AI Researchers
Embedding Responsible AI in Research

AI READINESS TOOLKIT

UC5-01 Use case 5: Paper presentations

Designing a research presentation with AI

Presenter guide for a 15-slide session, using a published meta-analysis as the worked example

AT A GLANCE

A ready-to-present 15-slide session on turning a published paper into a research presentation with AI support. Using an open-access meta-analysis as the worked example, it shows how to request alternative talk structures, shorten slide messages and check every statistic against the source. A presenter guide gives the objective, cues and a full script for each slide. Content, claims and delivery remain the researcher's.

TYPE
Slides with presenter guide

FORMAT
Slides (PPTX and PDF), presenter guide (PDF)

CONTRIBUTING PARTNER
University of Lodz

WHO IT IS FOR
Researchers preparing a research presentation, and the trainers and supervisors who run the session

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

HUMAN-ONLY BOUNDARY
The scientific content, claims and delivery remain the researcher's.

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.

About this guide

This guide accompanies the 15-slide AI Researchers presentation. It gives the facilitator the objective, cue and key message already contained in the slide notes, followed by a complete suggested narration in italics. The narration is written to be read aloud or adapted to the audience.

Before the session

> Allow approximately 30–35 minutes for the narration and short audience pauses. Keep the published article open at pages 1, 4, 6, 7 and 8.
>
> Distinguish the explainable 30-to-29 study-selection step from the unresolved numerical difference between Results section 5.2 and Figure 2. Do not announce one exact pooled estimate as verified.
>
> The prompts, draft statements and alternative talk structures in the deck are teaching examples. They are not outputs documented by Dong, Tang and Wang. The italic paragraphs below are the words the facilitator may speak.

Slide by slide guidance

Slide 1: Designing a research presentation with AI

OBJECTIVE	Introduce the responsible slide workflow through one published meta-analysis.
FACILITATOR CUE	Show the article citation at the end. The paper is open access under CC BY 4.0. Tell participants that the deck deliberately preserves a discrepancy between the Results text and Figure 2 as a verification task.
KEY MESSAGE	The article is the evidence source; AI suggestions remain provisional.
ARTICLE LOCATION	Title page and copyright statement on p. 1.

> **TEXT TO READ ALOUD**
>
> *Welcome. Today we will practise turning a research paper into a presentation with the help of AI. The goal is not to press a button and accept a finished deck. We will see which parts of the work AI can help with, where we need to read the paper ourselves, and how to catch a claim that sounds convincing but goes beyond the evidence. We will use one published meta-analysis by Dong, Tang and Wang about AI and students' academic achievement. First, we will decide who the audience is and what this particular talk should achieve. Then we will select evidence from the paper, ask AI for two possible structures, and turn sections into messages that work on slides. Finally, we will check wording, visuals, speaker notes and the final scientific claim. There is also a real numerical inconsistency in this article. We will use it to practise a useful response: pause, mark what is uncertain, and return to the source. You do not need to know every statistical term at the start. I will explain each one as we reach it.*

Slide 2: What AI can and cannot do

OBJECTIVE	Set the human and AI boundary using an actual inconsistency.
FACILITATOR CUE	Invite participants to say whether a model should choose the more attractive of two effect estimates. Both positive findings are in the paper, but the mismatch in the exact pooled estimates needs a source-level resolution.
KEY MESSAGE	The researcher checks and authorises all claims, figures and oral explanations.
ARTICLE LOCATION	Results §5.2, p. 4; Figure 2, p. 6.

TEXT TO READ ALOUD

Let us begin with the division of responsibility. AI can give us possible slide orders, shorter titles or ideas for a diagram. Those suggestions may save time, especially when a paper is long. But a neat sentence does not tell us whether the sentence is true to the source. The researcher must check the study design, the numbers, the interpretation and the words they will say aloud. This paper gives us a concrete reason. Its Results section and its forest plot both suggest a positive average effect, but they print different numerical summaries. If we ask AI to make a slide quickly, it might copy one number, mix the two, or invent an explanation for the difference. None of those actions would resolve the issue. Our task is to notice it, keep a record of where each value appears, and avoid presenting one exact estimate as settled. Keep this boundary in mind throughout the session: AI can propose a way to communicate; we decide what the evidence permits us to communicate.

