

Algorithmic Recommendation Feedback and Mental Health Self Diagnosis

A Theoretical Model Inspired by the Looping Effect

Anna Podolak¹, Agnieszka Zarotyńska² and Robert Sito²

¹ University of Applied Sciences in Nowy Targ
Nowy Targ, Poland

² Wyższa Szkoła Biznesu – National Louis University
Nowy Sącz, Poland

Abstract— This article presents a model of how recommender systems shape diagnostic self-identification among young adults. Drawing on Ian Hacking's looping effect, it examines how classifications influence self-understanding, accounts of experience and behaviour, and how people's responses feed back into expert knowledge, institutions and the social meanings of categories. A focused narrative review of psychological, sociological, clinical and media studies literature was completed on 8 August 2026. Research documents feed personalisation, uneven mental health content quality and associations between exposure and users' beliefs. However, the evidence does not establish a single causal chain from exposure to self-diagnosis, identity change or symptom intensification. The model therefore distinguishes a well-established short recommendation feedback loop from a hypothesised long loop involving institutional knowledge and practice. The former follows from the architecture of personalisation; the latter remains a research task. Depending on context, diagnostic labels can organise experience and facilitate help-seeking or acquire excessive centrality in the self-concept. The article outlines implications for clinical practice and proposes longitudinal studies linking exposure histories, changes in self-description and institutional responses.

Keywords— recommendation feedback, looping effect, self-diagnosis, diagnostic identity, social media, mental health

I. INTRODUCTION

The language of psychiatry and clinical psychology has entered everyday conversations on social media. Distractibility is sometimes described as a symptom of attention-deficit/hyperactivity disorder (ADHD), social overload as autistic masking, and rapid emotional changes in terms associated with borderline personality disorder. These

statements serve different functions. They organise experience, signal a need for difficulties to be taken seriously, express belonging or formulate a provisional hypothesis before a consultation. Only some self-descriptions assert a definite diagnosis.

Public debate categorises these cases too quickly. For one person, online material may provide the first clue that finally brings longstanding difficulties to the attention of an appropriate specialist. Another viewer may interpret behaviour within the broad range of normality as evidence of a disorder after watching a similar video. These differing experiences underpin two dominant accounts: psychoeducation as a route to recognising a problem, and an online trend towards diagnostic labelling. Each captures part of the phenomenon. Platforms participate in both trajectories by changing the availability and ordering of content. Discussions of expanding psychological concepts and possible "prevalence inflation" provide the cultural context for this debate (Haslam, 2016; Foulkes and Andrews, 2023).

The neurodiversity movement complicates this debate by shifting attention from diagnostic accuracy alone to the social function of a label. In a study of 657 people with differing relationships to autism, self-identification and familiarity with neurodiversity were associated with viewing autism as a positive component of identity (Kapp et al., 2013). Participants also identified common ground with the medical approach. This finding concerns autism and illustrates how a label operates simultaneously in clinical and community contexts.

Ian Hacking's concept of the looping effect connects these levels of analysis. We propose describing the relationship between user behaviour and the selection of subsequent content as a short recommendation feedback loop. The critical

transition in the model is from momentary self-recognition in a piece of content to the consolidation of a label through repeated recommendations. The term describes a distribution mechanism and its hypothesised consequences; it does not establish a new psychological effect. The literature documents individual links in this process. The complete causal sequence has not yet been tested in a single study.

The analysis focuses mainly on young adults, in practice people of university age and in early adulthood, because they predominate in research on online self-diagnosis. In this article, “self-diagnosis” means firmly assigning oneself a diagnosis without a completed clinical assessment. A preliminary suspicion, a screening questionnaire result and identification with the experiences of a particular community are distinct phenomena. Combining them into a single category overstates the scale of the problem and obscures differences in their consequences.

II. SCOPE AND METHOD OF THE REVIEW

The article draws on a focused narrative review completed on 8 August 2026. Searches covered PubMed/MEDLINE, Crossref metadata, publishers’ websites and references cited in the retrieved publications. English search terms concerned self-diagnosis, mental health labels, TikTok and other social media, recommender systems, illness identity, labelling effects, cyberchondria and Hacking’s concepts. Peer-reviewed empirical studies, reviews and theoretical works needed to construct the model were included. Classification manuals were treated as normative documents.

