

Updating the AI Assessment Scale

Updating the AI Assessment Scale

1	NO AI	The assessment is completed entirely without AI assistance in a controlled environment, ensuring that students rely solely on their existing knowledge, understanding, and skills You must not use AI at any point during the assessment. You must demonstrate your core skills and knowledge.
2	AI PLANNING	AI may be used for pre-task activities such as brainstorming, outlining and initial research. This level focuses on the effective use of AI for planning, synthesis, and ideation, but assessments should emphasise the ability to develop and refine these ideas independently. You may use AI for planning, idea development, and research. Your final submission should show how you have developed and refined these ideas.
3	AI COLLABORATION	AI may be used to help complete the task, including idea generation, drafting, feedback, and refinement. Students should critically evaluate and modify the AI suggested outputs, demonstrating their understanding. You may use AI to assist with specific tasks such as drafting text, refining and evaluating your work. You must critically evaluate and modify any AI-generated content you use.
4	FULL AI	AI may be used to complete any elements of the task, with students directing AI to achieve the assessment goals. Assessments at this level may also require engagement with AI to achieve goals and solve problems. You may use AI extensively throughout your work either as you wish, or as specifically directed in your assessment. Focus on directing AI to achieve your goals while demonstrating your critical thinking.
5	AI EXPLORATION	AI is used creatively to enhance problem-solving, generate novel insights, or develop innovative solutions to solve problems. Students and educators co-design assessments to explore unique AI applications within the field of study. You should use AI creatively to solve the task, potentially co-designing new approaches with your instructor.



Perkins, Furze, Roe & MacVaugh (2024). The AI Assessment Scale

leonfurze.com

It's been over 12 months since the [first blog post](#) about the AI Assessment Scale, and a lot has changed, both with the technology and with our understandings of how it impacts assessments in K-12 and higher education across a range of disciplines. The AIAS has been [adopted by schools and universities worldwide](#), and will [feature in this year's UNESCO digital week](#). It has also been promoted as a tool for assessment transparency by the Australian [Tertiary Education Quality and Standards Agency](#), as well as adapted and translated around the world.

But we (the authors, [Mike Perkins](#), [Jasper Roe](#), [Jason MacVaugh](#) and me) acknowledge that there is still a lot of work to be done in understanding the implications of generative and attentional intelligence for assessment and learning. The original AIAS and its subsequent formal version ([published in JUTLP](#)) represents a moment in time where educational institutions across the world were reaching for something to help with the immediate problems of AI, such as the perceived threat to academic integrity.

1	NO AI	<p>The assessment is completed entirely without AI assistance. This level ensures that students rely solely on their knowledge, understanding, and skills.</p> <p>AI must not be used at any point during the assessment.</p>
2	AI-ASSISTED IDEA GENERATION AND STRUCTURING	<p>AI can be used in the assessment for brainstorming, creating structures, and generating ideas for improving work.</p> <p>No AI content is allowed in the final submission.</p>
3	AI-ASSISTED EDITING	<p>AI can be used to make improvements to the clarity or quality of student created work to improve the final output, but no new content can be created using AI.</p> <p>AI can be used, but your original work with no AI content must be provided in an appendix.</p>
4	AI TASK COMPLETION, HUMAN EVALUATION	<p>AI is used to complete certain elements of the task, with students providing discussion or commentary on the AI-generated content. This level requires critical engagement with AI generated content and evaluating its output.</p> <p>You will use AI to complete specified tasks in your assessment. Any AI created content must be cited.</p>
5	FULL AI	<p>AI should be used as a 'co-pilot' in order to meet the requirements of the assessment, allowing for a collaborative approach with AI and enhancing creativity.</p> <p>You may use AI throughout your assessment to support your own work and do not have to specify which content is AI generated.</p>

The AI Assessment Scale as published in JUTLP:

<https://open-publishing.org/journals/index.php/jutlp/article/view/810>

[Jason Lodge at University of Queensland and TEQSA refers to these as the acute problems of AI](#), but we recognise the need for robust frameworks that also tackle the chronic problems brought on in some ways by how we approach ideas of assessment and academic integrity in education.

So we have reflected on all of the [versions of the AIAS we have seen across the world in K-12 and higher education](#). We have sought out critique and engaged with diverse perspectives, from school teachers to students, university lecturers, to disability activists, experts in fields including assessment security, cognitive sciences, and pedagogy.

And over the past months, we have refined and invigorated the AI Assessment Scale to bring it up to speed with our current understandings of generative AI and learning.