Slide 3: Audience and presentation objective

OBJECTIVE	Make the audience and speaking objective explicit.
FACILITATOR CUE	With doctoral researchers, define Hedges' g briefly, but keep the main narrative focused on what the meta-analysis found and how strongly it varies across settings. For practitioners, place even more weight on the limits to transfer and less on moderator statistics.
KEY MESSAGE	A useful objective guides both the story and what to leave in speaker notes.
ARTICLE LOCATION	Research purpose and questions, p. 2; Results §5.2, p. 4; Limitations, p. 8.

TEXT TO READ ALOUD

Before we make a single content slide, we need to know who is listening. For this exercise, imagine an eight-minute presentation to doctoral researchers. They know what a meta-analysis is, but they have not read this particular paper. In eight minutes we cannot retell every paragraph, every study and every subgroup result. We need one clear objective. Here it is: by the end, listeners should understand what the authors report on average, why the results vary across studies, and why we cannot automatically apply that average to every classroom. That objective affects the level of explanation. I will briefly define the effect measure when it appears, but I will spend more time on what the result means and what remains uncertain. If our audience were school leaders, we might give less statistical detail and more context about the limits of using the findings for decisions. The evidence itself would stay the same. A useful test is whether someone could leave the room and state the finding and one important limitation in their own words.

Slide 4: The paper's central contribution

OBJECTIVE	Identify the contribution without overstating the data.

FACILITATOR CUE	The abstract and §5.1 report 29 included studies and 2,657 participants. Section 5.2 reports g = 0.924, 95% CI [0.612, 1.235], and I^2 = 93%. Figure 2 prints an overall g = 0.89 with CI 0.55–1.24 and Q df = 24. The case card explicitly labels the first value as a text result to avoid suppressing the discrepancy.
KEY MESSAGE	The paper reports a positive mean effect with substantial heterogeneity. Exact pooled values need reconciliation before final delivery.
ARTICLE LOCATION	Abstract p. 1; Results §5.1–5.2 p. 4; Figure 2 p. 6.

TEXT TO READ ALOUD

Now let us identify the paper's contribution before asking AI to summarise it. The authors bring together studies that compared educational uses of AI with approaches without those AI interventions. Their final analysis reports 29 studies and 2,657 learners. They use a meta-analysis to estimate the average result and to examine differences across settings. In the Results text, the average is expressed as Hedges' g equal to 0.924. Hedges' g is a standardised difference, so it lets researchers compare results measured with different tests. It is not a percentage and it does not mean that each student improved by 92.4 per cent. The reported 95 per cent confidence interval runs from 0.612 to 1.235 in that text. The authors also report I-squared of 93 per cent, a sign of very large variation in study results beyond what we would expect from sampling alone. We must label these as values from the text because Figure 2 gives another pooled value. The safe core message is a positive reported average accompanied by substantial variation; the exact number needs checking.

Slide 5: Selecting verified content

OBJECTIVE	Show a concrete source-to-slide selection process.
FACILITATOR CUE	The PRISMA flow diagram ends at 30 studies; §5.1 explains that one outlier was later excluded, leaving 29 for analysis. That transition should be narrated clearly, not treated as an unexplained error. By contrast, §5.2 and Figure 2 disagree in the pooled value and Q degrees of freedom, and the PDF does not give enough information to resolve that discrepancy. The between-group test for AI type is p = .78, so a ranking of type-specific estimates is not evidence of a difference between types.
KEY MESSAGE	Retain traceable items and mark conflicts rather than smoothing them away.
ARTICLE LOCATION	Methods §4.5 p. 3; Figure 1 and §5.1–5.2 p. 4; Figure 2 p. 6; Table 3 p. 7.

