ASCILITE 2026

The Future is Connection:

Collaboration, Creativity and Compassion in a Digital Society

Primed but sceptical: Student engagement, performance, and perceptions of AI-generated lecture primers

Miriam Edwards, Patrick Kelly, Sean Pinder

University of Melbourne

This study investigates whether written, audio, human discussion, or AI-generated lecture primer formats differentially affect student engagement, academic performance, and perceptions of efficacy among first-year undergraduate finance students. Engagement was universally low declining sharply after week one. Whilst no differences in engagement were observed in week one, Audio attracted significantly greater engagement than AI Discussion by week two, consistent with evidence that students remain sceptical of AI-generated content. Primer group allocation was not associated with differences in academic performance, supporting the integrity of group assignment. Sustained engagement beyond week one was associated with better performance: those engaging with AI Discussion and Audio primers after week one performed better on the mid-semester test and final examination respectively, suggesting these formats appealed disproportionately to higher-performing students. Perceptions of accuracy were highest for Audio, whilst AI Discussion users were least confident in content reliability. Importantly, for educators considering AI-generated preparatory materials all primer formats were rated equally and positively across key efficacy dimensions, suggesting students valued priming itself over delivery format. This suggests that AI-generated primers represent a viable and scalable alternative to human-produced equivalents, though transparency about their provenance remains advisable given residual student scepticism about reliability.

Keywords: Lecture primers, generative AI, student engagement, AI scepticism, student perceptions

Introduction

The positive impact of student engagement with written pre-lecture review material has been well established in the literature in the areas of academic performance (Gammerdinger & Kocher, 2018; Moravec et al., 2010) as well as student perceptions of their preparedness for success in a tertiary education setting (Balakumar et al., 2024). In recent years there has been significant evolution in the format that this pre-lecture material, hereafter referred to as "lecture primers," has taken. For example, Popova et al. (2014) analyse the impact of using audio primers in universities, and report that students perceive significant benefits that arise via three conduits: assistance in linking new material with existing knowledge, the development of understanding of lecture topics, and as an aide to stimulating deeper thinking about lecture content and subsequent applications.

In recent years we have seen a proliferation of opportunities to utilise generative AI-facilitated materials in university courses in different settings. In terms of the impact of this tool on student attainment, there have been a number of recent meta-analyses that have documented the generally positive effect that generative AI can have on learning outcomes (Ma & Zhong, 2025; Pallant et al. 2026). A quite separate issue though is student perceptions of the efficacy of generative AI-produced materials. Here the results are a little more mixed. For example, Er et al. (2025) report that students perceive human-instructor feedback as more useful than AI feedback and that students receiving human-instructor feedback performed significantly better than students receiving AI feedback even after accounting for initial knowledge levels. Similarly, Thomas et al. (2025) report significantly higher levels of student criticism of AI scoring of assessment relative to human-instructor scoring. With respect to generative AI-generated content, as opposed to feedback, students seem a little more optimistic about the role of generative AI. For example, students in an English for Academic Purposes course rating highly materials generated by Google Notebook LM (Woo et al., 2026), medical students responding favourably to question sets generated with ChatGPT (Tate et al. 2025) and Hungarian and Spanish university students valuing the adaptability and continuous guidance provided by AI tutors over their human counterparts (Hajeer et al. 2024).

In the current study, we focus on lecture primers to enable us to leverage off an extant literature that has documented their efficacy in student learning (e.g., Moravec et al., 2010; Popova et al., 2014) and investigate the following three research questions. First, are students more likely to engage with lecture primers that are (a) written primers written by the lecturer, (b) an audio version of the written primer read by the lecturer (c) a discussion between two lecturers covering the topics in the primer, or (d) a discussion generated by AI? Second, does the link between performance and engagement with lecture primers differ across primers? Third, how do student perceptions of accuracy, reliability and efficacy vary across the different lecture primer formats? We find that sustained engagement with primers beyond week one was positively associated with academic performance, with some cross-sectional difference in primer type reported although these relationships likely reflect self-selection by higher-performing students rather than a direct effect of the primers themselves. Whilst the "AI Discussion" users were least confident in content reliability, all primer formats were rated equally and positively across key efficacy dimensions, suggesting students valued priming itself over delivery format.

