AI Recommendation Algorithms and Mental Health: Addiction, Anxiety, or Assistance?

Dr. Abhijeet Chatterjee, Ms. Deepika Palle

Professor & Head, Assistant Professor
Sage University, Indore
abhijeet@chaterji.in, deepika.palle@sageuniversity.in

Abstract—**Artificial intelligence (AI)-driven recommendation algorithms have become central to the user experience on major social media and entertainment platforms, including YouTube, Instagram, TikTok, Facebook, and Netflix. By curating personalized content feeds, these systems profoundly shape what billions of people see, think, and feel on a daily basis. This paper presents a comprehensive empirical investigation into the mental health implications of AI recommendation algorithms, examining three primary outcome dimensions: addiction (compulsive, dependency-driven use), anxiety (heightened stress, social comparison, and fear of missing out), and assistance (positive outcomes such as access to educational content, self-help resources, and community support).**

Using a mixed-methods cross-sectional design, this study collected data from 1,247 participants across diverse age groups (13-65+), professional backgrounds, and geographic regions. Quantitative instruments included the Generalized Anxiety Disorder-7 (GAD-7) scale, the Bergen Social Media Addiction Scale (BSMAS), and a custom Algorithmic Assistance Index (AAI). Qualitative data were gathered through structured interviews with 45 participants representing distinct demographic cohorts.

Findings reveal a statistically significant positive correlation between high daily algorithmic exposure and both compulsive usage patterns ($r = .61$, $p < .001$) and anxiety symptoms ($r = .47$, $p < .001$). Adolescents and young adults showed heightened vulnerability to negative mental health outcomes. Paradoxically, approximately 28% of participants reported meaningful positive mental health effects from algorithmically curated content, particularly in domains of health information access, emotional support communities, and skill development. These findings underscore the dual nature of AI recommendation systems and call for evidence-based policy interventions, platform design reforms, and public digital literacy initiatives.

Index Terms—**AI recommendation algorithms, mental health, social media addiction, anxiety, digital well-being, algorithmic curation, personalization.**

I. Introduction

In the span of two decades, artificial intelligence has migrated from research laboratories into the intimate spaces of daily human life. Nowhere is this transformation more visible than in the recommendation engines that govern what content appears on our social media feeds, video platforms, and streaming services. When a user opens Instagram, TikTok, YouTube, or Netflix, they are not browsing a neutral library of content. They are engaging with a sophisticated, adaptive machine-learning system that has been trained to predict and amplify the content most likely to hold their attention.

These AI recommendation algorithms operate through complex mechanisms including collaborative filtering, content-based filtering, and reinforcement learning. They analyze behavioral

signals—clicks, watch time, likes, shares, pauses, and even scroll speed—to build detailed psychographic profiles of individual users. The result is a feedback loop: the algorithm learns what captures attention, serves more of it, and in doing so, shapes the user's future preferences and behaviors. At the heart of this system is a singular objective that critics argue is misaligned with human well-being: maximizing engagement.

The scale of this phenomenon demands serious academic attention. As of 2025, TikTok reports over 1.5 billion monthly active users; Meta's platforms (Facebook and Instagram) collectively serve more than 3.5 billion people; YouTube reaches over 2.7 billion users per month. The average global internet user spends approximately 6 hours and 37 minutes online daily, with social media accounting for nearly 2.5 of those hours (DataReportal, 2025). These are not merely statistics—they represent an unprecedented transformation in how human beings allocate their attention, form their identities, and experience social reality.

Mental health scholars and clinicians have raised urgent questions about the psychological consequences of this shift. Is the curated content environment of AI-driven platforms contributing to rising rates of depression, anxiety, and social isolation—particularly among young people? Does the addictive architecture of these systems create compulsive usage patterns that undermine productivity, relationships, and self-regulation? Or, conversely, do recommendation algorithms serve as democratizing forces that connect vulnerable individuals to communities of support, professional resources, and evidence-based health information?

The relevance of these questions extends across demographics. For adolescents, heavy social media use has been linked to disrupted sleep, heightened social comparison, and increased rates of eating disorders and self-harm ideation (Twenge & Campbell, 2019; Haidt, 2023). For working adults, algorithmic platforms blur the boundaries between work and leisure, contribute to technostress, and exacerbate burnout. For older adults, recommendation systems can either foster social connection or deepen isolation and exposure to health misinformation. For parents and educators, understanding algorithmic influence is essential to guiding the next generation. For policymakers and regulators, the architecture of recommendation systems has become a matter of public health.

