MUNI FSS

JOURNAL OF PSYCHOSOCIAL RESEARCH ON CYBERSPACE
CYBERPSYCHOLOGY

Fung, J. M. T., McMichael, S. L., Kadotani, K. M., & Kwan, V. S. Y. (2026). Users vs. independent judges: Comparing perceived authorship and ownership in AI-supported collaborative writing. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 20*(5), Article 6. https://doi.org/10.5817/CP2026-5-6

Users vs. Independent Judges: Comparing Perceived Authorship and Ownership in AI-Supported Collaborative Writing

Joseph M. T. Fung, Samantha L. McMichael, Kamryn M. Kadotani, & Virginia S. Y. Kwan

Department of Psychology, Arizona State University, Tempe, Arizona, USA

Abstract

With the rapid integration of Artificial Intelligence (AI) in academic and creative settings, understanding how this technology impacts fundamental concepts of authorship and ownership is important to protect learning, maintain academic integrity, and establish clear policies. This study (N = 320; 18 to 44 years old, M = 19.55, SD = 2.39) utilized a 2 (user vs independent judge) x 4 (completely user generated; AI-generated and user edited 4 sentences; user chooses from 3 AI-generated options; completely AI generated) between-subjects design to examine how levels of Large Language Model (LLM) input affect attributions of authorship and ownership by both the users and the independent judges in a collaborative writing task. Findings reveal a significant inverse relationship between the level of LLM input and users' reported ownership and authorship and a positive relationship between their attributions of ownership and authorship. Additionally, discrepancies emerged between users' and judges' attributions, particularly at high and low levels of LLM input, suggesting that third-party evaluators and users may use different factors in allocating credit. Furthermore, this research investigated the role of background factors such as users' experience with AI, their attitudes toward AI, and their writing level in moderating the relationships between (1) AI input level and attributions of ownership and (2) attributions of ownership and authorship. The findings reveal the need to identify optimal levels of AI integration that support learning without diminishing students' sense of authorship and to develop assessment strategies that address differences in attributions of ownership by users and independent judges. For policy, the observed differences in attribution demonstrate the complexities of assigning credit and responsibility in AI-assisted creative works.

Keywords: AI; authorship; ownership; collaborative; human-AI teams; LLM

Editorial Record

First submission received:
June 2, 2025

Revisions received:
March 5, 2026
July 15, 2025
October 12, 2026

Accepted for publication:
October 13, 2026

Editor in charge:
Lenka Dedkova

Introduction

Since its introduction in 2023, Turnitin's new Artificial Intelligence (AI) detection tool has identified over 6 million student papers in which generative AI produced at least 80% of the content (Chechitelli, 2024). While this statistic highlights AI's rising impact in academia, it also demonstrates AI's evolving role as a collaborator capable of producing content alongside its user. These quickly advancing output capabilities may begin to affect conventional notions of authorship and ownership, which were traditionally reserved for human contributors.

Credit allocation is a nuanced process in any collaborative work, with both internal team members and external observers or independent judges assigning credit for a task's success or failure based on various factors (e.g., perceived pivotality of an individual's contribution, Lagnado et al., 2013; Zultan et al., 2012; level of effort individuals are willing to invest, Xiang et al., 2025). Research demonstrates that the processes involved in distributing credit are complex, and certain team members may receive more credit than others (Bikard et al., 2015; Savitsky et al., 2012). The rise of generative AI adds another layer of complexity, as its versatility and efficiency in executing subtasks reduces reliance on human teammates (Bozkurt, 2023). Despite its capabilities, AI remains dependent on user prompting, distinguishing it from fully autonomous human collaborators. As human-AI partnerships increasingly contribute to the creation of intellectual works, it becomes crucial to evaluate whether the extent of AI involvement in collaborative tasks influences perceptions of authorship and ownership within such teams.

These questions are emerging in real-world contexts where AI-assisted work is becoming commonplace. The following vignette illustrates the growing issue in higher education: Jane entered her engineering capstone with an idea to utilize soil composition for localized energy regeneration. As she lacked advanced computational coding skills, she engaged in an intense, highly iterative dialogue with a generative AI tool to bridge the technical gap. Jane provided the foundational chemical constraints and real-world parameters, while the AI drafted the complex molecular configurations, simulated the structural integrity of the compounds, and continuously refined the digital models based on her critique. The resulting prototype—a highly novel, viable approach to energy regeneration—emerged as a true 50/50 hybrid creation, seamlessly blending human engineering intuition with machine-driven optimization. Jane submitted the project in partial fulfillment of her engineering degree. Is Jane the sole author of her capstone project?

The present research begins with a review of relevant literature, including studies on the allocation of credit for collaborative tasks, writing with the assistance of AI, and the distribution of credit for AI-human partnerships. Following the literature review, we detail the current research that explored how authorship and ownership was attributed when a user employed a Large Language Model (LLM; i.e., an AI system trained on textual datasets to enable human-like responses based on user prompting) to complete a collaborative writing task.

As self-perception often differs from external evaluation, this research further investigated whether these attributions differed depending on whether they are assessed by the users themselves or by an independent judge. We also considered potential factors that could affect these perceptions, specifically experience with AI, attitudes toward AI, and baseline writing proficiency. The findings offer insight into how human-AI collaboration is reshaping traditional practices of credit distribution, with important implications for psychological sciences, policy, and academia.

Attribution of Credit by Independent Judges in Collaborative Tasks

The process of assigning credit in collaborative tasks is complex. Independent (i.e., uninvolved) judges may distribute credit unevenly among the collaborating group members. To understand this phenomenon, researchers have explored the mechanisms underlying credit allocation judgements. One factor in credit attribution is the perceived pivotality of an individual's contribution (Lagnado et al., 2013; Zultan et al., 2012). Specifically, judges allocate a greater share of responsibility to individuals whose contributions were deemed essential for the group's success. For instance, in a soccer match, a player who was expected to score the winning goal would face a majority of the blame if he failed to do so. Alternatively, if many other teammates were viable scorers, the initial player's role becomes less pivotal, and he would accordingly face less blame if the team lost (Zultan et al., 2012). Another study adds that judges also weigh the *expected* impact of each member's contributions against their actual pivotality (Allen et al., 2015). Using the same example, beyond solely assessing whether a player's performance was essential, judges would also consider if the players performance aligned with the judges' pre-match predictions. Similarly, another study found that credit attribution involves the consideration of actual and counterfactual factors (Xiang et al., 2023). Actual factors include comparing a member's effort with their maximum potential effort, while counterfactual factors consider the necessity of their contributions for the team's outcome (Xiang et al., 2023).

Distinguishing Authorship and Ownership in Collaborative Work

To fully understand the nuance of credit allocation, it is crucial to differentiate between authorship and ownership. Authorship, though ambiguous in the literature, often refers to the public recognition of who contributed to the intellectual content (Bozkurt, 2024). In personal writing, a declaration of authorship represents an individual judgment which may be shaped by a sense of control and personal influence over the work product (Draxler et al., 2024). Ownership, on the other hand, can involve legal rights, such as the intellectual property rights and control over a work (Simone, 2019). Importantly, ownership can take multiple forms. While legal ownership focuses on legal rights, psychological ownership involves a relationship between the person and the object and a perception that the object is "mine" (Pierce et al., 2003). Legal ownership may grant formal legal benefits, but it does not necessarily imply recognition of contribution. Psychological ownership reflects an individual's personal sense of contribution and responsibility. Crucial to the current research, the distribution of both authorship and ownership can be uneven for members of teams, and they can reflect varying levels of contribution or legal agreements. This distinction becomes especially relevant in collaborative settings where multiple individuals may contribute to a work (Dreyfuss, 2000).

As AI output is increasingly used alongside human work, questions arise regarding who should be recognized as an author or the owner of the work. Under current U.S. law, copyright protection extends only to original works created by a human author. Courts in the U.S. have repeatedly rejected AI as an author. For instance, in Thaler v. Perlmutter, the court ruled that an AI system (The Creativity Machine) cannot hold copyright because the law requires a human mind to be the creative source (Thaler v. Perlmutter, 2025).

While AI cannot be the owner, there is ongoing debate about whether the human user of the AI can claim the invention if the AI did most of the work. For instance, the U.S. Copyright Office partially revoked protection for the comic book Zarya of the Dawn, determining that while the human-authored text and layout were protectable, the AI-generated images themselves were not (Kasunic, 2023). At the other end of the spectrum, the Beijing Internet Court granted copyright to an AI-generated artwork, reasoning that the user's specific prompts and parameter adjustments constituted sufficient intellectual investment to qualify the human user as the author (Li v. Liu, 2023).

While AI cannot legally be owners of human-AI collaborative work, the involvement or use of AI in collaborative work may impact the attribution of authorship and ownership for the human user. The attribution of authorship and ownership is shifting toward a sliding scale of human involvement. Understanding lay beliefs about concepts such as authorship and ownership—independent of legal definitions—is crucial because the law does not operate in a vacuum; it relies on the cooperation and perceived legitimacy of the public it governs. When we study what people naturally believe authorship and ownership mean, we may uncover the psychological reality that drive jury behavior, compliance, and social trust. By examining these unfiltered perceptions, we can identify areas where the law is out of step with societal values, ultimately helping to shape a legal system that is more intuitive, accessible, and democratically grounded.

Importantly, the present research does not provide participants with explicit definitions of authorship or ownership. Instead, it allows them to rely on their own lay judgments when making these attributions. Given this study design, the present research examines participants' perceptions and judgments of authorship and ownership, rather than providing evidence about the legal status of human-AI collaborations. This method aligned with previous AI credit allocation work (e.g., Draxler et al., 2024; Lehmann et al., 2022) which allowed participants to assess authorship and ownership based on their own understanding of the terms (e.g., providing a text box allowing participants to fill in an author; asking "To whom does this text belong?" (Draxler et al., 2024); asking participants to indicate agreement with "I feel like I am the author of the text" (Lehmann et al., 2022). Examining these unstructured perceptions is important because beliefs about authorship and ownership may differ from formal legal definitions and reflect the psychological processes that guide real-world evaluations.

