AI Researchers
Embedding Responsible AI in Research

AI READINESS TOOLKIT

UC5-03 Use case 5: Paper presentations

Would you use this AI-generated research presentation?

A peer-led discussion pack for evaluating AI-supported research slides

AT A GLANCE

A peer-led discussion built around an AI-generated slide deck that mixes accurate content with oversimplifications, unsupported claims and misleading visuals. In pairs or groups, participants review six slides against a short evidence summary of an open-access study and decide whether to retain, revise, reject or verify each one. The pack includes the sample slides, a review sheet and facilitator notes.

TYPE
Peer-led discussion

FORMAT
Discussion pack (PDF), sample slides (PPTX and PDF)

CONTRIBUTING PARTNER
University of Lodz

WHO IT IS FOR
Early-career researchers, supervisors and research groups

FRAMEWORK BARRIERS
Trust in AI, Know-how, This task

HUMAN-ONLY BOUNDARY

AI may help suggest slide outlines, summaries and visual structures. Scientific content, evidential claims, interpretation, limitations and delivery remain the researcher's responsibility.

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Purpose and learning objectives

This activity asks participants to review six deliberately flawed AI-generated slides based on a real published article. The task is not to judge whether the slides look good, but whether they faithfully represent the paper.

> distinguish visual polish from scientific accuracy;
>
> recognise causal overstatement, sample distortion, model simplification, visual distortion and limitation loss;
>
> evaluate presentation claims against a source paper;
>
> decide whether each AI-generated slide should be Retained, Revised, Rejected or Verified;
>
> explain why final scientific responsibility cannot be delegated to AI.

Source paper and licensing note

The exercise is based on the following open-access article: van Niekerk, J., Delport, P. M. J., & Sutherland, I. (2025). Addressing the use of generative AI in academic writing. Computers and Education: Artificial Intelligence, 8, 100342. https://doi.org/10.1016/j.caeai.2024.100342

> Reuse note: The article is published under a CC BY-NC-ND licence. This toolkit resource therefore uses a short factual evidence summary and original training slides created for this activity. It does not reproduce or modify the article's original figures, tables or extended text.

Evidence sheet for participants

Participants use this evidence sheet to check the AI-generated slides. The slide deck should not be approved from memory or appearance alone.

ELEMENT	EVIDENCE FROM THE PAPER
Research focus	Whether an active-learning intervention could reduce students' behavioural intention to rely inappropriately on ChatGPT for academic writing.
Theoretical frame	Technology Acceptance Model: Perceived Usefulness (PU), Perceived Ease of Use (PEoU) and Behavioural Intention (BI).
Setting	Academic writing skills course at Noroff University College in Norway, involving undergraduate computing students.
Student cohort	173 students were enrolled in the course and participated in the educational intervention.
Survey participation	Pre-intervention survey: N = 159. Post-intervention survey: N = 95. Survey participation was voluntary and anonymous.
Intervention design	Students generated ChatGPT essays, completed a voluntary pre-survey, evaluated AI-generated writing against academic-writing criteria, reflected on the activity and completed a voluntary post-survey.
Writing criteria	Structure/flow, evidence/references, critical insight, balanced/unbiased writing, conciseness, objectivity and formal tone.

ELEMENT	EVIDENCE FROM THE PAPER
Analysis	Cronbach's alpha for internal consistency and McNemar-Bowker tests for changes in pre/post Likert-scale response distributions.

FINDING AREA	WHAT CHANGED
Evidence and references	BI to use ChatGPT for evidence/references dropped from 56.46% to 15.39%; this was reported as statistically significant.
Critical insight	BI to use ChatGPT for critical insight dropped from 48.33% to 20.88%; this was reported as statistically significant.
Structure and flow	BI dropped from 65.58% to 45.65%, but students continued to see value in ChatGPT for structure and flow.
Balanced and unbiased writing	BI changed from 51.03% to 31.87%, but the change was reported as not statistically significant.
Objectivity	BI to use ChatGPT for objective presentation dropped from 49.29% to 20.46%; this was reported as statistically significant.
Formal tone	BI dropped from 67.67% to 44.82%, but many students still valued ChatGPT for formal language support.

> Important limitations: The study used one institution, one cohort and one academic year; pre/post survey attrition was substantial; detailed demographic and prior-AI-exposure data were not collected; long-term persistence of the effect was not measured; the intervention used ChatGPT 3.5, so transfer to newer tools requires caution.

Activity route

TIME	ACTIVITY
0-5 min	Introduce the purpose, the source paper and the human-only boundary.
5-10 min	Participants read the evidence sheet.
10-25 min	Pairs/groups review the six AI-generated slides and complete the review sheet.
25-35 min	Groups compare Retain / Revise / Reject / Verify decisions.
35-40 min	Facilitator reveals the benchmark diagnosis.
40-45 min	Complete the presentation-integrity checklist and discuss transfer to participants' own slides.

Decision options

DECISION	USE WHEN
RETAIN	The slide accurately represents the paper and needs only minor editing.

DECISION	USE WHEN
REVISE	The slide idea is useful, but wording, scope, evidence or visual design must be corrected.
REJECT	The slide is fundamentally misleading or presents a model, claim or visual that should not be retained.
VERIFY	The slide may be acceptable, but the source summary is insufficient to approve it without checking the paper, data or analysis.

