

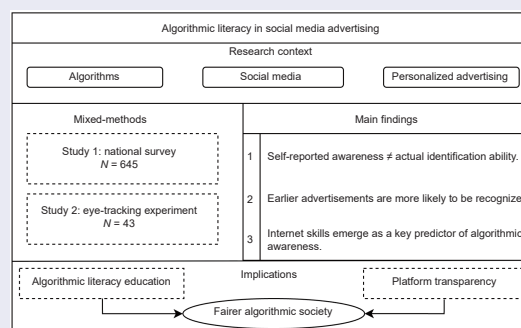
Measuring literacy in the algorithmically infused society: A mixed methods study on social media algorithmic advertising

Yunci Tian, Pu Yan✉

Cite this article: Tian YC, Yan P. *iLibrary* 2026, 1(1): 9690002. <https://doi.org/10.26599/ILIB.2026.9690002>

ABSTRACT: This study addresses the challenges of information equality by investigating algorithmic literacy and the capacity to understand and navigate recommendation algorithms. It responds directly to the core concern of building an algorithmic society by examining how users perceive and identify algorithmically generated social media ADs. A mixed-methods design was used in this study. A nationwide survey ($N = 645$) assessed internet use and algorithmic literacy, followed by an eye-tracking experiment ($N = 43$, Tobii Pro Fusion) that recorded users' AD identification behavior in simulated social media feeds, providing behavioral validation of the survey findings. Quantitative data from both eye-tracking metrics and the survey were statistically analyzed using R. This included correlation tests and regression modeling to identify the key factors and pathways influencing the ability to recognize ADs. The analysis revealed that algorithmic awareness and age were significant factors in the ability to identify algorithmic advertisements. Furthermore, proficiency in internet skills was identified as a crucial factor for enhancing algorithmic awareness. The findings provide crucial user insights and empirical evidence for the design of a fairer and more transparent algorithmic system and provide strategies for improving algorithmic literacy education that could bridge the digital divide. This is essential to foster an equitable and digitally enlightened society.

KEYWORDS: algorithmic literacy; social media; mixed methods; eye-tracking



1 Introduction

1.1 The rise of algorithmic society

Social media, comprising a wide range of applications built upon Web 2.0 technologies, is characterized by user-generated content (UGC) (Kaplan and Haenlein, 2010). It has become an integral part of modern life, with 5.24 billion users worldwide (63.9% of the global population) (Meltwater, 2025). Within this data-saturated “black-box society” (Pasquale, 2015), algorithms have become central to analytical and decision-making processes, a phenomenon that is particularly pervasive on social-media platforms (Gillespie, 2014). These systems have evolved from simple sorting mechanisms to complex predictive models that personalize content delivery and thus enhance user experience.

However, algorithms also entail substantial risks, including commercial bias, digital inequality, and diminished transparency and accountability of online systems (Just and Latzer, 2017; Koren et al., 2009). These problems have been conceptualized through frameworks such as *filter bubbles*, *echo chambers*, and *selective exposure* (Arguedas et al., 2022). Algorithmic biases may reinforce

demographic-based filter bubbles, thereby constraining informational diversity and reducing content richness (Ekstrand et al., 2018; Pariser, 2011), which in turn aggravates preexisting social inequalities (Melchiorre et al., 2021; Rekabsaz et al., 2021). Moreover, recommendation systems are vulnerable to privacy breaches through inference attacks (Ganhör et al., 2022; Zhang et al., 2024).

As algorithms become deeply embedded in all dimensions of social operation, algorithmic society should be understood not only as a technological formation but also as a sociocultural paradigm. In this context, China has gradually established a comprehensive algorithmic governance system that emphasizes transparency, accountability, and social stability. Since the E-Commerce Law (2018) first prohibited “big data-driven price discrimination”, a series of landmark regulations have built the foundation for algorithmic oversight. The Anti-Monopoly Guidelines on the Platform Economy (2021a) explicitly addressed algorithmic collusion, while the Personal Information Protection Law (2021b) incorporated the right to object to automated decision-making, echoing the General Data Protection Regulation (GDPR) framework. The Provisions on the Administration of

Department of Information Management, Peking University, Beijing 100871, China.

✉ Corresponding author. E-mail: puyan@pku.edu.cn

Received: February 25, 2026; Revised: April 20, 2026; Accepted: June 7, 2026

© The Author(s) 2026. This is an open access article under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0, <http://creativecommons.org/licenses/by/4.0/>).

Algorithmic Recommendation for Internet Information Services (2022) formally introduced the transparency and accountability requirements for recommendation algorithms. Together with Cybersecurity Law, Data Security Law, and additional regulations on critical information infrastructure and data sharing, these efforts form a multilayered governance architecture for the algorithmic environment in China's digital economy.

Within the field of library and information science (LIS), the rise of an algorithmic society has brought new disciplinary challenges. Traditional models of information organization, retrieval, and access are increasingly mediated by algorithmic systems, resulting in algorithmic explainability, information fairness, and knowledge visibility, which are central issues of concern. Scholars have begun to examine how algorithmic mechanisms shape users' information behavior and have emphasized the importance of maintaining human agency in algorithmic knowledge mediation.

Recent LIS studies have explored the characteristics and consequences of algorithmic recommendations for digital knowledge systems. Zhang et al. (2021) conducted a content analysis of 23 academic resource platforms and demonstrated that recommendation mechanisms differ in terms of form, interpretability, and temporal relevance, suggesting that user-centered optimization is essential for equitable access. Yu and Dong (2023) further argued that algorithmic bias in digital reading platforms can result in pseudopersonalization and suppression of individual rights, calling for governance approaches grounded in human-centered values.

With the recent expansion of artificial intelligence and large language models, algorithmic mediation has entered a new phase in library service innovation. Chen (2020) highlighted the potential of knowledge graphs to enhance recommendation precision by linking user profiles with semantic document relationships. Chen and Pan (2023) integrated topic modeling and reader profiling to construct reader-book knowledge graphs for context-aware book recommendations. More recently, Guo et al. (2025) combined ChatGPT and BERT to develop an encoding-classification strategy that improves semantic understanding and interpretability of book recommendations. Similarly, Niu et al. (2025) constructed a large-model-based personalized resource recommendation system for university libraries capable of capturing multimodal and temporal user preferences while addressing privacy and reliability concerns. Together, these developments signal the emergence of intelligent, adaptive library systems that utilize generative AI to provide personalized and ethically responsible knowledge services.

Collectively, these transformations mark a critical juncture in the evolution of both algorithmic governance and information service design. As algorithms increasingly shape how individuals access, interpret, and value information, their transparency, fairness, and interpretability have become central to the mission of libraries and academic institutions. However, technological regulation and design innovation alone are insufficient. Users must possess the awareness and critical understanding necessary to navigate algorithmic environments responsibly. In this context, algorithmic literacy emerges as a vital civic and informational competence that determines whether individuals can meaningfully engage with, question, and make informed decisions within algorithmically curated knowledge systems.

1.2 Algorithmic literacy and digital literacy

Currently, the cultivation of algorithmic and digital literacy has become a strategic priority in public policy. In 2021, the Cyberspace Administration of China issued the Action Plan for

Improving National Digital Literacy and Skills, positioning digital literacy as a central pathway toward national modernization, bridging the digital divide, and fostering inclusive development (Cyberspace Administration of China, 2021). The plan emphasizes that strengthening citizens' literacy and competencies in digital environments is essential for building a "strong cyber nation". Similarly, the Opinions on Accelerating the Promotion of Education Digitalization, jointly released by nine ministries including the Ministry of Education, call for a vertically integrated digital literacy framework spanning primary, secondary, and higher education, embedding digital and algorithmic competencies into national evaluation systems (Ministry of Education, 2022). Collectively, these policy initiatives demonstrate China's growing recognition that understanding and critically engaging with algorithmic systems are fundamental to civic participation and social equity in the digital era.

Within the LIS community, scholars have explored how public service institutions can contribute to enhancing citizens' algorithmic literacy. Wu (2022) argued that libraries, while advancing intelligent and personalized services, must also shoulder the responsibility of fostering public algorithmic literacy. This involves both internal innovation, by prioritizing transparency, explainability, and accountability in algorithmic design, and external education, by incorporating algorithmic awareness into user-oriented cultural and educational programs. Similarly, Wang et al. (2020) emphasized that in mobile reading contexts shaped by algorithmic recommendation, readers should cultivate comprehensive "mobile reading literacy", encompassing awareness of information environments, digital habits, operational strategies, and personal information security. These perspectives reframe libraries not only as service providers but also as institutions of critical algorithmic education. In sum, algorithmic literacy is not merely a subset of digital competence but a capacity essential for equitable participation in the algorithmic society.

