aejlogo_h.png

Astronomy Education Journal, 2026, Vol. 06, No. 1

doi: 10.32374/AEJ.2026.6.1.16rpw
ISSN: 2004-2981 · Published in Sweden

RESPONSE TO PRIOR WORK

Algorithmic Detection of "Pseudoscience" and the Demarcation Problem: Educational Lessons from Impey et al. (2025)

James Houran[1,*] and Mark Rodeghier[2]

[1]Integrated Knowledge Systems, Southlake, TX, 76092, USA ; [2]Center for UFO Studies, Chicago, IL, 60631 USA

[*]Jim@drjimhouran.com

Abstract

This essay critically examines Impey et al.'s (2025) application of machine-learning to detect ostensible scientific misinformation, arguing that their operational definition of pseudoscience is philosophically inadequate. While the authors achieve impressive accuracy rates (> 95%) in flagging content across domains including astrology and UFO/UAPs, their reliance on a binary "real versus fake" classification ignores decades of philosophy and sociology of science literature demonstrating that demarcation between science and pseudoscience is contested, context-dependent, and socially negotiated rather than algorithmically determinable. Drawing on foundational work by Laudan, Hansson, Pigliucci and Boudry, and Gieryn's boundary-work theory, we show that encoding oversimplified demarcation criteria into neural networks risks creating what might be called "algorithmic pseudo-demarcation"—technologically sophisticated tools lacking epistemological grounding. The essay examines methodological limitations including the black-box problem, synthetic training data concerns, and educational implications. It is argued that effective misinformation detection requires integrating philosophical rigor with technological innovation, acknowledging epistemic ambiguity, and fostering critical thinking rather than algorithmic deference.

Keywords: demarcation problem, machine-learning, pseudoscience, boundary-work, scientific misinformation

1 Introduction

Impey et al. (2025) presented an ambitious study applying machine-learning and large language models (LLMs) to the detection of ostensible "scientific misinformation," with a particular focus on astronomy related pseudoscience such as astrology and unidentified flying objects (UFOs). Their project unfolded in two phases. First, undergraduate science majors collected and classified 800 short, non-technical articles across a range of scientific and pseudoscientific topics, achieving high inter-rater reliability in distinguishing "real" from "fake" science. In the second phase, the authors scaled up to tens of thousands of articles and trained recurrent neural networks and other LLMs to flag misinformation with reported accuracies exceeding 90% in testbeds on evolution and climate change, and over 95% across ten domains including astrology and UFOs. The study culminated in the development of a prototype browser extension designed to provide users with Bayesian, color-coded probabilities of whether online science content is legitimate.

To their credit, Impey et al. (2025) explicitly acknowledged

that demarcation is not straightforward and that no single litmus test cleanly separates science from pseudoscience. Our concern is therefore not that the authors are unaware of this complexity, but that their binary operationalization—and especially the probability-to-color interface of the browser extension—risks obscuring that very complexity in practice. Even when researchers recognize philosophical nuance, design choices can unintentionally encourage users to treat algorithmic output as authoritative rather than provisional, thereby obscuring the epistemic ambiguity that the demarcation literature emphasizes.

The broader claim of this essay is that astronomy educators and students benefit from understanding why the science–pseudoscience boundary is historically contingent, socially negotiated, and often ambiguous. Our critique is therefore primarily educational: to show how insights from philosophy of science and science studies can strengthen the pedagogical framing of tools like those proposed by Impey et al., ensuring that such tools cultivate critical thinking rather than algorithmic deference.

The authors concluded that astronomy provides a particularly salient lens for studying misinformation because of the cultural prominence of astrology and UFO beliefs, both of which dominate online search traffic and social media discourse. They argued that machine-learning offers a scalable solution to the erosion of trust in science and the growing prevalence of pseudoscience in civic life. Their broader claim is that pseudoscience can be operationally defined as "beliefs, statements, or practices that claim to be factual, but which are incompatible with the scientific method" (Impey et al., 2025, p. 138aer-2). This definition underpins their classification system and the training of their algorithms. While the technological innovation is certainly commendable, the conceptual foundation of their work requires greater engagement with the philosophy of science literature and science studies scholarship. Without such grounding, the project risks conflating contested categories with settled science, undermining both its credibility and its utility in addressing the complex ecosystem of misinformation.

