Abstract

This study examines the shifting research on social media and psychological well-being. Prior work is split between a connection discourse (social support, belonging) and an anxiety discourse (social comparison, FoMO, problematic use), but remains largely cross-sectional. Building upon this context, this study employs BERTopic dynamic topic modeling on 7,254 Web of Science articles (2010–2025), identifying 110 topics and revealing three thematic clusters: anxiety, connection, and contextual/methodological themes. The findings indicate that anxiety-related topics are more semantically cohesive, whereas connection-related topics are more dispersed. Notably, the field experienced a turning point around 2016–2017, marking the rise and sustained dominance of anxiety-related research.Taken together, these results provide a longitudinal, computational perspective on the field and demonstrate the value of BERTopic for tracking knowledge evolution.

Keywords: digital anxiety; BERTopic; Fear of Missing Out

---

1. Introduction

Social media is one of the most remarkable transformations of the early twenty-first century. In less than two decades, platforms like Facebook, Instagram, and TikTok have become indispensable to billions of people in daily life. People now build relationships, create identities, access information, and connect with others faster than ever (Valkenburg, 2022). This pervasive presence raises a key question: when the human need for recognition and belonging shifts into an algorithm-driven space, is that need truly met, or is it subtly changed?

This question has generated two parallel yet conflicting lines of inquiry. The first centers on a connectivity thesis. It begins with the idea that humans are innately motivated to form lasting, positive interpersonal bonds. Social media might offer a new channel to satisfy this need (Baumeister & Leary, 1995). Empirical work has also shown a significant association between online networking and the accumulation of social capital (Ellison, Steinfield, & Lampe, 2007). Yet, this research's most enduring contribution is its nuanced view of connection quality. Social media's positive effects come not from passive browsing but from active, personalized interaction with close ties (Burke & Kraut, 2016). It is this meaningful communication that produces measurable reductions in loneliness and improvements in well-being.

During the same period, a second line of inquiry gathered a different and increasingly convergent body of evidence. Its theory draws from Leon Festinger's (1954) social comparison theory. People are inherently motivated to evaluate themselves by comparing themselves with others. This tendency is notably stronger in social media environments. Highly visual platforms like Instagram show users idealized highlight reels of others' lives for long periods. This increases both how often and how intensely people compare themselves. The envy these comparisons generate is a major psychological mechanism that predicts depressive symptomatology (Appel et al., 2016; Tandoc et al., 2015).

 doi:10.20944/preprints202601.0149.v2

Building on these theoretical foundations, Fear of Missing Out (FoMO) has emerged as a key mediating construct. FoMO refers to a persistent concern that others are experiencing rewarding events in one's absence (Przybylski et al., 2013), capturing the psychological strain associated with constant social media connectivity. It has been shown to predict compulsive use and is embedded within a broader network of psychological vulnerabilities. For instance, insecure attachment may undermine self-evaluation and reduce self-esteem, thereby increasing sensitivity to social comparison. This heightened sensitivity fosters preoccupation with others' activities and anxiety about exclusion, which in turn intensifies FoMO. Over time, such anxiety can lead to repetitive engagement and eventually problematic social media use (Gori et al., 2023). In this sense, FoMO can be understood as a digital manifestation of social comparison, transforming self-evaluation into sustained emotional distress and behavioral dependence.

Despite this well-established theoretical linkage, it remains unclear whether these interrelated constructs have received consistent and proportionate scholarly attention or whether research focus has shifted over time. Existing reviews, including narrative (Valkenburg, 2022), systematic (Keles et al., 2020), and meta-analytic (Huang, 2020), primarily assess effect sizes and causal relationships at specific time points, offering limited insight into the temporal evolution of scholarly attention in this domain.

To address this gap, conventional literature reviews and meta-analyses alone cannot capture the dynamic evolution of academic discourse at the macro level. In contrast, integrating bibliometric analysis with topic modeling allows for systematic mapping of a field's knowledge structure and reveals its developmental trajectory over time. This study uses bibliometric analysis as its main methodology. It applies the BERTopic dynamic topic model (Grootendorst, 2022) to analyze 7,254 peer-reviewed journal articles on social media and well-being from 2010 to 2025, retrieved from the Web of Science database. Unlike conventional Latent Dirichlet Allocation (LDA) models, BERTopic uses pre-trained Transformer language models to capture contextual relationships among terms. This approach overcomes limitations of bag-of-words models and produces topics with greater semantic coherence and interpretability in diverse academic texts.

On the basis of the foregoing, this study advances the following three research questions:

RQ1: What are the main topics in the literature on social media and well-being from 2010 to 2025, and how are they connected?

RQ2: How has the relative scholarly attention devoted to connectivity-oriented versus anxiety-oriented research themes evolved over time, and are there identifiable paradigm-shift junctures?

RQ3: To what extent do the observed shifts in research focus correspond to identifiable technological developments and sociocultural changes in social media platforms?

2. Literature Review

2.1. The Evolution of Social Media Platforms

Before examining the impact of social media on well-being, it is essential to understand how its technological architecture has undergone fundamental transformations over the past two decades. Early forms of social media, represented by Facebook, were structured as relationship-oriented social network sites. Foundational definitions characterize these platforms by three core features: the construction of public or semi-public profiles, the articulation of explicit connection lists (e.g., "friends" or "followers"), and the ability to traverse these connections within the system (Boyd & Ellison, 2007). Within this framework, platforms primarily functioned to maintain existing social ties, with text-based communication and reciprocal interaction as the dominant modes (Ellison et al., 2007).

The launch of Instagram in 2010 marked a critical shift: the logic of social media moved from relationship maintenance to the presentation and consumption of visual content. Unlike text-based platforms such as Twitter, image-centric platforms like Instagram provide more immediate gratification of social connection needs. At the same time, they intensified visual comparison

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 April 2026 doi:10.20944/preprints202601.0149.v2

processes (Pittman & Reich, 2016). As communication shifted from words to images, the scope and intensity of social comparison grew substantially.

Around 2016, a further transformation emerged with the widespread adoption of algorithm-driven content recommendation systems. At this point, the core logic shifted toward algorithmic curation aimed at maximizing user engagement, privileging interaction data over social connections to determine content visibility (Valkenburg, 2022). TikTok represents the most advanced form of this logic; its "For You Page" enables highly personalized content delivery even without established social networks. As a result, user behavior has shifted from active social interaction to increasingly passive content consumption (Twenge, 2017).

These two structural shifts have profound psychological implications. The rise of visual self-presentation amplifies social comparison and body image anxiety. Algorithmic curation—through features like infinite scrolling—creates persistent triggers for Fear of Missing Out (FoMO) and compulsive use. The surge in anxiety-related research after 2017 can be seen as an academic response to these platform changes.