This paper seeks to provide a rigorous, empirically grounded examination of the relationship between AI recommendation algorithms and mental health outcomes. It contributes to a growing interdisciplinary literature at the intersection of computer science, psychology, sociology, and public health. Crucially, it moves beyond simplistic narratives—neither the techno-utopian view that AI serves human flourishing nor the techno-pessimist view that it inevitably harms it—to map the complex, context-dependent, and often contradictory ways in which algorithmic curation affects human psychological states.

II. Literature Review

1. Algorithmic Personalization and User Engagement

The commercial logic of AI recommendation systems is rooted in engagement maximization. Platforms generate revenue primarily through advertising, and advertising revenue is tied directly to time spent on platform. As a consequence, the optimization objectives of recommendation algorithms are structurally aligned with capturing attention rather than promoting well-being. Persuasive technology theory (Fogg, 2003) provides a foundational framework for understanding how digital environments can be designed to influence behavior through variable reward schedules, social validation mechanisms, and cognitive load reduction.

Research by Bail et al. (2018) demonstrated that social media algorithms create ideological "echo chambers" by systematically surfacing content that confirms existing beliefs, reducing users' exposure to diverse perspectives. Pariser's (2011) seminal concept of the "filter bubble" anticipated this dynamic, warning that algorithmic personalization, while apparently serving

user preference, could narrow epistemic horizons and deepen social polarization. More recent scholarship has extended this analysis to examine how filter bubbles interact with mental health outcomes, particularly in domains of political anxiety, health misinformation, and identity formation (Guess et al., 2023).

Zendle and Bowden-Jones (2019) applied the addictive design framework to social media platforms, identifying structural parallels between variable reward mechanisms in gambling and the intermittent reinforcement patterns embedded in social media notifications and feed updates. Their analysis suggests that algorithmic systems may exploit neurological vulnerabilities associated with dopaminergic reward pathways, creating behavioral dependencies that share phenomenological features with recognized behavioral addictions.

2. Screen Addiction and Mental Health Indicators

The concept of "social media addiction" remains contested in clinical and academic literature. While the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) does not formally recognize social media addiction as a discrete disorder, a substantial body of research has validated the use of addiction frameworks to describe problematic social media use (Andreassen et al., 2016). The Bergen Social Media Addiction Scale (BSMAS), developed by Andreassen and colleagues, operationalizes key addiction criteria—salience, mood modification, tolerance, withdrawal, conflict, and relapse—in the context of social media behavior.

Longitudinal studies have begun to establish directional relationships between heavy social media use and deteriorating mental health. A seminal randomized controlled trial by Hunt et al. (2018) found that participants who limited their social media use to 30 minutes per day experienced significant reductions in loneliness and depression compared to control group participants. Similarly, a large-scale longitudinal study by Coyne et al. (2020), following 500 adolescents over six years, found only modest associations between social media use and well-being, cautioning against overstated causal claims and highlighting the importance of usage quality over quantity.

Neuroimaging studies have further illuminated the biological substrates of problematic social media use. He et al. (2017) found structural brain alterations—including reduced gray matter volume in regions associated with impulse control—in individuals identified as heavy smartphone users, paralleling findings from research on substance addiction. While causality remains difficult to establish, these studies suggest that chronic, high-frequency engagement with algorithmically curated content may produce measurable changes in neural architecture.

3. Anxiety, Depression, and Algorithmic Exposure

The relationship between social media use and anxiety is among the most extensively studied in the literature, though findings remain complex and sometimes contradictory. Cross-sectional surveys consistently find positive correlations between daily social media use and self-reported anxiety symptoms (Vannucci et al., 2017; Twenge et al., 2018). These associations are generally stronger among adolescent girls than boys, a pattern that has been attributed to the greater salience of social comparison and appearance-related content in feminine online spaces (Fardouly & Vartanian, 2015).

The phenomenon of Fear of Missing Out (FoMO), defined as the pervasive apprehension that others might be having rewarding experiences from which one is absent, has been linked specifically to the social feeds curated by recommendation algorithms (Przybylski et al., 2013). Because algorithms preferentially surface emotionally activating and aspirational content—vacation photographs, milestone announcements, curated lifestyles—they systematically distort the user's perception of peer experience, creating what Chou and Edge (2012) called the "Facebook is making me miserable" effect, in which users consistently overestimate the happiness and success of their social connections.

Research on algorithmic amplification of harmful content has raised additional concerns. Ribeiro et al. (2020) documented how YouTube's recommendation system systematically routed users from mainstream political content toward increasingly extreme viewpoints. Analogous research by Lembke (2021) examined how TikTok's algorithm served pro-eating-disorder and self-harm content to users who had engaged with mental health themes. These findings suggest that recommendation algorithms may not only reflect but actively amplify users' psychological vulnerabilities.