Navigating this complex environment requires users to develop algorithmic literacy, the ability to understand and critically engage with content recommendation algorithms. Existing empirical studies have primarily relied on surveys and interviews, which risk over- or underestimating algorithmic literacy because they depend on self-reported data. To address this methodological limitation, the present study adopts a mixed-methods design that combines a survey with an eye-tracked click-through AD recognition task. This approach provides behavioral evidence to complement subjective self-reports, thereby offering a more accurate assessment of the gap between users' perceived awareness and their actual performance.

This study aimed to examine the level of algorithmic literacy among social media users by analyzing their observable behaviors and the factors influencing them. The ability to identify algorithmically generated advertisements serves as a direct behavioral proxy for algorithmic literacy. Analyzing task performance helps to reveal behavioral patterns and cognitive differences among diverse user groups, whereas the survey provides complementary insights into potential determinants.

2 Literature review

Algorithms have become deeply embedded in modern life, and digital societies are increasingly permeated by traces of algorithmic selection. Existing research on algorithmic literacy has primarily developed along three interrelated strands: (1) studies on algorithmic awareness and user understanding of algorithmic systems, (2) research on algorithm-related skills and user practices, and (3) conceptual work on users' perceptions and interpretations

of algorithmic processes. Together, these perspectives provide a foundation for understanding how individuals engage with algorithmically curated environments. Within this context, examining how users understand and respond to algorithms—referred to as algorithmic literacy—is crucial for understanding their digital behaviors.

The academic exploration of algorithmic literacy originated from the concept of algorithmic awareness. Eslami et al. (2015) conducted a pioneering study with 40 Facebook users to examine their awareness of the platform's algorithm and compare their participants' perceived control and satisfaction before and after learning about its existence. In the same year, Rader and Gray (2015) studied 464 Facebook users and revealed significant diversity in user awareness, contrary to Eslami's (2015) finding that more than half of participants were unaware of the algorithm's presence.

Refining the participant scope, a survey of 449 university students and faculty found that most respondents could accurately identify the basic mechanisms of Twitter's trending algorithms (Proferes, 2017). As algorithms continue to grow in importance across digital platforms, contemporary social media users, particularly TikTok users, exhibit greater awareness and sensitivity to the role algorithms play in shaping their online experiences (Bruns, 2019). These inconsistent findings indicate that algorithmic awareness is a complex, multidimensional construct that resists universal measurements.

Thus, early research largely focused on developing tools to measure algorithmic awareness and related concepts—such as algorithmic knowledge, algorithmic skills, and algorithmic imagination—within specific contexts. For instance, Klawitter and Hargittai's (2018) interview-based study demonstrated that varying levels of algorithmic skills affect a business's ability to reach potential customers. Through an analysis of tweets and user interviews, Bucher (2017) gained an in-depth understanding of user experiences with Facebook's algorithm and introduced the concept of "algorithmic imagination", which encompasses user perceptions, expectations, and interpretations of algorithmic operations. Using a random sample from a survey of American internet users, Cotter and Reisdorf (2020) identified structural social inequalities as the root cause of knowledge disparities in this domain. Although these distinct concepts and tools can be adequately explained within their respective research frameworks, their underlying connotations are often intertwined. As Hargittai observed in research on algorithmic skills, the mere awareness that algorithms determine the content presented to users is itself a distinct skill (Hargittai et al., 2020), suggesting that these concepts are difficult to fully disentangle in practice.

Consequently, scholars have attempted to refine this definition. Cotter and Reisdorf (2020) defined algorithmic awareness as users' understanding of algorithms, including their definition, functions, benefits to individuals, and potential negative impacts on both individuals and groups. Swart (2021) and Zarouali et al. (2021) further incorporated ethical dimensions such as fairness, transparency, accountability, and trust into its core. Although subtle differences exist among these definitions, there is a broad scholarly consensus that algorithmic awareness denotes users' understanding of how algorithms function and influence them. Concerns regarding transparency and fairness remain essential for cultivating authenticity in algorithmic systems and fostering a digitally enlightened society.

Therefore, this study adopted a more integrated perspective. Prior studies have provided important insights into users' awareness, skills, and perceptions of algorithmic systems. However, most existing research relies heavily on self-reported

measures, which may not fully capture users' actual behavioral abilities in interacting with algorithmic content. In line with the views of Bruns (2019) and Swart (2021), this paper conceptualizes algorithmic literacy as a comprehensive framework encompassing algorithmic awareness, knowledge, skills, strategies, and attitudes. This perspective emphasizes the holistic competence users need to navigate increasingly complex digital environments. Given the multidimensional nature of algorithmic literacy, employing a mixed-methods design that integrates subjective self-reports with objective behavioral measures is crucial for capturing its full scope. This study applies an integrative perspective to provide empirical evidence for enhancing public algorithmic literacy and promoting digital equality, ultimately contributing to a more transparent and digitally enlightened society.

Grounded in the theoretical and empirical literature reviewed above, this study formulates one research question and a set of hypotheses to examine algorithmic literacy from both behavioral and perceptual perspectives. The study first explores the discrepancy between users' subjective awareness and objective identification behavior, then analyzes behavioral mechanisms in the algorithmic AD identification task, and finally investigates the structural determinants of algorithmic awareness.

Research question 1 (RQ1). *Is there a discrepancy between users' self-reported algorithmic awareness and their actual behavioral ability to identify algorithmic advertisements?*

To further examine the behavioral mechanisms underlying algorithmic AD identification, the following hypotheses are proposed:

H1a. Age positively predicts the time required to identify the first algorithmic advertisement.

H1b. Algorithmic awareness positively predicts the time required to identify the first algorithmic advertisement.

H2. AD sequence position negatively predicts the likelihood of AD identification—earlier ADs are more likely to be identified.

H3. Educational level, male gender, and internet skills are significant positive predictors of algorithmic awareness.

In summary, the research design connects the perceptual and behavioral dimensions of algorithmic literacy. RQ1 bridges subjective cognition and behavioral performance; H1a–H2 identify the behavioral mechanisms and contextual factors underlying identification ability; and H3 highlights the broader social and individual determinants shaping algorithmic awareness. This research question and these hypotheses together provide a coherent framework linking self-perception, behavior, and structure in the study of algorithmic literacy.

3 Methods

This study focuses on the experiences of general social media users. It was organized into two sequential studies—study 1 (survey) and study 2 (behavioral experiment)—that build upon one another to address the research questions through a coherent progression of inquiry. The overall research workflow is illustrated in Fig. 1. Although study 2 was conducted earlier than study 1, the numbering follows the conceptual sequence of the research design rather than the chronological order of data collection.

3.1 Data collection

The present research was conducted using two sequential studies that together addressed the proposed research questions. Both studies were reviewed and approved by the ethics committee of the authors' department. Participants in both studies received detailed explanations and provided their consent forms at the beginning of the study. The private information of participants,

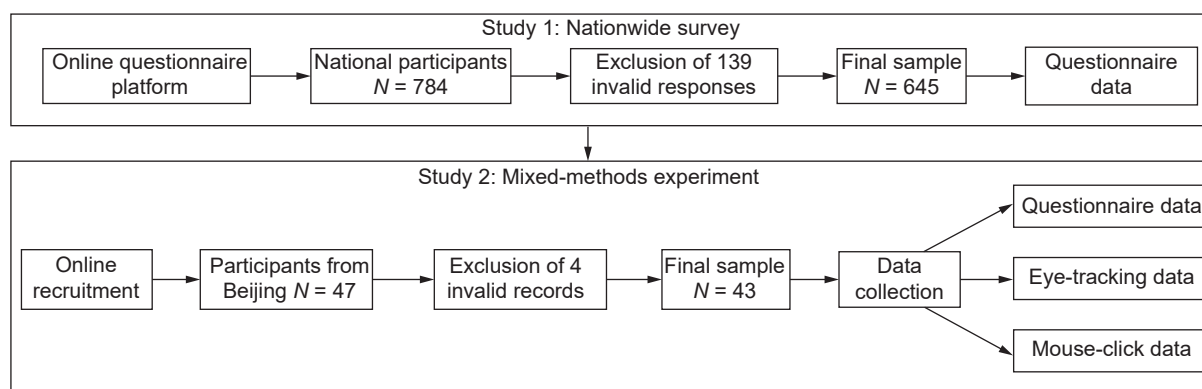


Fig. 1 Research workflow.

such as contact information, was not collected by the research team. All data reported in this study were anonymized.