Six AI-generated slide audit cases

AI-GENERATED SLIDE	MAIN ISSUE	BENCHMARK DECISION	DEFENSIBLE DIRECTION
1. "Active learning eliminates inappropriate ChatGPT use"	Overstates the intervention and implies elimination of behaviour. The study measured perceptions and BI, not actual long-term use.	REVISE	Rephrase: "The intervention was associated with reduced behavioural intention to use ChatGPT for several academic-writing tasks."
2. "173 students completed the pre/post study"	Confuses the course cohort with survey response numbers. There were 173 students in the course, 159 pre-survey responses and 95 post-survey responses.	REJECT / major REVISE	Use separate labels for cohort, educational intervention participation and voluntary survey responses.
3. "Training → Less ChatGPT use"	Oversimplifies TAM and removes PU, PEoU, BI and task-specific variation. The study was not a simple training-causes-behaviour design.	REJECT	Show the intervention as affecting perceptions of usefulness/ease of use and behavioural intention, with task-specific differences.
4. "Intervention impact (%)" chart	Treats change values or Likert shifts as percentages/effect sizes. It also omits the balanced/unbiased characteristic, making the chart look more comprehensive than it is.	REJECT	Use a cautious before/after distribution summary or show selected BI percentages with clear labels. If only selected characteristics are shown, label the selection and explain the omission.
5. "Students no longer want to use ChatGPT"	Ignores nuance: students remained comparatively positive about structure, concision and formal tone while becoming more sceptical about references and critical insight. It also selectively lists only five of the seven writing characteristics.	REVISE	Present a selective-use message: stronger caution for evidence/critical insight, continued value for structure and language support. Avoid implying comprehensive coverage if categories are omitted.
6. "Universities should adopt this intervention everywhere"	Makes a broad policy and long-term effectiveness claim. The study used one cohort and did not measure long-term persistence.	REVISE / VERIFY	Frame as a promising example of active-learning-based AI literacy, requiring adaptation and further evidence in other settings.

Participant review sheet

SLIDE	EVIDENCE CHECKED	DECISION	REASON	WHAT SHOULD CHANGE?
1		Retain / Revise / Reject / Verify		
2		Retain / Revise / Reject / Verify		
3		Retain / Revise / Reject / Verify		
4		Retain / Revise / Reject / Verify		
5		Retain / Revise / Reject / Verify		
6		Retain / Revise / Reject / Verify		

Error taxonomy

ERROR TYPE	TYPICAL FORM IN AI-GENERATED SLIDES
Causality inflation	associated with / was followed by -> caused / eliminated / proved
Sample distortion	course enrolment -> complete research sample
Construct drift	behavioural intention -> actual behaviour
Model simplification	TAM/intervention logic -> one-line causal chain
Visual distortion	ordinal shifts or test outputs -> percentage impact bars
Limitation loss	single cohort, attrition and tool version disappear
Evidence-to-policy leap	one study -> universal institutional recommendation
Selective omission / cherry-picking	selected writing characteristics -> implied comprehensive pattern across all tasks

Discussion questions

> Which slide looked most convincing before you checked the evidence sheet?
>
> Which error would be easiest to miss when preparing a conference presentation quickly?
>
> When is an attractive visual scientifically misleading?
>
> Which problems could be reduced by a better prompt, and which still require researcher judgement?
>
> What information about sample, method or limitations should never be simplified away?
>
> When should a slide be marked VERIFY instead of revised immediately?

Presentation-integrity checklist

CHECK AREA	CONFIRM BEFORE ACCEPTING AI-SUPPORTED SLIDES
Scientific accuracy	Numbers, sample details, constructs, method and findings match the paper.
Claims	Association, intention and measured perception have not been turned into causality, behaviour or universal adoption.
Visualisation	The chart type fits the data; scales and labels are meaningful; visuals do not imply unsupported precision.
Limitations	Sample, attrition, tool-version and long-term-effect limitations remain visible where relevant.
Human responsibility	The presenter has checked every slide against the source and can defend each claim.

Transfer task

Ask participants to choose one slide from their own research presentation and audit it using the same four questions:

> What claim does the slide make?
>
> What evidence supports the claim?
>
> What limitation or uncertainty is hidden?
>
> Should the slide be Retained, Revised, Rejected or Verified?

About this resource

TITLE	Would you use this AI-generated research presentation?
TOOLKIT REFERENCE	UC5-03, use case 5: Paper presentations
DEVELOPED BY	University of Lodz
SUGGESTED CITATION	AI Researchers (2026). Would you use this AI-generated research presentation?. AI Readiness Toolkit, UC5-03. University of Lodz. airesearchers.eu
QUALITY ASSURANCE	Reviewed by the AI Researchers consortium for quality, categorisation and alignment with the AI Engagement Framework. External peer review by senior academics is under way, and the item may be updated as a result.
VERSION	1.0, November 2026
PROJECT	AI Researchers: Embedding Responsible AI in Research. Erasmus+ Cooperation Partnership in Higher Education, KA220-HED-66835D2B. October 2025 to March 2028.
CONSORTIUM	University of Lodz (coordinator), FEP University of Porto, UPF Barcelona School of Management, ACEEU, NIPI (OPI PIB), European E-learning Institute and Momentum Marketing Services

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

AI Researchers
Embedding Responsible AI in Research