# When AI Blurs the Boundaries of Contribution: An Empirical Study of Authorship Calibration

Célina Treuillier[1], Denis Lalanne[1]

[1]Human-IST Institute, University of Fribourg, Boulevard de Pérolles 90, 1700 Fribourg, Switzerland

**Abstract**

The broad adoption of Artificial Intelligence (AI), especially Generative AI, raises pressing questions about how users interact with these systems to produce new content. In this paper, we introduce the concept of authorship calibration, defined as users' awareness of their actual authorship when interacting with AI. Using the CoAuthor dataset, we empirically examine how authorship calibration varies across users and how it relates to their frequency of AI use. Our results reveal high variability: users relying heavily on AI tend to misjudge their authorship, whereas those using AI less frequently exhibit more accurate authorship calibration. These findings suggest that AI can obscure users' perception of their own authorship. In learning contexts, miscalibration can affect metacognitive monitoring and learning strategies, ultimately impacting learning outcomes. Fostering authorship calibration then appears essential for promoting responsible and educationally meaningful AI integration.

**Keywords**

Generative AI, Learning Analytics, Metacognitive Calibration, Authorship

## 1. Introduction

The advent of sophisticated Artificial Intelligence (AI) algorithms, such as Large Language Models (LLMs), has significantly impacted our daily habits. Among the various tasks for which such generative AI models can be employed, writing stands out as one of the most prominent, as LLMs can produce human-like content rapidly and efficiently [1]. AI can then support writers by stimulating creativity, improving grammatical and orthographic accuracy, and enhancing the overall coherence and flow of a text [2]. However, these capabilities also raise fundamental questions regarding ownership and authorship [3], since parts of the output may be generated entirely by the AI without meaningful user contribution. More specifically, depending on how users interact with AI, the system can assume a range of roles, from simple spelling and grammar corrector to functioning as an undeclared ghostwriter [4].

The education field emerges as a particularly significant area of adoption of generative AI [5], with learners now relying on such models to support them in achieving educational tasks. While providing them with permanent, personalized, and on-demand assistance [6], AI integration into educational contexts also introduces important challenges, especially concerning its effects on metacognitive processes [7] that are essential for effective and autonomous learning [8]. Among metacognitive processes, *calibration*, or *metacognitive calibration*, plays an important role: it reflects the degree to which individuals' judgments about their understanding, capability, competence, or preparedness align with their actual demonstrated performance [9]. Given that AI usage in educational contexts may affect such metacognitive processes, it also can also potentially impact associated learning outcomes.

In this paper, we aim to address both the challenges related to users' authorship in AI-assisted writing tasks and the associated impact on metacognitive processes, especially in the educational context. To guide our research work, we formulate the following Research Questions (RQs):

- **RQ1**: How accurately do users evaluate their authorship when engaging in AI-assisted tasks?
- **RQ2**: Does the frequency of AI use influence authorship perception, and in what ways?

*Third International Workshop on Generative AI and Learning Analytics, 2026*
celina.treuillier@unifr.ch (C. Treuillier); denis.lalanne@unifr.ch (D. Lalanne)
https://celinatreuillier.github.io/ (C. Treuillier); https://lalanne.human-ist.ch/ (D. Lalanne)
0000-0003-1634-8856 (C. Treuillier); 0000-0001-7834-0417 (D. Lalanne)

To investigate these questions, and building on the established notion of calibration [10], we introduce the concept of **authorship calibration**, defined as users' accurate evaluation of their own contribution during an AI-assisted task. High authorship calibration indicates that users are aware of how AI is used and integrated into the final output, while accurately evaluating their own authorship. In contrast, *mis*calibration arises when users either overestimate or underestimate their actual contribution. In an educational context, as with traditional performance-related *mis*calibration, poor authorship calibration may hinder learning, potentially resulting in dissatisfaction, overconfidence, or reduced motivation.

We study authorship calibration in the context of writing tasks leveraging the CoAuthor dataset [11], and situates our findings within educational settings to discuss the possible implications for metacognitive processes and learning outcomes. This paper advances the field by (1) introducing the concept of authorship calibration in GAI environments, (2) offering insights into how GAI usage shapes this calibration, and (3) providing new evidence that accurate calibration can benefit educational contexts.