Authorship and Ownership in Human-AI Teams

Much of the past research on credit attribution has primarily focused on collaborations involving only humans. However, the advancing capabilities of generative AI are shifting the composition of teams. While AI is commonly referred to as a tool, its evolving sophistication allows it to assume the role of a group member that is capable of playing a pivotal role in its team's success (Seeber et al., 2020). This shift is evident across industries, where professionals in medicine, law, research, and content creation have delegated tasks that once required their

expertise to AI (Anantrasirichai & Bull, 2022; Rashid & Kausik, 2024; Wilder et al., 2020). Determining the threshold of human (or AI) contribution required to grant ownership over AI-assisted work, alongside the broader debate on AI authorship, represents a critical challenge in many fields.

In educational settings, there is a growing and evolving literature on AI and the use of generative AI (GenAI) in education (see Allison et al., 2025; Tu & Lu, 2025). Foundational research has demonstrated how AI presents both opportunities and challenges in educational settings, impacting both learners and educators (e.g., Fan et al., 2025; Fu et al., 2025; Liu et al., 2025;). In higher education, one pressing issue is traceable AI contribution. Course syllabi now often explicitly define the permissible boundaries of AI integration, requiring students to report the extent of AI assistance in their work. Perceptions of authorship and ownership for written assignments with varying degrees of AI collaboration is an emerging area of research. Recent research has begun to examine how individuals allocate credit within human–AI collaborations. For example, a study of human-AI writing teams found that users assigned different levels of credit depending on the type of contribution, the amount of contribution, and whether the AI's contribution was generated proactively or in response to a human prompt (He et al., 2025; Pennanen et al., 2026). Importantly, across conditions, users assigned AI collaborators less credit than human collaborators for equivalent contributions (He et al., 2025).

The concepts of authorship and ownership become increasingly complex when AI performs portions of the work in producing an academic output. Students enrolled at a university are generally expected to develop an expertise and contribute to existing knowledge through their academic work. However, students who utilize AI LLMs produce written outputs based on existing knowledge and data which may lead to homogenization of scholarly work (Yang et al., 2026). As a result, the use of LLMs in academic work has raised questions about how authorship and ownership should be attributed when students incorporate AI-generated content into their work.

In addition to producing original knowledge, students and instructors are expected to follow principles of academic integrity, working within the bounds of university policy to appropriately evaluate a student's scholarly progression. Moreover, written assignments are expected to be original work authored by the student for instructors to administer evaluations reflecting student mastery. The use of AI writing tools introduces an additional consideration for evaluations; the extent that a student uses AI may vary across assignments and may not always be readily apparent from their submitted work. Currently available AI-detection tools are often unreliable and are particularly unreliable for some fields of study (e.g., engineering) and for writers who are not native English speakers (Pratama, 2025). With this evaluative boundary unclear, perceived authorship over written academic work may not yet adequately capture the extent to which the student's work demonstrates understanding of the content without the use of AI writing tools.

Collaborative Writing With AI

An example of human-AI collaboration can be observed in academic and professional settings, where users work alongside non-personalized LLMs to improve their writing in various ways. Unlike personalized AI, which requires training on individual user data prior to use, non-personalized LLMs like ChatGPT and Gemini are readily accessible and widely used, making them a popular focus of research.

Such research has found that LLMs can assist a writer throughout different stages of the writing process. In the beginning stages, AI can help writers brainstorm, create rough drafts, and conduct background research (Arif et al., 2023). During later stages, AI can improve writers' existing drafts by making suggestions for grammar, style, and clarity (Lockwood & Castleberry, 2024).

Self-Attribution and Independent Judgments in Human-AI Writing

While independent judges consider factors like pivotality when making credit attributions, individuals often use different processes to assess their own authorship and ownership of a work. In general, (i.e., psychological ownership) is derived from creation, control, and personal investment (Pierce et al., 2003). In the context of AI-assisted writing, users felt a stronger sense of authorship when they actively selected and implemented AI-generated phrases themselves, while authorship was reduced when the AI produced text without the user's input (Lehmann et al., 2022). A study focusing on psychological ownership in AI-assisted writing found similar results where user editing and iterative prompting predicted greater psychological ownership than AI-assisted writing

produced from a single user prompt (Pennanen et al., 2026). Importantly, users who produced writing without AI assistance reported more psychological ownership than all AI conditions.

A recent study explored how varying levels of collaboration with a personalized LLM influenced users' attributions of authorship and ownership (Draxler et al., 2024). The researchers began by collecting writing samples from participants to train the LLM on their individual writing style and personality. Following personalization, participants collaborated with the LLM to create four postcards, each with a different level of interaction: (1) entirely participant written, (2) written by LLM with participant editing up to 30 words, (3) participant selection from three LLM-generated options, and (4) entirely LLM written. Participants completed each postcard with the help of either a genuinely personalized LLM or a placebo-personalized LLM, which received no prior training. Participants were then asked to identify the author of the postcard, rate their perceived ownership, and report their sense of control and authority over the task. The study revealed that receiving LLM-generated text led to lower perceived ownership compared to self-written postcards. Interestingly, participants rarely credited the LLM as the author, regardless of the LLM's input on the text and their perceived ownership (Draxler et al., 2024). The researchers coined this phenomenon the ghostwriter effect. Furthermore, no significant differences in authorship or ownership attributions were found between the personalized and placebo-personalized conditions at any level. Participants who believed they were collaborating with a human rather than the LLM perceived less self-ownership and credited their collaborator with authorship more often (Draxler et al., 2024).

Building on this literature, the present research examines whether varying levels of AI contribution alter users' attributions of ownership and authorship. Specifically, we ask: Does the level of AI input influence users' attributions of ownership and authorship of a collaborative writing assignment **(Research Question 1)**? Based on prior research (Pennanen et al., 2026; Pierce et al., 2003), we hypothesize:

H1: Users report greater personal ownership when the AI contributes less to the task.

However, consistent with past findings showing that users tend to credit themselves as the author regardless of the amount of AI involvement (Draxler et al., 2024), we also hypothesize:

H2: Users' authorship attributions do not significantly vary by level of AI input.

The relationship between users' attributions of ownership and attributions of authorship further warrants examination. Although these constructs are conceptually distinct, individuals may treat them as psychologically linked or they may not. To examine that relationship we tested the following: Do users' attributions of ownership predict reported authorship; that is, do users who attribute greater ownership to the AI credit greater authorship to the AI **(Research Question 2)**? Consistent with prior research (Draxler et al., 2024), we hypothesize:

H3: Users' ownership attributions do not significantly predict users' authorship attributions.

Although prior research has focused primarily on users' self-perceptions, credit allocation in collaborative settings often extends beyond the contributors themselves. Independent judges may rely more heavily on observable contribution levels and perceived pivotality when assigning ownership and authorship. Research on credit attribution demonstrates that external evaluators weigh both actual and counterfactual contribution factors, as well as perceived necessity for success (Xiang et al., 2023; Zultan et al., 2012). As increasing AI input reduces the apparent pivotality of the human contributor, judges may systematically decrease attributions of ownership and authorship to the user as AI contribution increases. Thus, it is important to determine whether the patterns observed in users' self-attributions generalize to uninvolved observers or whether independent judges apply different evaluative standards. To examine this issue, we tested the following: Do the patterns observed for users also hold for independent judges' evaluations of ownership and authorship? In other words, at each level of AI input, do independent judges attribute ownership and authorship similarly to the user themselves and do judges' attributions of ownership predict their attributions of authorship **(Research Question 3)**? Based on prior research (Pierce et al., 2003; Zultan et al., 2012), we hypothesize:

H4: Similar to the users, judges' attributions of ownership decrease as AI input increases.

H5: In contrast to users, judges' attributions of the user as author decline as AI contribution increases.

H6: At the highest levels of AI input, judges assign lower ownership and authorship to the user than users assign to themselves

H7: Similar to users, judges' ownership attributions do not significantly predict judges' authorship attributions.

Human-AI Teams and Moderating Factors

Importantly, human-AI teams have shown potential to outperform humans or AI working independently, particularly in optimizing manufacturing, administrative, and academic tasks. This demonstrates the complementary potential of such teams (Bansal et al., 2021; Chui et al., 2016; Hemmer et al., 2021; Schleiger et al., 2024). To optimize the effectiveness of these teams, social science researchers are exploring strategies that go beyond simply maximizing AI's individual performance potential. For instance, they found that a critical factor in a user-AI team's success is the user's pre-existing attitudes towards AI. Research indicates that positive attitudes toward AI led to increased willingness to adopt AI and improved team performance (Kumar et al., 2025; Prentice et al., 2020; Zahlan et al., 2023). Further, more experience with AI improves users' trust in the AI, which consequently enhances team performance (Bansal et al., 2021; De Visser Ewart et al., 2020; Zhang, 2022). Developers can also improve human-AI partnership synergy by training AI to seek user guidance for tasks where it may lack precision, rather than guessing or working around uncertain information (Cooray, 2026; Green & Chen, 2019; Liu et al., 2021; Wilder et al., 2020).

Specific to human-AI collaborative writing, the effectiveness of the human-AI partnership may be related to the users' baseline writing proficiency. Though potentially useful to all writers, AI impacts writers of varying skill levels and experience with AI differently. Specifically, users that are infrequent writers and those with less technology experience benefit more from highly independent (less user input) AI collaboration than other users (Dhillon et al., 2024). Furthermore, human-AI teams achieve superior results when the human has complementary strengths, rather than overlapping strengths, with the AI (Bansal et al., 2021; Zhang et al., 2022). Applied to collaborative writing, this suggests that less proficient writers may observe greater improvements than those with more writing experience when using AI.

