Journal of Learning Development in Higher Education
ISSN: 1759-667X
Issue 39: March 2026
https://doi.org/10.47408/jldhe.vi39.1809

OPINION PIECE

When AI writes the doctoral thesis: reclaiming the oral defence as a learning development intervention

Valerie Storey

Franklin University, USA

ABSTRACT

Large language models have fundamentally challenged traditional methods of verifying doctoral competency as AI-generated text becomes increasingly difficult to distinguish from human scholarship. This paper argues that thesis committees and doctoral supervisors must reclaim the oral defence as a critical checkpoint for assessing authentic threshold crossing rather than a ceremonial rite of passage. Drawing on historical examples from medieval oral disputations through to the rise of written theses, this paper asserts the necessity of returning to rigorous oral assessment. Given the limitations of detection technologies and the growing use of AI in thesis writing, oral defences must move from confirmatory questions that permit regurgitated responses to exploratory inquiry that demonstrates genuine conceptual transformation. This requires developing assessment literacy among examiners to distinguish candidates who have achieved deep disciplinary understanding from those who have merely assembled AI-generated text, thereby revealing human capacities for critical thinking, spontaneous reasoning, and scholarly judgment.

KEYWORDS: AI detection, academic integrity, large language models, learning development, oral examination, thesis defence.

ARTICLE HISTORY: Received 8 January 2026. Accepted 6 February 2026.

The doctoral thesis marks the culmination of doctoral education, demonstrating research competence and readiness for an academic position (Mirick et al., 2020). In 2024, approximately 277,000 doctoral candidates worldwide successfully defended their research,

with the United States leading the way at about 71,000 per year (World Population Review, 2025). Failing the defence (viva) is rare. Statistics suggest that around 1-3% of defences result in failure and that the defence is predominantly a ritual, functioning mostly as a ceremonial event with limited academic impact on research quality (Nir & Bogler, 2021). From a learning development perspective, this ceremonial approach represents a missed opportunity. Rather than functioning as a summative gatekeeping ritual, the oral defence could serve as what Boud and Dawson (2023) term a 'consequential assessment', one that both evaluates competency and develops the candidate's capacity for scholarly judgment. This reframing becomes urgent in the GenAI era, in which the oral defence may represent the only reliable means of distinguishing genuine doctoral-level thinking from GenAI text.

Doctoral examination procedures vary significantly across international contexts, with fundamental differences in assessment models, supervisor involvement, and submission gatekeeping between North American, European, and Commonwealth systems (Barnett et al., 2017). Similarly, terminology for those who guide doctoral candidates also differs across systems: 'supervisor' in UK and Commonwealth contexts, 'advisor' or 'dissertation chair' in North American institutions (doctoral theses are referred to as dissertations in North America), and variations such as 'promotor' in some European countries (Barnett et al., 2017). This paper refers to the 'oral defence' to encompass all forms of oral doctoral examination across institutional contexts and uses 'supervisor' to refer to the primary faculty member guiding the doctoral candidate, acknowledging that equivalent roles may be designated differently across institutions.

To effectively conduct oral defences that distinguish GenAI text from authentic scholarship, committee/supervisor development is critical. Yet research reveals that many institutions provide minimal training, ranging from non-existent support to single workshops (Hill & Vaughan, 2018; Wichmann-Hansen et al., 2020). This gap in preparation becomes particularly problematic when committee members must assess not just content mastery but also the authenticity of intellectual engagement. For example, Mirick et al.'s (2020) single-institution study of social work dissertation (thesis) chairs in the United States revealed considerable variation in the specific tasks identified as part of the chair's responsibility, such as reading and commenting on thesis drafts and providing 'substantive' or 'consistent' feedback. These inconsistencies suggest that chairs may lack a shared understanding of assessment

standards—a deficiency that becomes acute when evaluating whether candidates can articulate genuine understanding versus regurgitating GenAI content. At some institutions, a faculty member must first serve as a member of a thesis committee before assuming the role of chair; in other programmes, faculty members with a doctorate are permitted to take on the chair role without any prior experience in serving on a doctoral committee (Carlin & Perlmutter, 2006). Without systematic preparation in assessment literacy and oral examination techniques, committee members cannot reliably distinguish authentic doctoral competency from sophisticated GenAI assistance.

Doctoral candidates regard the oral defence as a stressful rite of passage rather than a pass/fail situation. However, a fundamental change in approach is necessary, and if a committee has even the slightest doubt regarding academic integrity, then a range of options should be available. For example, if there is clear evidence in the oral defence that the doctoral candidate lacks an understanding of their own research, then this is a serious situation with significant ethical implications. Clearly, the doctoral candidate has not met the fundamental requirements for a doctoral degree. The committee/examiners then have a moral obligation to investigate thoroughly and even deny the degree. If concerns arise in the deliberation stage, any decision should be delayed while a follow-up investigation is conducted. Awarding a doctorate to a person who did not complete the work not only harms academic integrity but also produces incompetent professionals in the field, which ultimately impacts not only the programme but also the institution.