3.1.1 Study 1 (survey)

Study 1 was conducted between July 15 and August 11, 2024, using a large-scale nationwide online questionnaire. This study aimed to examine how algorithmic literacy can be measured and to identify sociodemographic and internet-use factors influencing users' self-reported algorithmic awareness. Participants were recruited via the Tencent Questionnaire, an established online Chinese survey platform, yielding 645 valid responses. Each verified respondent received automatic compensation of 5 CNY upon completion.

The platform employed a simple random sampling mechanism within its respondent pool, allowing data to be collected across a broad range of regions, age groups, and occupations. This design enhances the diversity and representativeness of the dataset among active Chinese social media users. However, as participation remained voluntary and limited to individuals with reliable internet access, the sample may overrepresent digitally literate and urban participants, potentially limiting the generalizability of the findings to populations with lower digital access or media engagement.

3.1.2 Study 2 (behavioral experiment)

Study 2 was implemented from March 30 to April 7, 2024, employing a mixed-methods design that combined a brief questionnaire with an eye-tracked click-through AD-identification task. This study focused on users' behavioral identification performance, exploring group differences and the influence of AD-sequence order. A total of 43 valid participants were recruited via social-media platforms and screened through an eligibility interview ensuring they were at least 18 years old with corrected visual acuity $\geq 20/40$. Each participant received 50 CNY per hour as compensation.

The findings from study 1 (survey) provided an initial understanding of how algorithmic literacy could be measured and which demographic and digital-use variables were associated with users' perceived algorithmic awareness. However, self-reported measures alone may not fully reflect users' actual comprehension or behavioral capacity. To address this limitation, study 2 (behavioral experiment) employed eye-tracked click-through data to objectively assess the participants' ability to identify algorithmic advertisements. The behavioral results revealed that subjective algorithmic awareness did not always match actual identification performance, suggesting a potential cognitive-behavioral gap in algorithmic literacy. Building upon this finding, the experiment further examined which factors influence users' true identification ability, thereby extending the analysis beyond self-perception toward measurable behavioral evidence.

3.1.3 Descriptive statistics of participants in both studies

In study 1 (survey), the 645 valid respondents had a mean age of 30.5 years ($SD = 9.7$, range = 18–71), with a nearly balanced gender distribution (59.5% female, 39.2% male, 1.2% nonbinary). Most participants reported monthly incomes between 0 and 10,000 CNY (84.7%), and 76.3% resided in urban areas. The sample was relatively well educated, with the majority holding at least a college degree (55.5% rated 6 on the 7-point education scale, mean = 5.5, $SD = 1.0$). Table 1 presents the demographic characteristics of the participants in study 1. These distributions suggest a diverse but moderately young and digitally active population suitable for investigating patterns of algorithmic literacy among general social media users.

Table 1 Demographic characteristics of study 1 participants ($N = 645$)

Variable	Category	N	(%)
Gender	Female	384	59.5
	Male	253	39.2
	Nonbinary	8	1.2
Age	Mean (SD)	30.5 (9.7)	—
	Range	18–71	—
Residence	Urban	492	76.3
	Rural	153	23.7
Income	< 10,000 CNY	546	84.7
	$\geq 10,000$ CNY	99	15.3
Education	Mean (SD)	5.5 (1.0)	—

Note: Percentages are rounded to one decimal place and the total may not equal exactly to 100 percent.

In study 2 (behavioral experiment), the 43 participants had a mean age of 31 years old ($SD = 8.9$, range = 20–61) and were predominantly female (65.1%), with 34.9% male. The majority came from rural areas (83.7%), representing a contrast to the urban-heavy survey sample. Income levels were concentrated below 20,000 CNY (approximately 62.8% of participants), and the educational level was high (mean = 5.9, $SD = 1.2$). Table 2 presents the demographic characteristics of the participants in study 2. This smaller yet homogeneous experimental group provided a focused dataset for analyzing behavioral identification performance.

Together, the two datasets offer complementary demographic structures: the large-scale survey (study 1) ensures representativeness and external validity, while the controlled behavioral experiment (study 2) supports internal validity and deeper interpretation of algorithmic identification behavior.

3.2 Survey design

A unified questionnaire structure was applied across both studies

Table 2 Demographic characteristics of study 2 participants ($N = 43$)

Variable	Category	N	(%)
Gender	Female	28	65.1
	Male	15	34.9
Age	Mean (SD)	31.0 (8.9)	—
	Range	20–61	—
Residence	Urban	7	16.3
	Rural	36	83.7
Income	< 20,000 CNY	27	62.8
	≥ 20,000 CNY	16	37.2
Education	Mean (SD)	5.9 (1.2)	—

to ensure conceptual comparability between self-reported and behavioral measures of algorithmic literacy. The survey was designed to assess multiple theoretical dimensions of users' digital and algorithmic competence by integrating validated scales from previous research. The main body of the questionnaire consisted of a Likert-scale section that captured six conceptual dimensions related to algorithmic literacy and general digital literacy: algorithmic access, algorithmic awareness, mental health, digital skills, privacy awareness, and information channels.

Algorithmic access was measured using items from the Oxford Internet Institute (2019) to assess the participants' ability to encounter, retrieve, and evaluate algorithmically curated content across online platforms, particularly within China's mainstream algorithm-driven digital ecosystems.

Algorithmic awareness was evaluated through the algorithmic media content awareness (AMCA) scale developed by Zarouali et al. (2021). This scale has been validated across Facebook, YouTube, and Netflix contexts and includes four subdimensions: awareness of content filtering, automated decision-making, human–algorithm interaction, and ethical considerations. Awareness of content filtering reflects users' understanding that online information exposure is selectively shaped by algorithmic curation rather than chronological order. Awareness of automated decision-making captures the recognition that algorithms, not humans, automatically determine what content appears on platforms. Awareness of human–algorithm interaction measures users' understanding that their behaviors and personal data actively influence algorithmic recommendations. Awareness of ethical considerations concerns users' sensitivity to privacy, bias, and social inequality issues arising from algorithmic personalization. In this study, all items were carefully contextualized for Chinese social-media environments and adapted linguistically for cultural relevance.

Mental health was assessed using the Warwick–Edinburgh mental well-being scale (Tennant et al., 2007), which captures self-perceived psychological well-being, emotional resilience, and life satisfaction.

Digital skills were measured based on van Deursen and van Dijk's (2009, 2010) multidimensional framework, which includes operational, informational, and strategic skills. In this study, a fourth dimension—participatory skills—was added to reflect contemporary user engagement with social media algorithms.

Privacy awareness items were drawn from the Oxford Internet Institute (2019) dataset, focusing on respondents' understanding of online privacy risks, consent mechanisms, and data-sharing practices.

Finally, the *information channels* section captured users' primary sources for technology and media information, including traditional news outlets, social media feeds, and interpersonal communication. The purpose of this section was to describe users' general media-exposure environments and to provide contextual

background for understanding their overall digital engagement.

In addition to these conceptual scales, the questionnaire collected sociodemographic variables (e.g., age, gender, education, income, urban/rural residence), technology-related factors (Yan and Schroeder, 2020), and gender role attitudes (Peking University Open Research Data Platform, 2010; Renmin University of China, 2021). Together, these variables provided a comprehensive profile of respondents' personal characteristics and digital environments, supporting a nuanced understanding of algorithmic literacy in the Chinese social media context.

To enhance transparency of measurement, representative example items from key constructs are provided below. For instance, algorithmic awareness was measured using items such as “The content I see on online platforms is pushed by algorithms” and “Recommendation algorithms automatically determine what I will see on online platforms”. Digital skills were assessed with items such as “I know how to find specific information online efficiently” and “I can strategically use digital tools to achieve my goals”. Privacy awareness included items such as “I am concerned about how my personal data are used by online platforms” and “I understand how to control the use of my personal information online”.

All multi-item constructs demonstrated high internal consistency. Cronbach's α values for the main scales exceeded the commonly accepted threshold ($\alpha > 0.80$), indicating satisfactory reliability of the measurement instruments.

3.3 Experimental design

The experimental session for study 2 (behavioral experiment) was designed to capture participants' real-time behavioral performance in algorithmic AD identification tasks while also collecting corresponding self-reported measures.