Individual differences between users may influence human-AI teams and their outcomes. Positive attitudes toward AI increase willingness to adopt AI and improve team performance (Kumar et al., 2025; Prentice et al., 2020; Zahlan et al., 2023;). Greater experience with AI enhances trust and collaboration effectiveness (Bansal et al., 2021; De Visser Ewart et al., 2020; Glikson & Woolley, 2020; Zhang, 2022). Users with lower writing proficiency benefit more from higher levels of AI contribution (Dhillon et al., 2024). As these factors shape human-AI collaboration, we examine whether they moderate the relationships between levels of AI input and users' attributions of ownership and authorship and the relationship between ownership and authorship for users. Accordingly, we test the final research question: Do users' experience with AI, attitudes toward AI, or baseline writing level moderate (a) the relationship between AI input level and attributions of ownership and (b) the relationship between attributions of ownership and authorship **(Research Question 4)**? Based on prior research, we hypothesize:

H8: Users with greater AI experience, more positive attitudes toward AI, and lower baseline writing levels are more likely to attribute ownership to the AI at all levels of AI input.

We further hypothesize that:

H9: For users with greater AI experience, more positive attitudes toward AI, and lower baseline writing levels, ownership attributions more strongly predict authorship attributions (i.e., greater attributions of AI ownership predict greater AI authorship) at all levels of AI input.

Overview of Research

Existing research has examined credit allocation and provided insight into users' perspectives when collaborating with AI, but gaps remain. Draxler et al. (2024) explored how perceived ownership and reported authorship was influenced by different levels of AI input and whether users believed they are working with AI or a human. However, this past study focused on AI systems personalized to individual users, leaving credit considerations for interactions with non-personalized AI largely unexplored. This distinction is crucial, as non-personalized LLMs, such as ChatGPT, represent a widely used form of AI writing assistance in both professional and academic settings (Kasneci, 2023). Additionally, existing AI research has focused mainly on users' self-perceptions, leaving open research questions regarding the perspectives of independent judges. A more comprehensive understanding of AI's impact on authorship and ownership requires comparing both perspectives to address the different factors each uses when attributing credit for collaborative work.

To bridge the gaps in the existing literature, this study explored the impact of non-personalized AI input on credit attribution. By comparing the perspectives of users and independent judges, this research aimed to investigate credit attribution beyond the user's perception to understand how independent judges may credit ownership and authorship similarly or differently to the user themselves. This understanding may provide important foundations for real world applications of Human-AI collaborative work (e.g., legal contexts of ownership and authorship). This study also investigated the effect of three potential moderating factors that may affect interactions between humans and AI, namely experience with AI, attitudes toward AI, and writing proficiency level.

Methods

Study Design and Statistical Methods

This study utilized a 2 (user versus independent judge) x 4 (Level 1- completely user generated, Level 2- AI-generated and user edited 4 sentences, Level 3- user chooses from 3 AI-generated options, Level 4- completely AI generated) between-subjects design. Based on a power analysis to detect a medium sized interaction effect (f = .25) with eight conditions, we aimed to recruit at least 280 participants (G*Power, 3.1.9.7; Faul et al., 2009). Participants were randomly assigned to one of the eight condition groups. We conducted moderation analyses using PROCESS (V.4.2). All other analyses were conducted using SPSS and SPSS AMOS (V.28.0). Prior to conducting analyses, we assessed whether the assumption of homogeneity of variance was met (e.g., using Levene's test). If the assumption was not met (e.g., Levene's test was significant), we used robust alternative tests and indicated them in the results (e.g., Welch's ANOVA; Dunnet T3 post-hoc test). When noted in the results below, we applied a Bonferroni-adjusted significance threshold to the p-values to control family-wise error rate (Dunn, 1961).

The research questions, hypotheses, and data analysis methods for this study were preregistered on Open Science Framework.[1] Several deviations from the preregistration should be noted. First, Hypotheses 2 and 9 were not included in the preregistration due to an oversight during the preregistration process. Accordingly, these hypotheses should be considered exploratory, and future research is needed to replicate these findings. Second, although the preregistration proposed testing Hypotheses 8 and 9 for both users and independent judges, these analyses were ultimately restricted to users as the independent judges evaluated only the final written products and did not participate in the writing task. Finally, we altered our analysis plans when the assumption of homogeneity of variance in a between-subjects design was not supported. In those cases, rather than proceeding with the standard parametric tests as originally planned, we employed appropriate robust statistical procedures (e.g., Welch's T).

Participants

Students at a large public university in the Southwestern United States participated in this study to earn credit for their Psychology 101 class in the Spring semester of 2025. After signing up and providing consent, participants completed an online Qualtrics survey, which took approximately 30 minutes. A total of 351 students participated in this study. Thirty-one participants were excluded because they did not answer any of the questionnaire items. The final sample size was 320, with ages ranging from 18 to 44, M = 19.55, SD = 2.39. The sample was 53.8% female, 45.8% male, and 0.3% chose not to say. The sample was 40.6% Caucasian (n = 130), 24.7% Latino (n = 79), 13.1% East Asian (n = 42), and 21.6% other (n = 69). Notably, 41.6% of participants were pursuing a business degree (n = 133). Sample size by the eight conditions ranged from 37 to 42.

Measures

Attribution of Ownership

Attribution of ownership was measured using a 7-point Likert item: "Please determine who owns the response." Participants in the independent judge conditions selected from the response options with low scores indicating greater AI ownership and high scores indicated greater user ownership (i.e., ranging from: (1) *The AI/ChatGPT completely owns the response* to (7) *The participant completely owns the response*). Similarly, in all other conditions participants selected between the same response options.

Attribution of Authorship

Attribution of authorship was measured using a 7-point Likert item: "Please determine the author of this response." Again, low scores indicated greater AI authorship and high score indicated greater user authorship. Participants in the independent judge conditions selected from the seven response options (i.e., ranging from: (1) *AI/ChatGPT is the sole author* to (7) *The participant is the sole author*). In all other conditions, participants selected between the same response options.

Experience With AI LLMs

Experience with LLMs was assessed through a 5-point self-reported frequency measure: "How frequently do you use AI Large Language Models (e.g., ChatGPT, Gemini, Google's AI Overviews, etc.)?" Participants selected from the following: (1) *Never: I have never used AI LLMs*, (2) *Rarely: I have used AI LLMs at least once*, (3) *Occasionally: I use AI LLMs 1-2 times per week*, (4) *Frequently: I use AI LLMs 3-4 times per week* and (5) *Very Frequently: I use AI LLMs 5 or more times per week*.

Attitudes Toward AI

Attitudes toward AI were measured using the AI Attitude Scale (AIAS; Grassini, 2023), a four item, 7-point Likert scale. The items included: "I believe that AI will improve my life," "I believe that AI will improve my work," "I think I will use AI technology in the future," and "I think AI technology is positive for humanity." Response options ranged from (1) *strongly disagree* to (7) *strongly agree*. An average of the responses to these four items was computed to form a composite of attitudes toward AI. Internal consistency was excellent (Cronbach's $\alpha = .92$). To confirm the one factor structure of the Attitudes toward AI scale, we conducted a confirmatory factor analysis (CFA) using Maximum Likelihood (ML) estimation through IBM SPSS AMOS (V.28.0). The CFA results supported the one factor structure, $\chi^2 (2) = 4.973$, $p = .083$, CFI = .997, NFI = .994, RMSEA = .068, 90% CI [.000, .143].

Writing Level

Writing level was assessed by having participants write 5-7 sentences about something interesting they learned in their favorite class this semester. We assessed baseline writing level using EssayGrader.ai, an AI essay grading platform that provides standardized writing evaluations (Tech Nest Ventures, 2024). Automated scoring was used to reduce variability associated with human raters and provide an assessment of relative skill differences. Each response was given three separate scores for (1) clarity and coherence, (2) grammar and mechanics, and (3) engagement and style. The scores ranged from (1) *Needs Improvement*, to (4) *Excellent*. The platform added these three scores together to give each response an overall score, which was used as writing level.

Procedure

Participants accessed the survey through the university's SONA system and consented to participating in the study. After providing consent, participants read a brief section that included an introduction to the study and instructions to answer questions truthfully. Participants were randomly assigned to one of the eight conditions.

In the four *user* conditions, users/participants were presented with the following prompt: "In the United Kingdom, some food producers include traffic light nutrition labels on the front of their packages. These labels display the number of calories, fat, saturates, sugars, and salts contained in one serving of the food. Each nutrient is also color-coded to indicate their percentage of an adult's recommended intake. Green demonstrates that the percentage is low, yellow demonstrates that the percentage is moderate, and red demonstrates that the percentage is high. Should food producers in the United States be required to include these labels? Why or why not? Your answer must contain 5 to 7 sentences."

Users in the first level were instructed to complete the assignment without any AI assistance. In the second level, users were provided with a seven sentence pre-generated response from a LLM and were instructed to edit four of the sentences before submitting. In the third level, users were provided with three pre-generated responses from a LLM and were instructed to select one for submission. In the fourth level, users were provided with a seven sentence pre-generated response from a LLM and were instructed to submit it without any editing. Following the

completion of the writing task, users completed a survey assessing their attributions of authorship and ownership, their experience with LLMs, their attitudes toward AI, and their writing level.

Participants in the *independent judge* conditions received a pre-made user response, purportedly generated by a user who had completed the writing task under one of the four LLM input conditions. The complete prompts by condition are available in the supplementary materials. Importantly, all independent judges observed the same pre-made response under different conditions and did not need to answer the prompt themselves. Informed of the specific level of LLM input into the task, independent judges were instructed to assess the response. They then completed the same survey as the users, answering items that evaluated authorship and ownership of the user's completed task and the items that measured experience with LLMs, AI attitudes, and writing level.

The study was IRB approved (#STUDY00021688) and conducted in accordance with best ethical practices outlined by the American Psychological Association (APA).

Results

Table 1 provides the descriptive statistics for the study measures by condition. Table S1 in the supplementary materials provides bivariate correlations between the relevant study variables by condition. Randomization checks revealed no significant differences across the eight conditions in participant sex (χ^2 (14, N = 319) = 16.03, p = .311), age (F(7, 308) = 6.37, p = .348), race/ethnicity (χ^2 (49, N = 319) = 40.95, p = .786), or experience with AI (F(7, 312) = 1.37, p = .233). The results for each research question are discussed below.