Experimental design and data

This study received ethics approval from the university's Human Research Ethics Committee (ID number: 33864). In July 2025, prior to the beginning of semester, students enrolled in the large first year Bachelor of Commerce subject Principles of Finance were allocated to one of five sample groups on the basis of the final digit of their student ID, where each group was given access to a different lecture primer prior to the weekly lecture for the first six weeks of semester. The "Written Primer" group (*n* =301) were given access to a one-page description of the background and motivation for the upcoming lecture written by the subject coordinator; the "Audio Primer" group (*n*=324) access to an audio recording of the coordinator reading the written primer, and the "Human Discussion" group (*n* =351) were provided a 3-4 minute audio recorded discussion involving the coordinator of the subject and an experienced academic colleague where their discussion was prompted by the written primer. The fourth group was the "AI Discussion group" (*n* =314) who were given access to a 2-4 minute podcast-style audio recording generated using the Google Notebook LM AI interface that had been provided a copy of only the written primer and a short list of prompts (e.g., sound enthusiastic, limit time to four minutes, do not directly refer to the document). A fifth group of students formed a control sample (*n* =330) and were not provided with any primer at all. Students could access the primers by logging into the university's Learning Management System (LMS). Students could not directly access any primers other than the one for the group to which they were assigned. For each student in the sample, we subsequently tracked statistics relating to weekly engagement with the platform as well as performance in the two substantive pieces of assessment in the semester: the mid-semester test and final exam contributing 25% and 60% to the students' final mark in the subject. We followed this up with a survey of students in the four subsamples who had access to a lecture primer each week to gauge their thoughts on the reliability of the information provided in the primer as well as different dimensions of the efficacy of the primers (Understanding/Relevance/Connections/Excitement/Focus/Familiarity with lecturer/Preparation).

Results

Engagement

In Table 1 we report the proportion of students that accessed a primer in the week described in the column title. We then conduct pairwise tests between all groups except the No Primer group (e), testing for a difference in engagement levels and where a difference is found, the superscript refers to the relevant primer that the significant statistic relates to. The only significant within-week difference is in week 2, where engagement with the AI Discussion (d), at 9.6%, is significantly lower (at the 5% level) than 15.4% for the Audio Primer (b).

Table 1
Weekly engagement statistics

		Week					
Primer type	**Count**	**1**	**2**	**3**	**4**	**5**	**6**
a. Written Primer	301	19.6	12.6	3.3	6.0	3.3	3.0
b. Audio Primer	324	20.7	15.4[d]	3.1	3.4	3.1	1.9

c. Human Discussion	351	19.7	13.7	4.6	4.3	2.8	1.4
d. AI Discussion	314	17.5	9.6[b]	2.9	4.8	4.1	1.0
e. No Primer	330	0.0	0.0	0.0	0.0	0.0	0.0

As can be seen from Table 1, a minority of students accessed any of the primers made available with initial engagement ranging from 17.5% of a sub-group (AI Discussion) to 20.7% of a sub-group (Audio). While there was no statistical difference in initial engagement statistics across the four sub-groups, by the end of the second week of class the proportion of students accessing audio primers was statistically greater than the proportion of students accessing the AI Discussion primers. One possible interpretation of this result is that it is consistent with other evidence of students having greater initial scepticism about AI-produced relative to human produced resources (Cabrera-Arnau, 2024); although this is a theme that will be considered more deeply in light of student survey responses relating to accuracy and reliability of the primers provided.

Performance

Table 2 shows that gains associated with engaging with primers are concentrated among students who took up a primer after week 1 rather than among those who merely had access or who started in week 1. Table 2 reports this as OLS regressions of the mid-semester test score (out of 25) and the final exam score (out of 100) on dummies for primer access (rows a–d), engagement from week 1 (rows e–h) and engagement from a later week (rows i–l), with students who had no primer as the omitted group. Average progressive-assessment scores to control for skill and unreported lecture-stream dummies are included because the four streams sat different mid-semester tests. *, **, *** indicate the statistical significance of a coefficient with a p-value of less than or equal to 10%, 5%, or 1%, respectively. Any coefficients within common groupings are tested for significant differences from other coefficients using a Wald test. For example, for coefficient g (Human Discussion engaged – week 1), the superscript “e5” indicates that it is significantly different to coefficient e (Written engaged – week 1) at the 5% level.