## 2. Related Work

### 2.1. Generative AI and Education

The advent of AI and its spread to the general public has highly transformed our daily lives, and the education sector is highly impacted [5]. Built upon AI algorithms that leverage existing content (text, audio, images, etc.), generative AI systems are able to automatically generate new content [12]. Their integration in educational contexts presents promising opportunities for learners to access permanent, personalized, and on-demand feedback [6], thereby potentially enhancing their overall learning experience [13]. AI is sometimes described as a transformative force capable of empowering learners and revolutionizing pedagogical practices [14]. However, beside these potential benefits, AI in education also raises challenges, especially related to the development of cognitive processes underlying learning [7, 15]. Of particular concern is the risk of reduced cognitive effort and learner engagement when AI is used for educational tasks [16] as learners can potentially become passive participants rather than active actors of their learning process.

To better understand how AI can impact learning, researchers have devoted considerable attention to identifying and describing user-AI interaction patterns. Gomez *et al.* [17] propose a taxonomy of seven human-AI collaboration patterns, differentiated by the temporal evolution of interactions and by which agent (human or AI) initiates the interaction. Complementing this framework, vanBerkel *et al.* [18] define three types of human-AI interaction based on the trigger mechanism for AI input: intermittent (explicit user request), continuous (implicit ongoing integration), and proactive (condition-dependent AI initiation). Within educational contexts, *Memarian et al.* [19] described a multidimensional taxonomy of learner-AI interaction centered on learning alignment and compatibility alignment, referring to the coherence among learning activities, learning experiences, and learning outcomes.

Beyond conceptualizing learner-AI interactions, a growing body of research examines how these interaction patterns impact learning performance. Specifically about AI-support for writing tasks, Nguyen *et al.* [20] demonstrated that learners who engage more actively in the interaction process achieve higher performances. Similarly, Yang *et al.* [21] revealed that active engagement in higher-order cognitive processes, such as critical evaluation, synthesis, and revision, is consistently associated with improved essay quality. Addressing this concern about reduced learner engagement, Arnold *et al.* [22] demonstrated that by modifying the writer-AI interaction design, AI can remain supportive without replacing the writer. However, the role AI plays for writers can vary a lot: from editor supporting writers in the reviewing of their output, to co-author involving collaborative patterns, to ideas source providing inspiration and direction, to ghostwriter generating content that human authors may not transparently declare [4]. This raises fundamental questions about authorship and ownership in AI-assisted writing tasks [23].

## 2.2. Authorship in AI-assisted Writing

The notion of authorship, traditionally referring to the state or fact of being the writer of a document or the creator of a work, becomes highly questionable when AI is integrated into the writing or creation process. As AI models now produce human-like text using powerful natural language models, considering AI as a co-author represents a legitimate conceptual question [24]. This raises complex issues regarding how to appropriately declare AI usage in a work, and what constitutes legitimate authorship in human-AI collaboration. Closely related is the concept of ownership, referring to the legal rights to control and profit from a work, which can be separate from its creator. Considering these related concepts of authorship and ownership, Draxler *et al.* [3] introduced the "AI Ghostwriter Effect", describing situations where AI users do not attribute authorship to the AI, although they attribute ownership to it. Declaration of authorship by AI users then become skewed, as they avoid acknowledging the AI's contribution.

These authorship and ownership concepts are particularly interesting in educational settings, where academic tasks may be jointly completed by learners and AI. This makes it highly difficult for teachers to evaluate knowledge acquisition and competence development as traditional assessment methods focus primarily on final products (i.e. product-based assessment) rather than the processes through which they were created (i.e. process-based assessment) [25]. Cheng *et al.* [26] develop a process-based assessment approach, where the final product is not the unique object being evaluated; instead, the interaction between learners and AI becomes equally essential. Complementing this growing assessment process, technical approaches relying on natural language processing models are able to automatically classify the work as belonging to the learner or the AI [27, 28], then informing about authorship and contribution in AI-assisted works.

Finally, an alternative approach focuses on making interaction patterns and text provenance visible [29, 30, 31]. In educational contexts, this represents the opportunity to raise learners' awareness about their engagement in the learning process. For teachers, this gives important insights that can inform evaluation process or adaptations of teaching practices. This awareness-oriented approach connects authorship concerns directly to metacognitive processes, suggesting that learners understanding their own contribution can then adapt their cognitive mechanisms to foster a more effective learning.