2 Literature Context

Before turning to specific philosophical debates, it is important to clarify the aim of this essay. Our primary contribution is educational: to show how the science–pseudoscience boundary has been understood in philosophy of science and science studies, and why this history matters for astronomy educators, especially for those deploying machine-learning tools. The goal is not to fault Impey et al. for not incorporating insights from a specialized body of scholarship, but to demonstrate how longstanding debates about demarcation, expertise, and boundary-work can enrich the pedagogical framing of their project.

The demarcation problem—the challenge of distinguishing science from pseudoscience—has long been recognized as complex and unresolved. Karl Popper (1959/2002) famously proposed falsifiability as the criterion, but subsequent scholars have shown that this is neither sufficient nor necessary. For example, string theory and multiverse hypotheses are not currently falsifiable yet are considered legitimate topics in science, while astrology makes falsifiable claims but fails in empirical testing (Pigliucci and Boudry, 2013). Hansson (2013; 2025) notably emphasized that pseudoscience cannot be defined by a single criterion but requires a nuanced, context sensitive analysis. Laudan (1983) went further, arguing that the demarcation problem itself may be misguided, since science is a historically evolving practice rather than a fixed set of rules. These debates underscore the difficulty of operationalizing pseudoscience for algorithmic classification.

The sociological literature complicates these philosophical debates by revealing that demarcation is not merely an epistemological problem but also a social practice. Thomas Gieryn (1983; 1998) introduced the concept of "boundary-work", showing how scientists actively construct and patrol the boundaries between science and non-science for rhetorical, professional, and institutional purposes. This perspective suggests that what counts as legitimate science is negotiated through community practices, institutional structures, and power relations rather than determined by fixed methodological criteria. Scientific practice is inherently provisional and often involves frontier areas where evidence is partial or contested (Longino, 2002). Similarly, work in Science and Technology Studies has shown that science in practice looks very different from idealized "textbook science" (Latour, 1987), and that expertise itself is contextual and socially constructed (Collins and Evans, 2007). These insights raise important questions about whether algorithmic classification can capture the dynamic, context-dependent nature of scientific legitimacy.

Historical examples further illustrate the dangers of rigid demarcation. Continental drift was ridiculed by mainstream geology for decades before plate tectonics vindicated Wegener's hypothesis. The bacterial cause of ulcers (Helicobacter pylori) was dismissed as implausible before Marshall and Warren's Nobel Prize-winning work. Ball lightning was long treated as folklore before receiving serious scientific attention. Gordin (2012) documented how the "pseudoscience wars" often reflect institutional politics as much as epistemological disagreements. These cases demonstrate that bright-line distinctions between science and pseudoscience are historically contingent and can impede rather than advance scientific progress.

Recent empirical work also highlights the consequences of anomalous and nonordinary experiences, which are often labeled as pseudoscientific but can have profound psychological effects. Maraldi et al. (2024) conducted a systematic review of 157 studies and found that such experiences can foster well-being, spiritual growth, and resilience, but also create stress, confusion, or exacerbate vulnerabilities. These findings complicate simplistic dismissals of anomalous phenomena (and their study) as pseudoscience, showing that they have real psychological and cultural significance that cannot be reduced to questions of empirical validity.

3 Methods Critique

The methodological design of Impey et al.'s (2025) study is innovative, combining human classification with machine-learning algorithms to detect misinformation. Their use of undergraduate coders to establish a training set is pragmatic, and the reported inter rater reliability also suggests reasonable consistency in judgments. However, the reliance on a simplified binary classification of "real" versus "fake" science raises important questions. Philosophers of science have long argued that scientific practice is not reducible to such dichotomies, as the boundaries between legitimate science, speculative hypotheses, and pseudoscience are porous and context dependent (Laudan, 1983). By embedding this binary into their algorithms, the authors risk encoding philosophical oversights into technological systems. Moreover, the reported accuracy rates, while impressive, may reflect the ease of distinguishing overt pseudoscientific claims rather than the nuanced challenges of borderline cases. This methodological limitation underscores the importance of conceptual clarity in defining pseudoscience before operationalizing it in machine-learning frameworks.

