This is a preprint of an article accepted for publication in *the Journal of the Association for Information Science and Technology (JASIST)*: Choi, W., Bak, H., An, J., Zhang, Y., & Stvilia, B. College students' credibility assessments of GenAI-generated information for academic tasks: An interview study. *JASIST*.

College Students' Credibility Assessments of GenAI-Generated Information for Academic Tasks: An Interview Study

Wonchan Choi[1], Hyerin Bak[1], Jiaxin An[2], Yan Zhang[2], and Besiki Stvilia[3]

[1] School of Information Studies, University of Wisconsin-Milwaukee

[2] School of Information, University of Texas at Austin

[3] School of Information, Florida State University

Author Note

Correspondence concerning this article should be addressed to Wonchan Choi, School of Information Studies, University of Wisconsin-Milwaukee, Northwest Quadrant Building D, Room 2890, 2025 East Newport Avenue, Milwaukee, WI 53211, USA. Phone: (414) 251-5753. Email: wchoi@uwm.edu

Hyerin Bak, School of Information Studies, University of Wisconsin-Milwaukee, Northwest Quadrant Building D, Room 2855, 2025 East Newport Avenue, Milwaukee, WI 53211, USA. Email: hbak@uwm.edu

Jiaxin An, School of Information, University of Texas at Austin, 1616 Guadalupe St, Suite #5.202, Austin, TX 78701, USA. Email: jiaxin.an@utexas.edu

Yan Zhang, School of Information, University of Texas at Austin, 1616 Guadalupe St, Suite #5.202, Austin, TX 78701, USA. Email: yanz@utexas.edu

Besiki Stvilia, 269 Louis Shores Building, 142 Collegiate Loop, Tallahassee, FL 32306, USA. Email: bstvilia@fsu.edu

Abstract

The study explored college students' use of generative artificial intelligence (GenAI) tools, such as ChatGPT, for academic tasks and their perceptions and behaviors in assessing the credibility of GenAI-generated information. Semistructured interviews were conducted with 25 college students in the United States. Interview transcripts were analyzed using the qualitative content analysis method. The study identified various types of academic tasks for which students used ChatGPT, including writing, programming, and learning. Guided by two models of credibility assessment (Hilligoss & Rieh, 2008; Metzger, 2007), six factors influencing students' motivation and ability to assess the credibility of GenAI-generated information were identified (e.g., task salience, social pressure). We also identified nine constructs (e.g., refinedness, explainability), five heuristics (e.g., inter- and intrasystem consistency heuristics), and 10 cues (e.g., version and tone) used by students to assess the credibility of GenAI-generated information. This study provides theoretical and empirical findings regarding students' use of GenAI tools in the academic context and credibility evaluation of the system outputs using rich, qualitative interview data.

Keywords: information credibility, generative artificial intelligence, generative artificial intelligence credibility, college students, LLM

College Students' Credibility Assessments of GenAI-Generated Information for Academic Tasks: An Interview Study

Generative artificial intelligence (GenAI) is a form of AI that operates probabilistically, producing content that resembles but does not duplicate its training data (Peres et al., 2023). In higher education, rapid adoption of GenAI tools such as ChatGPT that leverage large language models (LLMs) has occurred. Surveys conducted in March 2023, shortly after ChatGPT's public launch, showed that 27% of college students regularly used GenAI tools for academic tasks, such as writing assistance, homework help, concept comprehension, and data analysis (Shaw et al., 2023). By September 2023, the adoption rate surged to 49%, with 75% of participants expressing intent to continue using GenAI for their academic work despite potential bans by their instructors or institutions (Shaw et al., 2023).

Although GenAI tools have the potential to facilitate students' performance in academic tasks, such as improving the speed and effectiveness of writing and coding processes, they are not without limitations, which could lead to risks. Specifically, GenAI tools can make factual errors—"inaccuracies in information or statements that are not in accordance with reality or the truth" (Borji, 2023, p. 5)—because they generate information based on statistical patterns in the training data, rather than retrieving data from an external memory or database. Thus, they are susceptible to fabricating facts. One common case of factual errors reported in academic writing is inaccurate or fabricated references (Athaluri et al., 2023; Emsley, 2023). Another substantial concern regarding GenAI tools pertains to bias—"systematic inaccuracies or stereotypes in the generated language output" (Borji, 2023, p. 10)—which are influenced by societal and cultural prejudices present in the training data. Biased GenAI outputs can spread false and harmful information and cause misunderstandings. These issues seriously challenge students' academic integrity (Anderson & Rainie, 2023; C. Wu et al., 2024).

Technical efforts have been made to mitigate these issues. For example, OpenAI has made significant efforts to identify and remove misinformation and biases from its systems' responses, such as soliciting user feedback on the usefulness of the systems' responses and using reinforcement learning through feedback from human indexers and coders (Ouyang et al., 2022). More effective mitigation of this information challenge requires in-depth understanding of people's information behavior in the new environment. However, understanding of users' information behaviors when using GenAI tools, including when and how they evaluate system outputs, is limited. In the present study, we collected rich qualitative data from semistructured interviews with 25 college students to start addressing this gap in the literature.

Theoretical Background

Conceptualization of Credibility in the GenAI Context

A two-dimensional model that identifies trustworthiness and expertise as the key underlying dimensions of credibility (Hovland et al., 1953) has been widely accepted in many fields, including library and information science (Rieh, 2017) and human–computer interaction (Danielson, 2006). Expertise captures the perceived knowledge, skill, and experience of the source relative to providing high-quality information, whereas trustworthiness captures the perceived goodness and morality of the source willing to provide high-quality information (Danielson, 2006; Rieh, 2017). Adopting this conceptualization of credibility, the credibility of GenAI-generated information (GenAI credibility, for short) can be defined as users' perceptions of GenAI-generated information objects—such as an individual answer or thread of conversations—as being accurate and insightful (i.e., expert) and free from bias and deceptive intention (i.e., trustworthy).

Dual Processing of Credibility Assessments

According to the dual processing model of web credibility assessment developed by Metzger (2007), a user may choose an "analytic," in-depth, and more involved evaluation strategy (i.e., central route) or a more casual, "experiential" strategy (i.e., peripheral route) based on how motivated and capable the user is in terms of evaluating information. The dual processing model also stipulates that to use a credibility heuristic with a particular cue, the user needs to perceive it as relevant to the cue, be able to recall the heuristic, and apply it with little cognitive effort (Bellur & Sundar, 2014; Dehnert & Mongeau, 2022).

Three Layers of Credibility Judgments

The unifying framework of credibility assessment created by Hilligoss and Rieh (2008) delineates three distinct layers of credibility evaluation: construct, heuristics, and interaction. The first layer, construct, deals with how individuals define credibility. In this layer, evaluators conceptualize the abstract notion of credibility using related terms such as believability, reliability, and objectivity. In the second layer, heuristics, evaluators identify intuitive rules to assess the credibility construct they settled on. Heuristics are applicable across diverse contexts rather than limited to specific instances. The third layer, interaction, is informed by specific attributes of a particular information object, which serve as cues in people's credibility judgments. Evaluators can find these cues through interactions with the information object. Hilligoss and Rieh (2008) specified three types of interactions: interactions with content cues (e.g., errors in content), interactions with source peripheral cues (e.g., author identity), and interactions with information object peripheral cues (e.g., presentation of information).