2.2. The Connectivity Mechanism of Social Media

The connectivity perspective is primarily grounded in a widely accepted principle in behavioral science, namely the *need to belong*. Individuals possess an inherent motivation to establish and maintain enduring and positive interpersonal relationships, and the fulfillment or deprivation of this need profoundly influences their cognitive processes, emotional states, and overall well-being (Baumeister & Leary, 1995). With the rise of digital social platforms, this perspective has gradually extended into online contexts. Social media is theoretically conceptualized as a novel form of social connection, enabling individuals to maintain or expand their social networks beyond temporal and spatial constraints. Particularly for those who experience geographic isolation or social limitations in offline settings, social media provides a compensatory interactive space that helps address unmet belonging needs.

However, subsequent research has moved beyond the simplistic assumption that mere usage is inherently beneficial, instead emphasizing the differential effects of various usage patterns. For example, the use of Facebook has been associated with both *bridging social capital*—the flow of information and opportunities derived from weak ties—and *bonding social capital*—the emotional support and obligation-based relationships derived from strong ties (Ellison, Steinfield, & Lampe, 2007). These findings further establish a key insight in social media research: the value of platforms lies in their potential to facilitate the formation of social capital, rather than in the amount of time spent using them.

Compared to passive consumption (i.e., browsing others' content), users who engage in active and personalized interactions with strong ties are more likely to experience observable increases in well-being and reductions in loneliness (Burke & Kraut, 2016). This distinction has shifted the focus of subsequent research from usage quantity to usage quality. Within this framework, *perceived social support* has increasingly been recognized as a critical positive mediating variable linking social media use to life satisfaction (Huang, 2020).

2.3. The Anxiety Mechanism of Social Media

The anxiety perspective is grounded in *social comparison theory*, which posits that individuals tend to evaluate their abilities and opinions by comparing themselves with others (Festinger, 1954). In earlier contexts, such comparisons were naturally constrained by limitations on time, space, and the size of social networks. However, contemporary social media—particularly platforms centered on visualized and idealized self-presentation, such as Instagram—has significantly expanded the scope of comparison contexts. This enables *upward social comparison* to operate at a larger scale and more continuously. Existing research indicates that repeated exposure to others' carefully curated "highlight reels" is significantly associated with increased feelings of envy, as well as the onset and exacerbation of depressive symptoms (Appel et al., 2016; Tandoc et al., 2015). Moreover, these effects

 doi:10.20944/preprints202601.0149.v2

are moderated by individual differences, with individuals who exhibit a higher tendency toward social comparison being more susceptible to declines in self-esteem and negative emotional responses (Vogel et al., 2015).

Building on the above, the psychological impact of social media extends beyond comparison mechanisms and is closely tied to internal psychological states arising from persistent connectivity. In particular, *Fear of Missing Out (FoMO)* is conceptualized as a pervasive concern that others may be having rewarding experiences from which one is absent (Przybylski et al., 2013). FoMO is a key psychological driver of repetitive platform checking and difficulty disengaging from social media. Within algorithm-driven environments where content is continuously pushed to users, this motivation is further reinforced, forming cyclical usage patterns. FoMO has been found to be positively associated with the need to belong, neurotic personality traits, and anxious attachment styles (Beyens et al., 2016; Blackwell et al., 2017) and contributes to problematic social media use through its negative impact on self-esteem (Gori et al., 2023).

Under the combined influence of these mechanisms, sustained social media use gradually accumulates into observable psychological outcomes. A substantial body of research consistently demonstrates stable associations between excessive use and increased levels of anxiety, depression, and social media fatigue (Dhir et al., 2018; Lin et al., 2016; Primack et al., 2017). Building on these findings, research methodologies have evolved from correlational analyses toward causal inference. For instance, randomized controlled trials have shown that limiting social media use can significantly reduce loneliness and depressive symptoms (Hunt et al., 2018), while quasi-experimental studies further support a causal relationship between the diffusion of social media and the deterioration of mental health (Braghieri et al., 2022).

2.4. COVID-19 as an Amplifier of Connectivity and Anxiety

The COVID-19 outbreak in 2020 has significant methodological and theoretical implications for the development of social media research. Large-scale physical distancing measures led to a rapid increase in individuals' reliance on digital platforms. This *exogenous shock* simultaneously amplified both the positive connectivity mechanisms and the negative anxiety mechanisms of social media within a short period of time.

During lockdowns and social distancing, digital platforms became essential channels for maintaining interpersonal connections, seeking emotional support, and alleviating loneliness (Geirdal et al., 2021). At the same time, social media served as a critical medium for disseminating public health information and mobilizing communities, demonstrating its positive role in times of crisis (Abbas et al., 2021).

However, within the same context, the negative mechanisms emphasized by the anxiety perspective were further intensified. Empirical evidence indicates that, particularly in the early stages of the pandemic, higher levels of social media exposure were positively associated with increased anxiety and depression (Gao et al., 2020). Moreover, *information overload* and the spread of misinformation further exacerbated uncertainty and psychological distress (Ni et al., 2020).

2.5. Research Gap

Existing review studies, including narrative reviews (Patti M. Valkenburg, 2022), systematic reviews (Keles et al., 2020), and meta-analyses (Huang, 2020), primarily adopt a cross-sectional analytical logic. These approaches focus on synthesizing effect sizes and identifying causal relationships from prior studies, yet they pay limited attention to the broader knowledge-development context in which these studies are produced. This limitation raises several issues that have not yet been systematically examined.

First, how has the relative prominence of the connectivity perspective and the anxiety perspective evolved over the past fifteen years? Second, are shifts in scholarly focus driven primarily by internal theoretical developments or by external social and technological changes, such as the rise of Instagram, the widespread adoption of TikTok, or the impact of COVID-19? Third, as research attention increasingly shifts toward anxiety-related issues, has this trend influenced the direction of policy discussions and intervention-oriented research?

To address these questions, this study employs a BERTopic-based dynamic topic modeling approach to conduct a longitudinal analysis of literature on social media and well-being from 2010 to 2025. Compared with traditional literature review methods, a bibliometric approach enables systematic and data-driven tracking of the evolution of research topics, thereby providing a dynamic academic landscape that reflects the trajectory of knowledge development (Zupic & Čater, 2015).

3. Research Methodology

To systematically examine the knowledge structure of the research domain on social media and well-being, and its dynamic evolution over time, this study adopts a bibliometric analysis approach combined with computational topic modeling as its core research design. Existing narrative reviews and meta-analyses are constrained by a cross-sectional analytical logic and are therefore unable to systematically capture the temporal rise and decline of research topics. In contrast, bibliometric analysis uses metadata from large-scale academic publications as the unit of analysis, enabling a data-driven, objective, and reproducible approach to uncover the knowledge structure, thematic evolution, and research frontiers within a specific domain (Zupic & Čater, 2015).