Timeline of the AIAS

To look at how we have arrived at this updated version, it's useful to consider where the AIAS came from:

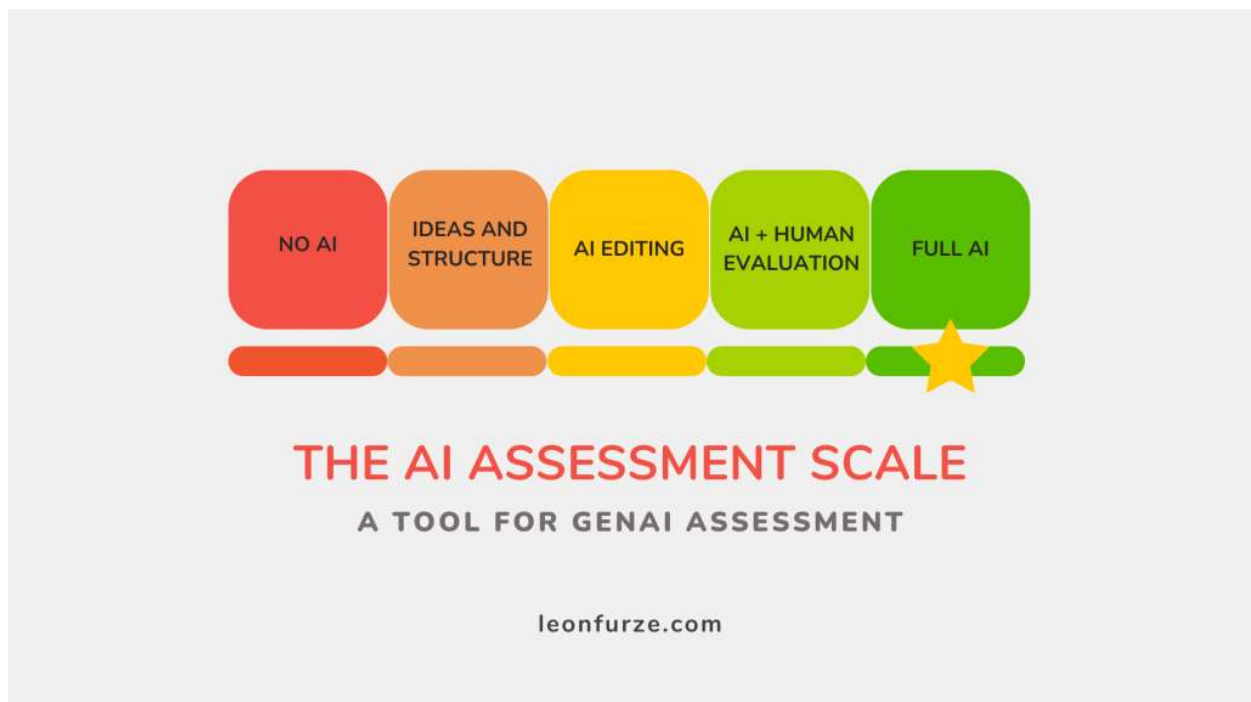
- It began as a discussion with teaching and learning staff at Edith Cowan University about the need for more than a binary “use or don't use” solution for AI and assessment during a time of peak AI threat in early 2023.
- The first version, which I will call Version 0, was largely based on my understanding of assessment as a teacher of English and literacy in K-12 and higher education.
- Version 0 was adapted by Mike Perkins, Jason MacVaugh (British University Vietnam), Jasper Roe (James Cook University Singapore) and me to broaden the applicability across other disciplines and make it more relevant to the international higher education context.
- This version, which I will call Version 1, is the most popular and widely used iteration of the AIAS, and was published in JUTLP and a forthcoming peer-reviewed publication based on the pilot study at British University Vietnam. It's also the version that has been most widely adopted around the world.

In between the academic articles on the AI Assessment Scale, the authors have been taking the work in various directions:

Mike has been presenting the work, seeking more feedback and supporting HE institutions worldwide in implementing the AIAS in various international contexts.

My focus has largely remained on K-12 with occasional forays into higher education, particularly following TEQSA's call for action plans requiring all Australian universities to outline an approach to artificial intelligence. I produced a free ebook containing examples and activities aligned to the levels of the AIAS, and wrote various other blog posts including discussions of the AI Assessment Scale and cheating.