Conceptual Model of the Study and Research Questions

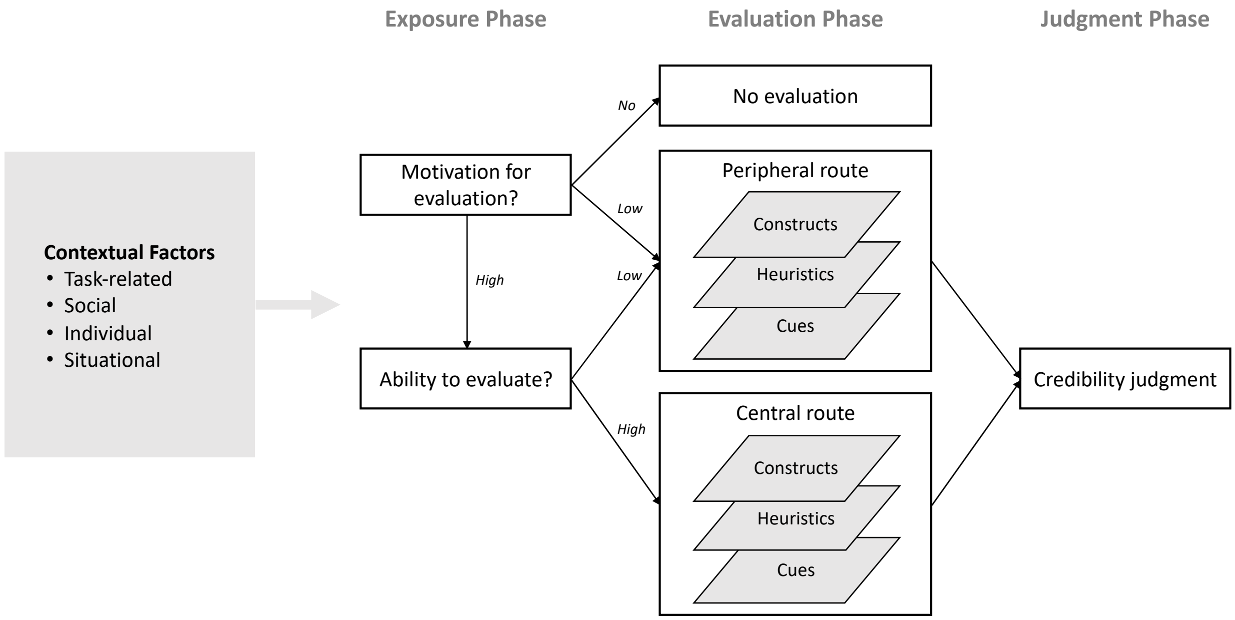
Building on the dual processing model (Metzger, 2007) and the unifying framework of credibility assessment (Hilligoss & Rieh, 2008), our synthesized model (Figure 1) delineates students' GenAI credibility assessment process into three phases: exposure, evaluation, and judgment. In the exposure phase, students decide whether and to what extent they will engage in the process of assessing the credibility of an encountered information object (e.g., chatbot response), based on their motivation and ability. To accommodate the context-dependent nature of credibility assessments, our synthesized model also incorporates various factors that influence students' motivation and ability to evaluate GenAI credibility, including task-related, social, individual, and situational factors.

In the evaluation phase, students' credibility evaluation process may follow a peripheral or central route depending on their motivation and ability to judge credibility. Both routes involve three layers of evaluation: the construct layer, where students conceptualize credibility by identifying relevant criteria to assess credibility in context; the heuristics layer, where students employ general rules applicable to the identified criteria; and the cues layer, where students gather evidence to apply the rules (heuristics) by examining aspects of a given information object. Given the relatively higher level of cognitive load expected for interacting with content cues, it is likely that only those on the central route will interact with content cues (e.g., facts), beyond superficial cues related to source (e.g., references), information object (e.g., length), and interface design (e.g., aesthetics), in their credibility evaluations.

In the judgment phase, students integrate the criteria, cues, and heuristics applied during the evaluation to form an overall judgement about the expertise and trustworthiness of the information. Thus, the judgement phase culminates in a coherent assessment of the information credibility.

Figure 1

Conceptual Model of the Study



Based on this framework, we intended to answer the following questions:

- RQ1: When do students evaluate the credibility of the information from ChatGPT or similar tools and what factors influence their motivation and ability to do so?
- RQ2: How do students evaluate the credibility of the information from ChatGPT or other similar tools?
    - RQ2-1: What criteria do they apply to evaluate ChatGPT responses?
    - RQ2-2: What heuristics do they apply to evaluate the responses?
    - RQ2-3: What cues do they use to support the evaluation?
- RQ3: What challenges do students face when evaluating GenAI credibility in academic contexts?

Literature Review

College Students' Evaluation of Information Credibility

Students differ in their awareness or recognition of information quality issues on the web. Wiley et al. (2009) found that some students lacked sensitivity to information quality on websites when completing scientific inquiry tasks. Although some students recognize that information on the internet is not particularly credible relative to traditional information sources (Hilligoss & Rieh, 2008), this awareness does not always translate into evaluation behavior. Metzger et al. (2003) reported that students in their study tended to verify online information diligently only rarely or occasionally. Their evaluation strategies tended to follow the least effort principle (Metzger, 2007) and often were compromised by factors such as speed and convenience (Rieh & Hilligoss, 2008).

Students' efforts to assess information credibility are situational and may be shaped by many factors. One factor is their motivations or search goals. Students exert greater efforts in assessing and verifying information credibility when searching relative to decisions with more importance or critical consequences, such as voting, academic achievement, finance, and health (Head & Eisenberg, 2011; Rieh & Hilligoss, 2008), and less for entertainment purposes (Flanagin & Metzger, 2000; Rieh & Hilligoss, 2008). Another contextual factor is the intended use of the information. Students were more concerned about the credibility of information when it involved someone they cared about (Head & Eisenberg, 2010).

Studies have shown that college students heavily rely on peripheral cues to assess information credibility (Metzger et al., 2010). These cues can be categorized into three domains: design, source, and content. Common design cues include site appearance, such as layout, graphics, colors, logo, and ease of navigation (Breakstone et al., 2021; McGrew et al., 2018). Common source cues include site sponsor, site type, and disclosure (e.g., About Us; Chang et al., 2021). Common content cues include spelling, tone of speech, and references, including the number and source of references (Breakstone et al., 2021; K.-S. Kim et al., 2011; Lim, 2013; Liu et al., 2021; McGrew et al., 2018; Metzger et al., 2003). These cues reflect values that students expect to see in online information, such as accuracy, trustworthiness, completeness, and expertise (Chang et al., 2021; Lee & Sundar, 2013).

Heuristics are at work when these cues are applied to assess online information credibility. There is no universally agreed-upon taxonomy of heuristics. Scholars have identified six main heuristics that students applied in assessing online information credibility: reputation, endorsement, consistency, confirmation, expectancy violation, and persuasive intent heuristics (Metzger et al., 2010). These cues may overlap, and people may apply multiple heuristics at the same time (Metzger & Flanagin, 2013). Sundar (2008) proposed a different set of more than 30 heuristics such as bandwagon, novelty, coolness, responsiveness, contingency, and similarity.

Credibility evaluation often involves social engagements. Students often turn to their social networks, such as friends, family members, classmates, and instructors, for help evaluating sources (Head & Eisenberg, 2010; Lim, 2013). The social nature of credibility evaluation is further fostered by the emergence of social media. Metzger and colleagues (2010) identified four social and group-based evaluation strategies: social information pooling (using user-generated content, such as reviews), social confirmation of personal opinion (seeking similar beliefs or perspectives), enthusiastic endorsement (relying on presumed but uncredentialed users who offer guidance via social platforms, such as forums and blogs), and resource sharing via interpersonal exchange (consulting interpersonal sources, such as friends and family, or recommender systems).

Students also rely on algorithmic endorsement to assess information credibility. A typical example is a source's rank on a search engine results page. Sources ranked higher were often assumed to be more reliable (Hargittai et al., 2010; Pan et al., 2007). Some students also perceived credibility from system-generated recommendations, assuming they were built on big data or authoritative expertise.

College Students' Evaluation of GenAI Content

Recent studies have indicated that college students perceive potential credibility issues with GenAI-generated information but lack the necessary skills to verify its accuracy. Amoozadeh et al. (2024) reported that nearly half of students surveyed in the United States and India did not trust the output of GenAI tools. Due to the risks of biases and errors from AI algorithms, these students suggested the need for human supervision to ensure the accuracy and reliability of GenAI responses. Similarly, a survey with postsecondary students in Hong Kong found that they struggled to verify the accuracy or validity of GenAI-generated information (Chan & Hu, 2023). The opaque nature of GenAI algorithms, absence of traditional cues signifying information credibility and quality, and occurrence of hallucinations in GenAI-enabled systems can complicate the evaluation process (Schuetzler et al., 2024; Shin et al., 2024).