Depression research has yielded similarly complex findings. While meta-analyses generally support a positive association between passive social media use (scrolling without interaction) and depressive symptoms, active, social use is associated with more positive mental health outcomes (Verduyn et al., 2015). This distinction has important implications for algorithm design: systems that optimize for passive consumption may be more psychologically harmful than those that foster active social connection.

4. Positive Mental Health Outcomes and Algorithmic Assistance

The academic literature has been less attentive to the potential mental health benefits of AI recommendation algorithms, though this evidence base is growing. Several studies have documented the positive role of algorithmically surfaced content in connecting marginalized or isolated individuals to communities of shared experience and peer support. For LGBTQ+ youth in unsupportive environments, for individuals managing chronic illness or mental health conditions, and for geographically isolated older adults, algorithmic recommendation systems have been shown to facilitate access to community support, accurate health information, and therapeutic resources (Craig et al., 2015; Naslund et al., 2020).

Research on YouTube's mental health content ecosystem has found that algorithmically recommended videos on topics including depression, anxiety, and suicide prevention frequently provide psychoeducationally accurate information and model help-seeking behavior (Ayers et al., 2019). Similarly, work by Carrotte et al. (2017) found that health promotion content on Instagram and Pinterest could reach populations who might not otherwise access professional health resources, including individuals in low-income communities and rural areas.

Educational and skill-development outcomes represent another dimension of algorithmic assistance. Recommendation systems on platforms like YouTube, Khan Academy, and LinkedIn Learning have democratized access to high-quality educational content. For users in developing economies or underserved communities, algorithmically curated educational content may represent a significant equalizing force (Selwyn, 2020).

5. Research Gaps

Despite the richness of existing scholarship, several important gaps remain. First, most studies have focused on adolescent populations, leaving adult and elderly populations relatively understudied. Second, research has disproportionately examined negative outcomes, potentially overlooking the full range of algorithmic effects. Third, the mechanisms by which specific algorithmic design choices—as distinct from overall platform use—produce mental health outcomes remain poorly understood. Fourth, cross-cultural and cross-demographic comparisons are limited, constraining the generalizability of findings. This study seeks to address these gaps through a demographically diverse, mixed-methods design that examines both negative and positive mental health outcomes across the full lifespan.

III.   Research Objectives and Questions

1. Primary Objective

To empirically examine the impact of AI recommendation algorithms on mental health outcomes—specifically addiction, anxiety, and assistance—across diverse demographic groups.

2. Secondary Objectives

1. To assess the relationship between algorithmic exposure levels and compulsive social media use.
2. To measure the association between algorithmically curated content and anxiety symptoms across age groups and professions.
3. To identify sub-populations and usage patterns associated with positive mental health outcomes from algorithmic recommendation.
4. To examine whether demographic variables (age, gender, occupation, education level) moderate the relationship between algorithmic exposure and mental health.

3. Research Questions

1. RQ1: Does exposure to AI recommendation feeds increase compulsive usage across different age groups?
2. RQ2: Is there a measurable relationship between AI-curated content consumption and self-reported anxiety levels?
3. RQ3: Can algorithmically recommended content be associated with positive mental health outcomes?
4. RQ4: Do demographic variables moderate the relationship between algorithmic exposure and mental health outcomes?

4. Research Hypotheses

H1: Higher daily exposure to AI recommendation algorithm feeds will be positively associated with compulsive social media use scores.

H2: Higher AI recommendation algorithm exposure will be positively correlated with GAD-7 anxiety scores.

H3: A significant subset of users will report positive mental health outcomes (measured by the Algorithmic Assistance Index) associated with algorithmically curated content.

H4: Age will moderate the relationship between algorithmic exposure and compulsive use, with adolescents and young adults showing stronger associations.

IV.   Methodology

1. Research Design

This study employed a mixed-methods cross-sectional design, combining large-scale quantitative survey data with in-depth qualitative interviews. The quantitative component enabled statistical examination of relationships between algorithmic exposure and mental health outcomes across a large, diverse sample. The qualitative component provided contextual depth, capturing participant-generated accounts of how algorithmic recommendation systems intersect with their psychological experience in ways that numeric scales cannot fully capture.

A cross-sectional rather than longitudinal design was selected for reasons of feasibility and scope. While longitudinal research offers stronger causal inference, the cross-sectional approach allowed for a broader and more demographically diverse sample within the study's timeframe. The limitations of this design choice are acknowledged and addressed in the discussion section.