Table 1. *Descriptive Statistics.*

	Condition: User											
	Level 1			Level 2			Level 3			Level 4		
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Authorship	40	6.75	0.81	38	3.97	1.57	40	2.25	1.74	37	1.95	1.49
Ownership	40	6.73	0.91	38	4.08	1.75	40	2.45	1.85	37	2.05	1.58
AI Use	40	2.92	0.94	38	3.34	0.99	40	3.40	1.03	37	3.32	1.16
AI Attitude	39	4.59	1.40	38	4.96	1.33	40	5.04	1.47	37	4.99	1.27
Writing	36	7.92	1.59	33	8.03	1.47	37	8.41	1.89	30	7.77	1.70
	Condition: Independent Judge											
	Level 1			Level 2			Level 3			Level 4		
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Authorship	41	5.10	1.85	40	4.18	1.34	42	3.69	1.87	42	3.38	1.91
Ownership	41	4.78	2.13	40	4.00	2.20	42	4.05	1.98	42	3.74	2.17
AI Use	41	3.20	1.03	40	3.47	1.09	42	3.45	0.92	42	3.45	0.94
AI Attitude	41	4.69	1.54	40	4.95	1.57	42	4.84	1.43	42	4.97	1.38
Writing	33	8.30	1.74	35	8.46	1.90	36	7.86	1.53	37	9.00	1.53

Note. Authorship = Attribution of authorship to AI or user; Higher scores indicate greater attribution of authorship to the user; Ownership = Attribution of ownership to AI or user; Higher scores indicate greater attribution of ownership to the user; Writing = Writing level score.

Research Question 1

Within the user condition, there were four levels of AI input: (1) user generated the response entirely themselves, (2) user edited four sentences of an LLM-generated response, (3) user selected from three LLM-generated options, or (4) user submitted an LLM-generated response without providing any input. For both attributions of authorship and ownership, Levene's test for homogeneity of variance was significant, indicating that the assumption of equal variances across groups was violated, Authorship: F(3, 151) = 8.89, p < .001; Ownership: F(3, 151) = 10.15, p < .001. As such, to test H1 and H2, we used Welch's ANOVA to compare the effect of AI input level on participants' attributions of authorship and ownership. There were significant differences in user attributions of both

ownership (Welch's $F(3, 78.34) = 150.57$, $p < .001$, $\eta^2 = .85$) and authorship (Welch's $F(3, 78.85) = 118.04$, $p < .001$, $\eta^2 = .82$) by level of AI input. See Table 2 for the Dunnett T3 post-hoc results between the levels of AI input. The results revealed that as AI input levels rose, users' perceptions of ownership and authorship decreased, though the difference between selecting from three LLM-generated options (Level 3) and submitting an LLM-generated response without editing (Level 4) was not significant. The ownership finding was in line with our hypothesis that ownership decreases as AI contributes more to the task (H1). However, the authorship finding was contrary to our hypothesis that authorship attributions do not vary by AI input (H2). We will return to this finding in the discussion.

Table 2. *Dunnett T3 Post-Hoc Results by AI Input Level for User Condition.*

AI Input	Level 1	Level 2	Level 3	Level 4
Level 1		***p* < .001**	***p* < .001**	***p* < .001**
Level 2	*p* < .001		***p* < .001**	***p* < .001**
Level 3	*p* < .001	*p* < .001		***p* = .892**
Level 4	*p* < .001	*p* < .001	*p* = .956	

Note. Dunnett T3 post-hoc results comparing attributions of authorship and ownership by AI input level. Attributions of authorship = Lower diagonal; Attributions of ownership = Upper diagonal (in **bold**).

Research Question 2

Pearson correlations were run to test H3 and determine the linear relationship between ownership and authorship. Significance was assessed based on a Bonferroni corrected p-value to account for five correlations ($p = .05/5 = .01$). Within the user conditions (i.e., all four levels of user conditions combined), participants' attributions of ownership strongly and positively predicted their attribution of authorship, $r(153) = 0.90$, $p < .001$, $R^2 = .82$. In other words, users who attributed greater ownership to AI (rather than themselves) were more likely to also attribute authorship to the AI. This finding held for all levels of AI input and was contrary to our hypothesis that users' ownership attributions do not significantly predict their authorship attributions (H3; see Table 3 for all Pearson correlations).

To explore this finding, we compared the strength of the relationships by AI input level using Fisher r-to-z transformations to determine the significance of the difference between two independent sample Pearson correlations (Fisher, 1915). With increased AI input levels, the strength of the relationship between attributions of ownership and attributions of authorship decreased. In fact, for users who were instructed not to use AI in their response (i.e., Level 1), the relationship between their attribution of ownership and authorship was significantly stronger than for users who used any amount of AI input in their response (i.e., Levels 2, 3, 4; all *p*'s were less than .001; see Table 3 for the full results). The differences between Levels 2, 3, and 4 were not significant. This finding suggests that with any amount of AI input, users may be less likely to see ownership and authorship as directly linked.

Table 3. *Strength of the Relationship Between Attributions of Ownership and Attributions of Authorship by AI Level.*

	Level 2	Level 3	Level 4
Level 1: $r(38) = .99$, $p < .001$	***z* = 7.19, *p* < .001**	***z* = 7.20, *p* < .001**	***z* = 7.73, *p* < .001**
Level 2: $r(36) = .74$, $p < .001$		$z = -0.10$, $p = .920$	$z = 0.58$, $p = .562$
Level 3: $r(38) = .75$, $p < .001$			$z = 0.68$, $p = .497$
Level 4: $r(35) = .67$, $p < .001$			

Note. Z-values represent the comparison of strength of the linear relationship (Pearson r) between attribution of ownership and attributions of authorship by AI input level using Fisher r-to-z transformations. To correct for multiple testing, we applied the Bonferroni correction, using a revised significance threshold of $p = .05/6 = .008$.

Research Question 3

To determine if these relationships held for independent judges, we tested H4, H5, and H6, by first testing the relationships between level of AI input and attributions of ownership and authorship for judges and then comparing the mean independent judges' attributions of authorship and ownership with mean users' attributions

for each level of AI input. Levene's test for homogeneity of variance was significant for authorship and not ownership, Authorship: $F(3, 161) = 3.65$, $p = .014$; Ownership: $F(3, 161) = 0.92$, $p = .435$. There were significant differences in judge's attributions of authorship (Welch's $F(3, 88.74) = 6.55$, $p < .001$, $\eta^2 = .12$) but not ownership (Welch's $F(3, 89.22) = 1.78$, $p < .001$, $\eta^2 = .03$) by level of AI input. See Table 4 for the Dunnett T3 post-hoc results between the levels of AI input. The results revealed that judges' attributions of authorship for Level 1 (no AI use) differed significantly from all other levels of AI input. Once AI use was introduced (Levels 2-4), judges' attributions of authorship did not differ across levels. This finding partially supported our hypothesis that judges' attributions of the user as the author decrease as AI contribution increases (H5). The findings for attributions of ownership were contrary to our hypothesis (H4). Regardless of level of AI input, there were no significant differences in judges' attributions of the user as the owner of the writing product.

Table 4. *Dunnett T3 Post-Hoc Results by AI Input Level for Judge Condition.*

AI Input	Level 1	Level 2	Level 3	Level 4
Level 1		***p* = .490**	***p* = .489**	***p* = .163**
Level 2	*p* = .070		***p* = .999**	***p* = .995**
Level 3	*p* = .005	*p* = .686		***p* = .982**
Level 4	*p* < .001	*p* = .174	*p* = .972	

Note. Dunnett T3 post-hoc results comparing attributions of authorship and ownership by AI input level. Attributions of authorship = Lower diagonal; Attributions of ownership = Upper diagonal (in **bold**).

Comparing user to judge attributions, results for the Levene's tests were significant at some levels of AI input indicating that the assumption for homogeneity of variance was not met. As a result, we conducted Welch's *t*-tests. Significance was assessed using a Bonferroni adjusted p value: $p = 0.05/8 = .006$. The full Welch's *t*-test results are available in Table 5. When users were instructed to use no AI in their response (i.e., Level 1), users were significantly more likely than independent judges to attribute greater levels of both ownership and authorship to themselves. At Level 2, when users edited four sentences of an AI response, that difference in attribution of both ownership and authorship disappeared. As AI input increased (i.e., Level 3 and 4 where users selected an AI response or used an AI response without editing, respectively), the trend reversed for both ownership and authorship. At those levels of AI input, independent judges attributed significantly more ownership and authorship to the user than the user did to themselves. This finding was unexpected and contrary to our hypothesis that at the highest levels of AI input, judges assign lower ownership and authorship to the user than users assign to themselves (H6). We will return to this finding in the discussion.

Table 5. *Mean Differences in Authorship and Ownership by AI Input Level and Condition (User vs. Independent Judge).*

		User		Independent Judge		*Welch's t*	*df*	*p*	Cohen's *d*
Variables	AI Input	*M*	*SD*	*M*	*SD*				
Authorship	Level 1	6.75	0.81	5.10	1.85	5.22	54.96	< .001	1.16
	Level 2	3.97	1.57	4.18	1.34	−0.61	72.82	.545	−0.14
	Level 3	2.25	1.74	3.69	1.87	−3.62	79.96	< .001	−0.80
	Level 4	1.95	1.49	3.38	1.91	−3.74	75.92	< .001	−0.83
Ownership	Level 1	6.73	0.91	4.78	2.13	5.38	54.31	< .001	1.19
	Level 2	4.08	1.75	4.00	2.20	0.18	73.75	.861	0.04
	Level 3	2.45	1.85	4.05	1.97	−3.78	79.98	< .001	−0.84
	Level 4	2.05	1.58	3.74	2.16	−3.98	74.5	< .001	−0.89

Note. Welch's *t*-test mean comparisons of authorship and ownership between users and independent judges by AI input level. Significance was assessed using a Bonferroni adjusted p value: $p = 0.05/8 = .006$. Authorship = Attribution of Authorship; Ownership = Attribution of Ownership.