Sources were selected purposively. Particular weight was given to studies of actual published content, research involving social media users, reviews addressing diagnosis and identity, and studies explaining how recommendations work. The scope corresponds to the areas identified in a published review protocol on mental health labels among students; its authors also highlighted the lack of research on changes in self-description associated with social media (Alexander et al., 2024). News reports were used only to locate sources. The narrative review procedure did not involve duplicate screening, a formal risk-of-bias assessment or a complete record of excluded items. It provided the basis for a conceptual model rather than an exhaustive synthesis of the literature.

The article distinguishes four levels of inference. A study finding is an observation within the boundaries of the study’s design and sample. The key distinction is between an association and an interpretation: the former denotes the co-occurrence of variables, whereas the latter connects findings from several studies within a theoretical framework. A hypothesis is introduced where an expected relationship awaits direct testing. This distinction is particularly important in a literature dominated by content analyses and cross-sectional designs.

III. THE LOOPING EFFECT AND CLASSIFICATIONS OF PEOPLE

Hacking emphasised that classifications applied to people are not neutral for those they describe. People learn the name of a category, begin to read their biographies through it and encounter institutions organised around that classification. Their responses subsequently enter clinical observation, research, public debate and professional practice. The object of knowledge therefore changes in the course of being understood (Hacking, 1996, 2007; Tsou, 2007).

In his later account, Hacking identified five interacting components: classifications, classified people, institutions, experts and knowledge. This knowledge includes popular understandings of a category alongside specialist findings. The loop runs through textbooks, patients, clinics, schools, patient organisations, the media, law, families and everyday language for describing difficulties. Platforms introduce an intermediary that both observes responses and organises subsequent encounters with content.

Diagnosis also operates institutionally. It provides access to treatment, benefits, educational accommodations and particular forms of recognition. It facilitates communication between specialists and shapes the expectations of both the diagnosed person and those around them (Jutel, 2009; Sims et al., 2021). DSM-5-TR (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision) and ICD-11 (International Classification of Diseases, Eleventh Revision) assess severity, duration, symptom onset, functional impairment and alternative explanations. A short list of readily recognisable behaviours captures only a fragment of this logic (American Psychiatric Association, 2022; World Health Organization, 2024).

Platforms do not replace any of Hacking’s components. They add a layer of distribution, selection and response measurement. A patient’s account, a clinician’s video, an advertisement for a diagnostic service and a joke may appear side by side in the same feed. Their order depends on user behaviour and the platform’s optimisation objectives. Attributing intentions to an algorithm would obscure this arrangement. It involves creators, audiences, moderation rules, interface design and a system that predicts engagement.

IV. THE RECOMMENDATION LAYER WHAT RESEARCH SHOWS

A. Personalisation and repeated exposure

TikTok’s recommender system rapidly adjusts content to traces of user behaviour. Boeker and Urman (2022) demonstrated the influence of personalisation factors on the content presented. Their experiment examined platform behaviour outside the mental health domain. Klug and colleagues (2021) showed that users develop their own, usually incomplete, theories of how the algorithm works and attempt to “train” it deliberately. Viewers do not sit motionless in front of a feed. They search, pause videos, return to creators, comment or scroll past. The system collects these small decisions and rearranges subsequent recommendations.

For this analysis, the route of exposure matters more than the platform's name. In search mode, users formulate a question. A recommendation feed can deliver content to someone who was not looking for information about a particular diagnosis. Both modes now operate on many platforms. Contrasting a "passive" TikTok with an entirely "intentional" YouTube oversimplifies how they work. Research on influence needs to reconstruct how a specific piece of content reached a specific viewer rather than assign a single logic to an entire platform (Boeker and Urman, 2022; White and Horvitz, 2009).

In interviews conducted by Milton and colleagues (2023), users described recognising their own experiences, gaining access to a community and the gradual adjustment of their content feeds. Algorithmic curation felt personal, sometimes even intimate. The concept of the "algorithmized self", self-knowledge formed in relation to a platform's predictions, points in a similar direction (Bhandari and Bimo, 2022). Both studies introduce user experience into the model and explain why a recommendation may be interpreted as a message about oneself.