This section proceeds by describing the data sources and corpus construction process, the analytical procedures and parameter settings of the BERTopic-based dynamic topic modeling approach, and the methods used for topic validation.

3.1. Data Sources

This study selects the Web of Science (WoS) Core Collection as the primary data source. The WoS database applies rigorous indexing standards, ensuring that all included publications are peer-reviewed, thereby effectively excluding grey literature that may introduce noise into topic modeling. In addition, WoS provides broad interdisciplinary coverage, encompassing fields highly relevant to this study, including psychology, communication, information science, public health, and sociology. This reduces the risk of disciplinary bias that may arise from relying on a single-domain database. Furthermore, WoS offers comprehensive metadata, including titles, abstracts, author keywords, and publication years, which meet the input requirements for subsequent BERTopic analysis.

The search period was defined from January 1, 2010, to January 31, 2025, covering a total of fifteen years. The year 2010 was selected as the starting point because Instagram was officially launched in 2010, marking the emergence of a new generation of social media ecosystems centered on visual content and mobile devices. Moreover, the volume of academic publications prior to 2010 remained in an early accumulation stage; thus, using 2010 as the baseline allows for a comprehensive capture of the field's development from its inception to maturity.

To achieve an optimal balance between comprehensiveness and precision, this study adopts a two-layer Boolean search strategy, as follows:

TI = ("social media" OR "social networking sites" OR "SNS" OR "Facebook" OR "Instagram" OR "Twitter" OR "TikTok") AND TS = ("well-being" OR "wellbeing" OR "happiness" OR "life satisfaction" OR "anxiety" OR "depression" OR "loneliness" OR "stress" OR "mental health" OR "social comparison" OR "FOMO" OR "fear of missing out" OR "social support" OR "belonging" OR "connection")

Following the initial retrieval, documents were filtered by publication type. Conference papers, book chapters, editorials, and short communications were excluded, leaving only original research articles. A total of 7,254 peer-reviewed journal articles were included in the final dataset. For each document, the title, abstract, author keywords, and publication year were extracted and integrated into a JSON-format corpus for subsequent analysis.

Notably, titles are often too concise and lack sufficient semantic density to support contextual understanding by Transformer-based models, while author keywords may vary in standardization. In contrast, abstracts contain research questions, methodological information, and key findings,

 doi:10.20944/preprints202601.0149.v2

thereby providing the most comprehensive and balanced semantic representation for topic modeling (Grootendorst, 2022).

3.2. BERTopic Dynamic Topic Modelling

This study employs the BERTopic model (Grootendorst, 2022) as the core analytical approach. Compared to traditional Latent Dirichlet Allocation (LDA), BERTopic offers several methodological advantages. It leverages pre-trained Transformer-based language models for semantic embedding, enabling the capture of contextual relationships that bag-of-words approaches cannot identify. In addition, BERTopic uses a density-based clustering algorithm that does not require predefining the number of topics, allowing the topic structure to emerge from the data. Furthermore, its built-in dynamic analysis function enables the tracking of topic prevalence over time, directly supporting the longitudinal design of this study. The overall analytical procedure is illustrated in Figure 1.

The analysis begins by transforming all document abstracts into dense semantic vectors using the pre-trained language model *all-MiniLM-L6-v2*. As part of the Sentence-BERT family, this model demonstrates strong performance on semantic similarity tasks while maintaining computational efficiency, making it suitable for large-scale corpora with over 7,000 documents (Reimers & Gurevych, 2019). Compared with bag-of-words approaches, such contextualized embeddings capture nuanced semantic differences across varying contexts.

Given that high-dimensional embeddings are not suitable for direct clustering due to the curse of dimensionality, dimensionality reduction is performed using the Uniform Manifold Approximation and Projection (UMAP) algorithm. The parameters are set as follows: *n_neighbors* = 15 to balance local and global structure preservation, *n_components* = 5 to define the target dimensionality, and *min_dist* = 0.0 to allow tight clustering in the reduced space. Unlike linear techniques such as PCA, UMAP preserves nonlinear structures in high-dimensional space, ensuring that semantically similar documents remain close in the reduced representation.

Clustering is subsequently conducted using the HDBSCAN algorithm, a hierarchical density-based method that identifies clusters based on data distribution. The parameter *min_cluster_size* is set to 30 to ensure that each topic contains a sufficient number of documents, while *min_samples* is set to 10 to control sensitivity to noise. The Euclidean distance metric is applied. Importantly, HDBSCAN does not require a predefined number of clusters and can identify outliers, which are assigned to Topic -1 rather than being forcibly included in any cluster. This helps maintain the internal coherence of the resulting topics.

Following clustering, topic representation is generated using the class-based TF-IDF (c-TF-IDF) approach. All documents within each topic are aggregated into a single document, and term frequencies are weighted by inverse document frequency across topics. This process highlights words that are distinctive to each topic while suppressing commonly shared terms. During preprocessing, English stop words are removed, and terms shorter than three characters are excluded using a CountVectorizer, thereby enhancing the interpretability of the topic representations.

To capture the temporal evolution of research topics, BERTopic's dynamic topic modeling function (*topics_over_time*) is applied. The publication year of each document serves as the temporal index, and the relative frequency of each topic is calculated annually. This relative measure reflects the proportion of each topic within the yearly corpus, thereby controlling for variations in total publication volume and ensuring comparability across time. The results are visualized in time-series plots that illustrate the dynamic evolution of core research themes from 2010 to 2025.

 doi:10.20944/preprints202601.0149.v2

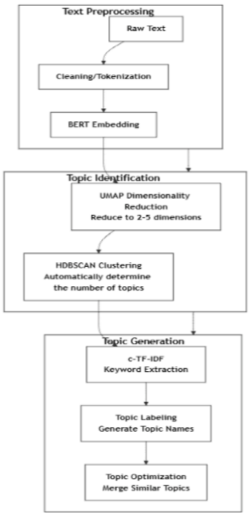

Figure 1. The flowchart of the BERTopic algorithm.

3.3. Topic Validation

The semantic validity of topics generated by computational topic modeling must be carefully examined to support subsequent academic interpretation. This study evaluates topic quality from three complementary perspectives.

First, a quantitative assessment is conducted using topic coherence. For each topic, the top 10 representative keywords are used to calculate the coherence score using the CV metric, which measures semantic consistency among words within a topic. Higher CV scores indicate stronger semantic co-occurrence and clearer topic structure. Following prior studies, a CV threshold of≥ 0.5 is adopted as a reference criterion for determining acceptable topic quality (Röder et al., 2015).

Second, manual content validation is performed. The researchers examine the top 15 major topics by reviewing their keyword sets and a random sample of representative documents (ten documents per topic). This process ensures that topic labels and keywords accurately reflect the substantive content of the literature, and adjustments are made when necessary to address potential semantic inconsistencies.