## 2.3. Metacognition, Self-Regulated Learning and Calibration

In the educational field, metacognition refers to learners' awareness and regulation of their own cognitive processes, encompassing the planning, monitoring, and evaluating of their learning activities [32]. Metacognition plays a crucial role by enabling learners to become more strategic, reflective, and autonomous [8]. Closely aligned with metacognition, Self-Regulated Learning (SRL) constitutes a well-established theoretical framework describing how learners actively control their own learning processes through a cyclical model involving planning, monitoring, and self-reflection phases [33, 34]. SRL has been extensively studied in educational research, with empirical evidence consistently demonstrating that it serves as a robust predictor of academic success [35]. Recently, Xu *et al.* [36] specifically examined the relationship between SRL and metacognition in AI environments, revealing that metacognitive support enhances SRL abilities, thereby improving the overall learning experience.

Within the broader metacognitive and SRL frameworks, the concept of *calibration*, or *metacognitive calibration*, represents an interesting construct [10]. Calibration reflects the degree of correspondence between individuals' subjective judgments about their understanding, capability, competence, or preparedness and their objectively measured performance across these dimensions [9]. Well-calibrated learners demonstrate accurate awareness of what they know and what they do not know, enabling effective allocation of study resources and appropriate help-seeking behaviors. Conversely, miscalibrated learners exhibit poor metacognitive awareness and may be either overconfident, overestimating their competencies, or underconfident, underestimating their competencies. Calibration is related to SRL, with calibration accuracy influencing the efficiency of self-regulatory behaviors [37]. Well-calibrated learners can accurately adapt their learning strategies throughout the SRL cycle, thereby enhancing

learning outcomes, whereas miscalibrated learners may implement ineffective strategies, resulting in suboptimal performance [10].

In today's rapidly evolving educational landscape, ensuring that learners remain aware of their own abilities and of the role AI plays in their learning is a major challenge. Yet, to the best of our knowledge, metacognitive calibration remains poorly explored within the context of AI usage in education. In particular, studies exploring how calibration relates to authorship in AI-assisted work are lacking. We therefore propose to investigate how learner–AI interactions influence what we refer to as authorship calibration, meaning a user's awareness of his own contribution during an AI-assisted task.

## 3. The Concept of Authorship Calibration

In this work, we conceptualize calibration through the lens of authorship, introducing the concept of **authorship calibration**. This authorship calibration reflects the extent to which an individual's judgment about his contribution in an AI-assisted task aligns with his actual contribution in the final output. Authorship calibration is operationalized using Equation 1.

$$\text{Authorship calibration} = \text{Declared Authorship} - \text{Actual Authorship} \tag{1}$$

Declared and actual authorship represent, respectively, the proportion of the final output that the user believes he has authored and the proportion he has actually authored. These proportion are expressed with continuous values ranging in $[0; 1]$ (e.g., if the user authored half of the output ($50\%$), the corresponding value is $0.5$). The resulting authorship calibration score ranges in $[\breve{}1; 1]$, with a score of $0$ indicating a perfect calibration: the user's perceived contribution matches his actual contribution. Negative scores reflect underestimation, meaning the user contributed more than declared; positive scores reflect overestimation, meaning the user contributed less than declared. The absolute magnitude value of the calibration score ($|$Authorship calibration$|$) indicates the degree of *mis*calibration, with larger deviations from zero corresponding to poorer calibration. Figure 1 illustrates how authorship calibration scores vary as a function of declared and actual authorship.

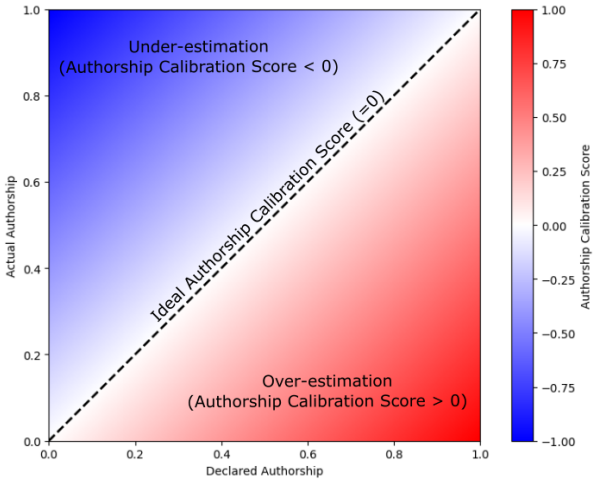



**Figure 1:** Variations in Authorship Calibration Scores as a Function of Declared and Actual Authorship

# 4. Methodology

## 4.1. Dataset

This study relies on the CoAuthor Dataset[1] [11], which includes data about 1,445 AI-assisted writing sessions from 60 authors recruited on Amazon Mechanical Turk. The dataset encompasses two writing genres: 830 creative and 615 argumentative writing sessions. Each writing session began with an initial prompt providing an overview of the assigned topic. Users could then choose to write independently (without using GPT) or request AI assistance through five GPT-generated suggestions, which they could ignore, modify, or incorporate without modifications. The dataset captures comprehensive interaction data, including GPT calls and keystroke events recording text insertions and deletions, cursor movements, and interactions with GPT suggestions.