B. Quality of mental health content

The quality of popular ADHD videos varies substantially. Yeung, Ng and Abi-Jaoude (2022) assessed 100 videos and classified 52% as misleading. In a later study by Karasavva and colleagues (2025), two clinical psychologists judged that fewer than half of the claims about symptoms met DSM criteria. The survey component involved 843 university students: those without a diagnosis, those reporting a self-diagnosis and those with a formal diagnosis. More frequent viewing of ADHD content was associated with greater willingness to recommend material rated both highly and poorly.

A systematic review by Carter and colleagues (2026) provides a broader picture: 27 studies and 5,057 posts concerning mental health and neurodiversity. The proportion of content judged inaccurate ranged from 0 to 56.9%. This variation was related to the topic, platform, sampling approach and assessment criteria. Content produced by professionals generally performed better. In practical terms, this supports specialists' active presence on social media and the development of material explaining the full diagnostic process.

These studies answer different questions. Content analyses measure the consistency of posts with clinical knowledge, interviews describe viewers' experiences, and a cross-sectional survey captures the co-occurrence of exposure and beliefs. Together, they justify incorporating the recommendation layer into the model, but they do not establish the frequency of inaccurate self-diagnoses or causal ordering. We identify data linking exposure histories to subsequent clinical assessment as the most important gap. Without such a link, content quality, distribution and individual responses remain three adjacent levels of analysis.

C. The proposed recommendation feedback model

The proposed model operates at two tempos (Figure 1). In the faster process, a clinical category is simplified for a broad audience. Users relate it to their own experiences and respond,

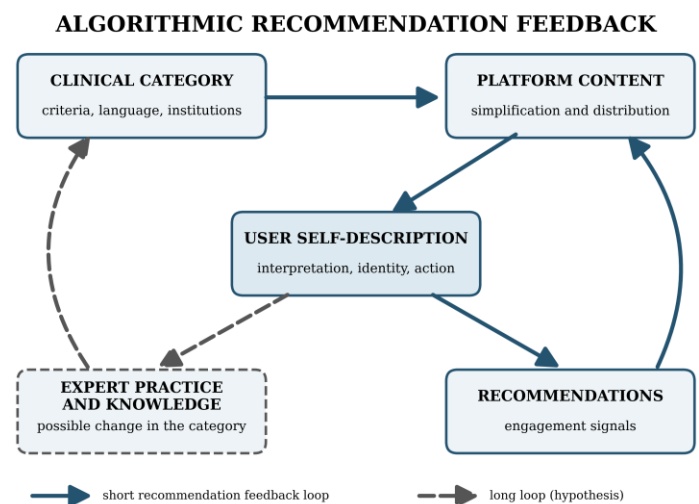
and traces of those responses influence the selection of subsequent material. Repeated exposure familiarises users with the vocabulary and presents ready-made ways of describing symptoms. Some users join a community organised around the label and begin sharing content or creating their own. The slower process emerges when popular representations of diagnoses return to consulting rooms in patients' accounts and enter expert debate.

We call the first process a short recommendation feedback loop. It occurs between user behaviour and subsequent recommendations and follows from the architecture of personalisation. The second, long loop concerns the social meaning of a category, institutional practice and clinical knowledge. It is slower and closer to Hacking's full conception of the looping effect. "Short feedback loop" and "long loop" serve as working terms in this article.

When change ends with users' self-descriptions and the selection of subsequent videos, a short recommendation feedback loop occurs. A looping effect in Hacking's sense begins only when people's responses reach expert knowledge or institutional practice and influence how the category is understood. The current literature provides no clear trace of this transition. The long loop therefore remains a hypothesis and a distinct object of research.

Routes into this process vary. For one person, the starting point may be an incidentally recommended video. Another may have been seeking an explanation for their difficulties for a long time, with the platform responding to an existing need. A mixed trajectory is probably common: symptoms prompt a search, after which a homogeneous feed narrows the range of interpretations considered. Longitudinal data and exposure histories will establish the relative prevalence of these routes.

FIGURE 1. GRAPHICAL ABSTRACT OF THE SHORT RECOMMENDATION FEEDBACK LOOP AND THE HYPOTHESISED LONG INSTITUTIONAL LOOP.