However, the integrity concerns that have traditionally focused on a candidate's mastery of content and research design now extend into new territory. Committee/supervisors must not only assess whether candidates understand their research but also monitor whether the written manuscript they are reviewing is human-generated or generated by a Large Language Model (LLM)—a form of generative artificial intelligence (GenAI)—as it is increasingly difficult to distinguish between the two (Sadasivan et al., 2023). GenAI text detectors, which purport to verify author authenticity, are unreliable when detecting LLM outputs, due to the emergence of improved paraphrasing models and GenAI tools that humanise GenAI text (Paullet et al., 2025; Sadasivan et al., 2023). As a consequence, the oral defence becomes an even more critical checkpoint. If GenAI has written the manuscript, then it is highly likely that GenAI developed the candidate’s defence presentation. Given the limitations in detecting GenAI text

through technological means alone, the oral defence becomes an even more essential means of evaluating authentic doctoral-level competency.

Eight hundred years ago, in medieval universities, students participated in public disputation, where they addressed a thesis or dissertation. According to Kruse (2006), the thesis was initially laid out orally but later took the form of a written poster. Students would defend or oppose the thesis, and an arbiter would decide the outcome. By the 16th and 17th centuries, the thesis was disseminated in a pamphlet that participants and the audience would read in advance. Russell (2002) argues that the tradition of using texts primarily as a starting point for oral discussion continued into the 19th century.

The invention and adoption of print led to the decline of the medieval oral system of learning. Cheap and reliable printing made it easier to claim authorship and thus ownership of knowledge. By the end of the 18th century, doctoral candidates were required to produce a written thesis (Malone, 1981), which was then presented and defended orally (Willis et al., 2010). The written thesis supplanted oral disputation as the definitive measure of doctoral scholarship. Just as the printing press fundamentally altered the relationship between oral and written forms of doctoral assessment, GenAI-assisted writing now challenges assumptions about authorship and intellectual ownership that have governed thesis scholarship for over two centuries, potentially necessitating a rebalancing toward oral examination.

The use of GenAI is not inherently problematic; rather, the ethical concern centres on doctoral candidates' misrepresentation of their scholarship through failure to properly disclose and attribute GenAI contributions. Given that GenAI assistance may have shaped the product, the oral defence assumes heightened significance as the primary way of evaluating the doctoral candidate's authentic intellectual engagement and comprehension of their unique research. In the GenAI era, interactive oral assessments have emerged as authentic, experiential forms of evaluation that reveal students' capacity to justify their reasoning, apply knowledge across contexts, and adapt to unscripted challenges in real-time (Sotiriadou et al., 2020). As doctoral assessment literacy becomes increasingly critical, committee members/supervisors require professional development to first probe beyond surface confirmations toward exploratory questions that reveal authentic threshold crossing (Carless & Boud, 2018) and second, to be cognisant of equity issues such as how anxiety, neurodivergence, language differences, and

cultural communication styles can affect oral performance, ensuring that assessments measure genuine scholarly competency rather than privileging particular modes of verbal expression (Pearson & Brew, 2002).This positions the oral defence not only as summative gatekeeping but also as the ultimate learning development checkpoint—distinguishing those who have achieved genuine conceptual transformation from those who have merely assembled GenAI-mediated text.

Committee members/supervisors require a mindset shift to that of a learning developer who actively scaffolds candidates' intellectual growth throughout the doctoral journey (Hilsdon, 2011). This developmental approach positions committee members/supervisors as 'critical friends' who should model scholarly thinking and create low-stakes opportunities for candidates to articulate, defend, and refine their reasoning (Kensington-Miller et al., 2015). Mock defences remain the most valued preparation method, with doctoral graduates consistently identifying practice defences as their most helpful preparation strategy (Lansoght, 2022). However, learning developers should extend practice defences beyond rehearsal, guiding candidates in developing metacognitive awareness—the ability to monitor, evaluate, and articulate their own thinking processes (Teng & Mizumoto, 2024). By explicitly teaching candidates to externalise their reasoning processes, committee members/supervisors as learning developers build the capacities that distinguish human scholarly judgment from GenAI text.

As GenAI reshapes academic writing, the written thesis can no longer stand alone as reliable evidence of doctoral capability. The oral defence must evolve from a ceremonial rite of passage to a rigorous intellectual examination that reveals not just what doctoral candidates know, but how they think, respond to challenges, and exercise scholarly judgment. Committee chairs/supervisors must reconceptualise their role from evaluator to learning developer, embedding formative oral assessments throughout the doctoral journey, which build candidates' evaluative judgement and metacognitive awareness. By positioning themselves as learning developers rather than gatekeepers, committee chairs/supervisors can simultaneously maintain rigorous standards while fostering candidates' capacity to distinguish their scholarly voice from GenAI content. In returning to the primacy of oral examination, we are not retreating to the past but adapting to a future where the human capacities revealed through dialogue become our most reliable measure of doctoral achievement.

Disclosure statement

The author used the following generative AI tool in the preparation of this manuscript: Grammarly. The tasks performed by Grammarly were limited to proofing of grammar and punctuation. The author has complied with the journal's principles of AI use.

Funding

The author received no funding for this research and declares no competing interests.