Beyond interaction logs, the dataset provides additional resources including metadata about system configuration, and user survey responses. The metadata includes GPT parameters (temperature and frequency penalty settings) and some quantitative metrics characterizing GPT usage (i.e. number of queries, number of accepted suggestions, proportion of final text written by user, *etc.*). The survey includes five sections assessing writer demographic and background information, perceived benefits of collaborative writing, user perceptions of GPT's capabilities and limitations, and overall writing experience. The CoAuthor Dataset is particularly well-suited to our research objectives, as it combines interaction data, essential for analyzing learner-AI collaboration patterns, with survey responses that enable assessment of authorship calibration.

## 4.2. Users Categorization Based on the Frequency of AI Use

To address our second research question (RQ2), we split users into two groups based on their AI usage. Following the methodology of Shibani et al. [29], we classify users into Low- and High-AI usage groups based on the median number of AI calls. This metric offers a comprehensive indicator of interaction intensity, capturing the full range of possible engagement with AI (e.g., requesting ideation suggestions, integrating suggestions with or without modification, *etc.*). Users whose number of AI calls exceeds the median are assigned to the High-AI usage group, while those below the median are assigned to the Low-AI usage group. The full code is accessible on our github repository[2].

## 4.3. Authorship Calibration Evaluation

To operationalize our introduced concept of authorship calibration, we leverage survey responses provided in the CoAuthor dataset. Importantly, participants were asked to answer the following question: "*—% of the essay/story is written by me (and the rest is written by taking the suggestions)*". Users then estimated the percentage ($[0 - 100\%]$) of their personal contribution in the final submitted text, as distinct from content originating from GPT suggestions. The actual percentage of text written by the human author is computed by comparing the number of user-written sentences with the number of sentences originating from GPT suggestions. This percentage is provided in the CoAuthor metadata. We then compute authorship calibration using Equation 1, which indicates the extent to which users accurately assess their authorship.

# 5. Results

## 5.1. Overall Authorship Calibration Scores

From the original 1,445 writing sessions of CoAuthor, we kept only those with both interaction data and complete survey responses. Filtering sessions with missing data, 1,252 sessions remained for analysis

[1] https://coauthor.stanford.edu/

[2] https://github.com/celinatreuillier/Authorship_Calibration_CoAuthor (will be publicly available upon acceptance).

(754 creative writing sessions and 508 argumentative writing sessions). For each session, authorship calibration is assessed by comparing declared authorship against actual authorship (See Section 3).