Source: Authors' own elaboration. Solid arrows indicate the short recommendation feedback loop; dashed arrows indicate the hypothesised long loop.

V. SELF DIAGNOSIS IDENTITY AND THE SOCIAL MEANING OF LABELS

Self-diagnosis may be an attempt to name genuine suffering. Rutter and colleagues (2023) found that people who believed they should have a diagnosis of depression, generalised anxiety disorder or insomnia had symptom severity comparable to participants with a formal diagnosis. From a clinical perspective, a stated suspicion therefore conveys information about distress. It should be incorporated into the assessment even when subsequent differential diagnosis leads to another explanation.

In an analysis of 1,195 Reddit comments, Underhill and Foulkes (2025) identified disputes over the right to diagnose, the need for self-understanding and barriers to care. Concerns included self-fulfilling processes and the loss of credibility of people with a diagnosis. These discussions reveal the institutional context of self-diagnosis: cost, waiting lists, a lack of specialists and previous invalidation of difficulties. Under such conditions, the readily accessible language of platforms becomes some people's first tool for describing their experiences.

Diagnosis enters the self-concept in different ways. A review of 38 qualitative studies of young people identified both beneficial and burdensome consequences (O'Connor et al., 2018). Diagnosis facilitated self-understanding, legitimised suffering and provided a reference group. It could also generate stigma, a sense of difference or a restricted view of one's own possibilities. A more recent review confirms the dynamic, multidimensional nature of illness identity (Eddington and Badillo-Winard, 2025). Differences in tools for measuring centrality, acceptance and rejection make findings difficult to compare.

A label does not diminish agency in a single, predictable way. Yanos, Roe and Lysaker (2010) related the significance of illness identity to factors including stigma, hope and self-esteem. Cruwys and Gunaseelan (2016) reported associations between strongly identifying depression with one's self, self-stigma and wellbeing. Within this configuration, the centrality of diagnosis intertwines with the course of the problem, self-evaluation and expectations of the future.

Adopting a label opens several possible pathways. Naming difficulties can reduce shame, support help-seeking and lead to concrete strategies. In other cases, the diagnosis becomes a dominant explanation under which very different experiences are subsumed. A broad continuum lies between these situations. The place of diagnosis in a person's life depends, among other things, on how it was obtained, responses from close others, information quality, access to treatment, symptom course and stigma.

The effects of labelling extend beyond self-evaluation. Diagnosis provides an explanation and access to resources; it also influences behaviour, relationships and professionals' decisions (Sims et al., 2021). In an experiment, labels changed social evaluations of cases with borderline symptom severity (Altmann et al., 2024). A meta-analysis by Kvaale, Haslam and Gottdiener (2013) found that biological explanations of disorders reduced blame, accompanied in some studies by pessimism about prognosis. The evidence in this section

combines qualitative reviews, cross-sectional research and experiments. It describes the social and identity-related consequences of labels but does not reconstruct a single causal sequence within an individual biography.

VI. SYMPTOM PERCEPTION AND THE REINTERPRETATION OF EXPERIENCE

Learning about a diagnostic category shifts attention. People begin to notice behaviours they previously considered unimportant, without necessarily simulating anything consciously. Models of symptom perception describe an experienced sensation as the product of interactions between bodily signals, expectations, context and prior knowledge (Van den Bergh et al., 2017). Research on symptom perception, nocebo, cyberchondria and social learning serves here as a set of mechanistic analogies. A direct test of the model would require exposure histories to be linked to clinical assessment and changes in self-description.

Nocebo research provides a well-developed account of how information and expectations affect reported physical symptoms. In an experiment by Withhöft and Rubin (2013), alarmist media material increased the likelihood of symptoms attributed to electromagnetic exposure. Within the proposed model, this finding supports including attention and expectations as one mechanism through which experience is reinterpreted. Its significance for mental health diagnoses should be measured separately.

Cyberchondria is a related phenomenon with a different typical emotional tone. White and Horvitz (2009) described the escalation of concerns during health information searches, while Starcevic and Berle (2013) synthesised knowledge about excessive use of online medical content. Fear of illness is central to this research. TikTok users' accounts more often emphasised recognition of their own experiences and relief arising from community (Milton et al., 2023). For an individual, anxiety and relief may sometimes occur at successive stages of engagement with content. The emotional function of that engagement therefore becomes a variable in the model, as does the route of exposure: deliberate searching or material delivered through a recommendation feed.