TEXT TO READ ALOUD

This is where we build a small evidence register for the talk. For each candidate claim, I want a location in the article. The Results section says that 29 studies were analysed after one extreme outlier was excluded. The same section reports 2,657 participants. The methods explain the random-effects model, which allows the true effects to differ across studies. These items can go into a draft, each with its section reference. Now notice two different kinds of apparent mismatch. The study-selection figure ends at 30 studies, and the later text explains why the final analysis has 29: one outlier was removed. We can describe that step. The Results text and Figure 2, however, differ in the exact pooled result and in the degrees of freedom attached to the heterogeneity statistic. The PDF does not explain that difference. We should mark the numerical claim for verification, not make up a bridge between the two. We should also hold back any claim that one type of AI is superior just because its subgroup estimate looks largest. Table 3 has a separate test for that question.

Slide 6: Requesting alternative narratives

OBJECTIVE	Demonstrate a prompt that makes uncertainty visible.
FACILITATOR CUE	The article is published open access under CC BY 4.0. The prompt asks for structures, not permission to alter results. Read the source locations in the returned outline and compare every proposed slide with the PDF. The model could still misread a forest plot or merge two different estimates, so the researcher remains the verifier.
KEY MESSAGE	Good prompts narrow the work and expose unresolved source conflicts.
ARTICLE LOCATION	Article sections and figures throughout; licence statement on p. 1.

TEXT TO READ ALOUD

Here is an example of a prompt we could give to an approved AI tool using this published paper. Notice how specific the task is. We say who the audience is and how long the talk will be. We request two outlines so that we have a choice: one follows the order of the paper, while the other begins with the research question. We name the topics that cannot disappear, including study selection, the main result, variation and limitations. We ask for a section or figure location beside every number. Most importantly, we tell the model to flag the mismatch between the Results text and Figure 2 and not to choose a number for us. This prompt is an illustrative teaching example; it is not a record of a model actually producing the outlines shown later. When we receive a real AI response, our next step is to check every proposed slide against the PDF. A precise prompt can make errors easier to see, but it cannot certify its own answer.

Slide 7: Comparing two structures

OBJECTIVE	Compare two illustrative outlines based on the same paper.
FACILITATOR CUE	These seven-slide routes are examples of a possible talk, not a recommendation to omit details from the paper. In the problem-oriented route, the open Figure 2 issue becomes a source audit before the conclusion. One can choose the conventional route for a methods-heavy audience or a question-led route for a mixed audience; in either case the evidence and uncertainty are identical.
KEY MESSAGE	Narrative order can change, evidence cannot.
ARTICLE LOCATION	Paper sections 3–7, pp. 2–8.

TEXT TO READ ALOUD

We now have two possible routes through the same article. The conventional route begins with the question and previous findings, then moves through search, sample, pooled results, moderator analyses and limitations. That structure is familiar to a research audience and makes the methods easy to follow. The problem-oriented route opens with the question the audience is likely to ask: does AI improve academic achievement? It then shows how the authors selected evidence, what the synthesis reports, why studies differ, and which subgroup claims the analysis supports. Before its conclusion it makes room for the Figure 2 discrepancy. We are not trying to create two different versions of the science. We are deciding when the audience encounters each piece of evidence. In either route, we must explain that the flow diagram ends with 30 studies and the final analysis uses 29 after an outlier exclusion. We must also show high heterogeneity and avoid presenting the inconsistent pooled number as though it were settled. Which route would best help this audience follow the argument?

Slide 8: Slide-level messages

OBJECTIVE	Demonstrate a defensible message when the exact numerical estimate is disputed within the source.
FACILITATOR CUE	A paper heading names a section; a slide message states the point for the audience. Both §5.2 and Figure 2 indicate a positive estimate and high I^2, but their printed pooled values and Q statistics differ. This slide therefore uses a qualitative description supported by both while flagging the exact value for verification.
KEY MESSAGE	The title should convey what the paper safely supports.
ARTICLE LOCATION	Results §5.2 p. 4; Figure 2 p. 6.