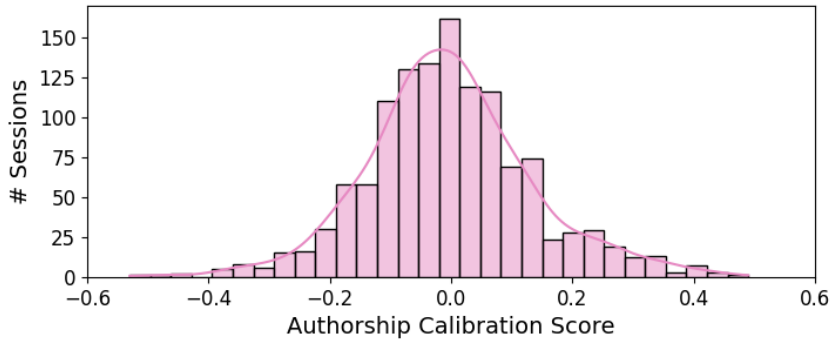


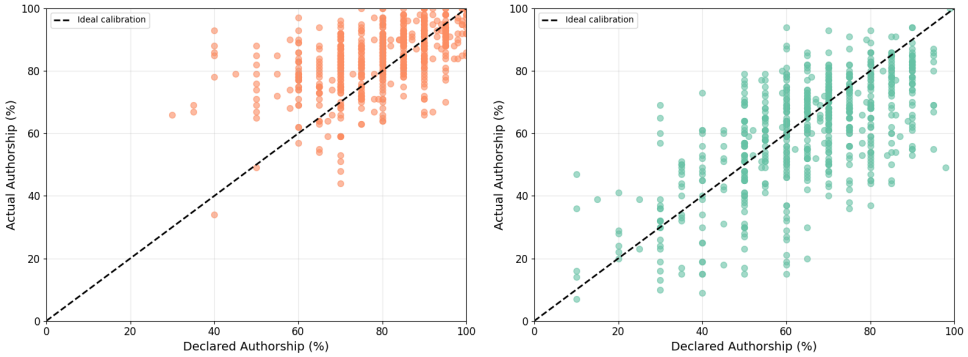

**Figure 2:** Distribution of Authorship Calibration Scores

The distribution of authorship calibration scores shows a roughly bell-shaped pattern centered around zero (mean value $= -0.003$), indicating accurate calibration for most writing sessions (See Figure 2). However, the Standard Deviation (SD) is high ($SD = 0.138$), with scores ranging from $-0.53$ to $0.49$, indicating the presence of both strongly under- and over-calibrated users. A limited skewness toward negative values is observed, suggesting a small tendency toward underestimation. Overall, these results highlight considerable variability, motivating further analysis of differences across Low- and High-AI usage groups.

## 5.2. Authorship Calibration Across Low- and High-AI Usage Groups

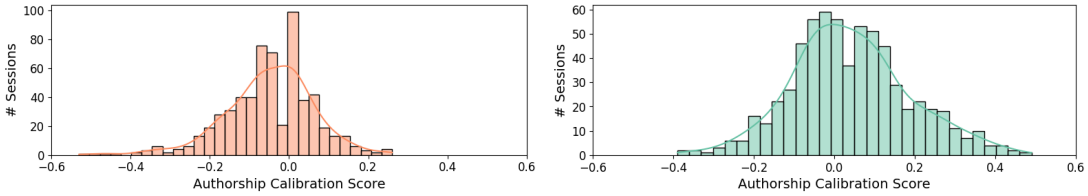
To deepen our analysis, we examine whether and how the frequency of AI use impact authorship calibration. As detailed in Section 4.2, we differentiate between Low-AI usage and High-AI usage comparing the number of calls made to AI (i.e. requests for AI-generated suggestions). In the filtered dataset of 1,251 writing sessions, the mean number of AI calls is 12.98 and varies greatly between users ($SD = 9.6$), ranging from 0 calls to 65 calls. Relying on the median value of 11 AI calls, we classified 647 sessions as High-AI usage and 605 sessions as Low-AI usage.

To compare calibration patterns across user groups, we plot calibration curves (Figure 3), where each point represents a user's declared authorship positioned against the corresponding actual authorship, and the distribution of corresponding authorship calibration scores is presented in Figure 4. To further analyze how users are distributed within the calibration space, we also provide density heatmaps in Figure 5, offering a more aggregated view of local concentrations and global patterns.




