EDITORIAL

Scientific judgment cannot be replaced by artificial intelligence: reflections on the education of future researchers

O pensamento científico não pode ser substituído por inteligência artificial: reflexões sobre a educação de futuros pesquisadores

Lucas Rodrigues Nascimento[1]

[1] Research Leader of the CAMINHAR Laboratory, Graduate Program in Physiological Sciences, Discipline of Physiotherap y, Universidade Federal do Espírito Santo, Brazil.

Correspondência:
lucas.r.nascimento@ufes.br

Opposition to the use of artificial intelligence in scientific writing is not uncommon. This resistance is not typically driven by technophobia, but by experience. For non-native English speakers, learning how to write and publish scientific papers in international peer-reviewed journals is a slow, demanding, and formative process that requires sustained intellectual effort. The emergence of new tools raises concerns that intellectual effort is being replaced with convenience, particularly among students and early-career researchers who are still developing their writing skills.

Despite these concerns, exploring the use of artificial intelligence in practice may help clarify where it can be useful and where it consistently fails. For instance, during the preparation of responses to reviewers for manuscripts submitted to international journals, artificial intelligence may, in most cases, provide only generic responses that are fluent in form but shallow in content. This observation is consistent with a recent systematic review showing that artificial intelligence can support academic writing in tasks such as structuring and editing, while remaining highly dependent on human guidance[1]. Two practical examples are presented below.

In this first example, when a reviewer questioned whether mobility measured in a cohort of individuals after stroke could truly represent stroke-induced mobility loss[2], artificial intelligence was asked to draft a response. The generic response agreed with the critique and framed it as a general limitation, but did not engage with how the construct had been defined or with the logic underlying the assessment timing. The author's final response followed a different logic: it clarified how mobility before and after stroke had been operationalized so that pre-existing mobility limitations could be accounted for, ensuring that the observed mobility loss was attributable to stroke.

In the second example, a statistical reviewer questioned how variables were selected and whether collinearity among them may have influenced

the results[3]. The artificial intelligence response followed familiar patterns: it acknowledged that choosing relevant outcomes based upon statistical significance is discouraged and suggested reporting collinearity diagnostics. However, it did not explain how the analysis was theoretically grounded, nor how this theoretical framework guided both variable selection and the construction of the regression analyses. The author's final response, edited only for clarity with the help of artificial intelligence, explained the rationale for grouping the clinical outcomes, why only certain predictors were retained, and included a table with real data from the study participants to assess collinearity.

The difference in both cases lay not in writing quality, but in the human sequence of reasoning that preceded the writing. Those decisions are necessarily human. This is in accordance with empirical evidence that supports this distinction. In a controlled comparison of human-written, AI-only, and AI-assisted scientific review articles, artificial intelligence reduced writing time but required extensive human fact-checking and revision, reinforcing that efficiency gains do not replace scientific judgment[4].

Scientific judgment, however, is not only about decisions made in isolation. Writing, too, is not stylistically neutral; over time, those decisions accumulate into something recognizable as authorial voice, reflecting how researchers frame questions, build arguments, and emphasize ideas. This voice is not innate, but developed gradually through practice, revision, supervision, and sustained engagement with one's own text.

The implications are particularly concerning for students and early-career researchers. At this stage, writing is not merely a vehicle for communication, but a central component of learning how to think scientifically. Struggling to articulate ideas, revising arguments, and confronting weaknesses in one's own reasoning are essential steps in intellectual maturation. For these researchers, the critical question is not whether to use artificial intelligence, but when and how. Used after the authorial voice is established, it can support clarity and efficiency; used too early, it risks interrupting the process

through which scientific independence, authorship, and intellectual identity are formed.

When a researcher has not yet developed a clear conceptual writing style, it becomes tempting to accept a text that is fluent and technically correct, even if it does not fully reflect their own thinking. In these cases, artificial intelligence may resolve surface-level problems while distancing the author from the manuscript. The issue is not textual quality, but the interruption of learning. Artificial intelligence cannot build a voice it did not help form.

Over time, the use of artificial intelligence by experts may evolve from responding to reviewers to drafting and revising manuscripts. However, the appropriate timing for using artificial intelligence is after a clear authorial voice has already been established. This will only be achieved after years of training shaped by mentorship and teaching from senior researchers. In Brazil, this process depends on effective supervision during master's and doctoral training, in which supervisors contribute to the development of ethical awareness and rigor in scientific writing, on institutional guidance provided by universities regarding ethical uses of artificial intelligence[5], and on governmental support that enables students and researchers to undertake immersive experiences in English-speaking countries. At this stage, artificial intelligence can help refine arguments and improve clarity in a non-native language, without influencing clinical or scientific positions beyond those supported by the data and grounded in theoretical, professional, and social views of the world. In this context, the author remains the mind behind the text; only the time required to put ideas on paper changes. Writing can be accelerated; thinking cannot. And thinking matters.

REFERÊNCIAS

1.   Khalifa M, Albadawy M. Using artificial intelligence in academic writing and research: an essential productivity tool. Comput Methods Programs Biomed Update. 2024;5:100145. doi:10.1016/j.cmpbup.2024.100145.

2.   Kuster E, Santuzzi CH, Correia TB, Ventura LM, Nascimento LR. Self-efficacy and pre-stroke activity participation combined with

functional independence best explain mobility limitations after stroke: a cross-sectional exploratory study. Top Stroke Rehabil. 2025 Oct 18:1-9. doi:10.1080/10749357.2025.2571566.

3. Correia TB. Preditores da participação social aos 3 e 6 meses após acidente vascular cerebral: coorte prospectiva [dissertação]. Vitória: Universidade Federal do Espírito Santo; 2025.

4. Kacena MA, Plotkin LI, Fehrenbacher JC. The use of artificial intelligence in writing scientific review articles. Curr Osteoporos Rep. 2024;22(1):115-121. doi:10.1007/s11914-023-00852-0.

5. Universidade Federal da Bahia. Guia para uso ético e responsável da inteligência artificial generativa na Universidade Federal da Bahia. Salvador: Universidade Federal da Bahia; 2025. Available from: https://ufba.br/sites/portal.ufba.br/files/guia_para_uso_etico_e_responsavel_da_inteligencia_artificial_generativa_na_universidade_federal_da_bahia.pdf

NOTES

Artificial intelligence statement

Artificial intelligence was used only for language refinement; all scientific decisions were made by the author.