TEXT TO READ ALOUD

A section heading from a paper is rarely enough for a projected slide. The heading 'Overall effect size' tells the reader which part of the article we are in, but it does not tell a listener what to take away. We can turn it into a message: the synthesis reports a positive mean effect, with substantial variation between studies. That message is more useful because it states the direction of the reported result and immediately preserves the most important qualification. Why do I avoid placing an exact g value in the headline? Because the Results text and Figure 2 print different values. They agree on the broad positive direction, but we cannot silently settle their numerical difference. If we later verify the correct estimate, we can revise the slide and its notes. Until then, we can communicate what both parts of the article support and clearly mark the outstanding issue. Try reading a draft slide title without its speaker notes. Would it remain accurate if someone encountered it on its own?

Slide 9: Reducing text without losing meaning

OBJECTIVE	Show the difference between a concise finding and a universal claim.
FACILITATOR CUE	The source synthesises selected experimental and quasi-experimental studies, not every student or AI application. The review reports a positive average effect; high heterogeneity means that individual contexts differ. The exact pooled estimate is not on this slide because the source's Results text and Figure 2 do not match. Make the limitation visible if the slide circulates without narration.
KEY MESSAGE	Shortened text must keep the scope of the source.
ARTICLE LOCATION	Eligibility Table 1 p. 3; Results §5.1–5.2 p. 4; Limitations p. 8.

TEXT TO READ ALOUD

When we shorten research for slides, small changes in wording can change the claim. The sentence on the left says, 'AI improves every student's results.' It sounds direct, but this paper did not study every student, every AI tool or every educational setting. It synthesised selected studies using particular eligibility rules. A more careful sentence says that the review reports a positive average effect across 29 studies. We can say that clearly without turning it into a universal guarantee. We also need to bring the variation back into view. I-squared of 93 per cent signals that the study estimates differ substantially. An average across diverse studies does not tell us precisely what will happen in any one classroom. Notice that the exact pooled g value is absent from this proposed headline while the figure and text disagree. A shorter slide is successful when it keeps the population, comparison and uncertainty intact, not simply when it has fewer words.

Slide 10: Diagrams and visual organisation

OBJECTIVE	Relate visual form to the article's actual evidence.
FACILITATOR CUE	Figure 1 shows 30 studies entering final analysis, while the later exclusion of an extreme outlier reduces the reported final set to 29 (§5.1). A flow redraw should annotate this step rather than silently present 29 as the figure's original endpoint. Figure 2 is a forest plot, but its pooled numbers differ from §5.2; do not redraw the diamond from one set of values without reconciliation. The journal PDF is CC BY 4.0; cite the authors and paper when reproducing or adapting its visuals.
KEY MESSAGE	AI can suggest the form; the researcher controls units, estimates and attribution.
ARTICLE LOCATION	Figure 1 p. 4; Results §5.1 p. 4; Figure 2 p. 6; Table 3 p. 7.

TEXT TO READ ALOUD

AI may also suggest how to organise evidence visually. For this paper, a study-selection flow can explain where the evidence came from. If we adapt the article's Figure 1, we need to show the extra step from 30 studies in the flow diagram to 29 in the final analysis after the outlier exclusion. A forest plot can show that individual study estimates and their confidence intervals spread across a range. But we should not redraw the pooled diamond from a value we have not reconciled with the Results text. A compact table might help an audience compare moderators, provided we show the test of differences between categories as well as subgroup estimates. A confidence interval shows the range of values compatible with an estimate under the statistical model; it is not a range containing 95 per cent of individual student results. The article is open access under CC BY 4.0, so we may reuse an appropriate figure with attribution. Permission to reuse a figure does not remove our duty to check its labels and numbers.

Slide 11: Checking visual accuracy

OBJECTIVE	Perform a line-by-line comparison of numerical text and figure.
FACILITATOR CUE	The Results section reports g = 0.924, 95% CI [0.612, 1.235], Q[28] = 233.384, I^2 = 93%, and describes 29 effect sizes. Figure 2's overall row prints 0.89, 0.55 and 1.24; beneath the plot it prints Q = 207.25 with df = 24 and I-squared = 0.93. These differences go beyond rounding. The PDF does not explain whether Figure 2 shows an earlier subset or a different calculation. Do not identify the correct value without access to underlying data or author clarification.
KEY MESSAGE	A chart can look convincing and still disagree with the main text.
ARTICLE LOCATION	Results §5.2 p. 4; Figure 2 p. 6.