Existing research has identified AI-specific criteria for credibility assessments, such as fairness, accountability, transparency, and explainability. Fairness ensures that the AI system does not amplify biases and discrimination in information (Shin, 2023; Starke et al., 2022). Accountability refers to whether information is shared in a way that holds a person or entity responsible for its impact (Diakopoulos & Koliska, 2017). For transparency, users must be able to see and understand the inputs and outputs of algorithms (Shin, 2023). Explainability requires that GenAI content is clearly explained for users (Rai, 2020; Shin, 2023). Shin et al. (2024) studied college students' credibility assessments of health-related information from GenAI tools using these criteria. They found that cues related to the source and content, such as information sources, fact-checked information, and brief explanations of how and why certain health information was generated, facilitated students' systematic evaluation of health information credibility. Additionally, students' level of confidence in checking against these criteria was positively associated with their likelihood of systematically evaluating the credibility of health information (Shin et al., 2024).

Methods

We adopted the semistructured interview method because it allowed us to elicit information about not only the tasks that students perform using GenAI tools but also the context and students' perceptions regarding the use of ChatGPT or similar tools. The study was approved by the institutional review boards at the three universities involved in the study (Florida State University, the University of Texas at Austin, and University of Wisconsin-Milwaukee).

Interview Guide

Our interview guide was developed based on findings of our prior research on people's use of and experience with ChatGPT (Choi, Zhang, et al., 2023) and consisted of four sections: (a) students' general understanding of and use of ChatGPT, including how they think ChatGPT works; (b) students' use of ChatGPT or similar GenAI tools for academic tasks; (c) students' evaluations of the information returned by ChatGPT; and (d) students' reflections on their use of ChatGPT and its impact on their studies. Questions included: Have you used ChatGPT recently for study purposes, which could be coursework or research activities? Can you recall and describe a specific memorable instance when you used ChatGPT and evaluated the content generated by it? What was the task or information you used ChatGPT for? Could you please describe the process of completing the task? Did you evaluate the credibility or quality of information that ChatGPT returned to you? Why did you evaluate the quality? How did you evaluate it? Did you feel that ChatGPT returned a good answer? What made you feel that was a good answer? What were some of the things you looked for to determine whether this might be good content? What made you feel that this was NOT a good answer? What were some of the things you looked for to determine whether this might NOT be good content? Did you modify or change the content returned by ChatGPT before you used it? Why did you change it? How did you change it? Why did you not change it? Have you encountered or observed any problems regarding your use of ChatGPT?

Procedure

Participants completed an online questionnaire reporting demographic information and GenAI tools that they used before the interview. They also reviewed the consent form and indicated their consent to participate online. The interviews were one-on-one and took place via Zoom or Microsoft Teams. The researchers began by greeting the participant and briefly introducing the study. Then they carried out the interview following the interview guide. At the end of the interview, participants had time to provide additional remarks and ask questions about the study. Each participant received a $20

Amazon gift card after the interview. The interview sessions were recorded and stored in the cloud; transcriptions were generated automatically by Zoom or Teams. Four researchers conducted the interviews independently, each interviewing five to eight participants. The interviews lasted from 17 minutes to 1 hour and 10 minutes. Data collection took place from April to June 2023.

Participants

Eligible participants were college students who were at least 18 years old and had used ChatGPT or other GenAI tools for academic purposes. Various recruitment methods were employed to recruit participants, including university and college listservs, flyers, and the snowball technique. Interested students provided their contact information to the researcher on their campus using the link provided in the recruitment message and flyers. We then contacted these students to schedule interviews. Twenty-five students participated in the study.

Data Analysis

The interview transcripts were analyzed following the qualitative content analysis method (Miles et al., 2014) and facilitated by NVivo 14 (NVivo, 2023). An initial codebook was developed based on our prior research on general people's application of and experience with ChatGPT (Choi, Zhang, et al., 2023) and the interview guide. It contained mostly high-level codes, such as knowledge of ChatGPT, use, user experience, barriers and challenges, and credibility evaluation.

Each researcher coded a subset of the transcripts. Before coding, we read the transcripts to become familiar with the content. When sentences or phrases were unclear or confusing (mostly due to errors in auto transcription), the audio recordings of the interviews were consulted to validate and correct the transcripts. Both new high-level codes that represented emerging categories and detailed codes that substantiated the high-level codes were added to the codebook. We held collective coding sessions to discuss new codes, compare them to existing codes, address disagreements, and update the codebook. Then we moved to the second stage of coding, where each member of the research team validated a subset of the codes. Multiple intense group meetings were held to discuss and address coding discrepancies. The codebook was further consolidated and solidified.

Findings

Participant Characteristics

Table 1 presents the characteristics of the participants. Most of them (76%) were aged 18 to 30. The gender composition was relatively balanced. More than half of them (64%) were White and the remaining included Asian, Hispanic or Latino, Black, and other (not reported). The sample included students pursuing three levels of academic degrees: undergraduate, master's, and doctoral. The majority (68%) majored in information studies, but the sample also included students majoring in humanities, business, and architecture. About 40% of the participants used ChatGPT monthly or less, and the remaining 60% used it weekly or daily.

Table 1

Participant Characteristics (N = 25)

Variable	***n***	**%**
Age		
18–22	12	48
23–30	7	28
31+	6	24
Sex		
Female	12	48
Male	12	48
Nonbinary	1	4
Race		
White	16	64
Asian	3	12
Hispanic or Latino	2	8
Black	2	8
Other	2	8
Academic status		
Undergraduate	13	52
Master’s	8	32
Doctoral	4	16
Major		
Information science	17	68
Humanities	4	16
Business	3	12
Architecture	1	4
Frequency of using ChatGPT or other GenAI tools for educational tasks		

Sometimes but not regularly	9	36
Once a month	1	4
Once a week	7	28
Two or more times a week	6	24
Every day	2	8

Use of GenAI Tools for Academic Tasks

Students reported using GenAI tools to support different academic tasks, mainly involving writing, programming, learning, presenting, and teaching. Examples of writing tasks are writing a report on a web design project, preparing application materials (e.g., resume, statement of purpose), and writing emails. Examples of programming tasks are creating a webpage and analyzing data. Learning-related tasks included studying course content and preparing for exams. Teaching-related tasks included creating lesson plans and explaining codes in programming. These academic tasks involved more fine-grained actions, such as summarizing texts, generating new text or code, editing text, debugging code, and searching for information. These actions were applied to different tasks. For instance, summarizing texts was performed for writing, learning, and presentation tasks.

Factors Influencing Motivation and Ability to Evaluate GenAI-Generated Information

Our analysis identified four factors influencing students' motivation and two related to their ability to evaluate GenAI credibility in the academic context (Table 2).

Table 2

Factors Influencing Students' Motivation and Ability to Evaluate GenAI Credibility

Elements	Factors	Definitions
Motivation	Task salience (+)	Perceived importance of the task based on the consequentiality of using GenAI-generated information
	Task verifiability (+)	Perceived potential for GenAI-generated information to be fact-checked or confirmed through reliable sources or methods

	Social pressure (+)	Influence of societal norms or authoritative figures on an individual or group to evaluate the credibility of GenAI-generated information
	Trust in GenAI (-)	Overall confidence or reliance on the capabilities of GenAI-enabled systems as sources of information
Ability	Domain knowledge (+/-)	Level of understanding, expertise, and familiarity with specific subject matter or the requirements of a particular task
	Time availability (+)	Amount of time and flexibility that an evaluator has for evaluating information

Note. The signs in the Factor column indicate how the factors influenced students' motivation or ability to evaluate GenAI-generated information based on our data: positive (+), negative (-), or both (+/-).