A new vocabulary can also reorganise the account of one's past. Someone learning the criteria for ADHD recalls earlier difficulties at school; someone reading about autism revisits longstanding social strategies. Such retrospection may be diagnostically valuable. Autobiographical memory is reconstructive, and concepts learned later influence the selection of events. In clinical practice, the account is therefore compared with records, information from close others and assessments of functioning at different life stages. Narrative coherence is evaluated alongside other evidence, and gaps in the narrative do not invalidate symptoms.

A particular example is functional tic-like behaviours, which became prominent during the COVID-19 pandemic. Tomczak and colleagues (2024) retrospectively analysed the course of 56 young patients in an observational study without an intervention involving withdrawal from TikTok. Marked improvement was recorded in 79% during follow-up. Given its

clinical distinctiveness, the study serves here solely as an example of the contribution of social learning and the media context to symptom trajectories.

VII. IMPLICATIONS FOR DIAGNOSIS AND PSYCHOTHERAPY

Patients increasingly begin consultations with a ready-made hypothesis. They know the criteria, have completed questionnaires and have compared their experiences with others' accounts. This knowledge becomes the starting point for discussion. It is useful initially to establish the source of the hypothesis, when it emerged, the material viewed and expectations of diagnosis. The exposure history describes the route to the consulting room. It does not, by itself, undermine the patient's account.

Clinical assessment then returns to the onset of difficulties, their persistence and their effects on everyday functioning. For neurodevelopmental disorders, childhood data and information from more than one setting are needed. Differential diagnosis considers factors including sleep, substance use, stress, physical illness, mood disorders and anxiety disorders. A screening questionnaire result indicates a direction for further assessment. The final conclusion follows from an interview and the integration of evidence.

How the outcome is communicated also matters. Dismissing a self-diagnosis in a single sentence can easily sound like a denial of suffering, especially to someone who has sought help for a long time. Unconditional confirmation, conversely, closes differential diagnosis prematurely. Stubbe (2025) proposes that physicians respond to platform messages through the quality of their engagement and explanations. In practice, this means acknowledging patients' observations, honestly describing the limits of the evidence and jointly agreeing on the next steps.

Questions about the history of difficulties and what the label does in the patient's life are helpful: "What does this hypothesis help you make sense of in your experience?" and "Which difficulties were present before you encountered this content?" Further discussion addresses explanations considered previously and what confirmation of the diagnosis would mean. As an illustrative example, a female university student reports ADHD after several months of watching videos on the subject. During the interview, she recalls difficulties at school, while their recent intensification coincides with shift work and chronic sleep deprivation. The label helped her name her experience; only the chronology reopened differential diagnosis. This approach assumes neither bad faith nor secondary gain.

Psychotherapy can broaden the self-concept without disputing the right to use a particular label. Diagnosis is one description of functioning alongside family and occupational roles, relationships, interests and goals. When it leads to effective strategies, it strengthens agency. If it begins to restrict expectations of one's own capabilities, therapeutic work can address the language of those expectations and behaviours through which they can be tested.

Clinicians, too, need to read online content critically. Foster and Ellis (2024) observe that forcefully discrediting self-

diagnosis makes working with young people more difficult. Some popular content emerged in response to symptoms having previously been overlooked, particularly in people whose presentation differs from stereotypes. The accuracy of a hypothesis is assessed using clinical evidence. The route to the consulting room describes the context and may indicate which kinds of explanation the patient has already encountered.

VIII. A RESEARCH AGENDA FOR RECOMMENDATION FEEDBACK AND THE LOOPING EFFECT

The most informative study would begin before self-diagnosis becomes consolidated. A longitudinal project could recruit people who have recently become interested in content about a particular diagnosis for the first time. Brief measurements in everyday settings would track exposure, symptom severity, diagnostic certainty, distress, label centrality and intentions to seek help. With participants' consent, self-reports could usefully be linked to viewing histories or downloaded activity archives. This would distinguish active searching from recommendations triggered by individual interactions.