To test H7 that the relationship between ownership and authorship would be similar for users' and judges' (i.e., the relationship would not be significant for either group), we first tested the relationship between ownership and authorship for judges. Similar to the findings for the users, within the independent judge condition, greater attributions of ownership to the user significantly predicted authorship attributions, $r(163) = 0.61$, $p < .001$, $R^2 = 0.37$. We then tested if condition (user versus independent judge) moderated that relationship in the full sample and at each level of AI input. The full moderation results are available in Table 6. Contrary to our hypothesis (H7), in the full sample (i.e., all levels of AI input combined), the interaction of attribution of ownership and condition was significant. Specifically, the positive relationship between attributions of ownership and authorship was stronger for the users themselves (i.e., those in the user condition) than for the independent judges (see Figure 1). To provide greater clarity on that relationship, we then tested the moderation at each level of AI. At Levels 1, 3, and 4, the interaction did not reach significance (see Table 6). The interaction was significant at Level 2 (i.e., users edited four sentences of the AI generated response). At that level of AI input, attribution of ownership was a stronger positive predictor of attribution of authorship for users than for independent judges (see Figure 2). We will return to this result in the discussion.

Table 6. *Attribution of Ownership, Condition (User vs. Independent Judge), and Their Interactions Predicting Attribution of Authorship.*

	df	*F*	*p*	R^2	Effect	*b*	*SE*	*p*	95% CI	
									LL	UL
All Levels	3, 316	191.75	<. 001	.65						
					Intercept	−1.25	0.45	.005	−2.13	−0.37
					Ownership	1.27	0.10	< .001	1.06	1.45
					Condition	1.56	0.29	< .001	0.98	2.13
					Interaction	−0.36	0.06	< .001	−0.49	−0.24
Level 1	3, 77	107.56	<.001	.81						
					Intercept	0.002	1.80	.999	−3.58	3.58
					Ownership	1.04	0.28	< .001	0.51	1.57
					Condition	0.81	0.93	.390	−1.05	2.66
					Interaction	−0.16	0.14	.278	−0.44	0.13
Level 2	3, 74	18.53	<.001	.43						
					Intercept	−0.40	1.00	.688	−2.40	1.59
					Ownership	1.02	0.23	< .001	0.57	1.47
					Condition	1.66	0.60	.007	0.48	2.85
					Interaction	−0.35	0.13	.010	−0.62	−0.09
Level 3	3, 78	27.64	<.001	.52						
					Intercept	−0.46	0.87	.603	−2.20	1.28
					Ownership	0.87	0.26	.001	0.35	1.39
					Condition	0.98	0.61	.112	−0.23	2.19
					Interaction	−0.16	0.16	.311	−0.48	0.16
Level 4	3, 75	17.91	<.001	.42						
					Intercept	−0.48	0.91	.598	−2.29	1.33
					Ownership	0.84	0.32	.011	0.19	1.48
					Condition	1.12	0.60	.065	−0.07	2.31
					Interaction	−0.20	0.19	.278	−0.57	0.17

Note. Moderated multiple regression analysis of the interaction between attributions of ownership and condition predicting attributions of authorship. Interaction = Interaction of Ownership and Condition (User versus Independent Judge).

Figure 1. *All AI Levels: Interaction Between Attribution of Ownership and Condition (User Versus Independent Judge) Predicting Authorship.*

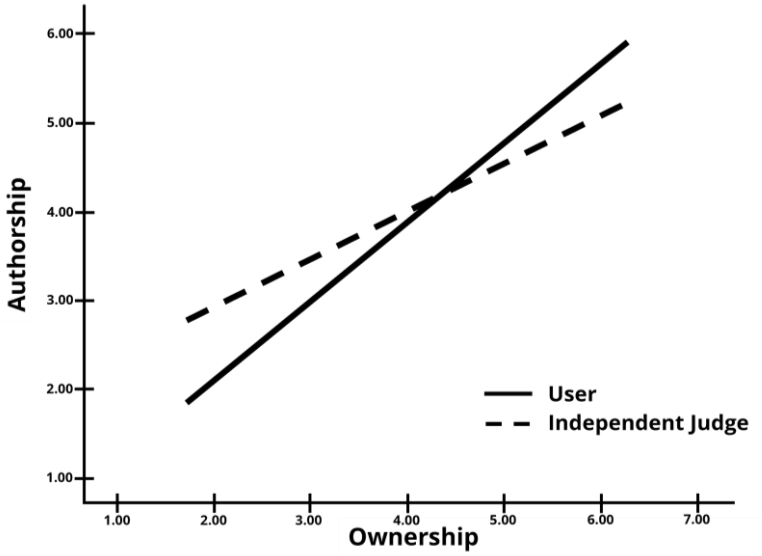
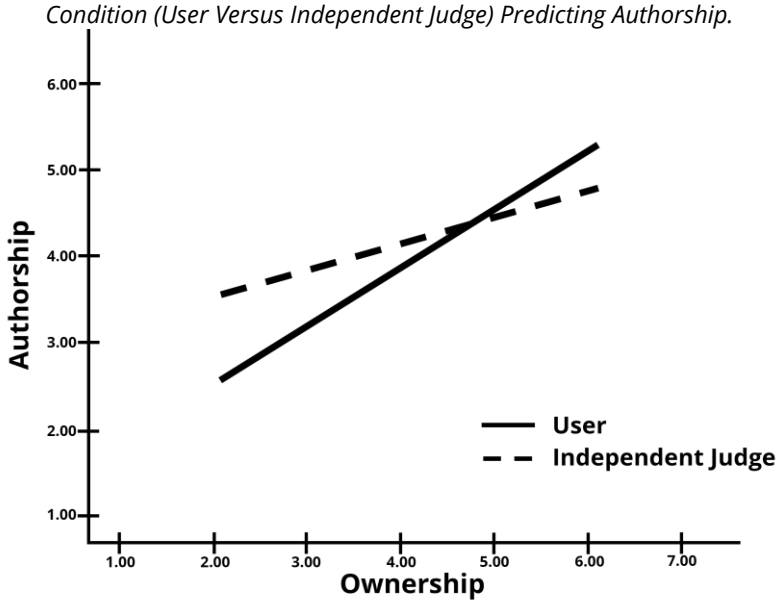


Figure 2. *All Users: Interaction Between Attribution of Ownership and Condition (User Versus Independent Judge) Predicting Authorship.*



Research Question 4

To test H8 and H9 and investigate whether users' experience with AI, attitudes toward AI, and writing proficiency moderated the relationship between (1) AI input level and attributions of ownership and (2) attributions of ownership and authorship, we conducted moderation analyses within the user condition. The full regression results are available in Tables S2 and S3 in the supplementary materials.

Contrary to our hypothesis (H8), experience with AI, attitudes toward AI, and writing proficiency did not predict users' attributions of ownership, Experience with AI: $r(152) = -.07$, $p = .388$, $R^2 = .01$, Attitudes Toward AI: $r(153) = -.15$, $p = .060$, $R^2 = .02$, Writing Proficiency: $r(134) = -.09$, $p = .303$, $R^2 = .01$. Considering the relationship between AI input level and attributions of ownership, a user's experience with AI, their attitude toward AI, and their writing level did not significantly moderate the relationship between the extent of AI input and how the user attributed ownership (see Table S2).

Considering H9 and the relationship between ownership and authorship attributions, within the user conditions, both experience with AI and attitudes toward AI were significant moderators. This was true at all levels of AI input, *p*'s ranged from <.001 to .044 (see Table S3 in the supplementary material). At Level 1 (i.e., users were instructed not to use AI), for users with lower levels of AI experience attributions of ownership were more predictive of attributions of authorship when compared to users with high levels of AI experience (see Figure S2 in the

supplementary material). Contrary to our hypothesis that this relationship would hold for all levels of AI input, as AI input increased (i.e., Levels 2-4), that moderation changed directions. At higher levels of AI input, users with greater levels of AI experience were more likely to have a strong association between attributions of ownership and authorship (see Figure S3-5 in the supplementary material). This same pattern of moderations held for attitudes toward AI (see Figure S6-9 in the supplementary material). Contrary to our hypothesis, writing level did not significantly moderate the relationship between attributions of ownership and authorship.[2]

Discussion

This research investigated how the level of AI input in a collaborative writing task influenced users' attributions of ownership and authorship. Our ownership results were largely consistent with our hypothesis and supported the previous research findings (Draxler et al., 2024). There was a significant inverse relationship between the level of AI input and users' perceptions of self-ownership, suggesting that as the AI held a more significant role in the creation of the response, users perceived their own involvement as less central and reported a diminished sense of personal ownership. This aligns with prior research (Lehmann et al., 2022; Pierce et al., 2003) suggesting that feelings of ownership are related to one's perceptions of control, effort, and investment.

Our pattern of findings for authorship attributions mirrored the ownership findings. Namely, as level of AI input increased, users were more likely to attribute authorship to the AI. This was contrary to our hypothesis and some of the prior research findings that suggested that users were more likely to attribute authorship to themselves regardless of level of AI input (i.e., the ghostwriter effect; Draxler et al., 2024). One potential explanation for this discrepancy is the nature of the writing task. Draxler et al. (2024) asked participants to write a personal postcard to a friend, a context in which individuals may be more likely to report themselves as the author regardless of the level of AI. In contrast, the current study asked participants to write a short opinion piece — a less personal writing task—that may have prompted participants to place greater emphasis on the source of the text itself when making authorship attributions. Our finding may align with previous research suggesting that users attribute greater authorship to themselves when they are more involved in actively implementing AI's suggestions than if they used AI continuously generated text (Lehmann et al., 2022). That study used two writing tasks (a birthday email to a friend and a short story about a vacation) but did not focus on if authorship attributions differed significantly by task (Lehmann et al., 2022). Note that this hypothesis (H2) was not preregistered. Future research should replicate these findings and continue to explore how ownership and authorship are attributed in different types of human-AI collaborative writing.

Importantly, while self-ownership and authorship decreased with increasing LLM input, the difference between the highest two levels of LLM input (i.e., Levels 3 and 4) was not significant. This finding was also observed in Draxler et al. (2024), suggesting that there is a generalizable floor effect where minimal user contribution leads to similarly low ownership and authorship perceptions.