Third, the stability of the topic modeling results is assessed. The BERTopic analysis is repeated three times with different random seeds, and the consistency in the number of topics and the composition of keywords for the major topics is compared across runs. This procedure ensures that the results are not overly sensitive to initialization conditions and reduces the influence of algorithmic randomness on the study’s conclusions.

4. Result

This study builds on BERTopic’s static clustering to identify dominant thematic clusters in literature on social media and well-being from 2010 to 2025. This step addresses RQ1. Next, the study uses a semantic topic space to uncover latent relationships among topics. Finally, it employs dynamic trend analysis to trace the evolution of the scholarly prominence of the connectivity and anxiety perspectives and to identify key inflection points linked to platform and technological shifts. These analyses address RQ2 and RQ3.

4.1. The Static Structure of Topic Clusters

Applying the BERTopic model to the abstracts of 7,254 publications, a total of 110 distinct research topics were identified (excluding Topic -1 outliers). To focus on the core knowledge structure, this study concentrates on the top 15 topics as the main units of analysis. Figure 2 shows these 15 core topics and their most representative c-TF-IDF keyword sets.

 doi:10.20944/preprints202601.0149.v2

These 15 topics are not isolated. They show clear patterns of semantic aggregation that can be divided into three functional clusters, as shown in Table 1. The distribution of topics across these clusters is significantly asymmetric. The anxiety-oriented cluster has six topics. The connectivity-oriented cluster includes only two topics. The contextual and methodological cluster contains seven topics. The anxiety-oriented cluster is three times as large as the connectivity-oriented cluster. This suggests that the research domain has shifted toward a focus on risks and negative mechanisms when examining the psychological effects of social media.

The anxiety-oriented cluster captures the multi-layered negative psychological impacts of social media use. Its structure reflects a sequence from cognitive mechanisms to emotional drivers, symptom manifestations, and behavioral outcomes. This cluster includes cognitive-level topics focused on social comparison (Topic 0). It also covers affective and attachment-related dynamics, such as Fear of Missing Out (FoMO) and problematic use (Topic 2). Vulnerability-related issues include body image and appearance anxiety (Topic 1) and problematic use among adolescents (Topic 4). These link to downstream behavioral consequences, including sleep quality (Topic 8) and cyberbullying (Topic 14).

In contrast, the connectivity-oriented cluster consists of Topic 5 (perceived social support in online contexts) and Topic 11 (online communities and sense of belonging), corresponding to the constructs of perceived social support and the need to belong discussed in Section 2. Although this cluster is relatively small, it represents earlier, foundational theoretical concerns within the field. Its relatively peripheral position in the static topic structure thus provides an initial structural indication of the shifting emphasis in scholarly discourse, which will be further examined through the dynamic analysis.

Table 1. Interpretation and Classification of the Top 15 Core Research Topics.

Topic	Topic Label	Representative Keywords	Cluster
Topic 0	Social Comparison and Envy on Facebook	Facebook, self, comparison, envy, disclosure	Anxiety-oriented
Topic 1	Body Image and Appearance-Related Anxiety	body, appearance, image, eating, dissatisfaction	Anxiety-oriented
Topic 2	Fear of Missing Out (FoMO) and Problematic Use	fomo, missing, fear, phubbing, psmu	Anxiety-oriented
Topic 4	Adolescent Problematic Use and Psychological Symptoms	adolescents, symptoms, use, smu, problematic	Anxiety-oriented
Topic 8	Social Media Use and Sleep Quality	sleep, quality, anxiety, use, psmu	Anxiety-oriented
Topic 14	Cyberbullying and Aggressive Online Behaviours	cyberbullying, bullying, aggression, victimization	Anxiety-oriented
Topic 5	Perceived Online Social Support	support, social, perceived, online, capital	Connection-oriented
Topic 11	Online Communities and Sense of Belonging	community, online, belonging, identity, support	Connection-oriented
Topic 3	Social Media Discourse During the COVID-19 Pandemic	covid, 19, pandemic, tweets, public	Contextual and methodological

Topic 6	AI-Based Depression Detection	depression, detection, learning, model, instagram	Contextual and methodological
Topic 7	Online Discourse on Mental Health Stigma	stigma, mental, health, twitter, tweets	Contextual and methodological
Topic 9	Student Learning and Educational Applications	students, learning, education, academic, teachers	Contextual and methodological
Topic 10	Personality Traits and Motivations for Use	self, esteem, personality, use, narcissism	Contextual and methodological
Topic 12	Political Participation and Information Diffusion	political, news, information, participation, fake	Contextual and methodological
Topic 13	Branding, Marketing, and Consumer Behaviour	brand, consumer, purchase, intention, instagram	Contextual and methodological

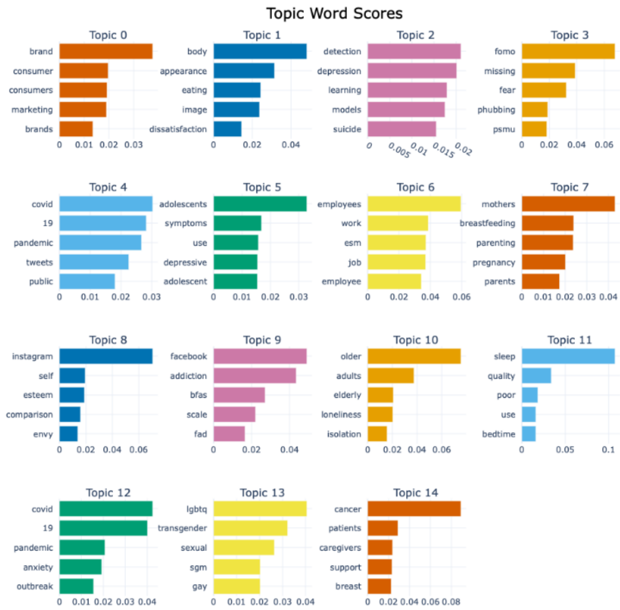

Figure 2. Top 15 BERTopic-Derived Topics and Their Representative c-TF-IDF Keywords.

4.2. Semantic Topic Space

While static clustering reveals the categorical structure of topics, it does not fully capture their semantic relationships. To address this limitation, Figure 3 visualizes the relative positions of topics in the high-dimensional semantic space by projecting them onto a two-dimensional plane using UMAP. This representation enables the examination of semantic distances between topics. The resulting configuration of the semantic space reveals three key features with important implications for the underlying knowledge structure.