Task Salience

Most participants considered academic tasks, such as homework assignments and research activities (e.g., literature review), critical enough to warrant spending time reviewing and evaluating the information from GenAI tools. One participant stated, "When asking it [ChatGPT] for research help, it most definitely needs to be evaluated on credibility" (P05). Conversely, when students considered a task to have low salience, they did not review and evaluate the content. In our study sample, such low-salience tasks typically involved casual or personal purposes: "Information about like cooking, I guess something lighthearted. I would not care to check" (P22).

Task Verifiability

Verifiable tasks included factual information (e.g., names, quotes, and references) and technical information that is testable (e.g., codes). Two quotes demonstrate this finding: *"*For accuracy, for completeness, yes, anything related to factual or logical solutions" (P13). "When it comes to the technical question, like the coding, … you can check it" (P18). However, students were less inclined to assess the credibility of information for creative tasks, such as writing a poem.

Social Pressure

Students' social contacts in the academic realm, such as their professors and cohort members, appeared to influence their motivation to evaluate GenAI credibility. For example, one participant mentioned that their professor highlighted the critical importance of discerning information validity when using ChatGPT (P21).

Trust in GenAI

Students showed different levels of trust toward GenAI. Some participants expressed a high level of trust in ChatGPT's capabilities for technical tasks, such as coding, and said additional steps for credibility assessments were unnecessary: "Most people trust ChatGPT and me specifically. … I do trust ChatGPT to give me these or just that it's correct" (P09). In contrast, some participants highlighted ChatGPT's limitations and therefore, were willing to scrutinize its content, especially for academic tasks: "You're verifying your answers, … not just following it plainly, because it is a tool at the end of the day" (P24).

Domain Knowledge

Prior knowledge of or familiarity with specific subject matter or the requirements of a particular task influenced when and how students examined GenAI credibility. One student (P09) noted that it is relatively straightforward to evaluate information credibility when dealing with a familiar topic or task. Additionally, highly knowledgeable users tended to evaluate the validity of the content rather than systematically verifying it (i.e., self-confirmation heuristics; see more about this evaluation approach in the Heuristics for GenAI Credibility Assessments section).

When interacting with GenAI tools on unfamiliar subjects, students highlighted the inherent uncertainty of relying on GenAI-generated content. When students were highly motivated, they cross-checked the information using available sources. For example, P14 recounted: "If I know the answers myself, I can do a quick look through and say, yep, that looks right. If it's something completely new to me, of course, I'm going to go back and verify that that is correct."

Time Availability

Several participants pointed out that time availability influenced their approach to information evaluation. One participant noted, "I believe it just depends on the assignment's deadline" (P18). Similarly, another participant commented, "If I'm in a hurry to do something and I just want like a fast answer, then [I will] probably not [evaluate it]" (P04).

Criteria for Evaluating GenAI Credibility

Students conceptualized credibility differently when using GenAI tools to complete academic tasks. These constructs served as criteria in guiding students' evaluation of GenAI-generated information. Our analysis identified nine constructs: correctness, pertinence, creativity, refinedness, structural accuracy, currency, unbiasedness, explainability, and human-likeness. Table 3 shows their definitions and specific terms used by the participants.

Table 3

Constructs Used to Define Information Credibility in the GenAI Context

Key dimensions	Constructs	Definition	Terms mentioned
Expertise	Correctness	Extent to which information is legitimate or valid and free from fault or error	• Accurate • Factual • Right
	Pertinence	Extent to which information is relevant and applicable in a given task	• Relevant • Focused
	Creativity	Extent to which information conveys original ideas or expressions	• Original • Novel • Generic
	Refinedness	Extent to which information is processed in a concise and straightforward form to address a given task	• Concise • Straightforward • Sophisticated
	Structural accuracy	Extent to which information is presented in a way that meets the expected formatting requirements for a given task	• Up to an academic standard
	Currency	Extent to which information is accurate and relevant at the time of evaluation	• Updated • Accurate today
Trustworthiness	Unbiasedness	Extent to which information is free pf favoritism or prejudice toward a particular viewpoint or outcome	• Opiniated • Equally represented
	Explainability	Extent to which the reasoning behind an output provided by a GenAI system is articulated and understandable	• Trustworthy sources • Help you understand
	Human-likeness	Extent to which information is expressed in human natural language	• Natural • [Not] robotic

Correctness

Students conceptualized credibility using correctness-related terms, such as accurate, factual, and right, especially in the context of erroneous outputs from GenAI tools, such as wrong answers to math problems and fabricated references for academic writing.

> *When I was taking [calculus] classes, … ChatGPT will essentially give me the wrong answer. (P25)*
>
> *It [ChatGPT] creates all the citation and the reference, the website which does not exist, … the people who didn't exist. For example, they can use some kind of like a Mr. Johnson who never exist or never say this kind of thing. (P18)*

Pertinence

Pertinence was mentioned as a critical criterion for judging information credibility in problem-solving tasks to find relevant and applicable solutions in a particular context. One participant stated, "Sometimes it doesn't give the best response. … Like, it may not be too relevant, or it may just limit" (P03).

Creativity

Creativity was considered mainly in creative writing tasks. One participant complained that ChatGPT's responses often contained clichés, which decreased the perceived ability to generate original content: "If I ask it to write a poem or a lyric or some sort of creative writing, I find that it pulls a lot of clichés. … I don't value it as much because it's not as original as it could be" (P13).

Refinedness

Those who mentioned refinedness or related terms, such as concise, straightforward, and sophisticated, expected GenAI to integrate information from relevant sources and process it to answer their questions directly. "I look for concise answers that make sense to me—not, you know, a bunch of technical mumbo jumbo" (P24).

Structural Accuracy

Content structural elements, such as the grammatical accuracy and writing style of ChatGPT's responses, were considered in credibility assessments. For example, P05 commented, "If the grammar was poor, punctuation was poor, I believe I would think differently about the credibility."

Currency

Those aware of the pretrained nature of ChatGPT paid attention to the currency of its information.

> *I mean, the world changes so fast every day that if I was asking about medical information, I think that's, like, really important after going through the pandemic. I don't want to ask an AI*

model like, "Should I wear a mask today?" And then they're like, "Yes" or "No"—whatever based on something that was decided. (P06)

Unbiasedness

Although multiple participants considered unbiasedness an important construct reflecting credibility, their views on potential biases in GenAI-generated information varied. One participant expressed trust in ChatGPT's objectivity:

When I'm Googling things, I will just find opinions of people. … It's all very opinionated, so I don't really trust it, whereas ChatGPT, I feel like it's giving me a nonbiased answer. So, I should—I use ChatGPT more for serious questions or serious work. (P22)

In contrast, one participant expressed wariness of biases derived from the algorithms behind GenAI tools and the datasets used to train them:

I think it was Bing AI, Microsoft's AI. … I think somebody was able to get it to say antisemitic information. … If you're going to [use] already problematic sources, then that information probably will not be the most credible. (P07)

Explainability

Some students valued ChatGPT's explanations, particularly for problem-solving tasks, that articulated the reasoning or sourcing behind an output it provides: "I think that's very credible when they try to explain to you why they did it this way" (P01).

Human-likeness

Some students highlighted that LLM-based GenAI chatbots are designed and trained to mimic human language. Therefore, how natural and fluent the language the chatbots use was considered an important benchmark. One participant noted, "If it's kind of doing it [using a human-like tone] wrong, it's gonna seem more [like a] robot. … I think psychologically, people like to trust something that's more human than something that's more robot or AI" (P09).

GenAI Credibility Assessment Heuristics

Our analysis identified five major heuristics—endorsement, intersystem consistency, self-confirmation, expectancy violation, and intrasystem consistency—and students' actions involved in applying the heuristics to evaluate GenAI-generated information, as shown in Table 4.