References

Barnett, J. V., Harris, R. A., & Mulvany, M. J. (2017). A comparison of best practices for doctoral training in Europe and North America. *FEBS Open Bio*, *7*(10),1444–1452. https://doi.org/10.1002/2211-5463.12305

Boud, D., & Dawson, P. (2023). What feedback literate teachers do: An empirically-derived competency framework. *Assessment & Evaluation in Higher Education*, *48*(2), 158–157. https://doi.org/10.1080/02602938.2021.1910928

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8),1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Carlin, D. B., & Perlmutter, D. D. (2006, September 5). Advising the new adviser. *The Chronicle of Higher Education*, *53*(3), 66. https://www.chronicle.com/article/advising-the-new-adviser/#:~:text=By%20Diana%20B.,unlimited%20time%20to%20helping%20you

Hill, G., & Vaughan, S. (2018). Conversations about research supervision – enabling and accrediting a Community of Practice model for research degree supervisor development. *Innovations in Education and Teaching International*, *55*(2), 153–167. https://doi.org/10.1080/14703297.2017.1406388

Hilsdon, J. (2011). What is learning development? In P. Hartley, J. Hilsdon, C. Keenan, S. Sinfield, & M. Verity (Eds.), *Learning development in higher education* (pp. 13–27). Palgrave Macmillan.

Kensington-Miller, B., Renc-Roe, J., & Morón-García, S. (2015). The chameleon on a tartan rug: Adaptations of three academic developers' professional identities. *International Journal for Academic Development*, *20*(3), 279–290. https://doi.org/10.1080/1360144X.2015.1047373

Kruse, O. (2006). The origins of writing in the disciplines: Traditions of seminar writing and the Humboldtian ideal of the research university. *Written Communication*, *23*(3), 331– 352. https://doi.org/10.1177/0741088306289259

Lansoght, E. O. L. (2022). Effectiveness of doctoral defense preparation methods. *Education Sciences*, *12*(7), Article 473. https://doi.org/10.3390/educsci12070473

Malone, T. L. (1981). *A history of the Doctor of Philosophy dissertation in the United States, 1861–1930* [Doctoral dissertation, Wayne State University]. ProQuest Dissertations & Theses. https://www.proquest.com/openview/d98263ea9b2641c14aec00114ae769f5/1?pq-origsite=gscholar&cbl=18750&diss=y

Mirick, R. G., Davis, A., & Wladkowski, S. P. (2020). Social work dissertation committee chairs' perceptions of their role. *Journal of Social Work Education*, *56*(1), 155–169. https://doi.org/10.1080/10437797.2019.1642273

Nir, A., & Bogler, R. (2021). International examiners' participation of the viva: A ritual or an actual indicator of research quality? *Quality Assurance in Education*, *29*(2–3),189–203. https://doi.org/10.1108/QAE-01-2021-0005

Paullet, K., Pinchot, J., Kinney, E., & Stewart, T. (2025). Precision check: A critical look at the reliability of AI detection tools. *Issues in Information Systems*, *26*(3), 401–412. https://doi.org/10.48009/3_iis_2025_2025_132

Pearson, M., & Brew, A. (2002). Research training and supervision development. *Studies in Higher Education*, *27*(2), 135–150. https://doi.org/10.1080/03075070220119986c

Russell, D. R. (2002). *Writing in the academic disciplines: A curricular history*. Southern Illinois University Press.

Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can AI-generated text be reliably detected? *arXiv*, Article 2303.11156. https://doi.org/10.48550/arXiv.2303.11156

Sotiriadou, P., Logan, D., Daly, A., & Guest, R. (2020). The role of authentic assessment to preserve academic integrity and promote skill development and employability. *Studies in Higher Education*, *45*(11), 2132–2148. https://doi.org/10.1080/03075079.2019.1582015

Teng, M. F., & Mizumoto, A. (2024). Validation of metacognitive knowledge in vocabulary learning and its predictive effects on incidental vocabulary learning from reading. *International Review of Applied Linguistics in Language Teaching*, *63*(3), 1941–1972. https://doi.org/10.1515/iral-2023-0294

Wichmann-Hansen, G., Eika, B., & Mørcke, A. M. (2020). What skills do PhD supervisors need? A systematic review. *The Clinical Teacher*, *17*(4), 396–402. https://doi.org/10.1111/tct.13122

Willis, J., Inman, D., & Valenti, R. (2010). *Completing a professional practice dissertation: A guide for doctoral students and faculty*. Information Age Publishing.

World Population Review. (2025). *Doctorates awarded by country*. https://worldpopulationreview.com/country-rankings/doctorates-awarded-by-country

Author details

Valerie Storey holds degrees from the University of Leeds, and the University of Manchester, and a PhD from Peabody College, Vanderbilt University. With 30 years of secondary school experience in the UK, including administrative roles at the school and district levels, she transitioned into higher education to support aspiring school leaders. She has served as president of the Florida Association of Professors of Educational Leadership and held roles within related professional organisations. Her research spans leadership preparation, andragogy, and EdD programme design, resulting in over 50 peer-reviewed articles, three books, and four edited volumes.

ORCID

Valerie Storey https://orcid.org/0000-0001-8868-0802