First, anxiety-related topics exhibit a highly clustered spatial structure. Topic 0 (social comparison), Topic 1 (body image), Topic 2 (FoMO), Topic 4 (problematic use among adolescents), Topic 8 (sleep quality), and Topic 14 (cyberbullying) form a dense cluster in the semantic space (Figure 4). The close proximity among these topics suggests that prior research addressing these negative psychological phenomena tends to rely on relatively consistent conceptual frameworks and analytical vocabularies. This pattern further implies a degree of homogeneity in theoretical

doi:10.20944/preprints202601.0149.v2

orientations and explanatory mechanisms, in which social comparison, attachment and dependency, and anxiety responses constitute a common explanatory pathway.

Second, compared to the dense distribution of anxiety-related topics, connectivity-related topics exhibit a more isolated spatial pattern. Topic 5 (perceived social support in online contexts) and Topic 11 (online communities and sense of belonging) are positioned close to each other and form a localized cluster. However, they remain semantically distant from the anxiety-related cluster (Figure 5). This separation suggests that research focusing on how social media facilitates social support and belonging may follow a relatively distinct conceptual and theoretical trajectory from studies examining its negative psychological effects.

Finally, Topic 10 (self-esteem and personality traits) occupies an intermediate position between the anxiety and connectivity clusters in the semantic space (Figure 6). This positioning indicates that individual difference variables, such as self-esteem, neuroticism, and attachment style, may simultaneously relate to both the positive and negative psychological effects of social media use and may serve as bridging or moderating mechanisms between the two perspectives. This observation is also consistent with the theoretical arguments discussed in Section 2. For example, insecure attachment has been shown to influence FoMO via self-esteem (Gori et al., 2023), further supporting the alignment between the semantic space analysis and existing theoretical frameworks.

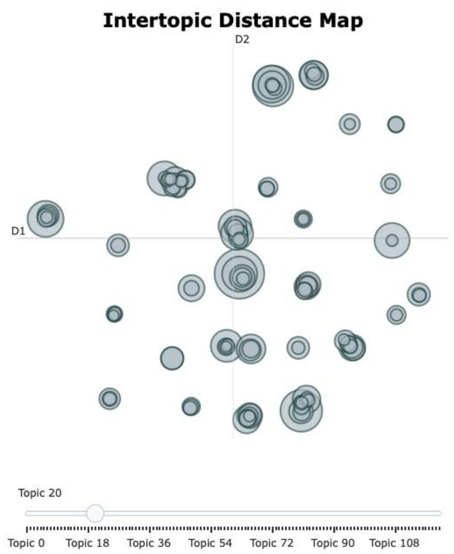

Figure 3. 15 Core Topics Identified by BERTopic and Their Most Representative Keywords.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 April 2026 doi:10.20944/preprints202601.0149.v2

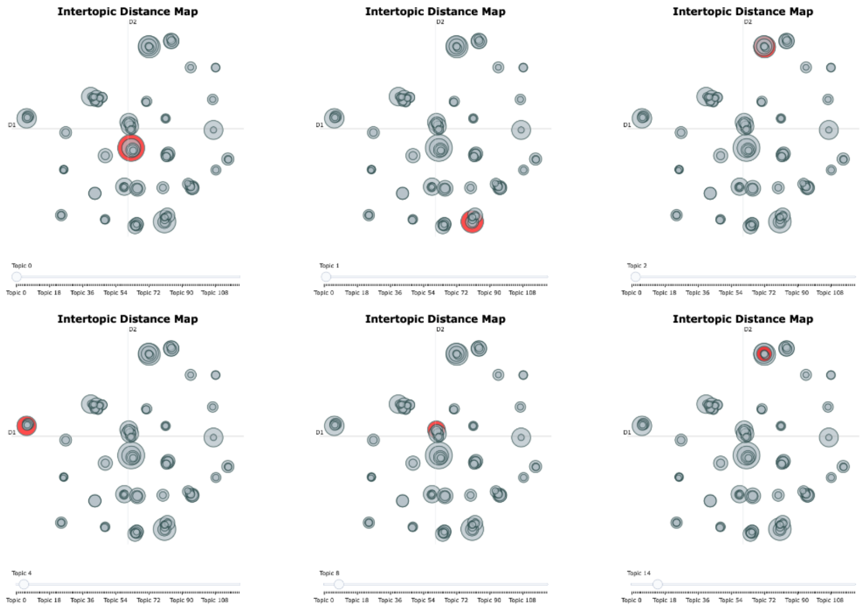

Figure 4. The Semantic Core of the Anxiety-Oriented Cluster.

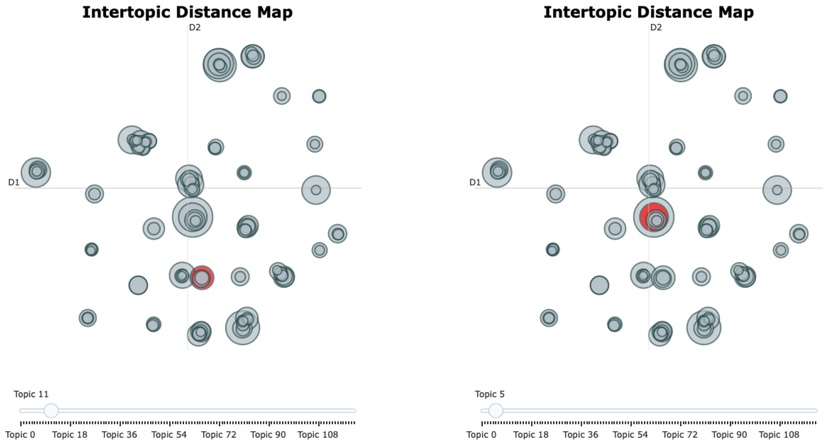

Figure 5. The Semantically Isolated Structure of the Connectivity-Oriented Cluster.

Figure 4. The Bridge between Methodology and Application.

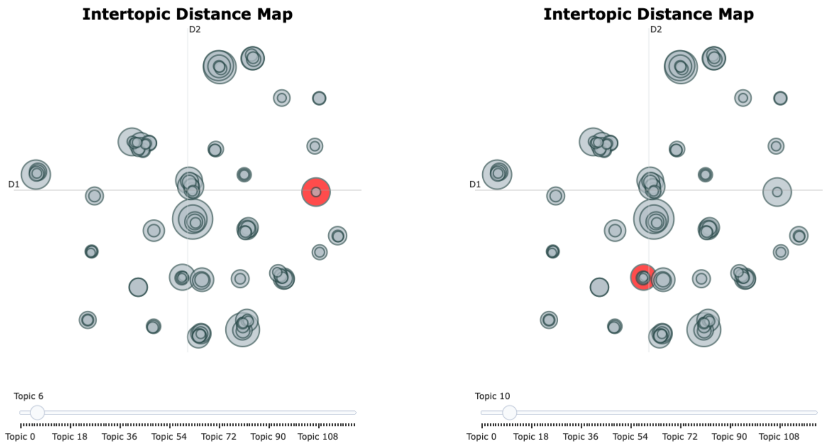

Figure 6. The Cross-Cluster Bridging Position of the Personality Traits Topic.

 doi:10.20944/preprints202601.0149.v2

4.3. Dynamic Evolution of Topics

While static analysis presents a cross-sectional map of the knowledge structure in this research domain, dynamic analysis reveals how this structure emerges and evolves over time. Figure 7 shows the annual relative frequency of core topics from 2010 to 2025. This analysis addresses RQ2 by examining changes in the scholarly prominence of the connectivity and anxiety perspectives, and addresses RQ3 by exploring whether these changes relate to shifts in social media platforms and related technologies. Based on observed trends, the development of this research domain can be divided into four sequential stages.