Taking up the audio primer later in the semester (row j) is worth 1.17 marks (4.68%) on the mid-semester test and 4.74 (4.74%) on the final exam, each significant at the 10% level with the 95% confidence interval also reported. Taking up the AI Discussion primer (row l) translated to 1.54 (6.16%) and 6.38 (6.38%) more marks, significant at the 5% level. This result is consistent with Julius and Nancy (2025) who also present evidence of higher-performing students more frequently engaging with AI-linked resources, although neither group significantly outperformed other students with primer access. Oddly, those that listened to the Human Discussion perform worse with a coefficient of -1.28; however, the positive and significant 1.59 coefficient on engagement in any other week tells us that this poor performance is concentrated among the roughly 6% of students who do not engage with subsequent primers (note the drop in Table 1 from 19.7% engagement to 13.7% engagement in week 2).

Table 2
Primer engagement and assessment performance

	Mid-semester test		**Final exam**	
	Coefficient	**95% CI**	**Coefficient**	**95% CI**
Intercept	11.64***	[10.14, 13.14]	18.03***	[12.05, 24.01]
a. Written group	0.28	[-0.55, 1.11]	1.31	[-1.72, 4.34]
b. Audio group	0.01	[-0.81, 0.84]	2.35	[-0.62, 5.31]
c. Human Discussion group	0.35	[-0.45, 1.16]	1.84	[-1.11, 4.79]
d. AI Discussion group	0.17	[-0.65, 0.99]	0.96	[-2.03, 3.95]
e. Written engaged – week 1	0.80[g5]	[-0.68, 2.29]	3.14	[-2.20, 8.49]
f. Audio engaged – week 1	1.03[g5]	[-0.35, 2.41]	2.31	[-2.66, 7.29]
g. Human Discussion engaged – week 1	-1.28*[,e5,f5]	[-2.63, 0.07]	4.07	[-0.90, 9.05]
h. AI Discussion engaged – week 1	-0.45	[-1.95, 1.06]	-0.80	[-6.20, 4.60]
i. Written engaged – other wk	-0.09[k10]	[-1.59, 1.41]	1.90	[-3.45, 7.25]
j. Audio engaged – other wk	1.17*	[-0.20, 2.54]	4.74*	[-0.18, 9.65]
k. Human discussion engaged – other wk	1.59**[,i10]	[0.28, 2.89]	2.96	[-1.75, 7.67]
l. AI Discussion engaged – other wk	1.54**	[0.10, 2.99]	6.38**	[1.23, 11.52]
Progressive task score	0.45***	[0.36, 0.55]	1.84***	[1.45, 2.23]

Adjusted R^2	0.06	0.06
N	1,617	1,539

Perceptions of primer accuracy, reliability and efficacy

We surveyed students at the end of the semester and found that overall, students felt "confident that the information provided by the primer was accurate and reliable" though significantly more were confident about the instructor read primer than either the two-person discussion or the AI generated primer. Regardless of primer, students believe that primers: helped them improve understanding, were relevant, connected new content to old, built excitement to learn, improved focus, made them feel better acquainted with the lecturer, and helped with preparation. We see this first result in the row labelled "Accuracy & reliability" (Table 3). For example, 78% thought that the written primer was accurate and reliable, significantly more than those who did not using a Wilson binomial test, with the 95% confidence interval ranging from 61.9% to 88.3% believing this primer is reliable. The chi-squared test suggests that we can reject the hypothesis that there is no difference between the distribution of responses amongst the primer groups. Further, pair-wise Fisher tests suggest that there are significant differences between the distribution of responses for Audio and Human Discussion primers and between Audio and AI Discussion primers. This suggests that students may find it more difficult to assess the accuracy and reliability of multi-party discussions as opposed to monologues.