To ensure the reliability and validity of the eye-tracking data, the experiment was conducted in a controlled laboratory environment with stable ambient lighting and minimal external distractions. Participants were seated approximately 60 cm from a 24-inch monitor (1920 × 1080 resolution), with the Tobii Pro Fusion eye-tracker (250 Hz) positioned at the base of the screen.

Participants were allowed to wear their regular glasses or contact lenses during the experiment to maintain natural viewing conditions. Prior to the formal task, a standard 9-point calibration procedure was conducted for each participant to ensure accurate gaze tracking. Calibration quality was checked and repeated if necessary, until acceptable accuracy was achieved.

During the experiment, both gaze trajectories (e.g., fixation duration and scan paths) and mouse click events were recorded simultaneously. This synchronized recording enabled the integration of visual attention data with behavioral responses in the AD-identification task, providing a robust basis for analyzing users' algorithmic recognition performance.

Upon arrival, participants first completed a short preliminary questionnaire consisting of demographic items and general attitudinal questions related to algorithmic awareness. This section served to establish individual background profiles before the behavioral task. Participants were then introduced to a simulated mobile social-media interface containing a curated feed of posts and advertisements. They were instructed to browse freely and to identify algorithmically generated advertisements by clicking on them during the session. Both gaze trajectories and click events were continuously recorded, forming the eye-tracked click-through dataset used for subsequent analysis. The experimental procedure is presented in Fig. S1 in the Electronic Supplementary Material (ESM).

After completing the behavioral task, participants completed a posttask survey that paralleled the structure of the main questionnaire described in Section 3.2, covering the key conceptual dimensions of algorithmic literacy and general digital literacy. This posttask survey allowed for a direct comparison between objective behavioral performance and self-reported literacy measures obtained within the same session.

To ensure data reliability, each participant completed a practice trial before the formal task to become familiar with the interface and instructions. The researcher monitored all sessions in real time to ensure compliance with procedures. Upon completion, the participants were debriefed and compensated for their time.

The resulting dataset combined objective behavioral measures (fixation and click logs) with subjective self-reports, enabling a comprehensive assessment of how users' perceived algorithmic awareness aligns—or fails to align—with their actual identification behavior. This integration supports the study's broader goal of bridging self-perceived literacy with measurable user performance.

3.4 Data analysis

All analyses were conducted using RStudio (version 2024.02). Although data collection proceeded sequentially from the survey (study 1) to the behavioral experiment (study 2), the following analyses are presented in reverse order for analytical clarity—beginning with the comparison of self-reported and behavioral awareness (RQ1), followed by behavioral performance analyses (H1a–H1b–H2), and concluding with survey-based predictors of algorithmic awareness (H3). This organization reflects the theoretical logic of the study: first identifying the discrepancy between perception and behavior, then examining behavioral mechanisms, and finally exploring the structural determinants of algorithmic literacy.

Integrated analysis (RQ1). The analysis first compared participants' self-reported algorithmic awareness (from the survey) with their behavioral AD-identification performance (from the eye-tracked click-through experiment) to address RQ1. Descriptive and correlational analyses were conducted to assess the degree of consistency between subjective awareness and actual recognition behavior. This integrated step provided empirical justification for adopting behavioral measures alongside self-reported metrics to capture the multidimensional nature of algorithmic literacy.

Study 2—Behavioral experiment analysis (H1a–H1b–H2). Behavioral data were derived from an eye-tracked click-through task simulating social media browsing. Each participant's mouse clicks and viewing durations were recorded through an interface that synchronized gaze tracking with click logging, enabling an objective assessment of algorithmic AD identification ability. Descriptive visualization illustrated the distribution of AD-click frequencies and sequential effects across AD positions. To test H1a and H1b, linear regression models were used to examine how age, education level, and algorithmic awareness influenced AD-identification latency (time to the first click). The time to first AD click was used as an indicator of participants' initial recognition of algorithmically generated advertisements. This measure captures the latency between first exposure and behavioral response, reflecting the cognitive processing required to identify algorithmic content. Compared with aggregate performance measures (e.g., total number of identified ADs), this metric is intended to capture participants' immediate response to algorithmic content at the early stage of exposure, before substantial learning or habituation effects occur during the task. To test H2, a binary logistic regression was employed, with AD identification (1 = clicked, 0 =

not clicked) as the dependent variable and AD sequence order as the key predictor. Model fit was evaluated using R^2 , AIC, and F -statistics, with statistical significance defined at $p < 0.05$ and $p < 0.001$. These analyses together provided a behavioral perspective on how individual traits and interface structure jointly shape algorithmic recognition performance.

Study 1—Survey analysis (H3). The large-scale survey dataset was used to identify sociodemographic and internet-use factors influencing algorithmic awareness. Descriptive statistics summarized respondents' profiles and overall awareness levels. Reliability was confirmed using Cronbach's α , and both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were performed to assess the internal consistency and dimensional validity of the adapted algorithmic awareness scale. Subsequently, multiple linear regression models tested H3, examining how education level, gender, and internet skills predicted users' self-reported algorithmic awareness. Continuous variables were mean-centered to reduce multicollinearity. Unless otherwise noted, significance was set at $p < 0.05$.

Throughout all analyses, informed consent was obtained from participants, and strict ethical standards regarding privacy and data confidentiality were maintained.

4 Findings

Building on the analytical framework described in Section 3, this section presents the empirical results in the order of behavioral analysis followed by survey analysis. The findings together provide both objective and subjective perspectives on users' algorithmic literacy.

4.1 Discrepancy between self-reported and behavioral awareness (RQ1)

To evaluate whether self-reported algorithmic awareness corresponds to users' actual behavioral identification, participants' self-reported measures—algorithmic awareness, internet skills, and platform-use frequency—were compared with their objective task performance in identifying algorithmic advertisements. These self-reported variables were derived from validated survey instruments administered during the experiment. Algorithmic awareness and internet skills were represented by exploratory factor scores, while platform usage was measured by average reported frequency.

Figure 2 presents three scatterplots illustrating participants' behavioral patterns in the identification task across these self-reported dimensions. In each subplot, the x -axis represents task duration (log-transformed), and the y -axis corresponds to one of the three self-reported measures. Point size reflects the number of ADs identified, while color indicates the average latency from first fixation to mouse click for each participant. "X" markers denote participants who did not identify any ADs.

The scatterplots reveal that 33 participants identified more than half of the ADs. Variation in point size demonstrates clear individual differences in identification counts, while the predominance of dark purple hues (indicating latencies below 4000 ms) suggests that most participants processed and responded to ADs relatively quickly. Participants with longer fixation-to-click times generally also exhibited longer total task durations. A negative association emerged between platform-use frequency and task duration—frequent users completed the task more efficiently. In contrast, distributions for algorithmic awareness and internet skills appeared scattered, showing no consistent patterns with respect to either identification count or task time.

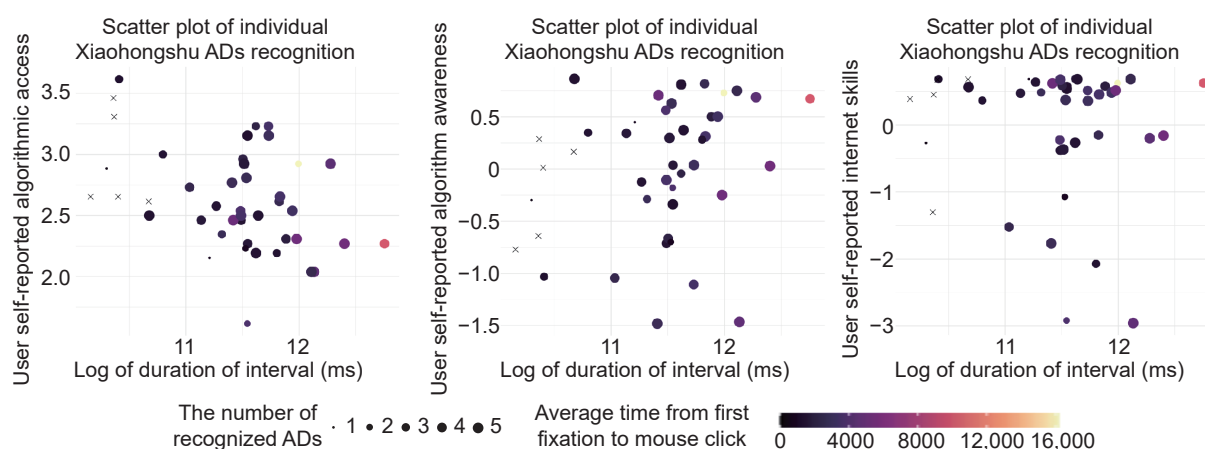


Fig. 2 Scatterplots of participant behavior patterns in the algorithmic AD identification task.