Furthermore, the opacity of neural network decision-making raises concerns about epistemic accountability. Even when achieving 95% accuracy, these models cannot explain why a particular text is classified as pseudoscientific, making it difficult to assess whether they are identifying genuine methodological flaws or merely detecting stylistic markers and linguistic patterns. This "black box" problem is well-documented in machine-learning literature (Noble, 2018; O'Neil, 2016) and has significant implications for systems designed to adjudicate truth claims. If the algorithm flags content based on surface features rather than substantive epistemological criteria, it may perpetuate biases present in the training data or misidentify legitimate scientific debate as "misinformation."

The use of ChatGPT to generate synthetic training data introduces additional issues. While Impey et al. (2025) acknowledged that large LLMs can "hallucinate" false information, they dismiss this risk as "statistically insignificant" based on limited sampling. This is problematic for two reasons. First, the irony of using a system prone to generating false information to detect false information deserves deeper consideration. Second, synthetic training data may introduce systematic biases that reflect ChatGPT's own training corpus rather than genuine epistemological criteria for distinguishing science from pseudoscience (Bender et al., 2021). If ChatGPT has learned to associate certain linguistic patterns with "legitimate science" based on its training data, the neural network may simply be learning to reproduce these associations rather than identifying substantive methodological differences.

Taken together, these concerns point to a broader and more durable issue that extends beyond the specific case of ChatGPT-generated data—as LLMs evolve rapidly, specific worries about "hallucinations" may become less salient, but the underlying problem remains. When training data reflect patterns learned from vast, uncurated corpora, models may inherit implicit assumptions about what scientific writing "looks like," rather than identifying substantive methodological criteria. This reinforces the need for transparency and for educational designs that help students interrogate why a system classifies a text in a particular way.

Given that the authors classified thousands of articles with impressive levels of reported accuracy—albeit using their particular definitions and methodological choices—it would be highly valuable, at this early stage of the project, to provide readers with concrete insight into how the algorithm behaves in practice. Specifically, scholars interested in the demarcation of pseudoscience, as well as those of us who share the concerns outlined above, would benefit greatly from a curated set of example texts or webpages, drawn from multiple topic areas, that the system classified as either scientific or pseudoscientific. Examples of edge cases—where classification is difficult, ambiguous, or sensitive to labeling assumptions—would be especially helpful for assessing the algorithm's performance, limitations, and underlying decision boundaries.

4 Conceptual Critique

The central issue lies in Impey et al.'s (2025) definition of pseudoscience. By characterizing it as "incompatible with the scientific method," the authors adopted a definition that is both circular and philosophically inadequate. The scientific method itself is not a monolithic entity but a diverse set of practices that vary across disciplines and historical contexts (Laudan, 1983). Popper's falsifiability criterion, while influential, has been shown to be insufficient, as many legitimate scientific concepts—such as string theory or multiverse cosmology—are not currently falsifiable, while some pseudoscientific claims—like astrology or homeopathy—are indeed falsifiable but often fail empirical tests when subjected to controlled studies (Pigliucci and Boudry, 2013; Popper, 2002). Pigliucci and Boudry (2013) argued that pseudoscience is best understood as a "family resemblance" concept, with overlapping features rather than strict boundaries. Hansson (2013; 2025) further critiqued the reliance on demarcation metaphors, noting that determining pseudoscience often requires specialized expertise rather than abstract criteria, a particularly relevant point with reference to Impey et al.'s study. Fasce (2019) proposed a metacriterion approach that acknowledges the inherent fuzziness of demarcation while still allowing for practical discrimination. By ignoring these debates, Impey et al. (2025) arguably presented an inadequate and incomplete definition of pseudoscience—one that lacks falsifiability, nuance, and philosophical grounding.