(a) Low-AI usage (b) High-AI usage

**Figure 3:** Calibration curves.

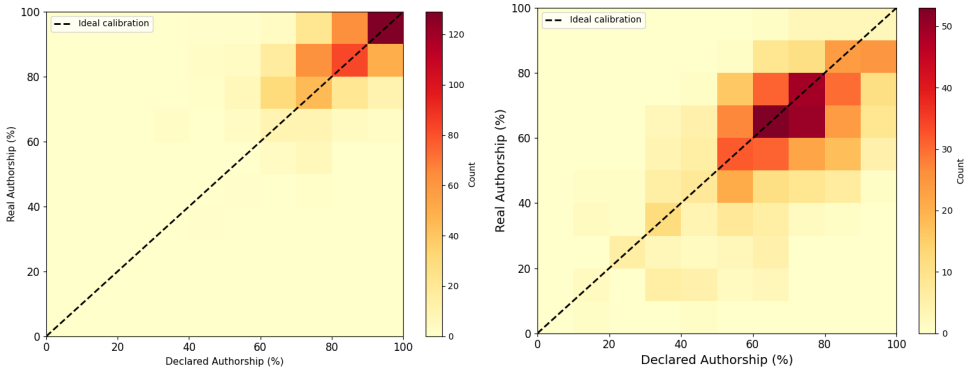



(a) Low-AI usage (b) High-AI usage

**Figure 4:** Distribution of Authorship Calibration Scores in Users Groups.

First, users from the the Low-AI usage group are predominantly concentrated in the upper-right quadrant, with the majority of both declared and real authorship values ranging between $70\% - 100\%$ (See Figure 3a). Low-GPT usage users are relatively evenly distributed on both sides of the ideal calibration line, though a slight underestimation tendency is observable (more user above the ideal calibration line). This is confirmed by the distribution of authorship calibration score that are skewed towards negative values (See Figure 4a), with a mean authorship calibration score of $-0.004$ ($SD = 0.112$), ranging from $-0.53$ to $0.26$. This negative bias indicates that users with limited AI interaction during writing sessions tend to slightly underestimate their authorship, declaring authorship lower than the actual one. The corresponding heatmap confirms this pattern with the highest density in the $90\% - 100\%$ range for both declared and real authorship (See Figure 3a). This tight clustering around maximal authorship values demonstrates that writers with less frequent AI usage maintain high authorship, while mostly accurately evaluating it.

Second, users from the High-AI usage group are broadly distributed in the calibration space, with both declared and actual authorship spanning from $10\%$ to almost $100\%$ (See Figure 3b). Specifically, there is considerable spread below the ideal calibration line (overestimation), particularly visible in the $30\% - 90\%$ declared authorship range. Corresponding authorship calibration values also show greater variability compared to the Low-GAI usage group, with a mean authorship calibration value of $0.003$, a higher standard deviation ($SD = 0.146$) and a wider range of $[-0.39, 0.49]$ (See Figure 4b). This high variability suggests that users experience an inconsistent awareness of their actual authorship, while having a slight tendency to overestimate it. The heatmap distribution reinforces these observations, with the high densities ($20 - 50$ users) occurring across the $50\% - 90\%$ range for both declared and real authorship (See Figure 5b). Importantly, a higher density appears below the ideal calibration line, particularly in regions where declared authorship of $60\% - 80\%$ corresponds to real authorship of only $40\% - 70\%$. Users relying more frequently on AI during the writing process then tend to overestimate their authorship. Finally, a statistical comparison further confirms that the distributions of authorship calibration scores differ significantly between Low-AI and High-AI users (Mann–Whitney U test, $p < 0.05$).




(a) Low-AI usage (b) High-AI usage

**Figure 5:** Heatmap of calibration curves.

## 6. Discussion and Future Work

Our experimental study reveals that while numerous users have accurate authorship calibration scores, an important proportion still show poor calibration, either overestimating or underestimating their authorship. This confirms that authorship calibration is not a trivial construct in AI-assisted contexts, and that important differences persist across users, as not all of them are able to accurately assess it (RQ1). Besides, our results confirms that the accuracy of authorship calibration vary depending on the frequency of AI use during the writing task (RQ2). Specifically, users in the Low-AI usage group show a more accurate authorship calibration, but a slight tendency to underestimate it. This may reflect stronger metacognitive awareness, as these users remain attentive to even minimal AI assistance and consequently minimize their own contribution. In contrast, users in the High-AI usage group show poorer authorship calibration, with a tendency to overestimate it. Extensive AI interactions then appears to blur the distinction between human-generated and AI-generated content. This may be the results of an "effort blend" phenomenon: the cognitive work involved in prompting, selecting, and integrating AI suggestions is subjectively experienced as equivalent to generating original text, resulting in a higher sense of authorship. These observations raise important concerns about authorship in AI-assisted learning environments. Given the conceptual links between authorship calibration and higher-order metacognitive processes, improving authorship calibration may represent a promising pathway to enhancing learning outcomes in AI-assisted contexts.

Despite introducing the concept of authorship calibration and offering promising insights, this study has several limitations. First, authorship is examined only within writing tasks rather than in authentic educational settings. Exploring a wider range of educational tasks would allow for richer experiments, ultimately providing a more comprehensive understanding of learner–AI collaboration. Second, the distinction between Low- and High-AI usage groups is based on a simple yet effective methodology. Applying more fine-grained Learning Analytics pipelines to educational datasets would allow for richer classifications of learner–AI interaction patterns, ultimately providing richer insights into how specific interaction patterns influence authorship calibration.