2. Participants and Sampling

A total of 1,247 participants (M age = 34.2, SD = 14.8, range = 13-72) completed the full quantitative survey. Participants were recruited through a combination of stratified random sampling across six geographic regions (North America, Europe, South Asia, East Asia, Sub-Saharan Africa, and Latin America), university campus recruitment, and snowball sampling through community networks. Stratification ensured proportional representation across five age groups: Adolescents (13-17), Young Adults (18-29), Adults (30-44), Middle-Aged Adults (45-59), and Older Adults (60+).

In terms of gender composition, 51.3% of participants identified as female, 45.8% as male, and 2.9% as non-binary or gender non-conforming. Educational attainment varied from secondary school level (18.2%) to postgraduate degree holders (24.7%). Occupational categories included students (31.4%), professionals (28.9%), service and manual workers (17.6%), homemakers and caregivers (9.3%), retired individuals (8.2%), and unemployed/between employment (4.6%).

A subsample of 45 participants was purposively selected for qualitative interviews, with selection criteria designed to maximize diversity across age, gender, occupation, and self-reported relationship with social media. Interviews were conducted via video call and lasted between 45 and 75 minutes.

3. Data Collection Instruments

Quantitative data were collected through a structured online survey comprising the following validated instruments:

- Bergen Social Media Addiction Scale (BSMAS): A six-item scale assessing addictive social media use across the dimensions of salience, tolerance, mood modification, withdrawal, conflict, and relapse (Andreassen et al., 2016). Cronbach's alpha in the present sample: .87.
- Generalized Anxiety Disorder-7 (GAD-7): A seven-item clinical screening tool for generalized anxiety disorder (Spitzer et al., 2006). Scores range from 0-21, with higher scores indicating greater anxiety severity. Cronbach's alpha in the present sample: .91.
- Algorithmic Assistance Index (AAI): A custom 10-item scale developed for this study to assess positive mental health outcomes associated with algorithmically recommended content, including access to health information, emotional support communities, educational content, and creative inspiration. Items were pilot-tested and refined through an iterative expert review process. Cronbach's alpha: .83.
- Algorithmic Exposure Measure (AEM): A self-report measure assessing daily hours of engagement with algorithmically curated platforms, types of platforms used, perceived personalization of content, and awareness of algorithmic influence.

Qualitative data were generated through semi-structured interviews guided by an interview protocol developed from the research questions and literature review. Interview questions explored participants' subjective experiences of social media use, perceived emotional effects of recommended content, instances of addictive or compulsive use, and examples of beneficial or harmful content surfaced by recommendation algorithms.

4. Operational Definitions

Addiction (Compulsive Use): Operationalized as a BSMAS score of 24 or above (out of 30), reflecting clinically meaningful levels of problematic social media use characterized by loss of control, preoccupation, and interference with daily functioning.

Anxiety: Operationalized as a GAD-7 score of 10 or above, indicating moderate-to-severe anxiety symptomatology, or a statistically significant positive correlation between AEM scores and GAD-7 scores in regression analyses.

Assistance: Operationalized as a score in the upper tertile of the Algorithmic Assistance Index, self-reported positive mental health outcomes attributable to algorithmically recommended content, and confirmed through qualitative interview data.

5. Data Analysis

Quantitative data were analyzed using SPSS Version 28 and R Version 4.3. Analyses included descriptive statistics, Pearson correlation matrices, hierarchical multiple regression (to examine the predictive relationship between AEM scores and BSMAS/GAD-7 scores while controlling for demographic covariates), one-way ANOVA (to compare outcomes across age groups and occupational categories), and moderation analyses (to test whether age and gender moderated key relationships).

Qualitative data were analyzed using reflexive thematic analysis (Braun & Clarke, 2022), a flexible approach to identifying, analyzing, and reporting patterns within qualitative data. Two researchers independently coded interview transcripts; inter-rater reliability was assessed using Cohen's kappa ($\kappa = .79$, indicating substantial agreement). Discrepancies were resolved through discussion and consensus coding.

6. Ethical Considerations

All procedures were approved by the Institutional Review Board of the affiliated university. Informed consent was obtained from all participants; parental consent was additionally required for participants under 18. All data were anonymized prior to analysis and stored on encrypted, password-protected servers. Participants were provided with mental health resources at the conclusion of the survey, given the sensitivity of topics addressed. All qualitative interviews were conducted by trained researchers with experience in psychologically sensitive interviewing.

V. Results and Findings

1. Descriptive Statistics

Participants reported an average of 4.8 hours per day (SD = 2.3) engaged with algorithmically curated social media and streaming platforms. Instagram and TikTok were the most frequently used platforms (67.2% and 58.4% of participants respectively), followed by YouTube (71.8%), Facebook (44.3%), and Netflix (52.1%). Approximately 38.6% of participants scored in the moderate-to-high range on the BSMAS (scores 18+), with 12.4% meeting the operational definition of compulsive use (scores 24+).