The sociological literature on "boundary-work" (e.g., Gieryn, 1983, 1998) suggests that demarcation is not merely an epistemological problem but a social practice through which scientists establish professional authority. By treating pseudoscience as an objective category detectable through algorithmic means, Impey et al. (2025) overlooked how scientific legitimacy is negotiated through institutional practices, peer review, and community consensus rather than through fixed methodological criteria. This oversight is particularly problematic when applied to contested domains that are often characterized as "maverick, fringe, or frontier science" (Houran and Bauer, 2022). For instance, parapsychology has institutional structures (journals, professional organizations, university positions) that follow mainstream science, yet it remains controversial precisely because boundary-work is ongoing rather than settled (Collins and Evans, 2007).

The treatment of UFOs is particularly challenging. Here, Impey et al. conflated two distinct claims: (1) some aerial phenomena remain unexplained, and (2) unexplained phenomena represent alien spacecraft. The former is empirically defensible—even the Pentagon acknowledges numerous unexplained cases—while the latter requires scientifically sufficient evidence. By categorizing all UFO-related content as potential pseudoscience, the algorithm risks flagging legitimate scientific investigation of anomalous aerial phenomena, especially given recent institutional interest from NASA and the Department of Defense (Yingling et al., 2023). Impey et al. further noted that the government has rebranded UFOs as UAPs (unidentified aerial phenomena) and that Congressional hearings have examined the topic, yet they maintain that "scientists find reports of extraterrestrials visiting the Earth to be unconvincing" (Impey et al., 2025, p. 138aer-3). This framing obscures an important distinction, i.e., scientists may find the extraterrestrial hypothesis currently unconvincing while still recognizing that some phenomena warrant serious investigation (e.g., Sturrock et al., 1998). If a topic is defined as pseudoscience, rather than allowing for the possibility of studying it with appropriate scientific methods, then the matter is settled a priori (see, for comparison, the study of a UFO artifact by (Powell et al., 2022)). An algorithm that cannot make this distinction may undermine rather than support scientific inquiry.

Similarly, the treatment of astrology overlooks important nuances. While "judicial astrology" (predicting specific events) and "natal astrology" (personality traits from birth charts) may lack strong or consistent empirical support, some research on circadian rhythms, seasonal birth effects, and chronobiology explores related questions through legitimate scientific frameworks. An algorithm trained on binary classifications may struggle to distinguish between these contexts, potentially flagging legitimate research on biological rhythms as pseudoscience or, conversely, giving unwarranted credibility to astrological claims that invoke

scientific-sounding language.

5 Implications

The negative consequences of oversimplification are threefold. First, it undermines the credibility of machine-learning tools by embedding contested categories into algorithms. If pseudoscience is defined too narrowly or simplistically, the resulting classifications may misrepresent legitimate scientific debates as misinformation or fail to capture the cultural and psychological dimensions of pseudoscientific beliefs. The algorithm may achieve high accuracy on clear-cut cases (creationism vs. evolutionary biology, climate denial vs. climate science) while performing poorly on borderline cases where epistemic status is genuinely ambiguous.

Second, it may make constructive engagement more difficult with publics who hold anomalous or extraordinary beliefs if those beliefs are framed as irrational rather than understood within their meaning-making frameworks. Research shows that anomalous and nonordinary experiences can have both positive and negative consequences for individuals, depending on context and interpretation (Maraldi et al., 2024). An algorithm trained to flag astrology or UFO beliefs as misinformation cannot distinguish between exploitative pseudoscience (e.g., fraudulent psychics charging vulnerable people for fake predictions) and sincere meaning-making practices (e.g., individuals using astrological frameworks for self-reflection or community building). This conflation may distance some publics and reinforce the perception that science is dismissive of human experience, potentially deepening the very distrust in scientific institutions that Impey et al.'s project aimed to address.

Third, the educational implications warrant deeper consideration. While Impey et al. framed their browser extension as a tool for combating misinformation, it risks promoting intellectual outsourcing rather than critical thinking. Lewandowsky et al. (2020), whose Debunking Handbook was cited by the authors, emphasizes the importance of teaching people how to evaluate claims, not merely what to believe. A browser extension that provides color-coded verdicts may undermine the development of epistemic virtues like careful reasoning, intellectual humility, tolerance for ambiguity, and appreciation for the provisional nature of scientific knowledge. Students who learn to defer to algorithmic authority may be less equipped to navigate the "messy" realities of scientific controversy, where experts disagree, evidence is incomplete, and conclusions are tentative. The classroom activities that Impey et al. proposed—having students classify articles and explain their reasoning—are pedagogically valuable, but the browser extension may work at cross-purposes by suggesting that classification can be automated.