In the first stage, from 2010 to 2016, the connectivity and anxiety perspectives exhibit a relatively balanced and parallel pattern. Perceived social support in online contexts (Topic 5) emerges as the most stable core topic during this period, reflecting an early usage context in which social media platforms were primarily oriented toward maintaining interpersonal relationships. This pattern is consistent with prior studies, such as Burke and Kraut (2016) and Ellison et al. (2007). In comparison, although social comparison (Topic 0) had already emerged, its scholarly prominence did not yet surpass that of connectivity-related topics, suggesting that researchers were still exploring the multifaceted effects of social media.

In the second stage, approximately between 2016 and 2018, a noticeable shift in topic development becomes evident. Several anxiety-related topics, including FoMO, body image, and problematic use among adolescents, exhibit simultaneous growth. Among these, Fear of Missing Out (Topic 2) shows particularly significant increases. This shift corresponds with several external developments that provide a potential explanation for RQ3. For example, the introduction of Instagram Stories in 2016 reinforced the trend toward real-time, ephemeral content consumption. At the same time, the iGen argument proposed by Jean Twenge (2017) heightened scholarly attention to adolescent mental health, while large-scale empirical studies examining the relationship between multi-platform use and mental health were published during this period (Lin et al., 2016; Primack et al., 2017). These factors may have collectively increased the visibility and momentum of anxiety-related research.

The third stage (2018–2020) shows increased scholarly focus. During this period, FoMO (Topic 2), social comparison (Topic 0), and body image (Topic 1) emerge as primary research topics. Perceived social support (Topic 5) remains but is less prominent than before. This shift marks anxiety-related topics as dominant, addressing the discourse change noted in RQ2.

Finally, since 2020, the research landscape has been characterized by the interplay between external events and methodological shifts. On the one hand, COVID-19 (Topic 3) rises sharply from 2020 onward, reaching a clear peak between 2020 and 2021 before gradually declining. This pattern reflects the academic community's concentrated attention on major public events, followed by a natural decrease in focus over time. Notably, related studies highlight the dual role of social media during the pandemic, both as a substitute for social connection and as a source that may intensify information overload and psychological distress. This further underscores the core tension between the connectivity and anxiety perspectives. On the other hand, AI-based depression detection (Topic 6) shows a steady upward trend after 2018, indicating a shift in research orientation from descriptive analysis toward computationally driven prediction and application. This trend reflects the growing integration of interdisciplinary methods within the field.

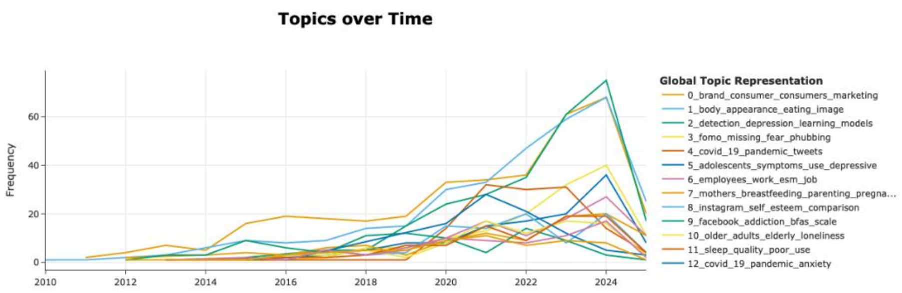

Figure 7. Temporal Evolution of Topic Trends.

4.4. Research Findings

By integrating the results of the static structural analysis and the dynamic trend analysis, this study identifies three key findings that correspond to the three research questions.

With regard to RQ1, the research domain of social media and well-being from 2010 to 2025 exhibits a tripartite structure consisting of the anxiety perspective, the connectivity perspective, and the contextual and methodological cluster. Among these, the anxiety perspective is more prominent, as indicated by both the number of topics and the degree of semantic cohesion. It not only encompasses a greater number of topics than the connectivity perspective, but also forms a more densely clustered structure in the semantic space. This structural pattern suggests that existing literature has increasingly shifted toward a focus on negative psychological mechanisms when examining the effects of social media, positioning anxiety-related issues at the core of the knowledge landscape.

With regard to RQ2, the relative scholarly prominence of the connectivity and anxiety perspectives shifts noticeably around 2016 to 2017. The early stage is characterized by a relatively balanced and parallel development between the two perspectives. However, this balance gradually transitions toward a research focus dominated by anxiety-related topics. After 2018, this asymmetrical pattern became stabilized, indicating that the visibility of the anxiety perspective remains consistently high within the overall research landscape. Notably, this dominance is not significantly reversed by subsequent external events, such as COVID-19.

For RQ3, the shift in scholarly attention coincides with several technological and social developments: the introduction of Instagram Stories in 2016, a surge of studies on Fear of Missing Out and multi-platform use from 2016 to 2017, and increased societal focus following Jean Twenge's 2017 publication. This timing suggests that changes in academic discourse may result not only from internal knowledge growth but also from technological advances in social media platforms and related socio-cultural issues.

5. Conclusion

This study employs a BERTopic-based dynamic topic modeling approach to conduct a longitudinal computational analysis of 7,254 publications on social media and well-being indexed in the Web of Science database from 2010 to 2025. Compared with traditional narrative reviews and meta-analyses, which primarily focus on static synthesis, this study further examines the evolutionary characteristics of the knowledge structure over time.

Based on the preceding findings, this section first synthesizes the key results for the three research questions and then discusses their theoretical and practical implications. It subsequently outlines the study's limitations and proposes directions for future research.

5.1. Addressing the Research Questions and Key Findings

First, regarding the overall knowledge structure, the literature shows a tripartite configuration. It consists of the anxiety perspective, the connectivity perspective, and the contextual and

 doi:10.20944/preprints202601.0149.v2

methodological cluster. Among these, anxiety-related topics not only outnumber those related to connectivity but also form a dense cluster in the semantic space. This pattern suggests that research has increasingly focused on negative psychological mechanisms when examining social media effects. As a result, anxiety-related issues have become more central within the knowledge landscape, while connectivity-related research remains stable but secondary.