Table 4

Heuristics Employed to Assess Information Credibility in the GenAI Context

Heuristics	Definitions	Actions

Endorsement	Viewing the GenAI system or information from the system as credible if others also do so, without much scrutiny of the system or its content	• Going to social media to learn what others say about GenAI tools
Intersystem consistency	Viewing GenAI-generated information as credible if the information across different sources is consistent	• Checking original sources cited • Crosschecking with other systems (e.g., search engines) • Comparing answers with peers working on the same task • Consulting with experts on the topic
Self-confirmation	Viewing GenAI-generated information as credible if it confirms preexisting beliefs and not credible if it counters preexisting beliefs	• Comparing to personal knowledge
Expectancy violation	Viewing GenAI-generated information as not credible if the AI system fails to meet expectations in some way	• Reviewing rewritten content • Comparing to instructions and norms • Testing suggested solutions
Intrasystem consistency	Viewing GenAI-generated information as credible if the information across different trials with slightly modified prompts is consistent	• Creating variations in prompt and comparing responses

Endorsement Heuristic

For college students, social media seemed to play an important role in shaping their perceptions of GenAI tools, such as ChatGPT, which then affected their perceptions of GenAI credibility. One participant said, "If I can go to LinkedIn, [people are saying] 'I have been using ChatGPT for days.' So, that kind of social recommendation or social endorsement about ChatGPT sometime makes you feel like, 'Oh, the answer is correct'" (P10).

Intersystem Consistency Heuristic

One common strategy for GenAI credibility assessments mentioned by our participants was checking if the information matched other sources. The most frequently used source in this triangulation approach was web search engines, primarily Google: “Ask Google the same concepts and if you also see generally the same response in Google, then you can consider ChatGPT to be mostly correct” (P05).

Students also used human sources, such as experts on given topics (e.g., their professors), acquaintances they believed to be better equipped with the topic under investigation, or peers working on the same type of tasks. P19, for example, stated, “I usually just discuss it with a collaborator or a stats consultant.”

Self-Confirmation Heuristic

Students employing the self-confirmation heuristic made immediate judgments of the credibility of GenAI-generated information when the topic was familiar to them: “I’ll just read it and then if I know for a fact that it’s just not making sense. … That’s when red flags will raise” (P21). Students also assessed the logical soundness of the content generated by ChatGPT. They evaluated whether the content provided convincing justifications and aligned with common sense: “I think that’s very credible when they try to explain to you why they did it this way” (P01).

Expectancy Violation Heuristic

Different from the self-confirmation heuristic, which is based on the evaluator’s beliefs and often associated with their cultural and social environments, the expectancy violation heuristic is outcome-based. When working on a task with specific guidelines, such as homework assignments, students checked the credibility of GenAI-generated content against the grading criteria set by their instructors. For problem-solving tasks, which usually have predetermined or correct answers, students tested the answers suggested by ChatGPT using external tools, such as a calculator for a math problem or coding validator (e.g., W3School) for a coding problem.

Intrasystem Consistency Heuristic

Several students employed an iterative testing approach to validate whether the tool produced consistent content across multiple trials using similar prompts. We labeled this approach intrasystem consistency to differentiate it from the intersystem consistency heuristic. One participant mentioned this heuristic: “Ask it again with a different prompt. If it gives me a similar answer, then that would be the first step [for credibility evaluation]” (P02).

Cues

Cues are elements of an information object that affect users' credibility assessments. Cues may reflect specific attributes of the content, the design of the medium through which the content is presented, or the creator of the content (Fogg, 2003). In the present study, we identified 10 cues: version, references, facts, codes, clichés, tone, length of information, presentation of information, interface type, and interface aesthetics. These cues clustered into three categories: application, content, and design (see Table 5).

Table 5

Cues of Information Credibility in the GenAI Context

Category	Description	Cues
Application	Cues related to the characteristics of the GenAI tools	• Version
Content	Cues related to the attributes of the GenAI-generated information objects or the messages they convey	• References • Facts • Codes • Clichés • Tone • Length of information • Presentation of information
Design	Cues related to the visual appearance, structure, and layout of the interface or usability of the interface	• Interface type • Interface aesthetics

Version

Some participants highlighted the pretrained nature of the GPT architecture, indicating that an older version is less capable of generating current information than a newer version. For example, one participant pointed out, "It [ChatGPT] is not able to give you the updated information more or less. … It is just a repository of information that it has learned, like a database, [an] already programmed database, so it's not able to go beyond 2021" (P10). Another participant commented on potential improvements in information credibility in a newer version: "So, maybe now the newer versions have [credible sources]. I've kind of taken information from past conversations, and that's been used. Give better answers, and you know, have less inaccuracies with the responses" (P24).

References

Several participants highlighted the significance of sources. They viewed the presence of sources as an important cue in evaluating the credibility of GenAI responses. For instance, P08 mentioned, “It [ChatGPT] definitely did not cite any sources, but I think if it did, it would definitely be a little more helpful because then you could definitely know if the information was trustworthy, ‘cause I could see where the information was compiled from.”

Beyond the mere presence of a source, students also considered the accuracy of references as a cue signaling information credibility. P18 stated, “When it comes to writings, when it comes to references, I’m always checking the references and trustworthy of in terms of the content, like the answer.”

Facts

For a task that involved factual or logical solutions, students checked statements of fact included in GenAI responses. Several students used Google to cross-check facts across multiple sources. P22 stated, “If it’s a fact-based answer, like ‘Who is John Smith?’ I can just check on Google or like check on the website and see if it matches that information.”

Codes

For programming tasks, specific pieces of code were used as cues to judge the validity of the solutions suggested by GenAI tools. For example, when debugging, P01 asked ChatGPT to highlight the changes made to his original code and explain its reasoning. By comparing differences in the original and revised codes, alongside the explanation for the rationale behind the changes, he judged the credibility of ChatGPT’s answers.

Clichés

In creative contexts, participants recognized that ChatGPT’s responses often contained clichés, which undermined the perceived originality and thus, the credibility of the content.

Tone

Participants preferred a human-like tone in responses, signifying more elaborate information. One student highlighted this point: “How natural it flows, like it doesn’t feel like it’s a computer-generated response. It’s not just it gave a good answer, but it also worded the answer well” (P03).

Length of Information

Some students mentioned the length of answers and favored conciseness. P19 mentioned, “It feels much more frustrating … the lengthy responses that it produces. … If they had more context, maybe it would do better.”

Presentation of Information

Students inspected the compliance of the organization and format of the content with a given academic standard (e.g., APA Style, instructor guidelines). For example, P19 stated, "I looked up like the APA Style reporting of a generalized regression model multilevel regression. I think the response wasn't really clear."

Interface Type

Participants appreciated the straightforwardness of ChatGPT as a chatbot, which aligns with the system's primary function (i.e., chatting) and avoids unnecessary complexity. One participant stated, "I think having a clean and simple and understanding design really does help with the credibility, because it's not overcomplicated" (P01).

Interface Aesthetics

One participant expressed greater trust in AI systems with a professional design. Specifically, P03 commented, "It would be easier to trust AI compared to like maybe a poorly designed website, where then it may feel like it wouldn't be as legitimate."

Barriers and Challenges in GenAI Credibility Assessments

Participants mentioned two major barriers in assessing GenAI-generated content. One was difficulty in establishing and validating the provenance or sources of system responses. One participant noted, "[ChatGPT] doesn't tell you the source. … So, I guess it is kind of hard to figure out how to validate it" (P21). The second barrier was the time cost of validation efforts. Some participants realized that assessing response quality could be time-consuming, requiring more effort than initially expected. For example, P22 commented that, "I think it can be pretty time consuming. … It is great because it's like immediate answers right then and there. But you do need to … really dig and see if it's like valid and true information."