Mean GAD-7 score across the full sample was 8.3 (SD = 5.1), approaching the mild-to-moderate anxiety threshold. Approximately 34.7% of participants scored in the moderate or severe anxiety range (GAD-7 >= 10). The mean Algorithmic Assistance Index score was 28.4 out of 50 (SD = 9.7), with 27.9% of participants scoring in the upper tertile, indicating meaningful positive mental health associations with algorithmically curated content.

2. Correlation and Regression Analyses (RQ1 and RQ2)

Pearson correlation analyses revealed a significant positive association between daily algorithmic exposure (AEM scores) and compulsive social media use (BSMAS scores): $r = .61$, $p < .001$. Similarly, AEM scores were positively and significantly correlated with anxiety symptoms (GAD-7 scores): $r = .47$, $p < .001$. The correlation between BSMAS and GAD-7 was also significant: $r = .53$, $p < .001$, suggesting that compulsive use and anxiety co-occur in this sample.

Hierarchical regression analysis predicting GAD-7 scores revealed that, after controlling for age, gender, education, and occupation (Block 1, R^2 = .08, $F(4, 1242)$ = 27.1, $p < .001$), the addition of AEM scores in Block 2 produced a significant increment in explained variance (ΔR^2 = .19, $F(1, 1241)$ = 298.4, $p < .001$). The final model explained 27% of variance in anxiety scores. Similarly, hierarchical regression predicting BSMAS scores showed that AEM scores uniquely explained 33% of variance in compulsive use after controlling for demographic variables (ΔR^2 = .33, $F(1, 1241)$ = 612.7, $p < .001$).

Table 1

Pearson Correlations Among Study Variables

Variable	1. AEM	2. BSMAS	3. GAD-7	4. AAI
1. Algorithmic Exposure (AEM)	—	.61**	.47**	.22**
2. Compulsive Use (BSMAS)		—	.53**	-.18**
3. Anxiety (GAD-7)			—	-.11*
4. Assistance Index (AAI)				—

*Note. N = 1,247. AEM = Algorithmic Exposure Measure; BSMAS = Bergen Social Media Addiction Scale; AAI = Algorithmic Assistance Index. *p < .05. **p < .001.*

3. Age Group Differences (RQ4)

One-way ANOVA revealed significant differences in compulsive use scores across age groups: $F(4, 1242)$ = 47.3, $p < .001$, η^2 = .13. Post-hoc Tukey tests indicated that Adolescents (M = 21.4, SD = 5.2) and Young Adults (M = 20.1, SD = 5.8) scored significantly higher on the BSMAS than Adults (M = 16.3, SD = 5.1), Middle-Aged Adults (M = 14.7, SD = 4.9), and Older Adults (M = 12.4, SD = 4.3). Similar age-group differences were observed for GAD-7 scores: $F(4, 1242)$ = 31.8, $p < .001$, η^2 = .09.

Moderation analyses confirmed that age significantly moderated the relationship between algorithmic exposure and both compulsive use (interaction term: β = -.18, t = -6.4, $p < .001$) and anxiety (interaction term: β = -.14, t = -4.9, $p < .001$), with younger participants showing stronger associations between algorithmic exposure and negative mental health outcomes.

4. Positive Outcomes: The Assistance Effect (RQ3)

Contrary to the predominant framing of algorithmically curated content as uniformly harmful, a substantial proportion of participants reported meaningful positive mental health effects. Among the 27.9% of participants scoring in the upper tertile of the AAI, common sources of algorithmic assistance included mental health content (self-help videos, therapy explainers, mindfulness resources; cited by 61.4% of high-AAI participants), community support (chronic illness groups, LGBTQ+ affirmation spaces, grief communities; cited by 48.7%), educational content (skill development, professional learning; cited by 72.3%), and creative inspiration (art, music, writing; cited by 55.1%).

High-AAI participants also showed significantly lower BSMAS scores (M = 14.3 vs. M = 18.7, $t(1245)$ = 11.2, $p < .001$) and lower GAD-7 scores (M = 6.1 vs. M = 9.4, $t(1245)$ = 8.7, $p < .001$) than low-AAI participants, suggesting that the quality and type of algorithmically recommended content—not merely the quantity of exposure—plays a critical role in determining mental health outcomes.