Notably, participants who failed to identify any ADs (“X” marks) clustered in the upper-left quadrants of the platform usage and internet skills subplots, meaning they self-reported moderate to high levels of these attributes. In the algorithmic awareness subplot, these same individuals were concentrated near the middle range rather than the low end. This indicates that participants with weak or absent behavioral performance still reported nontrivial levels of self-assessed competence, implying a possible overestimation bias in self-reported measures. Such overestimation may partially explain the nonsignificant relationship between self-reported traits and actual AD identification performance.

In summary, this analysis reveals a clear discrepancy between perceived and actual algorithmic literacy. Although participants generally reported moderate to high levels of algorithmic awareness, internet skills, and platform familiarity, these subjective assessments did not significantly predict behavioral identification performance. The observed divergence between self-perception and behavior, particularly the clustering of low performers with inflated self-reports, underscores the limitations of subjective measures and justifies using behavioral indicators as a more reliable reference for evaluating algorithmic literacy in subsequent analyses.

4.2 Performance in the algorithmic AD identification task

To examine participants’ behavioral performance in the algorithmic AD identification task, we first visualized the overall click distribution across the five algorithmic ADs presented during the experiment (see Fig. 3). The *x*-axis represents the number of ADs identified (clicked) by each participant (ranging from 0 to 5),

while the *y*-axis shows the number of participants corresponding to each click frequency. The percentages above the bars indicate the proportion of participants who identified a given number of ADs relative to the total sample.

The chart reveals that most participants engaged actively with the task, identifying multiple algorithmic ADs. The largest proportions of participants clicked four ADs (25.6%) and five ADs (39.5%), suggesting a generally high level of behavioral engagement and algorithmic AD identification accuracy. Notably, five participants failed to identify any ADs despite clear task instructions, indicating possible variations in task comprehension or attention.

Following the overall distribution analysis, a grouped bar chart was produced to further examine click behavior across individual ADs (see Fig. 4). The *x*-axis represents the sequential position of each AD (from the first to the fifth), while the *y*-axis shows the corresponding number of participants. For each AD, red bars indicate the number and percentage of participants who identified (clicked) the AD, and green bars indicate those who did not. Percentages labeled on each bar provide a visual reference for comparative click-through rates across different AD positions.

The visualization reveals a clear decline in click-through rates as the AD sequence progresses. The first two ADs achieved the highest identification rates—74.4% and 76.7%, respectively—indicating heightened participant attention and algorithmic AD identification accuracy at the beginning of the task. Identification performance decreased with later ADs, dropping to 67.4% for the third and further to 62.8% for the fifth, illustrating a gradual attenuation of recognition effectiveness with increasing exposure

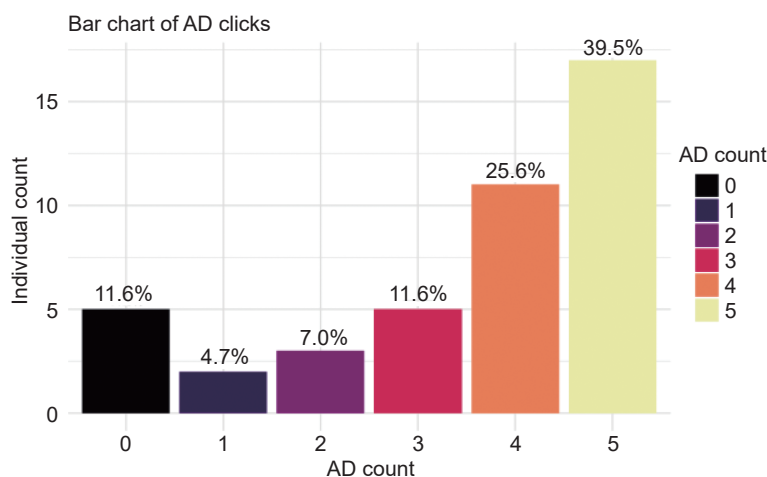


Fig. 3 Distribution of the number of ADs identified per participant.

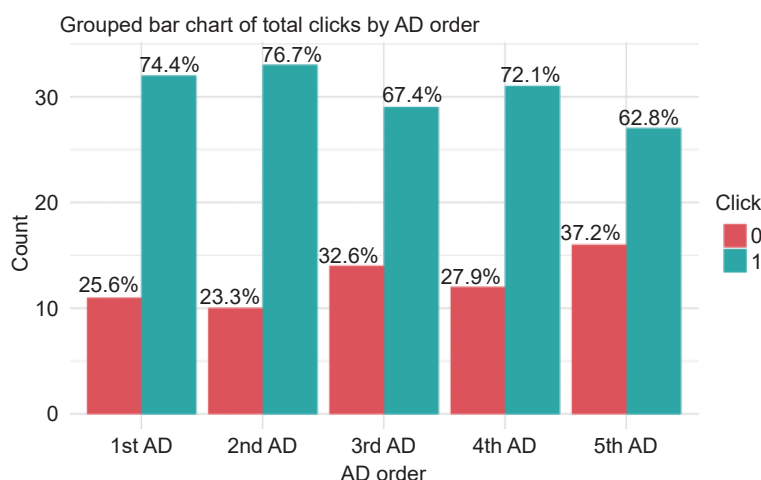


Fig. 4 Click-through rate for each algorithm AD by sequence position.

order.

Finally, a stacked bar chart was generated to visualize the distribution of individual participant clicking patterns (see Fig. 5). The x-axis displays participants (arranged by their total number of ADs identified), and the y-axis represents cumulative identification counts. Each stack segment corresponds to an AD's sequence position, color-coded as follows: black for the first AD, purple for the second, dark purple for the third, orange for the fourth, and yellow for the fifth. This visualization provides a participant-level overview of sequential clicking behavior, further supporting the observed influence of AD order on algorithmic AD identification performance.

4.3 Effects of age and algorithmic awareness on algorithmic AD identification (H1a, H1b)

Building upon the descriptive analysis above, this section examines whether age, education level, and algorithmic awareness are associated with participants' ability to identify algorithmic advertisements, measured by time to first AD click (in ms).

A multiple linear regression model was estimated with time to first AD identification as the dependent variable, while age, education level, and algorithmic awareness were entered as predictors. This metric captures participants' immediate recognition ability during their first exposure to the page—before habituation or learning effects can occur—thus serving as a sensitive indicator of algorithmic AD identification capacity.

Regression results show that age has a significant positive

association with time to first AD identification ($b = 1478.0$, $p < 0.001$), indicating that older participants required longer to identify the first AD. Thus, we accept **H1a**. Similarly, algorithmic awareness shows a significant positive relationship with identification latency ($b = 9412.2$, $p < 0.05$), suggesting that participants reporting higher algorithmic awareness also spent more time processing and identifying the AD content, which also means **H1b** is accepted. In contrast, education level shows a negative but nonsignificant effect ($b = -3711.1$, $p > 0.05$), indicating that formal education alone may not significantly predict behavioral algorithmic literacy. The model's overall fit is acceptable ($R^2 = 0.605$; adjusted $R^2 = 0.561$; $F = 13.79$, $p < 0.001$), explaining approximately 61% of the variance in identification time.

In summary, both age (**H1a**) and algorithmic awareness (**H1b**) are significant correlates of AD identification performance, while education level does not show a robust association. These findings indicate that older users and those reporting higher awareness demonstrate greater cognitive engagement with algorithmic content—reflected in longer processing time—whereas educational background exerts limited influence.