Discussion

Students' Use of GenAI Tools for Academic Tasks

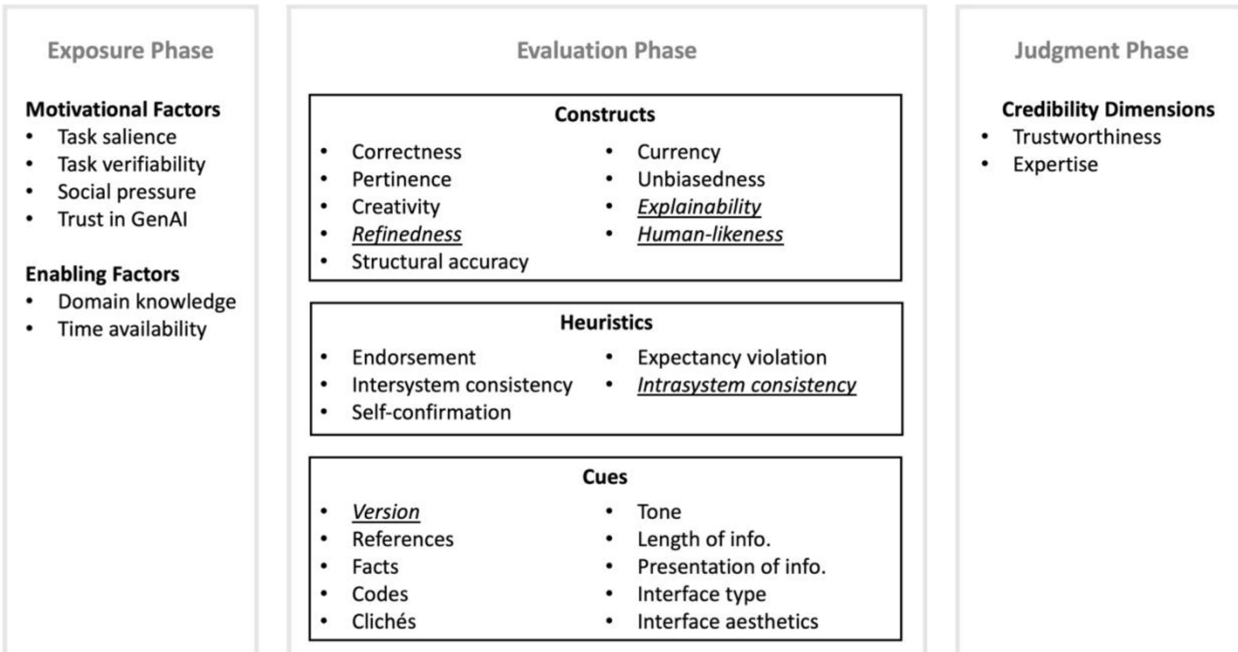
We identified that students used ChatGPT and other GenAI tools for three major work tasks: writing, programming, and learning. For each kind of task, we identified specific tasks performed by the students using ChatGPT, such as summarizing texts, understanding a concept, understanding codes, and debugging codes. These tasks were in line with the current literature. Nevertheless, we identified a new specific task, ideation, wherein students used ChatGPT as a starting point to develop project ideas. It is possible that as users learn more about GenAI, they will perform a wider spectrum of work tasks and specific academic tasks using the tools.

When and How Students Evaluate GenAI Credibility

Figure 2 provides a summary of our findings based on the key components of our synthesized model of GenAI credibility assessment.

Figure 2

Summary of Findings



Note. New constructs, heuristics, and cues identified in the current study are highlighted by underlining and italicizing.

We identified contextual factors that motivated or enabled students to evaluate the credibility of GenAI-generated information in the exposure phase. Students in our study prioritized evaluating GenAI-generated information for academic tasks, such as assignments and research activities, given that they perceived the consequences of adopting the information (i.e., task salience) as high. This finding aligns with the literature on college students' credibility assessments (S. U. Kim & Syn, 2016; Rieh & Hilligoss, 2008). Also, our study identified task verifiability as an important motivational factor. The possibility of checking whether the given information was correct influenced students' willingness to evaluate information, especially for solution-driven (e.g., coding) or fact-checking tasks, for which immediate feedback is applicable and useful in credibility judgments (Metzger et al., 2003).

Social pressure, such as professors' instructions or policies regarding the use of GenAI tools for coursework, can affect students' motivation to assess information credibility. This finding resonates with Lim's (2013) study, which found a significant positive correlation between professors' endorsement and students' perceptions of Wikipedia's credibility when Wikipedia was widely contested as a reliable source of information (e.g., Denning et al., 2005). In the early stages of emerging information technologies, such as GenAI today, educators such as instructors and faculty members can promote students' critical and ethical use of such tools.

Students' trust in GenAI was an individual factor associated with motivation for credibility evaluation, consistent with literature indicating that students' trust in a system or source may influence their evaluation behaviors (Rieh & Danielson, 2007). Trust in a system may be formed by prior experience with or presumption about the system, which then affects credibility assessment (i.e., earned and presumed credibility; Fogg, 2003). Our findings suggest that a high level of trust in GenAI might decrease students' willingness to examine the credibility of GenAI-generated information, resulting in accepting the information without thorough evaluation. If blind trust is one side of the coin, blind distrust is the other, which risks depriving students of opportunities to benefit from GenAI tools for academic tasks (Amoozadeh et al., 2024). These findings imply the necessity of GenAI literacy education for students to ensure an appropriate level of trust based on a good understanding of the strengths and limitations of GenAI tools and appropriate use contexts, especially for academic tasks (Amoozadeh et al., 2024).

Students' domain knowledge is another individual factor that influenced their ability to evaluate GenAI-generated information critically, because domain knowledge was associated with their familiarity with tasks and skills and enabled them to evaluate information or solutions provided by ChatGPT. The positive impact of the evaluator's familiarity with a given task in their credibility assessment has been reported in the credibility literature regarding other types of information systems, such as Wikipedia (Bidegain et al., 2024; Lucassen et al., 2013).

Time availability emerged as a situational factor affecting students' capacity to assess the credibility of GenAI-generated information. When students prioritize speed and convenience over information credibility, such as when facing tight assignment deadlines, they often skip or opt for quick, superficial evaluations, as reported in prior research (Metzger, 2007; Rieh & Hilligoss, 2008).

We identified nine constructs, five heuristics, and 10 cues used by students to assess the credibility of GenAI-generated information (Figure 2). Among the nine constructs, three were unique to GenAI systems, reflecting evolving user expectations in the GenAI context: refinedness, human-likeness, and explainability. Refinedness captures the expectation of polished responses from GenAI tools, going beyond mere relevance or completeness—common criteria in traditional web information credibility and quality assessments (Choi, Stvilia, et al., 2023; Liu et al., 2023)—and tailored to completing a particular task. The human-likeness criterion highlights the diverse roles that ChatGPT and other similar chatbots can fulfill for users in various contexts, such as writing partner, tutor, and coding assistant (Choi, Zhang, et al., 2023). Moreover, students expressed appreciation for insights into why and how a certain answer was generated, especially for solution-driven tasks such as debugging (i.e., explainability; Shin, 2023).

Of the five heuristics we identified, four aligned with those in the web credibility literature (Hilligoss & Rieh, 2008; Metzger & Flanagin, 2013)—endorsement heuristic, intersystem consistency heuristic, self-confirmation heuristic, and expectancy violation heuristic. The newly added heuristic, intrasystem consistency, seems useful in evaluating GenAI credibility, given the probabilistic content-generation behavior of GenAI tools (Peres et al., 2023).

Traditional web credibility cues, such as the presence of sources, voice tone, length, information presentation, and interface aesthetics, were useful in assessing the credibility of GenAI-generated information (Choi & Stvilia, 2015; Sun et al., 2019). One unique cue we found involved the version of the LLM used in the GenAI tool (e.g., GPT-3.5). In the traditional web environment, a timestamp attached to a content item (e.g., blog post, answer on a social Q&A site) indicates the currency of the content, which can influence its perceived credibility (Westerman et al., 2014). However, on ChatGPT or other similar GenAI-enabled systems, the version number or model iteration can serve as a cue indicating the currency of the training dataset, which is related to answer quality.