A formal consultation should constitute a separate measurement point. People whose hypotheses were confirmed, modified or rejected should be compared, followed by monitoring of label centrality and responses to uncertainty. This design would show whether consultation interrupts the previous cycle of interpretation, weakens it or redirects attention towards another category.

An experiment is suitable for examining the short-term influence of presentation format. Participants would view material differing in quality, categorical certainty and the inclusion of information about differential diagnosis. Subsequent measurements would assess recall, recognition of one's own behaviours, confidence in judgement and intentions to seek consultation. The study should not suggest to participants that they have just received a diagnosis. Responses immediately after a video should be analysed separately from lasting identity change.

A second necessary component consists of platform audits using controlled accounts. These would establish how many interactions, and of what kinds, are needed before a feed begins to narrow around mental health content. Results must be analysed separately for platforms, languages and diagnostic categories. Recommender systems change rapidly, so each audit describes a specific version of a service during a specific period.

The long loop requires a different methodology. Analysis of successive versions of guidelines, training materials and clinical documentation could identify traces of popular platform narratives in professional practice. Interviews with diagnosticians should address new forms of patient self-description, the course of consultations and the language of clinical notes. The absence of such traces would also be a finding: it would indicate confirmation of the short recommendation feedback loop without empirically completing Hacking's looping effect.

IX. LIMITATIONS OF THE MODEL

The model brings together literature from several fields that share neither a common method nor a single vocabulary. Content analyses assess posts, user studies rely substantially on self-report, and identity research often concerns people with formal diagnoses. Connecting these bodies of evidence is a theoretical step with a different status from the finding of an individual study.

Most evidence has been collected for TikTok and ADHD. Directly transferring these findings to autism, personality disorders, psychotic disorders or dissociative disorders would be mistaken. These categories differ in criteria, age of onset, the role of developmental data, stigma and the presence of online communities. Format is equally important. Short videos do not fully represent forums, podcasts or closed support groups.

The word “algorithm” is convenient shorthand for a complex technical and social arrangement. Recommendations emerge from data generated by people and objectives set by platform owners. Creators adapt their material to perceived rules of visibility, users attempt to reshape their feeds, and moderation sets the boundaries of available content. The outcome cannot be attributed to a single actor.

The narrative nature of the review is another limitation. Selection may have omitted works published in languages other than English or in disciplines using a different vocabulary. Further operationalisation of the model is necessary before empirical testing, particularly regarding label centrality, exposure homogeneity and changes in diagnostic certainty over time.

X. CONCLUSIONS

Recommender systems shorten the path from encountering a diagnostic category to relating it to oneself and receiving further similar material. Feed personalisation and the uneven quality of mental health content are well documented. The literature on diagnosis shows that a label can enhance self-understanding but, in other circumstances, restrict perceptions of one’s own possibilities. Connecting these findings in a single causal chain is a central task for future research.

Hacking’s framework places the platform within a broader arrangement of classifications, people, institutions, experts and knowledge. The recommendation layer increases the frequency of contact between some of these components, and its significance depends on users’ previous difficulties, how they use the platform and responses from those around them. At present, the evidence primarily concerns the distribution mechanism. The full looping effect remains a research agenda. This distinction explains why the title refers to recommendation feedback and presents Hacking’s concept as the model’s inspiration.

For clinicians, self-diagnosis provides information about how patients understand their problems and how they reached the consulting room. It should be assessed as carefully as any other diagnostic hypothesis. The most informative approach combines a chronology of difficulties, assessment of functioning, differential diagnosis and discussion of the label’s

meaning. This approach draws on the value of online psychoeducation while keeping the diagnostic process open.