A more philosophically-informed approach would be to integrate the demarcation literature, acknowledge the limits of operational definitions, and situate pseudoscience within broader epistemic contexts. This likely involves several key modifications. First, rather than binary classification, the system could provide multidimensional assessments that flag specific methodological concerns (e.g., "no peer review," "relies on anecdotal evidence," and "contradicts established findings") while acknowledging uncertainty. Second, the system should incorporate transparency mechanisms that explain why content is flagged, citing specific passages and methodological issues rather than merely outputting a probability score. This would enable users to evaluate the reasoning and develop their own critical thinking skills. Third, the training data should include borderline cases and explicitly address how the algorithm handles epistemic ambiguity. Finally, the system should be designed with epistemic humility, acknowledging that algorithmic classification cannot replace expert judgment and that scientific legitimacy is context-dependent and historically contingent.

6 Educational Design Recommendations

To support astronomy educators and to align misinformation-detection tools with established philosophy-of-science insights, we propose several concrete design principles for classroom or public-facing deployment:

- *Emphasize uncertainty and borderline cases.* Instead of presenting a single binary verdict, the interface could highlight when a text falls into an epistemically ambiguous region, prompting users to consider why classification is difficult.
- *Replace global verdicts with evidence-based indicators.* Rather than labeling content as "real" or "fake," the system could flag specific features—e.g., absence of citations, reliance on anecdotal evidence, contradiction with established findings—that students can evaluate using classroom rubrics.
- *Provide explanation prompts.* When the system flags content, it could offer short, transparent rationales (e.g., "This article relies on testimonial evidence without methodological detail"), which instructors can integrate into critical-analysis assignments.
- *Support metacognitive reflection.* The tool could include optional prompts asking students to articulate why they agree or disagree with the system's assessment, reinforcing epistemic virtues such as evidence-based justification, intellectual humility, and tolerance for ambiguity.

These recommendations shift the tool from an authority that delivers binary verdicts to a scaffold that cultivates critical thinking—aligning its use with both the demarcation literature and AEJ's educational mission.

7 Conclusion

Impey et al. (2025) contributed valuable technological tools for potentially addressing misinformation, but their definition of pseudoscience is philosophically inadequate. By treating pseudoscience as a fixed category rather than a contested problem, they risk encoding flawed demarcation criteria into algorithmic systems that may misidentify legitimate inquiry, distance some publics, and undermine critical thinking. Future work should pursue three directions: (1) engaging with demarcation literature and science studies scholarship to develop more nuanced classification frameworks that acknowledge borderline cases, context-dependency, and the social construction of scientific boundaries; (2) incorporating transparency mechanisms that explain why content is flagged, enabling users to develop critical thinking skills rather than defer to algorithmic authority; and (3) recognizing that some phenomena labeled "pseudoscientific"—from anomalous experiences to speculative but legitimate scientific hypotheses—occupy epistemically ambiguous spaces that resist binary classification.

Their project's ambition to combat misinformation is laudable, and machine-learning may indeed offer valuable tools for this purpose. However, these tools must be built on philosophically robust foundations that respect the complexity of scientific knowledge production. Only by integrating methodological innovation with conceptual rigor and epistemic humility can projects like this meaningfully address the challenges of

misinformation while avoiding the creation of new problems. The alternative—deploying powerful algorithmic systems based on oversimplified definitions—risks producing what might be called "algorithmic pseudoscience" or "pseudo-demarcation:" technologically sophisticated tools that lack the epistemological grounding necessary to adjudicate truth claims. In an information ecosystem already struggling with questions of authority and legitimacy, such tools may exacerbate rather than resolve the crisis of trust that motivates their development.

Framed in this way, Impey et al.'s (2025) core approach and initiative become not only a technological intervention but also an opportunity to teach students how scientific boundaries are drawn, contested, and revised—an educational outcome fully aligned with the mission of astronomy education.