Second, from a temporal perspective, the research focus shifts clearly over time. In the early stage, between 2010 and 2016, the literature reflects the parallel development of connectivity and anxiety perspectives (Burke & Kraut, 2016; Ellison et al., 2007). However, beginning around 2016 to 2017, anxiety-related topics—especially Fear of Missing Out (FoMO)—became more prominent. They remain at a high level after 2018 (Lin et al., 2016; Primack et al., 2017). In contrast, the connectivity perspective does not disappear, but its share gradually declines. This trend indicates a shift away from research on social connectivity toward the negative psychological impacts of social media.

Finally, this shift in research focus aligns with key technological and social developments. Instagram Stories, introduced in 2016, changed how content was presented and how it was engaged with. Between 2016 and 2017, several studies on Fear of Missing Out and multi-platform use were published (Lin et al., 2016; Primack et al., 2017). At the same time, Jean Twenge's (2017) work heightened attention to adolescent mental health. The timing of these changes and shifts in study topics suggests research evolution is linked to both knowledge accumulation and technological changes in social media. It is also closely associated with the socio-cultural issues these platforms generate.

5.2. Theoretical and Practical Implications

This study contributes to the existing literature in three ways: methodological advancement, theoretical integration, and theoretical development. First, methodologically, this study shows that computational bibliometric approaches can be used in psychological research. Specifically, it demonstrates that BERTopic-based dynamic topic modeling can track the evolution of knowledge. This data-driven and reproducible approach is more systematic than traditional narrative reviews and meta-analyses. It enables a clearer examination of how research topics evolve and offers a complementary perspective on how academic knowledge develops over time.

Building on this, the study offers new insights into how the connectivity and anxiety perspectives have long coexisted in social media research. It uses a semantic-structural perspective to highlight these perspectives. The distribution within the semantic space shows that these two approaches have distinct structural patterns. These patterns relate to how concepts and theories are used. This suggests prior research followed different conceptual systems and research paths. As a result, future integration efforts must address differences in empirical findings. Researchers will also need to resolve issues of conceptual alignment and theoretical language. This points to the importance of cross-theoretical integration.

Furthermore, the temporal patterns identified through trend analysis provide new insights into how research about the psychological effects of social media has developed. Shifts in scholarly attention seem to align with changes in platform functionality and usage patterns. This means that research topics emerge not only from knowledge accumulation. They are also shaped by technological changes and sociocultural concerns. This matches arguments about the role of platform characteristics in psychological outcomes (Valkenburg, 2022). It also extends this perspective by situating it within a broader framework of knowledge evolution.

Research attention has gradually shifted away from social media's connectivity functions and toward its possible negative psychological impacts. As a result, issues such as fear of missing out (FoMO) and social comparison are now more prevalent among adolescents and general users. In this setting, adding digital well-being literacy to education and counseling could help users better understand platform logics. It might also foster more reflective usage and better self-regulation. However, studies show that connectivity-focused research has not vanished. It remains a steady part of the knowledge landscape. This means that social media's positive roles, such as boosting social support and belonging, are still relevant. Therefore, instead of focusing only on risks, a constructive next step for platform design and policy is to create mechanisms that support meaningful interactions within current algorithmic and commercial frameworks.

5.3. Limitations and Future Research

First, the corpus of this study is entirely derived from the Web of Science Core Collection. Although this database ensures a high level of academic quality, it may exclude relevant studies indexed in other databases, such as Scopus, PsycINFO, or discipline-specific sources. As a result, certain disciplinary perspectives, such as clinical psychology or communication studies, may be underrepresented in the analysis. This limitation implies that the topic distribution identified in this study reflects a knowledge slice constrained by a specific data source. Future research could address this issue by integrating multiple databases to construct a more comprehensive corpus. In addition, further analyses could be conducted by differentiating across social media platforms, such as Facebook, Instagram, or TikTok, to examine how platform-specific design logics and content formats correspond to distinct psychological mechanisms and thereby deepen understanding of platform-specific effects.

Second, this study uses abstracts as the unit of analysis. While this approach enhances consistency and comparability across the corpus, abstracts provide only a limited representation of each study's full semantic content and may not fully capture theoretical depth or nuanced constructs. In particular, research on protective factors, such as digital literacy, mindfulness traits, or attachment styles, is often elaborated more extensively in full texts but may be only briefly mentioned in abstracts, leading to their underrepresentation in topic modeling results. This limitation is also consistent with the structural pattern observed in this study, in which risk-related mechanisms appear more concentrated, whereas protective factors remain relatively dispersed. Future research could incorporate full texts or specific sections, such as theoretical frameworks or methodological descriptions, to provide a more comprehensive analysis. Such efforts would enable a more systematic examination of how individual differences and usage contexts jointly moderate the psychological effects of social media and contribute to the development of differential susceptibility theory.

Finally, although BERTopic provides a computationally consistent approach to topic identification, the processes of topic labeling and cluster interpretation inevitably involve researcher judgment. Despite efforts to mitigate this issue through content validation procedures, interpretive bias cannot be entirely eliminated. Future studies could enhance robustness by involving multiple coders and conducting inter-rater reliability assessments. Moreover, the bibliometric approach adopted in this study focuses on the evolution of academic discourse at a macro level and does not directly capture psychological processes at the individual level. Future research could therefore integrate longitudinal designs, such as ecological momentary assessment (EMA) and behavioral data derived from smartphone usage, to more precisely examine the dynamic interactions among usage behaviors, contextual factors, and psychological states, thereby complementing the macro-level insights provided in this study with micro-level evidence.

References

1. Abbas, J., Wang, D., Su, Z., & Ziapour, A. (2021). The role of social media in the advent of COVID-19 pandemic: Crisis management, mental health challenges and implications. *Risk Management and Healthcare Policy*, *14*, 1917–1932. https://doi.org/10.2147/RMHP.S284313
2. Appel, H., Gerlach, A. L., & Crusius, J. (2016). The interplay between Facebook use, social comparison, envy, and depression. *Current Opinion in Psychology*, *9*, 44–49. https://doi.org/10.1016/j.copsyc.2015.10.006
3. Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