The Relationship Between Ownership and Authorship

Considering the relationship between attributions of ownership and authorship, the findings revealed a strong, positive relationship between users' attributions of ownership and authorship. This finding was contrary to our hypothesis and indicates that users generally aligned their attributions of who authored writing with who owns it. This was true at all levels of AI input. However, in line with previous research, with increased AI input, the strength of the relationship between attributions of ownership and attributions of authorship decreased (Draxler et al., 2024). For users who were instructed not to use AI in their response (i.e., Level 1), the relationship between their attribution of ownership and authorship was significantly stronger than for users who used any amount of AI input in their response (i.e., Levels 2, 3, 4). The differences in that relationship between Levels 2, 3, and 4 were not significant, suggesting that a less systematic relationship (i.e., more individual differences) existed between how users attributed their authorship and ownership with AI input in a collaborative task.

Comparing Users and Independent Judges

This research aimed to extend the research by understanding these relationships in independent judges and how those findings compare to the users themselves. An important consideration when interpreting these findings is that the users and independent judges evaluated the task under meaningfully different informational conditions.

Independent judges evaluated ownership and authorship based on the understanding that the user chose to use AI, created their own prompt, and determined how to edit or incorporate the AI's output. In contrast, the users knew that their AI decisions were determined by the experimental design. As a result of this difference, Judges may have inferred greater control on the part of the user, while users' authorship and ownership attributions were informed by their understanding of their more limited decision-making. Accordingly, the differences between users and judges presented below should not be interpreted solely as differences in how ownership and authorship were evaluated by self versus other, but also as reflecting differences in perceived control that was a result of the study design. Future research could manipulate perceived control over AI use directly to disentangle these effects.

To begin to understand the relationship between user and independent judge credit allocation, we compared independent judges' attributions of authorship and ownership with users' attributions at each level of AI input. When users were instructed to generate the response without using AI (i.e., Level 1), users attributed greater authorship and ownership to themselves compared to independent judges. This pattern is consistent with attribution theory, which suggests that individuals often interpret outcomes differently depending on whether they are evaluating their own actions or those of others (see Kelley & Michela, 1980). In particular, the findings may reflect a self-serving bias, with users attributing greater credit to themselves because they are more motivated to view their contributions positively (Greenberg et al., 1982; Miller & Ross, 1975). At this level, it is noteworthy that independent judges attributed a noticeable degree of ownership to the AI, despite being instructed that the response was generated entirely by the user. In fact, regardless of level of AI input (Level 1-4), there were no significant differences in judges' attributions of the user as the owner of the writing product. One possibility is that independent judges were influenced by introducing the concept of AI and believed that users may have used AI despite instruction not to do so or that users relinquish some degree of ownership whenever they are using an AI application. Interestingly, at Level 2, where users edited 4 sentences of the AI output, there was a convergence in how both the users and independent judges perceived the contribution and resulting credit, as both authorship and ownership were not significantly different. From an attribution theory perspective, this convergence may occur because the observable contributions of both the user and the AI are more balanced at this level, reducing the influence of self-serving bias and leading users and independent judges to reach similar conclusions about how credit should be allocated.

However, as AI input increased (i.e., Level 3 and 4 where users selected an AI response or used an AI response without editing respectively), judges attributed significantly more authorship and ownership to the users than the users reported themselves. Judges' attributions of authorship were significantly higher for no AI use (Level 1) than at all other levels of AI input, but once AI use was introduced (Levels 2-4), judges' attributions of authorship did not decrease significantly even as AI input increased. This could indicate that judges focus on AI's lack of autonomy, believing that the user is ultimately responsible for prompting the creation of a shared work, which could lead them to overestimate the user's contribution when the AI's role is substantial. When combined with the findings that users perceived a strong link between ownership and authorship, this discrepancy suggests that users might feel less control and less invested in work they perceive as mostly AI-generated. This finding may also reflect judges' perceptions of greater user control in their decisions to use AI. Due to the study design, users knew that they had no part in creating the AI content (e.g., they did not write the prompt for the AI themselves), while judges were not aware of that lack of control and, as a result, may attribute greater ownership and authorship to the user. From an attribution theory perspective, judges may attribute responsibility to the user because they view the user as the intentional agent who chose to employ the AI, even when the AI generated most of the content. In contrast, users have direct knowledge of the limited extent of their own contribution and therefore may attribute less authorship and ownership to themselves. Users' reduced authorship attributions are consistent with prior research on human–AI teams demonstrating that proactively generated AI contributions (i.e., without human prompting) are more likely to be attributed authorship credit (He et al., 2025).

Notably, in the full sample (i.e., all AI input levels combined) the relationship between authorship and ownership was weaker for independent judges than for users. This suggested that independent judges differentiated between ownership and authorship to a greater extent than users. To better understand that finding, we tested that relationship in each of the AI input levels individually. The interaction was significant only at Level 2 (i.e., users edited four sentences of the AI generated response).[3] This suggests that when users use AI but make a relatively significant effort to collaborate, attribution of ownership was a stronger positive predictor of attribution of authorship for users than for independent judges. This finding could be evidence that the factors independent judges use to distribute ownership, such as the perceived pivotality of an individual's contribution (Zultan et al.,

2012), led to different attribution patterns compared to users. Independent judges, without firsthand involvement in the writing collaboration, may have based their authorship attributions on the users' role in initiating the task, while their ownership attributions were more influenced by the perceived contribution of the AI to the final text. This weaker correlation suggests that external observers may consider ownership and authorship more independently than creators assisted by AI, an important distinction for informing policy. This may be particularly important given the differences in strength of this finding by level of AI input. For example, in legal and educational applications, independent judges are more likely not to have the benefit of knowing the extent to which user's employed AI. Future research should continue to explore the factors that may differentiate users' and independent judges' attributions of ownership and authorship and the strength of the relationship between those two types of credit allocation.

Moderating Role of AI Experience, Attitudes, and Writing Proficiency

The final aim of our research was to investigate whether prior experience with AI, attitudes toward AI, and writing proficiency moderated the relationship between AI input level and attributions of ownership and/or the relationship between ownership and authorship attributions. The results suggested that none of these three variables were significant predictors of ownership for users or significant moderators in the relationship between AI input level and how users attributed ownership. This indicates that the relationship between the level of AI input and participants' perceptions of ownership and authorship was relatively consistent across individuals with varying levels of AI experience, attitudes towards AI, and writing proficiency. While prior research has indicated that these factors can influence human-AI interactions (Bansal et al., 2021; Kumar et al., 2025; Zhang et al., 2022), their effect on the attribution of ownership in this collaborative writing task was not significant. Future research should consider other potential factors that may account for individual differences in how users attribute ownership at varying levels of AI input (e.g., perception of contribution and control; demographic factors).

Considering the relationship between ownership and authorship attributions, both users' experience with AI and their attitudes toward AI were significant moderators at all levels of AI input. This finding built on past research indicating that positive attitudes toward AI increased willingness to adopt AI, predicted greater user trust in AI, and improved team performance (Bansal et al., 2021; De Visser Ewart et al., 2020; Kumar et al., 2025; Prentice et al., 2020; Zahlan et al., 2023; Zhang, 2022). Interestingly, when users were instructed not to use AI, users with less experience with AI, and with less positive attitudes toward AI, showed a stronger relationship between ownership and authorship compared to those with greater AI experience and more positive AI attitudes. Even without direct input from AI, users with greater experience with AI and more positive attitudes toward AI might consider the implications of the terms of service within an AI platform in their response to authorship and ownership.

At increased levels of AI (i.e., Levels 2-4), the direction of the moderations changed. Users with greater levels of AI experience and more positive attitudes were more likely to have a strong association between attributions of ownership and authorship. This was in line with our original hypothesis and may suggest that users with more experience with AI and more positive attitudes toward AI may perceive a more direct link between crediting ownership and authorship. In other words, if the user is willing to attribute ownership to the AI they may also be willing to attribute authorship to it. Note that this hypothesis (H9) was not preregistered. Future research should seek to replicate these findings.

Notably, writing level did not significantly moderate the relationship between attributions of ownership and authorship. This finding in combination with the finding that writing level also did not moderate the relationship between AI level and ownership, may suggest that the relative benefit of using AI may not be a crucial factor in determining users' attributions of credit in human-AI collaboration.

Implications

The findings of this study may have implications for the design in educational settings. The advancement of AI-integrated classroom has made it challenging to determine "who did the assignment." Finding the right balance for AI in education is not only about whether to allow students to use it or not, but also about ensuring intellectual integrity. The identification of a point of agreement when users edited 4 of the 7 sentences of the LLM-generated output (e.g., Level 2 in this study), where users and judges showed similar attributions of ownership and authorship, suggests a shared level of understanding of authorship. Thus, writing tasks where students are allowed to edit and refine AI-generated content might serve as a reference level of AI integration that affords

students' perceived authorship and ownership while aligning with the external evaluations of their instructors. Importantly, this is a preliminary finding that warrants further theoretical development and empirical investigation. The ownership and authorship ratings in this study were based on the participants' lay judgments. As a result, the current research cannot directly draw conclusions about the optimal level of AI integration in education. Future research should continue to work toward that goal.

From a policy perspective, the weaker correlation between ownership and authorship for judges and the varying attribution tendencies for users and judges at both high and low AI input levels highlights the complexity of assigning credit in AI-generated content. These findings may have implications across a wide range of contexts in which AI-assisted writing is produced, including education, publishing, journalism, marketing, software development, and other creative industries. Additionally, these findings could raise concerns for intellectual property policymakers. For instance, the results showing that external judges attribute more credit to users than they claim themselves in high AI input collaborations could lead to legal ambiguities. To address these challenges, policymakers may need to consider new legal frameworks that reflect the realities of human-AI collaboration in creative works. This could involve quantifying and determining a level of human input required for legal ownership.