Table 3
Student perceptions of primer accuracy, reliability and efficacy

Dimension	**a. Written**	**b. Audio**	**c. Human Discussion**	**d. AI Discussion**	**Joint equivalence**
Accuracy & reliability	78%***	91%***c10,d10	81%***b10	75%***b10	χ^2 = 11.340*
[95% CI]	[61.9, 88.3]	[77.6, 97.0]	[65.0, 90.2]	[56.6, 87.3]	
i. Understanding	3.82***	3.86***	3.80***	3.88***	χ^2 = 0.619
[95% CI]	[3.39, 4.21]	[3.46, 4.21]	[3.50, 4.07]	[3.52, 4.20]	
ii. Relevance of topics	3.96***	4.07***	3.83***	3.92***	χ^2 = 1.893
	[3.57, 4.32]	[3.68, 4.43]	[3.53, 4.13]	[3.52, 4.28]	
iii. Connections with	4.04***	3.82***	3.67***	3.68***	χ^2 = 3.168
existing knowledge	[3.71, 4.36]	[3.43, 4.18]	[3.37, 3.93]	[3.28, 4.04]	
iv. Excitement to	3.96***	3.68***	3.67***	3.64***	χ^2 = 2.125
learn new things	[3.54, 4.36]	[3.29, 4.07]	[3.30, 4.00]	[3.20, 4.08]	
v. Focus on lecture	4.04***	3.89***	3.60***	3.92***	χ^2 = 4.115
content	[3.64, 4.39]	[3.46, 4.25]	[3.23, 3.93]	[3.56, 4.24]	
vi. Lecturer familiarity	3.86***	3.71***	3.57***	3.88***	χ^2 = 2.765
	[3.39, 4.29]	[3.29, 4.11]	[3.25, 3.87]	[3.56, 4.20]	
vii. Preparation for	4.18***	3.96***	3.77***	4.04***	χ^2 = 5.746
classes	[3.79, 4.54]	[3.57, 4.32]	[3.45, 4.07]	[3.68, 4.36]	
Number of responses	36	35	36	28	

In the second part of Table 3 we report the average student response score on a five point Likert scale across seven (i. to vii.) different aspects of the efficacy of each primer group. While a bootstrap test with 10,000 resamples for the difference of the mean score from 3, "neither agree nor disagree," indicated with *, **, and *** for significance at the 10%, 5%, and 1% levels, supports the rejection of the null hypothesis that the mean response to each prompt is "neither agree nor disagree" which implies that, on average, respondents responded positively to each prompt. At the same time, however, neither the Kruskal-Wallis test results nor Nemenyi pairwise comparisons (Liu & Chen, 2012) suggest statistically significant differences in the responses across the different primer groups.

Discussion and conclusion

Engagement with primers was low across all formats and declined sharply after week one, consistent with the broader literature on voluntary pre-class preparation (Gammerdinger & Kocher, 2018; Moravec et al., 2010). The modest differential drop-off between Audio and AI Discussion primers by week two is consistent with evidence of student scepticism towards AI-generated content (Cabrera-Arnau, 2024; Er et al., 2025), though this gap had largely disappeared by week three. Where performance associations emerged, they were tied to

sustained engagement beyond week one and most plausibly reflect self-selection by higher-performing students rather than a direct instructional benefit of any particular format (Julius & Nancy, 2025). Notwithstanding modest differences in perceived reliability, all primer formats were rated positively and equivalently across all seven efficacy dimensions, suggesting that the perceived benefits of pre-lecture preparation are robust to delivery mode. For educators considering AI-generated preparatory materials, this is an encouraging finding: AI-generated primers represent a viable and scalable alternative to human-produced equivalents, though transparency about their provenance remains advisable given residual student scepticism about reliability.