4. Beyens, I., Frison, E., & Eggermont, S. (2016). "I don't want to miss a thing": Adolescents' fear of missing out and its relationship to adolescents' social needs, Facebook use, and Facebook related stress. *Computers in Human Behavior*, *64*, 1–8. https://doi.org/10.1016/j.chb.2016.05.083
5. Blackwell, D., Leaman, C., Tramposch, R., Osborne, C., & Liss, M. (2017). Extraversion, neuroticism, attachment style and fear of missing out as predictors of social media use and addiction. *Personality and Individual Differences*, *116*, 69–72. https://doi.org/10.1016/j.paid.2017.04.039
6. Boyd, d. m., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x
7. Braghieri, L., Levy, R., & Makarin, A. (2022). Social media and mental health. *American Economic Review*, *112*(11), 3660–3693. https://doi.org/10.1257/aer.20211218
8. Burke, M., & Kraut, R. E. (2016). The relationship between Facebook use and well-being depends on communication type and tie strength. *Journal of Computer-Mediated Communication*, *21*(4), 265–281. https://doi.org/10.1111/jcc4.12162
9. de Hesselle, L. C., & Montag, C. (2024). Effects of a 14-day social media abstinence on mental health and well-being: Results from an experimental study. *BMC Psychology*, *12*(1), 141. https://doi.org/10.1186/s40359-024-01611-1
10. Dhir, A., Yossatorn, Y., Kaur, P., & Chen, S. (2018). Online social media fatigue and psychological wellbeing: A study of compulsive use, fear of missing out, fatigue, anxiety and depression. *International Journal of Information Management*, *40*, 141–152. https://doi.org/10.1016/j.ijinfomgt.2018.01.012
11. Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x
12. Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202
13. Gao, J., Zheng, P., Jia, Y., Chen, H., Mao, Y., Chen, S., Wang, Y., Fu, H., & Dai, J. (2020). Mental health problems and social media exposure during COVID-19 outbreak. *PLoS ONE*, *15*(4), e0231924. https://doi.org/10.1371/journal.pone.0231924
14. Geirdal, A. Ø., Ruffolo, M., Leung, J., Thygesen, H., Price, D., Bonsaksen, T., & Schoultz, M. (2021). Mental health, quality of life, wellbeing, loneliness and use of social media in a time of social distancing during the COVID-19 outbreak: A cross-country comparative study. *Journal of Mental Health*, *30*(2), 148–155. https://doi.org/10.1080/09638237.2021.1875413
15. Gori, A., Topino, E., & Griffiths, M. D. (2023). The associations between attachment, self-esteem, fear of missing out, daily time expenditure, and problematic social media use: A path analysis model. *Addictive Behaviors*, *141*, 107633. https://doi.org/10.1016/j.addbeh.2023.107633
16. Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*. https://doi.org/10.48550/arXiv.2203.05794
17. Huang, C. (2020). A meta-analysis of the problematic social media use and mental health. *International Journal of Social Psychiatry*, *68*(1), 12–33. https://doi.org/10.1177/0020764020978434
18. Hunt, M. G., Marx, R., Lipson, C., & Young, J. (2018). No more FOMO: Limiting social media decreases loneliness and depression. *Journal of Social and Clinical Psychology*, *37*(10), 751–768. https://doi.org/10.1521/jscp.2018.37.10.751
19. Keles, B., McCrae, N., & Grealish, A. (2020). A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents. *International Journal of Adolescence and Youth*, *25*(1), 79–93. https://doi.org/10.1080/02673843.2019.1590851
20. Kross, E., Verduyn, P., Demiralp, E., Park, J., Lee, D. S., Lin, N., Shablack, H., Jonides, J., & Ybarra, O. (2013). Facebook use predicts declines in subjective well-being in young adults. *PLoS ONE*, *8*(8), e69841. https://doi.org/10.1371/journal.pone.0069841
21. Lin, L. Y., Sidani, J. E., Shensa, A., Radovic, A., Miller, E., Colditz, J. B., Hoffman, B. L., Giles, L. M., & Primack, B. A. (2016). Association between social media use and depression among U.S. young adults. *Depression and Anxiety*, *33*(4), 323–331. https://doi.org/10.1002/da.22466

22. Naslund, J. A., Aschbrenner, K. A., Marsch, L. A., & 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. https://doi.org/10.1017/S2045796015001067
23. Ni, M. Y., Yang, L., Leung, C. M. C., Li, N., Yao, X. I., Wang, Y., Leung, G. M., Cowling, B. J., & Liao, Q. (2020). Mental health, risk factors, and social media use during the COVID-19 epidemic and Cordon Sanitaire among the community and health professionals in Wuhan, China: Cross-sectional survey. *JMIR Mental Health, 7*(5), e19009. https://doi.org/10.2196/19009
24. Pittman, M., & Reich, B. (2016). Social media and loneliness: Why an Instagram picture may be worth more than a thousand Twitter words. *Computers in Human Behavior, 62*, 155–167. https://doi.org/10.1016/j.chb.2016.03.084
25. Primack, B. A., Shensa, A., Escobar-Viera, C. G., Barrett, E. L., Sidani, J. E., Colditz, J. B., & James, A. E. (2017). Use of multiple social media platforms and symptoms of depression and anxiety: A nationally-representative study among U.S. young adults. *Computers in Human Behavior, 69*, 1–9. https://doi.org/10.1016/j.chb.2016.11.013
26. Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior, 29*(4), 1841–1848. https://doi.org/10.1016/j.chb.2013.02.014
27. Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.
28. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 3982–3992. https://doi.org/10.18653/v1/D19-1410
29. Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, 399–408. https://doi.org/10.1145/2684822.2685324
30. Tandon, A., Dhir, A., Talwar, S., Kaur, P., & Mäntymäki, M. (2022). Social media induced fear of missing out (FoMO) and phubbing: Behavioural, relational and psychological outcomes. *Technological Forecasting & Social Change, 174*, 121149. https://doi.org/10.1016/j.techfore.2021.121149
31. Tandoc, E. C., Ferrucci, P., & Duffy, M. (2015). Facebook use, envy, and depression among college students: Is Facebooking depressing? *Computers in Human Behavior, 43*, 139–146. https://doi.org/10.1016/j.chb.2014.10.053
32. Twenge, J. M. (2017). *iGen: Why today's super-connected kids are growing up less rebellious, more tolerant, less happy—and completely unprepared for adulthood—and what that means for the rest of us*. Simon & Schuster.
33. Valkenburg, P. M. (2022). Social media use and well-being: A review and synthesis. *Annual Review of Psychology, 73*, 219–242. https://doi.org/10.1146/annurev-psych-020821-112328
34. Vogel, E. A., Rose, J. P., Okdie, B. M., Eckles, K., & Franz, B. (2015). Who compares and despairs? The effect of social comparison orientation on social media use and its outcomes. *Personality and Individual Differences, 86*, 249–256. https://doi.org/10.1016/j.paid.2015.06.026
35. Weaver, J. L., & Swank, J. M. (2021). An examination of college students' social media use, fear of missing out, and mindful attention. *Journal of College Counseling, 24*(2), 132–145. https://doi.org/10.1002/jocc.12181
36. Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods, 18*(3), 429–472. https://doi.org/10.1177/1094428114562629