8 Acknowledgements

Not applicable

9 Declarations

The authors declares that there are no conflicts of interest with respect to the authorship and/or publication of this response essay, as well as that no funding was received in support of this work.

9.1 Ethical Considerations

Because this project involves no human participants, questions of informed consent do not arise.

9.2 Consent for Publication

All authors have consented to this being published.

9.3 Competing Interests

The authors declare that they have no competing interests.

9.4 Funding

No funding was received in support of this work.

9.5 Use of Generative AI

The authors declare no use of generative AI in the production of this work.

References

Bender, E. M., Gebru, T., McMillan-Major, A., and Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, pages 610–623.

Collins, H. and Evans, R. (2007). *Rethinking Expertise*. University of Chicago Press.

Fasce, A. (2019). Are pseudosciences like seagulls? a discriminant metacriterion facilitates the solution of the demarcation problem. *International Studies in the Philosophy of Science*, 32(3–4):155–175.

Gieryn, T. F. (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review*, 48(6):781–795.

Gieryn, T. F. (1998). *Cultural Boundaries of Science: Credibility on the Line*. University of Chicago Press.

Gordin, M. D. (2012). *The Pseudoscience Wars: Immanuel Velikovsky and the Birth of the Modern Fringe*. University of Chicago Press.

Hansson, S. O. (2013). Defining pseudoscience and science. In Pigliucci, M. and Boudry, M., editors, *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*, pages 61–78. University of Chicago Press.

Hansson, S. O. (2025). Demarcating, defining, and diagnosing pseudoscience. *Philosophy of Science*, 92(4):922–937.

Houran, J. and Bauer, H. H. (2022). "fringe science"—a tautology, not pariah [editorial]. *Journal of Scientific Exploration*, 36(2):207–217.

Impey, C., Danehy, A., Wenger, M., and Buxner, S. (2025). Using machine learning to detect scientific misinformation. *Astronomy Education Journal*, 5(1):138aer.

Latour, B. (1987). *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press.

Laudan, L. (1983). The demise of the demarcation problem. In Cohen, R. S. and Laudan, L., editors, *Physics, Philosophy, and Psychoanalysis: Essays in Honor of Adolf Grünbaum*, pages 111–127. D. Reidel.

Lewandowsky, S., Cook, J., Ecker, U., Albarracín, D., Amazeen, M., Kendeou, P., Lombardi, D., Newman, E., Pennycook, G., Porter, E., Rand, D., Rapp, D., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C., Sinatra, G., Swire-Thompson, B., van der Linden, S., Vraga, E., Wood, T., and Zaragoza, M. (2020). The debunking handbook 2020. https://www.climatechangecommunication.org/wp-content/uploads/2020/10/DebunkingHandbook2020.pdf.

Longino, H. E. (2002). *The Fate of Knowledge*. Princeton University Press.

Maraldi, E. d. O., Taves, A., Moll, J., Hartle, L., Moreira de Oliveira, M. E., Bortolini, T., and Fischer, R. (2024). Nonordinary experiences, well being and mental health: A systematic review of the evidence and recommendations for future research. *Journal of Religion and Health*, 63(1):410–444.

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Pigliucci, M. and Boudry, M., editors (2013). *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*. University of Chicago Press.

Popper, K. (2002). *The Logic of Scientific Discovery*. Routledge. Reprint of 1959 edition.

Powell, R., Swords, M., Rodeghier, M., and Budinger, P. (2022). Isotope ratios and chemical analysis of the 1957 brazilian ubatuba fragment. *Journal of Scientific Exploration*, 36(1):39–48.

Sturrock, P., Eshleman, V. R., Holzer, T. E., Jokipii, J. R., Papike, J. J., Reitz, G., Tolbert, C. R., and Veyret, B. (1998). Physical evidence related to ufo reports: The proceedings of a workshop held at the pocantico conference center, tarrytown, new york, september 29–october 4, 1997. *Journal of Scientific Exploration*, 12(2):179–229.

Yingling, M., Yingling, C., and Bell, B. (2023). Faculty perceptions of unidentified aerial phenomena. *Humanities and Social Sciences Communications*, 10:246.