XI. REFERENCES

- Alexander, E.D., Chung, V.-H.-A., Yacovelli, A., Sarmiento, I. and Andersson, N. (2024). Social media and postsecondary student adoption of mental health labels: Protocol for a scoping review. *BMJ Open*, 14(2), e078193. <https://doi.org/10.1136/bmjopen-2023-078193>
- Altmann, B., Fleischer, K., Tse, J. and Haslam, N. (2024). Effects of diagnostic labels on perceptions of marginal cases of mental ill-health. *PLOS Mental Health*, 1(3), e0000096. <https://doi.org/10.1371/journal.pmen.0000096>
- American Psychiatric Association (2022). *Diagnostic and statistical manual of mental disorders*. 5th ed., text rev. Washington, DC: American Psychiatric Association Publishing. <https://doi.org/10.1176/appi.books.9780890425787>
- Bhandari, A. and Bimo, S. (2022). Why’s everyone on TikTok now? The algorithmized self and the future of self-making on social media. *Social Media + Society*, 8(1), 20563051221086241. <https://doi.org/10.1177/20563051221086241>
- Boeker, M. and Urman, A. (2022). An empirical investigation of personalization factors on TikTok. In: *Proceedings of the ACM Web Conference 2022*. New York: Association for Computing Machinery, pp.2298–2309. <https://doi.org/10.1145/3485447.3512102>
- Carter, A., Gracey, F., Moody, J., Ovens, A. and Chatburn, E. (2026). Quality, reliability and misinformation in mental health and neurodivergence content on social media: A systematic review. *Journal of Social Media Research*, 3(1), pp.30–47. <https://doi.org/10.29329/jsomer.84>
- Cruwys, T. and Gunaseelan, S. (2016). ‘Depression is who I am’: Mental illness identity, stigma and wellbeing. *Journal of Affective Disorders*, 189, pp.36–42. <https://doi.org/10.1016/j.jad.2015.09.012>
- Eddington, K.M. and Badillo-Winard, E. (2025). Mental illness identity: A scoping review. *Identity*, 25(3), pp.412–427. <https://doi.org/10.1080/15283488.2024.2379412>
- Foster, A. and Ellis, N. (2024). TikTok-inspired self-diagnosis and its implications for educational psychology practice. *Educational Psychology in Practice*, 40(4), pp.491–508. <https://doi.org/10.1080/02667363.2024.2409451>
- Foulkes, L. and Andrews, J.L. (2023). Are mental health awareness efforts contributing to the rise in reported mental health problems? A call to test the prevalence inflation hypothesis. *New Ideas in Psychology*, 69, 101010. <https://doi.org/10.1016/j.newideapsych.2023.101010>
- Hacking, I. (1996). The looping effects of human kinds. In: D. Sperber, D. Premack and A.J. Premack, eds. *Causal cognition: A multidisciplinary debate*. Oxford: Clarendon Press, pp.351–394.
- Hacking, I. (2007). Kinds of people: Moving targets. In: P.J. Marshall, ed. *Proceedings of the British Academy*, Volume 151, 2006 Lectures. London: British Academy, pp.285–318. <https://doi.org/10.5871/bacad/9780197264249.003.0010>
- Haslam, N. (2016). Concept creep: Psychology’s expanding concepts of harm and pathology. *Psychological Inquiry*, 27(1), pp.1–17. <https://doi.org/10.1080/1047840x.2016.1082418>
- Jutel, A. (2009). Sociology of diagnosis: A preliminary review. *Sociology of Health & Illness*, 31(2), pp.278–299. <https://doi.org/10.1111/j.1467-9566.2008.01152.x>
- Kapp, S.K., Gillespie-Lynch, K., Sherman, L.E. and Hutman, T. (2013). Deficit, difference, or both? Autism and neurodiversity. *Developmental Psychology*, 49(1), pp.59–71. <https://doi.org/10.1037/a0028353>
- Karasavva, V., Miller, C., Groves, N., Montiel, A., Canu, W. and Mikami, A. (2025). A double-edged hashtag: Evaluation of #ADHD-related TikTok content and its associations with perceptions of ADHD. *PLOS ONE*, 20(3), e0319335. <https://doi.org/10.1371/journal.pone.0319335>