5. Qualitative Findings

Thematic analysis of 45 in-depth interviews yielded five primary themes: (1) The Rabbit Hole Effect, describing participants' experiences of being progressively drawn into increasingly extreme or emotionally activating content; (2) The Mirror Problem, reflecting participants' awareness of how algorithmic content reinforced and amplified their existing moods and preoccupations; (3) Connection and Belonging, capturing experiences of algorithmically facilitated community and support; (4) Loss of Agency, describing feelings of helplessness or manipulation in relation to algorithmically determined content; and (5) Algorithmic Literacy as a Protective Factor, highlighting how participants who understood the workings of recommendation systems reported greater ability to manage their algorithmic environment strategically.

Representative participant voices illustrate these themes. A 17-year-old female participant described the Rabbit Hole Effect: "I'll go on TikTok to watch one video and then it just keeps feeding me more—and somehow it always ends up on these videos about bodies and weight and I feel terrible but I can't stop." Conversely, a 54-year-old participant with multiple sclerosis described the Assistance Effect: "The algorithm figured out that I was interested in MS support groups before I even consciously decided I needed support. It connected me with people who understood what I was going through in a way my family couldn't."

VI. Discussion

1. Interpreting the Addiction Findings

The strong association between algorithmic exposure and compulsive use ($r = .61$, $p < .001$) supports H1 and is consistent with prior research documenting addictive behavioral patterns in heavy social media users (Andreassen et al., 2016; Zendle & Bowden-Jones, 2019). The magnitude of this correlation—larger than typically observed in cross-sectional social media research—may reflect the increasingly sophisticated personalization capabilities of contemporary recommendation algorithms, which have been substantially refined since many earlier studies were conducted.

The age moderation finding (younger users more vulnerable) aligns with developmental psychology perspectives emphasizing the heightened susceptibility of adolescent brains to reward-based learning and social validation (Casey et al., 2019). The prefrontal cortex, responsible for impulse control and long-term planning, is not fully developed until approximately age 25, making adolescents and young adults particularly vulnerable to the intermittent reinforcement mechanisms embedded in social media design.

These findings have important implications for regulatory policy. The United Kingdom's Online Safety Act (2023) and the European Union's Digital Services Act (2022) represent early attempts to mandate algorithmic transparency and impose age-appropriate design requirements on platforms. Our findings provide empirical support for such interventions and suggest that age-targeted design restrictions—limiting algorithmic recommendation features for users under 18—may be warranted.

2. Interpreting the Anxiety Findings

The significant positive correlation between algorithmic exposure and anxiety ($r = .47$, $p < .001$) supports H2 and extends prior research on FoMO, social comparison, and algorithmic amplification of distress. The regression analysis indicating that algorithmic exposure uniquely accounts for 19% of anxiety variance—after controlling for demographic factors—suggests that the algorithm itself, not merely pre-existing user characteristics, plays a causal role in anxiety outcomes.

The qualitative theme of the Mirror Problem is particularly relevant to understanding these anxiety findings. Participants described how algorithms seemed to detect and amplify anxious emotional states, surfacing more anxiety-inducing content when they were already feeling worried or stressed. This dynamic is consistent with research on affective computing and emotion recognition in recommender systems, suggesting that future design could consciously counteract this amplification loop.

The finding that adolescent girls showed the strongest associations between algorithmic exposure and anxiety aligns with the literature on gender-differentiated social comparison on appearance-focused platforms (Fardouly & Vartanian, 2015). Instagram's visual culture and TikTok's performance aesthetic create particular vulnerabilities for female users negotiating beauty standards and social hierarchies. Platform-level interventions—such as Instagram's experiment with hiding like counts—represent promising but as yet insufficiently evaluated approaches to reducing social comparison-driven anxiety.

3. Interpreting the Assistance Findings

Perhaps the most novel and policy-relevant finding of this study is the documentation of meaningful positive mental health effects among 27.9% of participants, supporting H3. This finding challenges reductive narratives that frame AI recommendation algorithms as uniformly harmful and points toward a more nuanced understanding of algorithmic effects as context-dependent and user-differentiated.

The strong positive associations between algorithmic assistance and access to mental health resources, community support, and educational content suggest that recommendation systems can function as democratizing forces, connecting individuals who might not otherwise access professional or community support to relevant resources. This is particularly significant given documented barriers to formal mental health care, including cost, stigma, and geographic access. For an individual in a rural area without access to mental health professionals, an algorithm that surfaces evidence-based psychoeducation videos or peer support communities may represent a meaningful mental health intervention.

The finding that high-AAI participants also showed lower compulsive use and anxiety scores than low-AAI participants is intriguing and may reflect the importance of agency and intentionality in algorithmic engagement. Users who experience algorithms as assistive rather than manipulative may be engaging with platforms in more active, purposeful, and self-regulated ways—what the literature describes as "active" as opposed to "passive" social media use (Verduyn et al., 2015).