TEXT TO READ ALOUD

Let us slow down and perform the visual check ourselves. In Results section 5.2, the paper describes 29 effect sizes and reports a pooled Hedges' g of 0.924, with a 95 per cent confidence interval from 0.612 to 1.235. It gives Q with 28 degrees of freedom as 233.384. Figure 2 prints 0.89 as its overall value, an interval from 0.55 to 1.24, and a Q value of 207.25 with 24 degrees of freedom. The confidence intervals are close but not identical; the different Q statistics and degrees of freedom make this more than ordinary rounding. Both presentations suggest a positive mean effect and high heterogeneity. The supplied PDF, however, does not tell us why their exact summaries differ. We cannot infer which is correct from the page alone. In a real final presentation we would consult underlying analyses or the authors before announcing one precise pooled estimate. For this teaching exercise, the discrepancy itself demonstrates why every chart deserves a line-by-line comparison with the text.

Slide 12: Unsupported simplifications

OBJECTIVE	Identify inferential leaps in an AI summary of Table 3.
FACILITATOR CUE	The type-of-AI between-group test is $\chi^2 = 1.09$, df = 3, p = .78. Although the chatbot subgroup has g = 1.05, one cannot claim it statistically outperforms the other types from these data. The duration between-group test is $\chi^2 = 8.85$, p = .12; a subgroup above eight weeks has a larger point estimate but this does not prove duration increases effectiveness. The learning strategy moderator is likewise nonsignificant (p = .49).
KEY MESSAGE	Read moderator significance and subgroup size before making comparisons.
ARTICLE LOCATION	Results §5.3.3 and §5.3.7 pp. 5–7; Table 3 p. 7.

TEXT TO READ ALOUD

A model can make another kind of mistake even when it copies numbers correctly: it can draw a conclusion those numbers do not support. Table 3 lists separate estimates for categories such as chatbots and intelligent tutoring systems. A draft might say that chatbots are the best option because their listed average is higher. But the between-category test for type of AI has p equal to 0.78. That does not prove the types are identical. It means this analysis does not provide evidence of a difference between the categories on that test. The duration test has p equal to 0.12. One duration group may have a larger point estimate, but we cannot turn that pattern into 'longer always works better'. Some groups are represented by very few studies, which adds further caution. Our repair is to report the moderator test, the size of the groups where relevant, and what the result does or does not establish. We should never promote a subgroup ranking into a proven recommendation simply because it makes a compelling slide.

Slide 13: Preparing speaker notes

OBJECTIVE	Show how notes preserve method, result and caveat while the projected slide stays short.
FACILITATOR CUE	The presenter should say the mismatch plainly in a teaching demonstration. In a final research talk, resolve it with the authors or data before announcing one exact pooled estimate. If the deck is shared without narration, add the Figure 2 caveat visibly on the slide. The note is a presenter aid, not a substitute for scientific review.
KEY MESSAGE	The researcher edits and owns every spoken qualification.

ARTICLE LOCATION Abstract p. 1; Results §5.1–5.2 p. 4; Figure 2 p. 6.

TEXT TO READ ALOUD

Here is how I would split the result between the projected slide and what I say aloud. The slide gives one short message: a positive average effect is reported, but the studies vary considerably. The speaker notes then carry details the audience needs in order to interpret it. I would say that the authors analysed 29 studies with 2,657 participants. The Results text reports g equal to 0.924 and I-squared of 93 per cent. I would also say that Figure 2 prints a different pooled estimate and that I am not treating the precise number as resolved. When the slide is projected during a live talk, that spoken context can support a short headline. If the deck will later circulate without narration, the qualification should also appear visibly on the slide. AI may draft speaker notes, but we have to read them aloud in rehearsal, check every factual phrase against the article, and make sure the spoken story does not claim more than the displayed evidence.