4.4 Influence of AD sequence on algorithmic AD identification

We further examined whether the sequential position of ADs affects participants' likelihood of identifying them (**H2**). Each AD's position in the simulated feed was standardized, with smaller

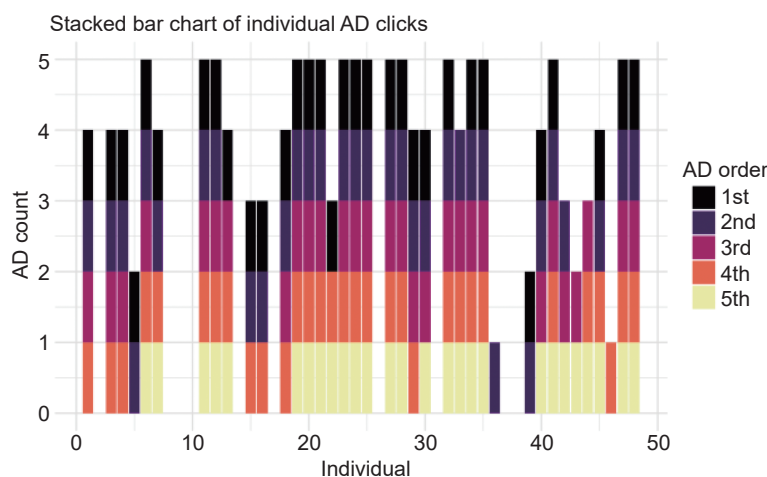


Fig. 5 Stacked distribution of AD clicks by sequence and participant.

values representing higher placement (i.e., ADs appearing earlier on the page). A binary dependent variable (1 = AD clicked, 0 = not clicked) was created to indicate whether each AD was identified. Logistic regression analysis was used to test the relationship between AD order and identification probability.

The results show that the AD sequence position has a significant negative association with the likelihood of AD identification. For the first AD, each unit increase in order decreases the probability of a click ($b = -0.484$, $p < 0.001$), with the model explaining 83.0% of the variance ($R^2 = 0.830$). Similarly, the second AD showed a comparable negative effect ($b = -0.352$, $p < 0.001$; $R^2 = 0.494$). In contrast, the third, fourth, and fifth ADs show no significant relationship with order ($b_3 = -0.057$, $p > 0.05$; $b_4 = -0.089$, $p > 0.05$; $b_5 = -0.026$, $p > 0.05$), indicating that position effects weaken as users progress through the feed.

In summary, these findings support **H2**, showing that earlier AD placement increases the probability of identification. ADs positioned higher on the feed are more likely to be recognized as algorithmic, while later ADs suffer from reduced attention and recognition accuracy. For social science researchers, this result highlights that users' algorithmic sensitivity may diminish as browsing continues, suggesting a potential attention decay, or fatigue, effect that complicates efforts to measure real-world algorithmic awareness.

4.5 Educational level, gender, and internet skills on algorithmic awareness

Having examined users' behavioral identification performance in study 2, this final section returns to study 1, analyzing the subjective dimension of algorithmic literacy. This transition from behavioral outcomes to survey-based predictors reflects the overall research logic: while the experiment captures users' objective

identification ability, the survey contextualizes these differences by exploring who tends to possess higher algorithmic awareness and why.

This section presents the results of the large-scale survey analysis (study 1) examining the demographic and skill-related predictors of algorithmic awareness. To ensure the robustness of the measurement, the adapted algorithmic awareness scale was first tested for reliability and validity within the Chinese social media context. The five measured dimensions—algorithmic awareness, online experience, internet skills, privacy awareness, and information channels—showed high internal consistency (Cronbach's $\alpha > 0.85$). The factor loadings from the exploratory analysis are reported in Table 3.

Exploratory factor analysis (EFA) extracted four major components, collectively explaining 61% of the total variance, while confirmatory factor analysis (CFA) verified the four-factor model with good fit indices ($CFI = 0.948$, $TLI = 0.937$, $RMSEA = 0.072$, $SRMR = 0.041$). The standardized factor loadings from the CFA are presented in Table 4. These results confirm that the adapted scale demonstrates strong construct validity and reliability across multiple platforms.

The structural model further indicated substantial covariance among the four dimensions—content filtering (CF), automated decision-making (ADM), human–algorithm interaction (HAI), and ethical concerns (ETH)—with correlations ranging from 0.60 to 0.91. As illustrated in Fig. 6, these consistent relationships highlight that users' awareness of algorithmic mechanisms is multifaceted, encompassing both functional and ethical understanding.

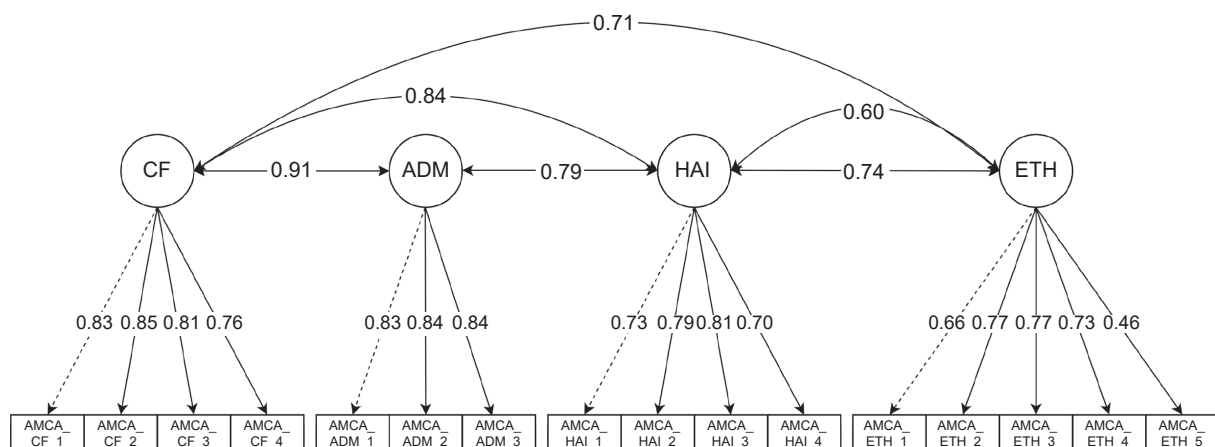
A multiple regression analysis was then conducted to test **H3**, examining how sociodemographic variables and internet skills predict users' algorithmic awareness. The results show that

Table 3 Factor loadings from exploratory factor analysis of algorithmic awareness variables

Item	ML1	ML2	ML3	ML4
I have heard of algorithmic recommendation mechanisms on online platforms: Algorithms are used to push posts, Weibo, videos, or other content to me.	0.27	0.05	0.12	0.51
The content I see on online platforms is pushed by algorithms.	0.28	0.03	0.00	0.69
The content recommended to me by online platforms' algorithms is tailored to my personal interests.	0.58	−0.06	0.15	0.18
Recommendation algorithms on online platforms lead to differences in the content seen by different people.	0.69	0.09	0.01	0.01
What I see on online platforms is automatically decided by recommendation algorithms.	0.85	0.02	−0.13	0.07
Recommendation algorithms, rather than human editors, are more likely to determine what I see on online platforms.	0.90	0.01	−0.05	−0.05
Recommendation algorithms automatically determine what I will see on online platforms.	0.89	−0.02	−0.01	−0.05
My behaviors on online platforms (e.g., sharing, liking, commenting) affect the content I see on those platforms.	0.46	−0.07	0.43	−0.06
The information I provide to online platforms (e.g., gender, location) and my device information affect the content I see on those platforms.	−0.06	0.02	0.87	0.00
The data I create on online platforms, which can be utilized, affect the content I see on those platforms.	0.00	0.01	0.79	0.05
I can adjust algorithm-recommended content by providing feedback to online platforms about what I am interested or not interested in.	0.17	0.08	0.45	0.06
Personalized algorithmic recommendations on online platforms exacerbate social inequality, discrimination, and bias.	0.03	0.51	0.17	−0.01
Under the influence of personalized algorithmic recommendations, what I see may not represent the full picture or diverse viewpoints.	0.08	0.56	0.14	0.05
Recommendation algorithms on online platforms excessively use my personal private data, such as age, gender, and location.	0.02	0.92	−0.21	0.06
I am worried that my personal private data may be leaked due to the use of personalized recommendation algorithms on online platforms.	−0.07	0.82	−0.07	0.06
I am able to proactively turn off the use of personal data that I consider unnecessary, such as location information, to prevent its use by personalized recommendation algorithms for content recommendations.	0.01	0.41	0.14	−0.11

Table 4 Standardized factor loadings from confirmatory factor analysis of algorithmic awareness variables