Finally, this study opens promising directions for future work. A key avenue lies in designing and evaluating methods that support authorship calibration, e.g. through adapted user interfaces or personalized feedback. By increasing learners' awareness about their actual contribution and fostering more accurate authorship calibration, such interventions may promote more appropriate metacognitive engagement, ultimately improving learning processes and outcomes. In addition, further research is needed to examine how authorship calibration relates to broader metacognitive processes, particularly within AI-assisted learning environments.

## 7. Conclusion

This paper introduces authorship calibration, the extent to which users accurately perceive their own authorship during AI-assisted tasks. Our results show clear variability across users, with authorship calibration being influenced by the level of AI usage. While some users maintain an accurate sense of authorship, others substantially misjudge their contribution, revealing how easily AI can blur boundaries between human- and AI-produced content. Importantly, authorship calibration offers a powerful lens for understanding learning in AI-assisted environments. Accurate authorship calibration may supports metacognitive monitoring and informed engagement, whereas miscalibration risks unproductive behaviors and weaker learning outcomes. As AI becomes broadly embedded in education settings, helping learners stay aware of their actual authorship appears essential. Designing tools and pedagogical strategies that strengthen authorship calibration will be key to fostering responsible, transparent, and educationally meaningful integration of AI in educational settings.

## Acknowledgments

The authors would like to thank the Hasler foundation for the financial support.

## Declaration on Generative AI

During the preparation of this work, the authors used GPT-5 in order to: Grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

## References

[1] A.-M. M. Gasaymeh, M. A. Beirat, A. A. Abu Qbeita, University students' insights of generative artificial intelligence (ai) writing tools, Education Sciences 14 (2024) 1062.

[2] G.-J. Hwang, N.-S. Chen, Exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions, Journal of Educational Technology & Society 26 (2023).

[3] F. Draxler, A. Werner, F. Lehmann, M. Hoppe, A. Schmidt, D. Buschek, R. Welsch, 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 (2024) 1–40.

[4] G. Kleiman, AI in writing class: Editor, co-author, ghostwriter, or muse?, Medium, 2024. URL: https://medium.com/@glenn_kleiman/ai-in-writing-class-editor-co-author-ghostwriter-or-muse-348532d896a6, accessed: September 10, 2026.

[5] I. C. Peláez-Sánchez, D. Velarde-Camaqui, L. D. Glasserman-Morales, The impact of large language models on higher education: exploring the connection between ai and education 4.0, in: Frontiers in Education, volume 9, Frontiers Media SA, 2024, p. 1392091.

[6] T. Adiguzel, M. H. Kaya, F. K. Cansu, Revolutionizing education with ai: Exploring the transformative potential of chatgpt., Contemporary educational technology 15 (2023).

[7] M. Giannakos, R. Azevedo, P. Brusilovsky, M. Cukurova, Y. Dimitriadis, D. Hernandez-Leo, S. Järvelä, M. Mavrikis, B. Rienties, The promise and challenges of generative ai in education, Behaviour & Information Technology 44 (2025) 2518–2544.

[8] J. Biggs, The role of metacognition in enhancing learning, Australian Journal of education 32 (1988) 127–138.

[9] S. Lichtenstein, B. Fischhoff, L. D. Phillips, Calibration of probabilities: The state of the art to 1980 (1981).

[10] P. A. Alexander, Calibration: What is it and why it matters? an introduction to the special issue on calibrating calibration, Learning and Instruction 24 (2013) 1–3.

[11] M. Lee, P. Liang, Q. Yang, Coauthor: Designing a human-ai collaborative writing dataset for exploring language model capabilities, in: Proceedings of the 2022 CHI conference on human factors in computing systems, 2022, pp. 1–19.

[12] Z. Epstein, A. Hertzmann, I. of Human Creativity, M. Akten, H. Farid, J. Fjeld, M. R. Frank, M. Groh, L. Herman, N. Leach, et al., Art and the science of generative ai, Science 380 (2023) 1110–1111.

[13] R. H. Mogavi, C. Deng, J. J. Kim, P. Zhou, Y. D. Kwon, A. H. S. Metwally, A. Tlili, S. Bassanelli, A. Bucchiarone, S. Gujar, et al., Exploring user perspectives on chatgpt: Applications, perceptions, and implications for ai-integrated education, arXiv preprint arXiv:2305.13114 (2023).

[14] Z. Bahroun, C. Anane, V. Ahmed, A. Zacca, Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis, Sustainability 15 (2023) 12983.