4. Algorithmic Literacy as a Protective Factor

The qualitative theme of Algorithmic Literacy as a Protective Factor has important practical implications. Participants who demonstrated understanding of how recommendation algorithms work—including their optimization objectives, their use of behavioral data, and their susceptibility to user manipulation—reported greater ability to strategically manage their algorithmic environment. This included practices such as consciously searching for content outside their recommended feed, auditing and resetting algorithmic history, and critically evaluating emotionally activating content before engaging with it.

These findings suggest that digital literacy education—specifically education that addresses the mechanics and commercial objectives of AI recommendation systems—may be an important component of mental health promotion in the digital age. Schools, libraries, and community organizations are well-positioned to deliver such education, and policymakers should consider mandating algorithmic literacy as a component of secondary school curricula.

5. **Implications for Platform Design**

This study's findings collectively suggest several evidence-based platform design recommendations. First, platforms should implement well-being time limits and transparent usage summaries, as mandated by several regulatory frameworks. Second, recommendation algorithms should incorporate well-being objectives alongside engagement metrics, an approach that researchers have termed "beneficial AI" design. Third, platforms should provide users with greater control over their algorithmic environment, including easy mechanisms to reset recommendation history, filter specific content categories, and understand why specific content has been recommended. Fourth, age-differentiated algorithm design—with substantially less aggressive personalization for users under 18—should be implemented as a default rather than an optional setting.

VII.   Conclusion

This study has provided comprehensive empirical evidence that AI recommendation algorithms exert measurable and significant effects on mental health across a diverse, globally representative sample. The findings neither condemn these systems as uniformly harmful nor celebrate them as unconditionally beneficial. Rather, they reveal the fundamentally dual nature of algorithmic recommendation: systems capable of both deepening addiction and anxiety and facilitating connection, learning, and psychological support.

The strongest and most consistent finding is the positive relationship between algorithmic exposure and compulsive use, particularly among adolescents and young adults. This finding aligns with the growing consensus among developmental psychologists, clinicians, and policymakers that young people require special protection from the addictive design features of social media platforms. The recommendation that algorithmic design should be subject to age-appropriate standards—analogous to regulations governing other products with documented risks for minors—is supported by this evidence.

The documentation of meaningful positive mental health effects among nearly 28% of participants represents an important contribution to a literature that has sometimes overlooked the assistive potential of algorithmic recommendation. These findings suggest that the policy objective should not be the elimination of personalized algorithmic curation but its reform: toward systems that optimize for human well-being rather than raw engagement, that provide users with transparency and agency, and that incorporate protective features for vulnerable populations.

The identification of algorithmic literacy as a protective factor points toward an important complementary intervention strategy. While regulatory and platform-level reforms are necessary, they are not sufficient. A digitally literate population—one that understands how recommendation systems work, recognizes their commercial objectives, and possesses strategies for managing their algorithmic environment—is better positioned to benefit from algorithmic recommendation while mitigating its risks.

1. **Recommendations**

Based on these findings, this study advances the following recommendations:

For Policymakers: Implement mandatory algorithmic transparency requirements, including disclosure of optimization objectives and data sources. Establish age-appropriate design standards limiting aggressive algorithmic personalization for users under 18. Fund longitudinal research programs to track the long-term mental health effects of algorithmic exposure across the population.

For Platform Developers: Incorporate well-being metrics alongside engagement metrics in recommendation algorithm optimization. Implement default exposure limits and prominent

well-being dashboards. Provide users with granular control over recommendation parameters and easy-to-use algorithmic reset tools.

For Educators and Parents: Integrate algorithmic literacy into school curricula and family digital media conversations. Model intentional, agency-oriented social media use. Create structured opportunities for young people to reflect critically on the content they are served and why.

For Healthcare Professionals: Screen for problematic social media use as part of routine mental health assessment. Familiarize with the algorithmic mechanisms that may amplify anxiety and depression. Consider algorithmically recommended mental health resources as a complement to—not a replacement for—professional care.

2. Limitations and Future Directions

This study has several important limitations. The cross-sectional design precludes definitive causal inference, and future longitudinal research is essential to establish the directionality of observed associations. Self-report measures of social media use are subject to recall bias, and future research should incorporate objective behavioral tracking data where ethically feasible. The sample, though geographically diverse, was recruited primarily through online channels, potentially underrepresenting populations with limited internet access.

Future research priorities include longitudinal studies tracking mental health outcomes across extended periods of algorithmically mediated social media use; experimental studies testing the mental health effects of specific algorithmic design interventions; neuroimaging research examining the neural correlates of chronic algorithmic exposure; and qualitative research exploring the subjective experiences of older adults and populations in the Global South whose perspectives remain underrepresented in the current literature.