Item	Content filtering (CF)	Automated decision-making (ADM)	Human–algorithm interplay (HAI)	Ethical concerns (ETH)
I have heard of algorithmic recommendation mechanisms on online platforms: Algorithms are used to push posts, Weibo, videos, or other content to me.	0.973	—	—	—
The content I see on online platforms is pushed by algorithms.	0.981	—	—	—
The content recommended to me by online platforms' algorithms is tailored to my personal interests.	0.873	—	—	—
Recommendation algorithms on online platforms lead to differences in the content seen by different people.	0.803	—	—	—
What I see on online platforms is automatically decided by recommendation algorithms.	—	0.912	—	—
Recommendation algorithms, rather than human editors, are more likely to determine what I see on online platforms.	—	0.938	—	—
Recommendation algorithms automatically determine what I will see on online platforms.	—	0.929	—	—
My behaviors on online platforms (e.g., sharing, liking, commenting) affect the content I see on those platforms.	—	—	0.812	—
The information I provide to online platforms (e.g., gender, location) and my device information affect the content I see on those platforms.	—	—	0.886	—
The data I create on online platforms, which can be utilized, affect the content I see on those platforms.	—	—	0.875	—
I can adjust algorithm-recommended content by providing feedback to online platforms about what I am interested or not interested in.	—	—	0.781	—
Personalized algorithmic recommendations on online platforms exacerbate social inequality, discrimination, and bias.	—	—	—	0.794
Under the influence of personalized algorithmic recommendations, what I see may not represent the full picture or diverse viewpoints.	—	—	—	0.884
Recommendation algorithms on online platforms excessively use my personal private data, such as age, gender, and location.	—	—	—	0.878
I am worried that my personal private data may be leaked due to the use of personalized recommendation algorithms on online platforms.	—	—	—	0.856
I am able to proactively turn off the use of personal data that I consider unnecessary, such as location information, to prevent its use by personalized recommendation algorithms for content recommendations.	—	—	—	0.522

**Fig. 6** Structural model of the algorithmic awareness scale with standardized path coefficients.

educational level ($\beta = 0.116, p < 0.01$), male gender ($\beta = 0.161, p < 0.01$), and internet skills ($\beta = 0.355, p < 0.001$) all have significant positive associations with algorithmic awareness, and hence, we accept H3. Meanwhile, income, age, and urban–rural background show no significant effects ($p > 0.05$). The model explains approximately 20% of the variance ($R^2 = 0.202$), indicating moderate but meaningful explanatory power. The regression results are reported in Table 5.

These findings suggest that higher education and advanced internet literacy enhance individuals' understanding of

algorithmic systems, while gender differences may reflect persistent digital inequality patterns.

In summary, this section confirms that the adapted algorithmic awareness scale is both reliable and valid in the Chinese context. Moreover, the results support H3, showing that educational attainment, male gender, and particularly internet skills are significant positive predictors of algorithmic awareness. This highlights that improving internet skill levels could be a practical avenue for enhancing public algorithmic literacy.

Table 5 Regression analysis of demographic factors and internet skills on algorithmic awareness

Dependent variable: Algorithmic awareness	Coefficient standard error
Income	−0.0004 (0.009)
Education level	0.116*** (0.035)
Age	−0.005 (0.004)
Gender (male)	0.161** (0.066)
Gender (nonbinary)	0.323 (0.280)
Urban–rural (urban)	−0.022 (0.076)
Internet skills	0.355*** (0.034)
Constant	−0.523** (0.219)
Observations	645
R^2	0.202
Adjusted R^2	0.193
Residual standard error	0.781 ($df = 637$)
F Statistic	22.975*** ($df = 7; 637$)

Note: Values represent regression coefficients, and values in parentheses indicate standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

5 Discussion

By integrating behavioral and self-reported data, this study seeks to bridge the gap between what users claim to know and what they actually do in algorithmically curated environments.

Through the combined use of survey-based measures (study 1) and eye-tracking behavioral data (study 2), we moved beyond traditional self-reporting approaches to capture a multidimensional portrait of algorithmic literacy. This mixed design reflects the research logic of the study—beginning from subjective awareness, validating it through behavioral evidence, and returning to structural predictors that shape this awareness. Together, these findings illuminate the complex interplay between user cognition, behavioral response, and algorithmic design, offering theoretical and practical implications for fostering a more critically informed digital society.

5.1 Theoretical contributions: Towards a multidimensional understanding of algorithmic literacy

This study advances the theoretical framing of algorithmic literacy as a multifaceted construct encompassing both awareness and behavioral competence.

First, the observed discrepancy between self-reported awareness and actual identification performance challenges the adequacy of perception-based measures. It underscores the necessity of mixed-methods approaches that combine attitudinal and behavioral data to capture authentic user understanding within algorithmic systems. This divergence between belief and action reshapes how digital literacy should be defined, measured, and taught.

Second, the findings demonstrate that user engagement is jointly influenced by intrinsic factors (e.g., age, internet skills) and extrinsic factors (e.g., AD sequence). These interactions reveal that literacy is not an isolated individual trait but a relational construct shaped by both personal capability and system design.

Such an expanded view of algorithmic literacy recognizes users as embedded actors in a dynamic sociotechnical environment, aligning with emerging perspectives on critical algorithm studies that emphasize transparency, reflexivity, and contextual literacy.

5.2 Practical implications: Informing authentic and user-empowering design

From a practical standpoint, this research highlights two

interrelated paths toward a more digitally enlightened society.

For platform designers and policymakers, the significant influence of the AD sequence on identification behavior reveals the behavioral consequences of algorithmic ordering. Transparent design cues—such as clearer labeling of sponsored or personalized content—could mitigate attentional decay and empower users to critically engage with algorithmic recommendations.

For digital literacy educators, the strong predictive power of internet skills points to the importance of expanding literacy training beyond basic navigation. Programs should incorporate algorithm awareness and data-driven decision-making literacy, ensuring that users not only operate within but also critically interpret algorithmic systems.

Moreover, the mixed-methods approach adopted here offers a validated model for bridging the gap between attitude and behavior, setting a precedent for future studies seeking to assess algorithmic understanding in realistic contexts. Recognizing that both user competence and system transparency are mutable provides a concrete basis for collaborative interventions that empower rather than manipulate digital citizens.

5.3 Limitations and future research

Despite its contributions, several limitations must be acknowledged.

The lab-based eye-tracking environment, although controlled and data-rich, cannot fully replicate the dynamic, distraction-laden context of real social media browsing. Future research could leverage mobile eye-tracking or in situ data collection to enhance ecological validity.

Additionally, the sample—skewed toward younger and highly educated users—may overrepresent populations with higher baseline algorithmic literacy. Subsequent work should pursue greater demographic and platform diversity, incorporating users from different socioeconomic backgrounds and platform cultures to reveal more nuanced inequalities in algorithmic understanding.

Finally, the use of time to first AD click as an indicator of identification ability may be influenced by individual differences in reaction speed, particularly across age groups. Although this measure captures early-stage recognition processes, it may partially reflect general cognitive or motor responses rather than algorithmic literacy *per se*. Future research could incorporate complementary behavioral indicators to further disentangle these effects.

6 Conclusions

In conclusion, this study demonstrates that the path toward a digitally enlightened society requires addressing both the literacy of users and the transparency of algorithmic systems.

By operationalizing algorithmic literacy as both self-reported awareness and observable behavior, we reveal a persistent gap between perception and action—a key challenge for digital governance and education.

Our findings confirm that strengthening internet skills significantly enhances algorithmic awareness, while the design of algorithmic environments (such as content sequencing) also shapes user perception and engagement. These insights jointly emphasize that literacy and design are two sides of the same coin: one resides in the user, and the other resides in the system.

Ultimately, a truly enlightened digital society is one where individuals are equipped with critical algorithmic literacy, and systems are engineered for transparency, accountability, and informed interaction—not merely for engagement.

Acknowledgements

The paper is funded by the National Natural Science Foundation of China (Grant No. 72304017) and Beijing Academy of Social Sciences (BASS) (Grant No. 22XCC013).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Electronic supplementary material

Supplementary material is available in the online version of this article at <https://doi.org/10.26599/ILIB.2026.9690002>.