Limitations and Future Directions

While the findings of the present study are informative, there are several limitations that should be considered when interpreting the findings and planning future directions. First, in educational settings, a current focus is traceable AI contribution. However, level of AI contribution is challenging to define objectively and quantitatively. As an initial attempt, the present research assessed how students lay understanding of authorship and ownership in human-AI collaborative tasks. By allowing participants to define these terms and make attributions based on their understanding we assessed their lay understanding of these concepts. Nevertheless, this conceptual distinction between authorship and ownership was not operationalized with sufficient clarity for participants, potentially confounding results. Future research should examine individual differences in understanding these concepts and how lay understanding of these concepts are different from legal definition of authorship and ownership.

Second, our experimental methods using an artificial AI writing task may open a question about the external validity of our findings. As an initial test of the relationship between levels of AI involvement and user's attribution of authorship/ownership in human-AI collaborative tasks, we presented participants with the same AI generated content based on the prompt created by the authors. When two people ask an AI the same question, they rarely get identical results. Tiny differences in how a question is presented—even if the core intent is the same—can create huge and unpredictable differences in the AI output (i.e., digital butterfly effect; Ferrara, 2024). By standardizing the procedure, we established a controlled baseline (e.g., controlling for difference in the prompt used). This isolation of variables was necessary to confirm the causal relationship before introducing the other factors present in ecological settings. Nevertheless, users usually create prompts themselves in real-life situations. These experimental manipulations are inherently artificial, our findings may not translate to the complex and dynamic context of AI-human collaborative tasks. Thus, the external validity of the present findings remains to be tested. Future research should test if our laboratory findings hold up under natural conditions and incorporate diverse data sources (e.g., natural behavioral data; multi-rater data). Relatedly, our sample was drawn from a single undergraduate population. As such, the results may not generalize to other populations (e.g., other schools, diverse education levels, or professional domains), where past experience with AI and understanding of the concepts of authorship and ownership may differ. Future research should examine more diverse populations (e.g., graduate students; professional and expert populations) to determine if the patterns in these findings replicate in different settings.

Third, normal human-AI interactions are more nuanced than the study portrayed. While typical user-AI interactions involve a cycle of prompting, reviewing, and editing, the study provided participants with pre-generated AI responses across different levels to limit confounds. This approach, while controlled, did not fully capture the interactive nature of many real-world AI writing collaborations and may introduce a question of external validity. Future studies can utilize additional types of tasks that tease apart the extent to which the user perceives the task itself is collaborative with an AI program. Additionally, following this study's findings, integrating a controlled LLM program into the study can allow the users themselves to engage in the continued cycle of

prompting, reviewing, and editing, and opens an avenue for researchers to observe and control for user behavior with the LLM program when prompted with a specific collaborative task.

Finally, we assessed participant writing level using an automated essay scoring system rather than a trained coding team (EssayGrader.ai; Tech Nest Ventures, 2024). Although automated scoring offers several advantages (e.g., scalability; objectivity), it may prioritize surface-level writing quality over other factors (e.g., creativity; argument quality). Future research should incorporate human raters to validate automated scoring systems.

To broaden the generalizability of this study's findings and gain a more comprehensive understanding of human-AI collaboration, future research should explore user interaction with AI in greater depth and in different contexts (e.g., in the workforce). While this study used pre-generated AI outputs to ensure consistency across conditions, future research could allow users to prompt and refine AI generated text throughout the writing process. Such studies could track the evolution of the written output with the participants' changing sense of contribution and control. This would allow researchers to observe how different prompting strategies and the nature of user revisions influence perceptions of credit attribution in a more natural collaborative setting. Future research should also test these findings and extensions in different contexts to determine if the findings can generalize beyond the education focus of the current study to other human-AI collaborative writing (e.g. journalism, marketing).

Furthermore, future research differentiating between specific forms of AI writing assistance, such as idea generation, writing, and editing, would be beneficial. This would help determine whether certain types of AI contributions elicit stronger or weaker attributions of ownership and authorship compared. Additionally, such research could offer insight into which stages of the writing process are most critical for AI-assisted writers to feel a sense of control and investment in the final work. While our study was not designed to evaluate learning outcomes or identify an optimal level of AI involvement for educational purposes, from a learning perspective, cognitive engagement is necessary for students to meaningfully process and learn from AI-assisted writing (Chi & Wylie, 2014). There may be educational benefits to encouraging levels of AI involvement that require students to actively evaluate and integrate AI-generated content rather than simply accepting it with minimal reflection. Education research may wish to extend the current research to consider not only how AI involvement affects perceptions of authorship, but also whether particular forms of AI-assisted writing promote the cognitive engagement necessary for learning.

Another potential avenue for future research is the role of AI transparency, specifically the explicit disclosure of the extent of AI contribution. Future research could manipulate transparency levels to assess whether informing users and independent judges about the degree of AI involvement alters how both perceive credit distribution. Understanding the impact of disclosure could have significant implications for ethical AI usage, particularly in the realm of creating intellectual works, where authorship and ownership attribution holds legal and reputational consequences.

Footnotes

[1] Preregistration link: https://archive.org/details/osf-registrations-j62kp-v1

[2] In the user condition at Level 1, there was not sufficient variability in the attribution of ownership and authorship variables to run the moderation analysis for the writing skill variable.

[3] Although the moderations at the other AI input levels were not significant, they were trending in the same direction.

Conflict of Interest

The authors do not have any conflicts of interest to report.

Use of AI Services

The authors declare they have not used any AI services to generate any part of the manuscript or data. AI was utilized as part of the experimental methodology to test peoples' perception of AI collaboration.

Data Availability Statement

The data used in this study can be accessed at the following: https://osf.io/8374u/overview

Authors' Contribution

Joseph Fung: conceptualization, data curation, formal analysis, investigation, methodology, resources, writing—original draft. **Samantha McMichael:** formal analysis, project administration, software, supervision, validation, visualization, writing—review & editing, revision. **Kamryn Kadotani:** conceptualization, investigation, methodology, project administration, supervision, writing—review & editing. **Virginia Kwan:** conceptualization, project administration, supervision, writing—review & editing.

References

Allen, K. R., Jara-Ettinger, J., Gerstenberg, T., Kleiman-Weiner, M., & Tenenbaum, J. B. (2015). Go fishing! Responsibility judgments when cooperation breaks down. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 84-85). Cognitive Science Society. https://cicl.stanford.edu/publication/allen2015fishing/

Allison, J., Hwang, G., Mayer, R. E., Pellas, N., Karnalim, O., De Freitas, S., Ng, O., Huang, Y., Hooshyar, D., Seidman, R. H., Al-Emran, M., Mikropoulos, T. A., Schroeder, N. L., Roscoe, R. D., & Sanusi, I. (2025). From generative AI to extended reality: Multidisciplinary perspectives on the challenges, opportunities, and future of educational computing. *Journal of Educational Computing Research, 63*(6), 1327–1363. https://doi.org/10.1177/07356331251359964

Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: A review. *Artificial Intelligence Review, 55*(1), 589–656. https://doi.org/10.1007/s10462-021-10039-7

Arif, T. B., Munaf, U., & Ul-Haque, I. (2023). The future of medical education and research: Is ChatGPT a blessing or blight in disguise? *Medical Education Online, 28*(1), Article 2181052. https://doi.org/10.1080/10872981.2023.2181052

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M. T., & Weld, D. S. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary team performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Article 81). Association for Computing Machinery. https://doi.org/10.1145/3411764.3445717

Bikard, M., Murray, F., & Gans, J. S. (2015). Exploring trade-offs in the organization of scientific work: Collaboration and scientific reward. *Management Science, 61*(7), 1473–1495. https://doi.org/10.1287/mnsc.2014.2052

Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education, 18*(1), 198–204. https://doi.org/10.5281/zenodo.7716416

Bozkurt, A. (2024). GenAI et al.: Cocreation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis, 16*(1), 1–10. https://doi.org/10.55982/openpraxis.16.1.654

Chechitelli, A. (2024, April 9). *Turnitin celebrates one year for AI writing detection.* Turnitin. https://www.turnitin.com/blog/turnitin-celebrates-one-year-for-ai-writing-detection

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chui, M., Manyika, J., & Miremadi, M. (2016). Where machines could replace humans—and where they can't (yet). *The McKinsey Quarterly*, 1–12. https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/where-machines-could-replace-humans-and-where-they-cant-yet

Cooray, S. (2026). Socio-technical machine learning – Putting the user in control. *Journal of Computer Information Systems*, 1–26. https://doi.org/10.1080/08874417.2025.2600955

De Visser, E. J., Peeters, M. M. M., Jung, M. F., Kohn, S., Shaw, T. H., Pak, R., & Neerincx, M. A. (2020). Towards a theory of longitudinal trust calibration in human–robot teams. *International Journal of Social Robotics, 12*(2), 459–478. https://doi.org/10.1007/s12369-019-00596-x

Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024). Shaping human-AI collaboration: Varied scaffolding levels in co-writing with language models. *Proceedings of the CHI Conference on Human Factors in Computing Systems, USA,* (pp. 1-18). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642134

Draxler, F., Werner, A., Lehmann, F., Hoppe, M., Schmidt, A., Buschek, D., & Welsch, R. (2024). The AI ghostwriter effect: When users do not perceive ownership of AI-generated text but self-declare as authors. *ACM Transactions on Computer-Human Interaction, 31*(2), Article 25. https://doi.org/10.1145/3637875

Dreyfuss, R. C. (2000). Collaborative research: Conflicts on authorship, ownership, and accountability. *Vanderbilt Law Review, 53*(4), 1159–1230. https://scholarship.law.vanderbilt.edu/vlr/vol53/iss4/2/

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association, 56*(293), 52–64. https://doi.org/10.1080/01621459.1961.10482090

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology, 56*(2), 489–530. https://doi.org/10.1111/bjet.13544

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Ferrara, E. (2024). The butterfly effect in artificial intelligence systems: Implications for AI bias and fairness. *Machine Learning with Applications, 15*, Article 100525. https://doi.org/10.1016/j.mlwa.2024.100525

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika, 10*(4), 507–521. https://doi.org/10.2307/2331838