Based on a web credibility framework categorizing various cues into three groups—application, content, and design (Fogg, 2003)—students primarily considered content-related cues when evaluating ChatGPT information credibility, followed by those related to design and application characteristics (Table 5). Previous research showed that younger adults tended to pay more attention to the appearance of the interface when judging credibility on the web than their older counterparts (Liao & Fu, 2014). On ChatGPT, however, due to its minimalistic and familiar interface as a chatbot, design elements did not seem to affect students' perceptions of information credibility as much. The only cue related to the application mentioned was its version. This may be due to the lack of understanding of the working mechanisms of LLMs, especially on which data they were trained and by which rules they generate responses. In other words, given the early adoption phase of the new technology as of the data collection (April–June 2023) and the "black box" nature of GenAI (Rai, 2020), students did not seem to have sufficient knowledge to judge whether LLMs (e.g., ChatGPT) or their developers (e.g., OpenAI) have good intentions and abilities to provide high-quality information.

Sample Instances of GenAI Credibility Assessments by College Students

Here, we present two memorable instances mentioned by participants, providing holistic descriptions of why and how they evaluated the credibility of GenAI-generated information based on the model we propose in Figure 2.

Instance 1

P01 used ChatGPT to revise his LinkedIn posts. In the exposure phase, P01's motivation for evaluating the credibility of the ChatGPT-generated content was low based on two factors. First, this participant had a high level of trust in the capabilities of ChatGPT, especially for writing tasks: "The grammar that it uses, so far, it's pretty credible. I don't see any major errors in that context." Second, the participant perceived the task verifiability to be low: "I don't really search things that need like a source. I'm not asking to have GPT give me like a source for proof." In addition, P01's ability to evaluate the credibility of the ChatGPT's response in this instance was deemed low because the participant was not familiar or confident with grammar: "I'm really, really bad with grammar, especially like the English grammar. … I like to copy the paragraph that I typed, put it in ChatGPT, and say, 'Clean this up or make it sound better.'"

In the evaluation phase, P01 conceptualized the credibility of LinkedIn posts written by ChatGPT with one construct: human-likeliness. To apply this criterion to the credibility evaluation of the rewritten content, P01 employed one heuristic, self-confirmation, determining the correspondence of the content to his norms by simply reviewing the content. In this evaluation process, he interacted with one peripheral content cue: tone.

In the judgment phase, when the rewritten content was not natural or appropriate for the purpose of the writing, P01 asked ChatGPT to regenerate it: "Sometimes it goes too fancy, like every word is like I don't know, it sounds like Shakespeare or some, you know, Leonardo da Vinci speaking, and that's too, you know—I didn't need it to do that. So sometimes I go like, 'Make it sound a little bit more simple, please. But still you use correct grammar or whatever.' So sometimes I just tell it to change exactly how it sounds."

Instance 2

P14 used ChatGPT to find information about an information technology product. In the exposure phase, P14's motivation was high because the task salience was high. However, P14's ability to examine the credibility was low because the participant was not familiar with the product.

In the evaluation phase, the participant used two criteria (constructs) to define the credibility of ChatGPT-generated information: correctness and refinedness. The participant employed one heuristic, intersystem consistency, searching to verify whether the same information was found on the web. Also, the participant consulted with colleagues who were familiar with the product under investigation. In this evaluation process, P14 interacted with source and content cues: references and fact. In the judgment phase, the participant triangulated the information from web and human resources.

Limitations

We acknowledge four limitations of our study. First, students in the sample were from R1 institutions and predominantly from information technology-related majors, which limits the generalizability of findings to broader populations. Second, the tasks that students mentioned during the interviews were limited to academic tasks. Therefore, our findings regarding motivational factors may not fully capture students' behaviors in assessing GenAI credibility in broader or nonacademic settings. Third, the findings were derived from interview data that relied on participants' self-reporting, which can introduce inaccuracies or gaps in information due to memory biases or misunderstandings. Fourth, data were collected from April to June 2023 and mostly reflected students' experiences with a basic version of ChatGPT (3.5). Current GenAI apps, such as ChatGPT-4 and Copilot, can offer a connection between the generated content and sources and thus, can facilitate source-based credibility evaluation. Our study did not capture students' credibility evaluation strategies and tactics in relation to these features.

To address these limitations, future research can diversify samples across demographics, majors, and institution types to enhance generalizability. Exploring a wider range of GenAI tools, particularly newer versions, and employing varied research methodologies, including controlled experiments, will provide a more comprehensive understanding of target user groups' credibility assessment behaviors.

Conclusion

We conducted semistructured interviews with 25 college students across three geographically dispersed campuses in the United States (Florida, Texas, and Wisconsin) to explore their use of GenAI tools. The participants reported using GenAI tools, primarily ChatGPT 3.5, to complete academic tasks, including writing, programming, and learning. Specific tasks performed with ChatGPT included summarizing texts, understanding new concepts, and simplifying complex information for presentations.

Guided by two existing models of credibility, we identified the structure and key factors influencing students' decision making on when and how deeply to engage in the credibility evaluation process. These factors included task-related, social, individual, and situational factors (Table 2). Our findings highlight the importance of considering the context when studying credibility assessments. We also identified several credibility evaluation criteria specific to GenAI systems, such as refinedness, explainability, and human-likeness, along with heuristics like intrasystem consistency and cues such as system version.

Our study emphasizes the importance of ongoing efforts to understand people's information evaluation behaviors and the contexts that shape these behaviors as new technologies are incorporated into our information environment. Our findings also highlight a need for literacy programs for college students in higher education. Such programs have the potential to increase both motivation and ability by improving students' understanding of the fundamental working mechanisms of GenAI and relevant strategies for evaluating information credibility for different tasks and remain updated with reliable tools that can facilitate their credibility assessments of GenAI-generated information for academic tasks.

Acknowledgements

This study was partially supported by the College of Community Engagement and Professions, University of Wisconsin-Milwaukee, Research Seed Award.

References

Amoozadeh, M., Daniels, D., Nam, D., Kumar, A., Chen, S., Hilton, M., Ragavan, S. S., & Alipour, M. A. (2024). Trust in generative AI among students: An exploratory study. *Proceedings of ACM Conference (Conference'17)*, 7. https://doi.org/10.48550/arXiv.2310.04631

Anderson, J., & Rainie, L. (2023). *The future of human agency*. Pew Research Center. https://www.pewresearch.org/internet/2023/02/24/the-future-of-human-agency/

Athaluri, S. A., Manthena, S. V., Kesapragada, V. S. R. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the Boundaries of Reality: Investigating the Phenomenon of Artificial Intelligence Hallucination in Scientific Writing Through ChatGPT References. *Cureus*. https://doi.org/10.7759/cureus.37432

Bellur, S., & Sundar, S. S. (2014). How can we tell when a heuristic has been used? Design and analysis strategies for capturing the operation of heuristics. *Communication Methods and Measures*, *8*(2), 116–137. https://doi.org/10.1080/19312458.2014.903390

Bidegain, E., Egaña, T., & Zuberogoitia, A. (2024). Explaining how Wikipedia deals with credibility to university students: The case of Wikipedia in the Basque language. *Journal of Information, Communication and Ethics in Society*, *22*(1), 159–173. https://doi.org/10.1108/JICES-08-2023-0104

Borji, A. (2023). *A categorical archive of ChatGPT failures* (arXiv:2302.03494). arXiv. https://doi.org/10.48550/arXiv.2302.03494

Breakstone, J., Smith, M., Connors, P., Ortega, T., Kerr, D., & Wineburg, S. (2021). Lateral reading: College students learn to critically evaluate internet sources in an online course. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-56

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Chang, Y.-S., Zhang, Y., & Gwizdka, J. (2021). The effects of information source and eHealth literacy on consumer health information credibility evaluation behavior. *Computers in Human Behavior*, *115*, 106629. https://doi.org/10.1016/j.chb.2020.106629

Choi, W., & Stvilia, B. (2015). Web credibility assessment: Conceptualization, operationalization, variability, and models. *Journal of the Association for Information Science and Technology*, *66*(12), 2399–2414. https://doi.org/10.1002/asi.23543

Choi, W., Stvilia, B., & Lee, H. S. (2023). Developing a platform-specific framework for web credibility assessment: A case of social Q&A sites. *Information Processing & Management*, *60*(3), 103321. https://doi.org/10.1016/j.ipm.2023.103321

Choi, W., Zhang, Y., & Stvilia, B. (2023). Exploring applications and user experience with generative AI tools: A content analysis of reddit posts on ChatGPT. *Proceedings of the Association for Information Science and Technology*, *60*(1), 543–546. https://doi.org/10.1002/pra2.823

Danielson, D. R. (2006). Web credibility. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 713–721). IGI Global. https://doi.org/10.4018/978-1-59140-562-7.ch107

Dehnert, M., & Mongeau, P. A. (2022). Persuasion in the age of artificial intelligence (AI): Theories and complications of AI-based persuasion. *Human Communication Research*, *48*(3), 386–403. https://doi.org/10.1093/hcr/hqac006

Denning, P., Horning, J., Parnas, D., & Weinstein, L. (2005). Wikipedia risks. *Communications of the ACM*, *48*(12), 152–152.