- Klug, D., Qin, Y., Evans, M. and Kaufman, G. (2021). Trick and please: A mixed-method study on user assumptions about the TikTok algorithm. In: Proceedings of the 13th ACM Web Science Conference 2021. New York: Association for Computing Machinery, pp.84–92. <https://doi.org/10.1145/3447535.3462512>
- Kvaale, E.P., Haslam, N. and Gottdiener, W.H. (2013). The 'side effects' of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review*, 33(6), pp.782–794. <https://doi.org/10.1016/j.cpr.2013.06.002>
- Milton, A., Ajmani, L., DeVito, M.A. and Chancellor, S. (2023). 'I see me here': Mental health content, community, and algorithmic curation on TikTok. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery, pp.1–17. <https://doi.org/10.1145/3544548.3581489>
- O'Connor, C., Kadianaki, I., Maunder, K. and McNicholas, F. (2018). How does psychiatric diagnosis affect young people's self-concept and social identity? A systematic review and synthesis of the qualitative literature. *Social Science & Medicine*, 212, pp.94–119. <https://doi.org/10.1016/j.socscimed.2018.07.011>
- Rutter, L.A., Howard, J., Lakhan, P., Valdez, D., Bollen, J. and Lorenzo-Luaces, L. (2023). 'I haven't been diagnosed, but I should be': Insight into self-diagnoses of common mental health disorders: Cross-sectional study. *JMIR Formative Research*, 7, e39206. <https://doi.org/10.2196/39206>
- Sims, R., Michaleff, Z.A., Glasziou, P. and Thomas, R. (2021). Consequences of a diagnostic label: A systematic scoping review and thematic framework. *Frontiers in Public Health*, 9, 725877. <https://doi.org/10.3389/fpubh.2021.725877>
- Starcevic, V. and Berle, D. (2013). Cyberchondria: Towards a better understanding of excessive health-related Internet use. *Expert Review of Neurotherapeutics*, 13(2), pp.205–213. <https://doi.org/10.1586/ern.12.162>
- Stubbe, D.E. (2025). Patient self-diagnosis: Physician engagement tools to compete with TikTok. *Focus*, 23(2), pp.212–216. <https://doi.org/10.1176/appi.focus.20250004>
- Tomczak, K.K., Worhach, J., Rich, M., Swearingen Ludolph, O., Eppling, S., Sideridis, G. and Katz, T.C. (2024). Time is ticking for TikTok tics: A retrospective follow-up study in the post-COVID-19 isolation era. *Brain and Behavior*, 14(3), e3451. <https://doi.org/10.1002/brb3.3451>
- Tsou, J.Y. (2007). Hacking on the looping effects of psychiatric classifications: What is an interactive and indifferent kind? *International Studies in the Philosophy of Science*, 21(3), pp.329–344. <https://doi.org/10.1080/02698590701589601>
- Underhill, R. and Foulkes, L. (2025). Self-diagnosis of mental disorders: A qualitative study of attitudes on Reddit. *Qualitative Health Research*, 35(7), pp.779–792. <https://doi.org/10.1177/10497323241288785>
- Van den Bergh, O., Withöft, M., Petersen, S. and Brown, R.J. (2017). Symptoms and the body: Taking the inferential leap. *Neuroscience & Biobehavioral Reviews*, 74, pp.185–203. <https://doi.org/10.1016/j.neubiorev.2017.01.015>
- White, R.W. and Horvitz, E. (2009). Cyberchondria: Studies of the escalation of medical concerns in web search. *ACM Transactions on Information Systems*, 27(4), article 23, pp.1–37. <https://doi.org/10.1145/1629096.1629101>
- Withöft, M. and Rubin, G.J. (2013). Are media warnings about the adverse health effects of modern life self-fulfilling? An experimental study on idiopathic environmental intolerance attributed to electromagnetic fields (IEI-EMF). *Journal of Psychosomatic Research*, 74(3), pp.206–212. <https://doi.org/10.1016/j.jpsychores.2012.12.002>
- World Health Organization (2024). ICD-11 for mortality and morbidity statistics, release 2024-01 [online]. Available at: <https://icd.who.int/browse/2024-01/mms/en> [Accessed 31 May 2026].
- Yanos, P.T., Roe, D. and Lysaker, P.H. (2010). The impact of illness identity on recovery from severe mental illness. *American Journal of Psychiatric Rehabilitation*, 13(2), pp.73–93. <https://doi.org/10.1080/15487761003756860>
- Yeung, A., Ng, E. and Abi-Jaoude, E. (2022). TikTok and attention-deficit/hyperactivity disorder: A cross-sectional study of social media content quality. *The Canadian Journal of Psychiatry*, 67(12), pp.899–906. <https://doi.org/10.1177/07067437221082854>