References

Balakumar, P., Fice, H., Richer, S., Western, T. L., & Yao, J. (2024). Students priming students for success: students as partners in development of pre-lecture resources for introductory genetics. *International Journal for Students as Partners*, *8*(1), 200-214. https://doi.org/10.15173/ijsap.v8i1.5471

Cabrera-Arnau, C. (2024). Understanding student perceptions on the use of generative AI tools in UK higher education. *preprints / SocArXiv / fs96d_v1* https://doi.org/10.31235/osf.io/fs96d

Er, E., Akçapınar, G., Bayazıt, A., Noroozi, O., & Banihashem, S. K. (2025). Assessing student perceptions and use of instructor versus AI-generated feedback. *British Journal of Educational Technology*, *56*(3), 1074-1091. https://doi.org/10.1111/bjet.13558

Gammerdinger, W. J., & Kocher, T. D. (2018). Understanding student perceptions and practices for pre-lecture content reading in the genetics classroom. *Journal of Microbiology and Biology Education*, 19(2). https://doi.org/10.1128/jmbe.v19i2.1371

Hajeer, A., Papp-Váry, Á., & Pólya, É. (2024). AI tutors vs. Human instructors: Perceptions of higher education students in Hungary and Spain. *Edutec, Revista Electrónica de Tecnología Educativa*, (89), 105-120. https://doi.org/10.21556/edutec.2024.89.3523

Julius, A., & Nancy, M. (2025). The digital crossroads: A comparative analysis of OpenAI and Google AI for enhancing learning among Gen Z in Ugandan private universities. *International Journal of Academic Pedagogical Research*, *9*(10), 84-92.

Liu, Y., & Chen, W. (2012). A SAS macro for testing differences among three or more independent groups using Kruskal-Wallis and Nemenyi tests. J*ournal of Huazhong University of Science and Technology [Medical Sciences]*, 32(1), 130-134. https://doi.org/10.1007/s11596-012-0023-9

Ma, N., & Zhong, Z. (2025). A meta-analysis of the impact of generative artificial intelligence on learning outcomes. *Journal of Computer Assisted Learning*, *41*(5), e70117. https://doi.org/10.1111/jcal.70117

Moravec, M., Williams, A., Aguilar-Roca, N., & O'Dowd, D. K. (2010). Learn before lecture: A strategy that improves learning outcomes in a large introductory biology class. *CBE—Life Sciences Education*, *9*(4), 473-481. https://doi.org/10.1187/cbe.10-04-0063

Pallant, J. L., Blijlevens, J., Campbell, A., & Jopp, R. (2026). Mastering knowledge: The impact of generative AI on student learning outcomes. *Studies in Higher Education*, *51*(4), 714-735. https://doi.org/10.1080/03075079.2025.2487570

Popova, A., Kirschner, P. A., & Joiner, R. (2014). Effects of primer podcasts on stimulating learning from lectures: How do students engage?. *British Journal of Educational Technology*, *45*(2), 330-339. https://doi.org/10.1111/bjet.12023

Tate, H., Bicknell, B. T., Fiore, P., Galloway, J., Borak, J., & Brooks, W. S. (2025). Medical students' perceptions of AI-generated practice questions as learning tools. *Journal of Investigative Medicine*, 74(7), 663-669. https://doi.org/10.1177/10815589251406265

Thomas, M. L., Yildirim-Erbasli, S. N., & Hariharan, S. (2025). Exploring undergraduate students' perceptions of AI vs. human scoring and feedback. *The Internet and Higher Education*, 101052. https://doi.org/10.1016/j.iheduc.2025.101052

Woo, D. J., Wang, D., & Guo, K. (2026). AI Slop or AI-enhancement? Student perceptions of AI-generated media for an English for Academic Purposes course. *arXiv preprint arXiv:2605.16275*. https://doi.org/10.48550/arXiv.2605.16275

ASCILITE 2026

The Future is Connection:

Collaboration, Creativity and Compassion in a Digital Society

Edwards, M., Kelly, P. & Pinder, S. (2026). Primed but sceptical: Student engagement, performance, and perceptions of AI-generated lecture primers. In Cochrane, T., Narayan, V., Poole, T., Lai, P., Lake, P., Moore, O., & Li, L. (Eds.), *The Future is Connection: Collaboration, Creativity and Compassion in a Digital Society*. Proceedings ASCILITE 2026. Brisbane (pp. 569–574). https://doi.org/10.65106/apubs.2026.3177

Note: All published papers are refereed, having undergone a double-blind peer-review process.
The author(s) assign a Creative Commons by attribution license enabling others to distribute, remix, tweak, and build upon their work, even commercially, as long as credit is given to the author(s) for the original creation.

© Edwards, M., Kelly, P. & Pinder, S. 2026