Diakopoulos, N., & Koliska, M. (2017). Algorithmic Transparency in the News Media. *Digital Journalism*, *5*(7), 809–828. https://doi.org/10.1080/21670811.2016.1208053

Emsley, R. (2023). ChatGPT: These are not hallucinations – they're fabrications and falsifications. *Schizophrenia*, *9*(1), 1–2. https://doi.org/10.1038/s41537-023-00379-4

Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet Information Credibility. *Journalism & Mass Communication Quarterly*, *77*(3), 515–540. https://doi.org/10.1177/107769900007700304

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

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust Online: Young Adults' Evaluation of Web Content. *International Journal of Communication*, *4*, 468–494.

Head, A. J., & Eisenberg, M. B. (2011). Truth Be Told: How College Students Evaluate and Use Information in the Digital Age. *First Monday*, *16*(4). https://firstmonday.org/ojs/index.php/fm/article/download/3484/2857

Head, A. J., & Eisenberg, M. B. (2010). Truth be told: How college students evaluate and use information in the digital age. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2281485

Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, *44*(4), 1467–1484. https://doi.org/10.1016/j.ipm.2007.10.001

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. Yale University Press.

Kim, K.-S., Yoo-Lee, E., & Joanna Sin, S.-C. (2011). Social media as information source: Undergraduates' use and evaluation behavior. *Proceedings of the American Society for Information Science and Technology*, *48*(1), 1–3. https://doi.org/10.1002/meet.2011.14504801283

Kim, S. U., & Syn, S. Y. (2016). Credibility and usefulness of health information on Facebook: A survey study with U.S. college students. *Information Research*, *21*(4). https://eric.ed.gov/?id=EJ1123245

Lee, J. Y., & Sundar, S. S. (2013). To tweet or to retweet? That Is the question for health professionals on Twitter. *Health Communication*, *28*(5), 509–524. https://doi.org/10.1080/10410236.2012.700391

Liao, Q. V., & Fu, W.-T. (2014). Age differences in credibility judgments of online health information. *ACM Transactions on Computer-Human Interaction*, *21*(1), 1–23. https://doi.org/10.1145/2534410

Lim, S. (2013). College students' credibility judgments and heuristics concerning Wikipedia. *Information Processing & Management*, *49*(2), 405–419. https://doi.org/10.1016/j.ipm.2012.10.004

Liu, J., Song, S., & Zhang, Y. (2021, March 17). Linguistic features and consumer credibility judgment of online health information. *iConference 2021 Proceedings*. iConference. http://hdl.handle.net/2142/109674

Liu, J., Zhang, Y., & Kim, Y. (2023). Consumer health information quality, credibility, and trust: An analysis of definitions, measures, and conceptual dimensions. *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*, 197–210. https://doi.org/10.1145/3576840.3578331

Lucassen, T., Muilwijk, R., Noordzij, M. L., & Schraagen, J. M. (2013). Topic familiarity and information skills in online credibility evaluation. *Journal of the American Society for Information Science and Technology*, *64*(2), 254–264. https://doi.org/10.1002/asi.22743

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education*, *46*(2), 165–193. https://doi.org/10.1080/00933104.2017.1416320

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, *58*(13), 2078–2091. https://doi.org/10.1002/asi.20672

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, *59*, 210–220. https://doi.org/10.1016/j.pragma.2013.07.012

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, *60*(3), 413–439. https://doi.org/10.1111/j.1460-2466.2010.01488.x

Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student Web use, perceptions of information credibility, and verification behavior. *Computers & Education*, *41*(3), 271–290. https://doi.org/10.1016/S0360-1315(03)00049-6

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook. 3rd* (3rd ed.). Sage Publications.

Ngo, T. T. A. (2023). The Perception by University Students of the Use of ChatGPT in Education. *International Journal of Emerging Technologies in Learning (Online)*, *18*(17), Article 17. Publicly Available Content Database. https://doi.org/10.3991/ijet.v18i17.39019

NVivo (Version 14). (2023). [NVivo]. Lumivero. www.lumivero.com

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). *Training language models to follow instructions with human feedback* (arXiv:2203.02155). arXiv. http://arxiv.org/abs/2203.02155

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google We trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, *12*(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*, *40*(2), Article 2. https://doi.org/10.1016/j.ijresmar.2023.03.001

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137–141. https://doi.org/10.1007/s11747-019-00710-5

Rieh, S. Y. (2017). Credibility and cognitive authority of information. In J. D. McDonald & M. Levine-Clark (Eds.), *Encyclopedia of Library and Information Science* (4th Ed., pp. 1113–1120). CRC Press. https://doi.org/10.1081/E-ELIS4-120044103

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, *41*(1), 307–364. https://doi.org/10.1002/aris.2007.1440410114

Rieh, S. Y., & Hilligoss, B. (2008). College students' credibility judgments in the information-seeking process. *Digital Media, Youth, and Credibility*, 49–72. https://doi.org/10.1162/dmal.9780262562324.049

Schuetzler, R. M., Giboney, J. S., Wells, T. M., Richardson, B., Meservy, T., & Sutton, C. (2024). Student interaction with generative AI: An exploration of an emergent information-search process. *Proceedings of the 57th Hawaii International Conference on System Sciences*, 7500–7509.

Shaw, C., Yuan, L., Brennan, D., Janson, N., Fox, K., & Bryant, G. (2023). *GenAI in higher education: Fall 2023 update of time for class study* (pp. 1–17). Tyton Partners. https://tytonpartners.com/app/uploads/2023/10/GenAI-IN-HIGHER-EDUCATION-FALL-2023-UPDATE-TIME-FOR-CLASS-STUDY.pdf

Shin, D. (2023). Embodying algorithms, enactive artificial intelligence and the extended cognition: You can see as much as you know about algorithm. *Journal of Information Science*, *49*(1), 18–31. https://doi.org/10.1177/0165551520985495

Shin, D., Koerber, A., & Lim, J. S. (2024). Impact of misinformation from generative AI on user information processing: How people understand misinformation from generative AI. *New Media & Society*, 1–31. https://doi.org/10.1177/14614448241234040

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society*, *9*(2), 20539517221115189. https://doi.org/10.1177/20539517221115189

Sun, Y., Zhang, Y., Gwizdka, J., & Trace, C. B. (2019). Consumer evaluation of the quality of online health information: Systematic literature review of relevant criteria and indicators. *Journal of Medical Internet Research*, *21*(5), e12522. https://doi.org/10.2196/12522

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. *Digital Media, Youth, and Credibility*, 73–100. https://doi.org/10.1162/dmal.9780262562324.073

Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social media as information source: Recency of updates and credibility of information. *Journal of Computer-Mediated Communication*, *19*(2), 171–183. https://doi.org/10.1111/jcc4.12041

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Wu, C., Wang, X., Carroll, J., & Rajtmajer, S. (2024). *Reacting to generative AI: Insights from student and faculty discussions on Reddit*. https://doi.org/10.35542/osf.io/h79ms

Wu, S., Zhang, J., & Du, L. (2022). I do not trust health information shared by my parents. *Health Communication*, *39*, 96–106.