References

- Arguedas, A. R.; Robertson, C. T.; Fletcher, R.; Nielsen, R. K. 2022. Echo chambers, filter bubbles, and polarisation: A literature review. Available at <https://doi.org/10.60625/risj-etxj-7k60>
- Bruns, A. 2019. *Are Filter Bubbles Real?* Cambridge, UK: Polity Press.
- Bucher, T. 2017. The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society* Vol. 20, No. 1, 30–44.
- Chen, L.; Pan, X. 2023. Book recommendation based on knowledge graph and reader profiling. *Data Analysis and Knowledge Discovery* Vol. 7, No. 12, 164–171.
- Chen, M. 2020. Research on library intelligent service toward innovation of universities scientific research. Harbin Institute of Technology. Ph.D. Thesis. (in Chinese)
- Cotter, K.; Reisdorf, B. C. 2020. Algorithmic knowledge gaps: A new dimension of (digital) inequality. *International Journal of Communication* Vol. 14, 745–765.
- Cyberspace Administration of China. 2021. Action plan for improving national digital literacy and skills. Available at https://www.cac.gov.cn/2021-11/05/c_1637708867754305.htm
- Ekstrand, M. D.; Tian, M.; Kazi, M. R. I.; Mehrpouyan, H.; Kluver, D. 2018. Exploring author gender in book rating and recommendation. In: Proceedings of the 12th ACM Conference on Recommender Systems, 242–250.
- Eslami, M.; Rickman, A.; Vaccaro, K.; Aleyasen, A.; Vuong, A.; Karahalios, K.; Hamilton, K.; Sandvig, C. 2015. “I always assumed that I wasn’t really that close to [her]”: Reasoning about invisible algorithms in news feeds. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 153–162.
- Ganhör, C.; Penz, D.; Rekabsaz, N.; Lesota, O.; Schedl, M. 2022. Unlearning protected user attributes in recommendations with adversarial training. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2142–2147.
- Gillespie, T. 2014. The relevance of algorithms. In: *Media Technologies: Essays on Communication, Materiality, and Society*. Gillespie, T.; Boczkowski, P. J.; Foot, K. A. Eds. Cambridge, MA, USA: The MIT Press, 167–194.
- Guo, L.; Yang, J.; Liu, Y.; Fu, Y. 2025. Exploring personalized, classification-based book recommendations by integrating ChatGPT with BERT. *Library Tribune* Vol. 45, No. 10, 109–117. (in Chinese)
- Hargittai, E.; Gruber, J.; Djukaric, T.; Fuchs, J.; Brombach, L. 2020. Black box measures? How to study people’s algorithm skills. *Information, Communication & Society* Vol. 23, No. 5, 764–775.
- Just, N.; Latzer, M. 2017. Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture & Society* Vol. 39, No. 2, 238–258.
- Kaplan, A. M.; Haenlein, M. 2010. Users of the world, unite! The challenges and opportunities of social media. *Business Horizons* Vol. 53, No. 1, 59–68.
- Klawitter, E.; Hargittai, E. 2018. “It’s like learning a whole other language”: The role of algorithmic skills in the curation of creative goods. *International Journal of Communication* Vol. 12, 3490–3510.
- Koren, Y.; Bell, R.; Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer* Vol. 42, No. 8, 30–37.
- Melchiorre, A. B.; Rekabsaz, N.; Parada-Cabaleiro, E.; Brandl, S.; Lesota, O.; Schedl, M. 2021. Investigating gender fairness of recommendation algorithms in the music domain. *Information Processing & Management* Vol. 58, No. 5, Article No. 102666.
- Meltwater. 2025. Digital 2025 global overview report. Available at <https://www.meltwater.com/en/global-digital-trends>
- Ministry of Education (MoE) and National Development and Reform Commission (NDRC) and Ministry of Industry and Information Technology (MIIT) et al. 2022. Opinions on accelerating the promotion of education digitalization. Available at http://www.moe.gov.cn/srcsite/A16/s3342/202203/t20220328_612225.html
- Niu, Y.; Hao, B. Q.; Zhao, Z. Y. 2025. Research on the construction and practice of university library personalized resource recommendation system based on Large Models. *New Century Library* No. 7, 66–73. (in Chinese)
- Oxford Internet Institute. 2019. 2019 OxIS questionnaire. Available at <https://oxis.oii.ox.ac.uk/wp-content/uploads/sites/16/2019/11/OxIS-Questionnaire-2019.pdf>
- Pariser, E. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. New York, USA: Penguin Press.
- Pasquale, F. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, USA: Harvard University Press.
- Peking University Open Research Data Platform. 2010. China family panel studies (CFPS). Available at <https://opendata.pku.edu.cn/dataverse/CFPS>
- Proferes, N. 2017. Information flow solipsism in an exploratory study of beliefs about Twitter. *Social Media + Society* Vol. 3, No. 1, Article No. 2056305117698493.
- Rader, E.; Gray, R. 2015. Understanding user beliefs about algorithmic curation in the Facebook news feed. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 173–182.
- Rekabsaz, N.; Kopeinik, S.; Schedl, M. 2021. Societal biases in retrieved contents: Measurement framework and adversarial mitigation of BERT rankers. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 306–316.
- Renmin University of China. 2021. Chinese general social survey (CGSS). Available at <http://www.cnsda.org/index.php?r=projects/view&id=65635422>
- Swart, J. 2021. Experiencing algorithms: How young people understand, feel about, and engage with algorithmic news selection on social media. *Social Media + Society* Vol. 7, No. 2, Article No. 20563051211008828.
- Tennant, R.; Hiller, L.; Fishwick, R.; Platt, S.; Joseph, S.; Weich, S.; Parkinson, J.; Secker, J.; Stewart-Brown, S. 2007. The Warwick–Edinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes* Vol. 5, No. 1, Article No. 63.
- The State Council of the People’s Republic of China. 2018. Electronic commerce law of the People’s Republic of China. Available at https://www.gov.cn/xinwen/2018-08/31/content_5318220.htm
- The State Council of the People’s Republic of China. 2021a. Anti-monopoly guidelines for the platform economy sector. Available at https://www.gov.cn/xinwen/2021-02/07/content_5585758.htm
- The State Council of the People’s Republic of China. 2021b. Personal information protection law of the People’s Republic of China. Available at https://www.gov.cn/xinwen/2021-08/20/content_5632486.htm
- The State Council of the People’s Republic of China. 2022. Provisions on the administration of algorithmic recommendation in Internet information services. Available at https://www.gov.cn/zhengce/zhengceku/2022-01/04/content_5666429.htm
- Van Deursen, A. J. A. M.; van Dijk, J. A. G. M. 2009. Using the Internet:

Skill related problems in users' online behavior. *Interacting with Computers* Vol. 21, Nos. 5–6, 393–402.

Van Deursen, A. J. A. M.; van Dijk, J. A. G. M. 2010. Measuring Internet skills. *International Journal of Human-Computer Interaction* Vol. 26, No. 10, 891–916.

Wang, Y.; Shen, J.; Jiang, Z. 2020. Mobile reading literacy under algorithmic recommendation in the mobile Internet era. *New Century Library*, No. 6, 24–29.

Wu, D. 2022. Algorithmic literacy: New trends in library intelligent services. *Library Construction* Vol. 316, No. 4, 4–5.

Yan, P.; Schroeder, R. 2020. Variations in the adoption and use of mobile social apps in everyday lives in urban and rural China. *Mobile Media & Communication* Vol. 8, No. 3, 318–341.

Yu, C. S.; Dong, J. 2023. On algorithmic bias and its governance in digital reading. *Publishing Research*, No. 6, 65–71. (in Chinese)

Zarouali, B.; Boerman, S. C.; de Vreese, C. H. 2021. Is this recommended by an algorithm? The development and validation of the algorithmic media content awareness scale (AMCA-scale). *Telematics and Informatics* Vol. 62, Article No. 101607.

Zhang, J.; Li, Y.; Li, D. 2021. Research on the characteristics of personalized recommendation service on the network academic resource platform. *Information and Documentation Services* Vol. 42, No. 5, 76–83.

Zhang, S.; Yuan, W.; Yin, H. 2024. Comprehensive privacy analysis on federated recommender system against attribute inference attacks. *IEEE Transactions on Knowledge and Data Engineering* Vol. 36, No. 3, 987–999.



Yunci Tian is a Ph.D. student in information management at Peking University. Her research interests lie at the intersection of people, technology, and society, with a particular focus on digital platforms, information behavior, and emerging technologies. She draws on behavioral experiments, computational social science, and quantitative methods in her work.



Pu Yan is an assistant professor at the Department of Information Management, Peking University. Her research bridges grounded fieldwork with big data analytics to explore the algorithmic society, specifically investigating digital divides, AI literacy, and cross-cultural communication mechanisms in global issues such as climate change. She leads international projects such as PKU-UCL seeds partnership fund, as well as projects funded by the National Natural Science Foundation of China (NSFC) and the China Association for Science and Technology (CAST), with publications in leading journals such as *Information, Communication & Society* and *Telematics and Informatics* and prestigious conferences such as ICWSM. She holds her Ph.D. degree on information, communication, and social science from the University of Oxford (Oxford Internet Institute), where she also completed her M.S. degree with distinction. Prior to Oxford, she obtained her B.S. degree from Tsinghua University.