Slide 14: Final scientific and ethical check

OBJECTIVE	Apply the final verification gate to a real article.
FACILITATOR CUE	The paper's limitations section identifies its restriction to peer-reviewed journal articles, the 2019–2024 window, possible publication bias and sparse longitudinal evidence. Check all statistics directly. The unresolved §5.2/Figure 2 inconsistency blocks the use of one exact pooled estimate in a final scientific talk without clarification. The article's CC BY 4.0 status permits reuse with proper credit; the project template uses CC BY-SA 4.0 separately. Follow the presenter's institutional disclosure policy for any AI assistance.
KEY MESSAGE	Source rights and scientific verification both require human decisions.
ARTICLE LOCATION	Title and CC BY statement p. 1; Methods p. 3; Results pp. 4–7; Limitations p. 8.

TEXT TO READ ALOUD

Before presenting, I would complete one last scientific and ethical check. First, I would follow every statistic back to the paper and confirm what was measured. Here, academic achievement is limited by the eligibility criteria, and the pooled number still needs reconciliation with Figure 2. Next I would look at the study's limits: it includes English-language peer-reviewed journal articles from 2019 to 2024, and the authors note possible publication bias and little evidence about long-term effects. Those limits affect how far we can take the conclusion. I would then check every adapted figure, its attribution and the terms under which I may reuse it. The article carries a CC BY 4.0 licence, while this project's educational resource has a separate CC BY-SA 4.0 mark. Finally, I would rehearse and follow the relevant policy on disclosing AI assistance. If a claim, figure or permission remains unclear, I would stop using that element until it is verified. A polished slide cannot substitute for a sound source.

Slide 15: Human-only responsibility checklist

OBJECTIVE	Close with a checklist participants can apply to their own paper.
FACILITATOR CUE	Pause on the precise source discrepancy. Figure 1 has 30 records at the final analysis stage; §5.1 describes an outlier exclusion and a final set of 29. The §5.2/Figure 2 mismatch is distinct and cannot be settled from the PDF alone. Participants should mark each unresolved slide element, consult data or authors when available, and refuse to let AI invent a reconciliation.

KEY MESSAGE	AI may draft a talk; the researcher verifies, authorises and presents it.
ARTICLE LOCATION	Figure 1 and §5.1 p. 4; Results §5.2 p. 4; Figure 2 p. 6; article licence p. 1.

TEXT TO READ ALOUD

Let us finish with a checklist we can use for any research presentation. Can I point to the passage, table or figure behind every statistic? Have I explained the step from 30 studies in the flow diagram to 29 in the final analysis? Have I either resolved or clearly flagged the different summaries in Results section 5.2 and Figure 2? Does the audience still hear that the studies vary and that the review has limits? If I reuse a figure, have I credited its authors and checked its licence? Finally, can I explain and defend every slide and every sentence I plan to say? If the answer to any of these questions is no, I would revise the talk before delivery. You can now apply the same process to your own paper. Begin with the audience and central contribution, let AI suggest alternatives, and then inspect each claim against the source. The final judgement, interpretation and presentation remain yours. Thank you. What part of this workflow would be most useful in your next talk?

Source and reuse

Dong, L., Tang, X., & Wang, X. (2025). Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. Computers and Education: Artificial Intelligence, 8, 100400. https://doi.org/10.1016/j.caeai.2025.100400.

The article is published under CC BY 4.0. Attribute the authors when reusing its text or figures. The AI Researchers presentation template bears a separate CC BY-SA 4.0 mark. The presenter should check institutional rules for disclosing AI assistance.

About this resource

TITLE	Designing a research presentation with AI
TOOLKIT REFERENCE	UC5-01, use case 5: Paper presentations
DEVELOPED BY	University of Lodz
SUGGESTED CITATION	AI Researchers (2026). Designing a research presentation with AI. AI Readiness Toolkit, UC5-01. 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