[15] U. Mittal, S. Sai, V. Chamola, D. Sangwan, A comprehensive review on generative ai for education, Ieee Access 12 (2024) 142733–142759.

[16] R. Abdelghani, H. Sauzéon, P.-Y. Oudeyer, Generative ai in the classroom: Can students remain active learners?, arXiv preprint arXiv:2310.03192 (2023).
[17] C. Gomez, S. M. Cho, S. Ke, C.-M. Huang, M. Unberath, Human-ai collaboration is not very collaborative yet: a taxonomy of interaction patterns in ai-assisted decision making from a systematic review, Frontiers in Computer Science 6 (2025) 1521066.
[18] N. Van Berkel, M. B. Skov, J. Kjeldskov, Human-ai interaction: intermittent, continuous, and proactive, Interactions 28 (2021) 67–71.
[19] B. Memarian, T. Doleck, A multidimensional taxonomy for learner-ai interaction, Education and Information Technologies 29 (2024) 18361–18378.
[20] A. Nguyen, Y. Hong, B. Dang, X. Huang, Human-ai collaboration patterns in ai-assisted academic writing, Studies in Higher Education 49 (2024) 847–864.
[21] K. Yang, M. Raković, Z. Liang, L. Yan, Z. Zeng, Y. Fan, D. Gašević, G. Chen, Modifying ai, enhancing essays: How active engagement with generative ai boosts writing quality, in: Proceedings of the 15th International Learning Analytics and Knowledge Conference, 2025, pp. 568–578.
[22] K. C. Arnold, A. M. Volzer, N. G. Madrid, Generative models can help writers without writing for them., in: IUI Workshops, 2021.
[23] S. Kim, H.-J. So, K. Park, Supporting learner agency in collaborative writing with generative ai, British Journal of Educational Technology (2025).
[24] A. Bozkurt, Genai et al. cocreation, authorship, ownership, academic ethics and integrity in a time of generative ai, 2024.
[25] Z. Swiecki, H. Khosravi, G. Chen, R. Martinez-Maldonado, J. M. Lodge, S. Milligan, N. Selwyn, D. Gašević, Assessment in the age of artificial intelligence, Computers and Education: Artificial Intelligence 3 (2022) 100075.
[26] Y. Cheng, K. Lyons, G. Chen, D. Gašević, Z. Swiecki, Evidence-centered assessment for writing with generative ai, in: Proceedings of the 14th learning analytics and knowledge conference, 2024, pp. 178–188.
[27] H. Pan, E. Araujo Oliveira, R. Ferreira Mello, Exploring human-ai collaboration in educational contexts: Insights from writing analytics and authorship attribution, in: Proceedings of the 15th International Learning Analytics and Knowledge Conference, 2025, pp. 903–909.
[28] A. Richburg, C. Bao, M. Carpuat, Automatic authorship analysis in human-ai collaborative writing, in: Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), 2024, pp. 1845–1855.
[29] A. Shibani, R. Rajalakshmi, F. Mattins, S. Selvaraj, S. Knight, Visual representation of co-authorship with gpt-3: Studying human-machine interaction for effective writing., International Educational Data Mining Society (2023).
[30] J. Torres, S. García, E. Peláez, Visualizing authorship and contribution of collaborative writing in e-learning environments, in: Proceedings of the 24th International Conference on Intelligent User Interfaces, 2019, pp. 324–328.
[31] M. N. Hoque, T. Mashiat, B. Ghai, C. D. Shelton, F. Chevalier, K. Kraus, N. Elmqvist, The hallmark effect: Supporting provenance and transparent use of large language models in writing with interactive visualization, in: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 2024, pp. 1–15.
[32] J. H. Flavell, Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry., American psychologist 34 (1979) 906.
[33] P. R. Pintrich, A conceptual framework for assessing motivation and self-regulated learning in college students, Educational psychology review 16 (2004) 385–407.
[34] B. J. Zimmerman, Becoming a self-regulated learner: An overview, Theory into practice 41 (2002) 64–70.
[35] H. Khiat, Using automated time management enablers to improve self-regulated learning, Active Learning in Higher Education 23 (2022) 3–15.
[36] X. Xu, L. Qiao, N. Cheng, H. Liu, W. Zhao, Enhancing self-regulated learning and learning experience in generative ai environments: The critical role of metacognitive support, British

Journal of Educational Technology (2025).
[37] N. J. Stone, Exploring the relationship between calibration and self-regulated learning, Educational psychology review 12 (2000) 437–475.