Fu, Y., Weng, Z., & Wang, J. (2025). Examining AI use in educational contexts: A scoping meta-review and bibliometric analysis. *International Journal of Artificial Intelligence in Education, 35*(3), 1388–1444. https://doi.org/10.1007/s40593-024-00442-w

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Grassini, S. (2023). Development and validation of the AI attitude scale (AIAS-4): A brief measure of general attitude toward artificial intelligence. *Frontiers in Psychology, 14*, Article 1191628. https://doi.org/10.3389/fpsyg.2023.1191628

Green, B., & Chen, Y. (2019). Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 90–99). Association for Computing Machinery. https://doi.org/10.1145/3287560.3287563

Greenberg, J. L., Pyszczynski, T., & Solomon, S. (1982). The self-serving attributional bias: Beyond self-presentation. *Journal of Experimental Social Psychology, 18*(1), 56–67. https://doi.org/10.1016/0022-1031(82)90081-6

He, J., Houde, S., & Weisz, J. D. (2025). Which contributions deserve credit? Perceptions of attribution in human-AI co-creation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. https://doi.org/10.1145/3706598.3713522

Hemmer, P., Schemmer, M., Vössing, M., & Kühl, N. (2021). Human–AI complementarity in hybrid intelligence systems: A structured literature review. *PACIS 2021 Proceedings*. https://aisel.aisnet.org/pacis2021/78/

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kasunic, R. J. (2023, February 21). *Re: Zarya of the Dawn (Registration # VAu001480196)*. U.S. Copyright Office. https://www.copyright.gov/docs/zarya-of-the-dawn.pdf

Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology, 31*(1), 457–501. https://doi.org/10.1146/annurev.ps.31.020180.002325

Kumar, R., Singh, A., Ahmed Kassar, A. S., Ismail Humaida, M., Joshi, S., & Sharma, M. (2025). Unlocking the power of AI: Healthcare workforce perception and its impact on their work performance in Saudi Arabia. *Pakistan Journal of Medical Sciences, 41*(3), 682–686. https://doi.org/10.12669/pjms.41.3.11014

Lagnado, D. A., Gerstenberg, T., & Zultan, R. (2013). Causal responsibility and counterfactuals. *Cognitive Science, 37*(6), 1036–1073. https://doi.org/10.1111/cogs.12054

Lehmann, F., Markert, N., Dang, H., & Buschek, D. (2022). Suggestion lists vs. continuous generation: Interaction design for writing with generative models on mobile devices affect text length, wording and perceived authorship. In *Mensch und Computer 2022* (pp. 192–208). Association for Computing Machinery. https://doi.org/10.1145/3543758.3543947

Li Yunkai v. Liu Yuanchun, (2023) Jing 0491 Min Chu No. 11279 (Beijing Internet Ct. Nov. 27, 2023). https://english.bjinternetcourt.gov.cn/2023-12/28/c_688.htm

Liu, H., Lai, V., & Tan, C. (2021). Understanding the effect of out-of-distribution examples and interactive explanations on human-AI decision making. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), Article 408. https://doi.org/10.1145/3479552

Liu, X., Guo, B., He, W., & Hu, X. (2025). Effects of generative artificial intelligence on K-12 and higher education students' learning outcomes: A meta-analysis. *Journal of Educational Computing Research, 63*(5), 1249–1291. https://doi.org/10.1177/07356331251329185

Lockwood, A. B., & Castleberry, J. (2024). Examining the capabilities of GPT-4 to write an APA-style school psychology paper. *Contemporary School Psychology, 29*(1), 145–152. https://doi.org/10.1007/s40688-024-00500-z

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin, 82*(2), 213–225. https://doi.org/10.1037/h0076486

Pennanen, N., Kanerva, P., & Guckelsberger, C. (2026). How prompting and editing shape psychological ownership in AI-assisted creative writing. *Computers in Human Behavior: Artificial Humans, 9*, Article 100381. https://doi.org/10.1016/j.chbah.2026.100381

Pierce, J. L., Kostova, T., & Dirks, K. T. (2003). The state of psychological ownership: Integrating and extending a century of research. *Review of General Psychology, 7*(1), 84–107. https://doi.org/10.1037/1089-2680.7.1.84

Pratama, A. R. (2025). The accuracy-bias trade-offs in AI text detection tools and their impact on fairness in scholarly publication. *PeerJ Computer Science, 11*, Article e2953. https://doi.org/10.7717/peerj-cs.2953

Prentice, C., Weaven, S., & Wong, I. A. (2020). Linking AI quality performance and customer engagement: The moderating effect of AI preference. *International Journal of Hospitality Management, 90*, Article 102629. https://doi.org/10.1016/j.ijhm.2020.102629

Rashid, A. B., & Kausik, A. K. (2024). AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications. *Hybrid Advances*. https://doi.org/10.1016/j.hybadv.2024.100277

Savitsky, K., Adelman, R. M., & Kruger, J. (2012). The feature-positive effect in allocations of responsibility for collaborative tasks. *Journal of Experimental Social Psychology, 48*(3), 791–793. https://doi.org/10.1016/j.jesp.2011.12.008

Schleiger, E., Mason, C., Naughtin, C., Reeson, A., & Paris, C. (2024). Collaborative intelligence: A scoping review of current applications. *Applied Artificial Intelligence, 38*(1), Article 2327890. https://doi.org/10.1080/08839514.2024.2327890

Seeber, I., Bittner, E., Briggs, R. O., De Vreede, T., De Vreede, G., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management, 57*(2), Article 103174. https://doi.org/10.1016/j.im.2019.103174

Simone, D. (2019). *Copyright and collective authorship: Locating the authors of collaborative work*. Cambridge University Press. https://doi.org/10.1017/9781108186070

Tech Nest Ventures. (2024). *Essay grader*. https://www.essaygrader.ai/

Thaler v. Perlmutter, No. 23-5233 (D.C. Cir. Mar. 18, 2025). https://law.justia.com/cases/federal/appellate-courts/cadc/23-5233/23-5233-2025-03-18.html

Tu, Y. F., & Lu, Y. C. (2025). Trends of generative AI applications in educational settings. *International Journal of Mobile Learning and Organisation, 19*(4), 442–467. https://doi.org/10.1504/IJMLO.2025.149024

Wilder, B., Horvitz, E., & Kamar, E. (2020). Learning to complement humans. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence* (pp. 1526–1533). https://doi.org/10.48550/arXiv.2005.00582

Xiang, Y., Landy, J., Cushman, F. A., Vélez, N., & Gershman, S. J. (2023). Actual and counterfactual effort contribute to responsibility attributions in collaborative tasks. *Cognition, 241*, Article 105609. https://doi.org/10.1016/j.cognition.2023.105609

Xiang, Y., Landy, J., Cushman, F. A., Vélez, N., & Gershman, S. J. (2025). People reward others based on their willingness to exert effort. *Journal of Experimental Social Psychology, 116*, Article 104699. https://doi.org/10.1016/j.jesp.2024.104699

Yang, L., Xin, T., Yu, Y., & Wu, Y. (2026). Human vs. LLM creativity: A comparative analysis of task-dependent asymmetry and linguistic mechanisms. *Journal of Intelligence, 14*(2), Article 27. https://doi.org/10.3390/jintelligence14020027

Zahlan, A., Ranjan, R. P., & Hayes, D. (2023). Artificial intelligence innovation in healthcare: Literature review, exploratory analysis, and future research. *Technology in Society, 74*, Article 102321. https://doi.org/10.1016/j.techsoc.2023.102321

Zhang, Q., Lee, M. L., & Carter, S. (2022). You complete me: Human-AI teams and complementary expertise. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (Article 114). Association for Computing Machinery. https://doi.org/10.1145/3491102.3517791

Zultan, R. I., Gerstenberg, T., & Lagnado, D. A. (2012). Finding fault: Causality and counterfactuals in group attributions. *Cognition, 125*(3), 429–440. https://doi.org/10.1016/j.cognition.2012.07.014

About Authors

Joseph M. T. Fung is a second year student at Sandra Day O'Connor College of Law pursuing a Juris Doctor Degree. He graduated from Arizona State University with his Bachelor's degree in Psychology in 2025. As an undergraduate, Joseph was a part of a research team studying Artificial Intelligence. After graduating, he intends to return to Honolulu, Hawaii to practice tax law at a full service law firm.

Samantha L. McMichael is a post-doctoral researcher in social psychology at Arizona State University. She earned her Ph.D. in Social Psychology and Master of Counseling from Arizona State University. Her research focuses on future time-perception and its relationship to behaviors and decisions.

https://orcid.org/0000-0003-0647-5982

Kamryn M. Kadotani is a Social Psychology PhD student at Arizona State University. Kamryn researches how intergenerational future oriented thinking and social resources shape decision-making in an increasingly digital and globally interconnected world. Budding lines of research investigate human-AI relationships, aiming to study its downstream impact on future generations. Kamryn is the Lab Manager for the Culture and Decision Science Network Lab at ASU where she mentors undergraduate research assistants in project management and statistical analyses.

https://orcid.org/0009-0001-4782-503X

Virginia S. Y. Kwan is a professor of social psychology in the Department of Psychology at Arizona State University. Her research focuses on the psychology of boundaries between perception of self and perception of others using multiple methods, multiple cultures, and multiple species. Virginia's recent research addresses the boundaries between perceived human and perceived nonhuman agent, and between cyberlife engagement and physical life engagement, and the impacts of these boundaries on how people attend to, understand, and interpret modern disasters, such as climate change and cyber-attacks. The ultimate goals of Virginia's research are to develop new ways to better serve people, making technology more user-friendly and safe.

Correspondence to
Virginia S. Y. Kwan, 950 S. McAllister Ave, Tempe, AZ 85287-1104, USA, virginia.kwan@asu.edu.

© Author(s). The articles in Cyberpsychology: Journal of Psychosocial Research on Cyberspace are open access articles licensed under the terms of the Creative Commons BY-SA 4.0 International License which permits unrestricted use, distribution and reproduction in any medium, provided the work is properly cited and that any derivatives are shared under the same license.