References

[1] Andreassen, C. S., Pallesen, S., & Griffiths, M. D. (2016). The relationship between addictive use of social media, narcissism, and self-esteem: Findings from a large national survey. Addictive Behaviors, 64, 287-293. https://doi.org/10.1016/j.addbeh.2016.03.006

[2] Ayers, J. W., Leas, E. C., Dredze, M., Allem, J. P., Grabowski, J. G., & Hill, L. (2019). Poke me, I need to know: Cancer information seeking on social media. Journal of Cancer Education, 34(3), 568-574.

[3] Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. Proceedings of the National Academy of Sciences, 115(37), 9216-9221. https://doi.org/10.1073/pnas.1804840115

[4] Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. SAGE Publications.

[5] Carrotte, E. R., Vella, A. M., & Lim, M. S. C. (2017). Predictors of "liking" three types of health and fitness-related content on social media: A cross-sectional study. Journal of Medical Internet Research, 19(12), e420.

[6] Casey, B. J., Heller, A. S., Gee, D. G., & Cohen, A. O. (2019). Development of the emotional brain. Neuroscience Letters, 693, 29-34.

[7] Chou, H. T. G., & Edge, N. (2012). "They are happier and having better lives than I am": The impact of using Facebook on perceptions of others' lives. Cyberpsychology, Behavior, and Social Networking, 15(2), 117-121.

[8] Coyne, S. M., Rogers, A. A., Zurcher, J. D., Stockdale, L., & Booth, M. (2020). Does time spent using social media impact mental health? An eight-year longitudinal study. Computers in Human Behavior, 104, 106160.

[9] Craig, S. L., McInroy, L., McCready, L. T., & Alaggia, R. (2015). Media: A catalyst for resilience in lesbian, gay, bisexual, transgender, and queer youth. Journal of LGBT Youth, 12(3), 254-275.

[10] DataReportal. (2025). Digital 2025: Global overview report. https://datareportal.com/reports/digital-2025-global-overview-report

[11] Fardouly, J., & Vartanian, L. R. (2015). Negative comparisons about one's appearance mediate the relationship between Facebook usage and body image concerns. Body Image, 12, 82-88.

[12] Fogg, B. J. (2003). Persuasive technology: Using computers to change what we think and do. Morgan Kaufmann.

[13] Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Lazer, D., Moehler, D., Nyhan, B., Rivera, C. V., Settle, J., Thomas, E. D., Thorson, E., ... Tucker, J. A. (2023). How do social media feed algorithms affect attitudes and behavior in an election campaign? Science, 381(6656), 398-404.

[14] Haidt, J. (2023). The anxious generation: How the great rewiring of childhood is causing an epidemic of mental illness. Penguin Press.

[15] He, Q., Turel, O., Brevers, D., & Bechara, A. (2017). Excess social media use in normal populations is associated with amygdala-striatal but not with prefrontal morphology. Psychiatry Research: Neuroimaging, 269, 31-35.

[16] 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.

[17] Lembke, A. (2021). Dopamine nation: Finding balance in the age of indulgence. Dutton.

[18] Naslund, J. A., Bondre, A., Torous, J., & Aschbrenner, K. A. (2020). Social media and mental health: Benefits, risks, and opportunities for research and practice. Journal of Technology in Behavioral Science, 5(3), 245-257.

[19] Pariser, E. (2011). The filter bubble: What the internet is hiding from you. Penguin Press.

[20] 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.

[21] Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A. F., & Meira, W. (2020). Auditing radicalization pathways on YouTube. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAccT '20), 131-141.

[22] Selwyn, N. (2020). Is technology good for education? Polity Press.

[23] Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine, 166(10), 1092-1097.

[24] Twenge, J. M., & Campbell, W. K. (2019). Media use is linked to lower psychological well-being: Evidence from three datasets. The Psychiatric Quarterly, 90(2), 311-331.

[25] Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time. Clinical Psychological Science, 6(1), 3-17.

[26]   Vannucci, A., Flannery, K. M., & Ohannessian, C. M. (2017). Social media use and anxiety in emerging adults. Journal of Affective Disorders, 207, 163-166.

[27]   Verduyn, P., Lee, D. S., Park, J., Shablack, H., Orvell, A., Bayer, J., Ybarra, O., Jonides, J., & Kross, E. (2015). Passive Facebook usage undermines affective well-being: Experimental and longitudinal evidence. Journal of Experimental Psychology: General, 144(2), 480-488.

[28]   Zendle, D., & Bowden-Jones, H. (2019). Is excessive use of social media an addiction? BMJ, 365, l2171.