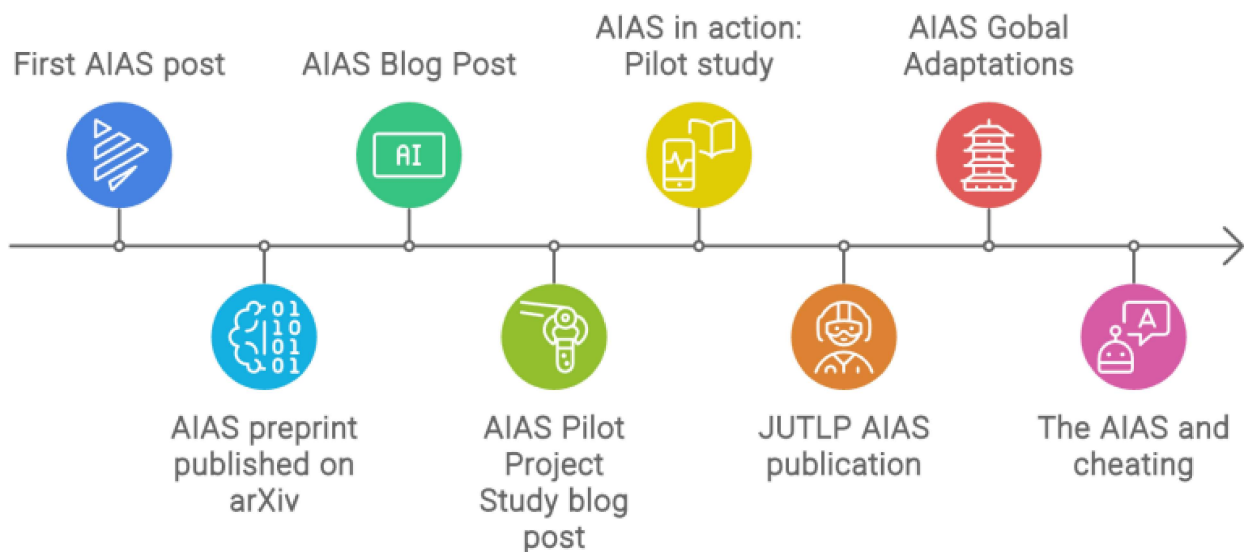
You can access a free ebook on the AIAS with over 50 activities for the 5 levels by [signing up for the mailing list here](#):



Mike and Jasper have continued to refine the AIAS in different contexts, including a recent preprint on how the AIAS can be reframed for English for Academic Purposes (EAP).

- April 2023: [First AIAS post](#)
- December 2023: [AIAS preprint published on arXiv](#)
- December 2023: [AI Assessment Scale Blog Post](#)
- March 2024: [AI Assessment Scale Pilot Project Blog Post](#)
- March 2024: [The AI Assessment Scale Pilot Study preprint](#)

- April 2024: [JUTLP Volume 21 Number 6: AI Assessment Scale Pilot](#)
- May 2024: [AIAS adaptations from across the world](#)
- August 2024: [AIAS for English for Academic Purposes preprint](#)
- August 2024: [The AIAS and "cheating" blog post](#)



Timeline image via Napkin.AI

Refining the AIAS

There are many reasons why we chose to update the AI Assessment Scale. The conversations we have had in the past few months have been long, winding, and complex. In a forthcoming article, we will articulate our main areas for revision in detail, including discussing the theoretical underpinnings of the new levels and our decisions and the discussions we have had with other academics along the way. In this blog post, I'm going to outline some of the main changes from my perspective and why I believe it has been necessary to update the scale.

Technology Changes

Even since the publication of the peer-reviewed version of the [AIAS in JUTLP in April 2024](#) (only four months ago), we have seen some significant advances in the technology which have impacted our thinking.

By the time of publication, GPT-4 was well established. [GPT-4o and 4o Mini](#) have now been released to the public as the base free version of ChatGPT. Similarly, [Claude 3.5 Sonnet](#) is available with limited credits for free users. That the most powerful models available on the market are now free has had huge ramifications. These tools are significantly more proficient in language, code, mathematical reasoning, and across a range of other benchmarks used to assess the quality of LLMs.

Multimodal generative AI has similarly progressed, and in just the past few weeks, we have seen the release of models like [Flux](#), which can generate images almost undetectable to the human eye, something which I proved recently with [a "real or fake" deep fake game, where out of almost 3000 plays, by the time of writing, fewer than a dozen people have hit 10 out of 10.](#) [Midjourney](#), another incredibly powerful and realistic photo generator announced on August 23 that it was officially moving beyond Discord to its [midjourney.com](#) platform with free credits for all users.

Although OpenAI touted their powerful video generation, Sora, in February, we still have not seen a public release, but models such as [Runway](#) and [Luma](#) are quickly pulling up in terms of quality, and have public versions available. And audio technologies, including the contentious [Udio and Suno](#) and [ElevenLabs](#)' voice generation models, have established themselves in the past two months.

As educators, we cannot hope to keep up with these technological changes, but we have adapted the AI Assessment Scale to account for them. Most notably, we have removed level four, formerly the "AI + human evaluation" level, and replaced it with the previous level five "Full AI". At this level, there are no restrictions on how a student chooses to use Artificial Intelligence to approach the task.

The new Level Five is an exploratory level where students, educators, and generative artificial intelligence work together to identify new ways of meeting learning outcomes. As the most recent advice from TEQSA points out, students may well have skills in generative AI that are advanced beyond those of their instructors, and there is no reason not to leverage those skills.

“Engaging in conversations and partnering with students and student groups is critical in developing immediate action. Students bring important perspectives to the discussion about AI as key contributors. In addition, many students are already sophisticated users of these technologies and contribute expert views about how they can and should be used in learning, teaching and assessment.”

[TEQSA \(2024\). The evolving risk to academic integrity posed by generative artificial intelligence: Options for immediate action](#)

The new level five also allows for experimentation with cutting-edge or near-horizon versions of this technology, such as artificial intelligence agents and more advanced multimodal technologies, or the convergence of technologies such as AR/VR and robotics with artificial intelligence.

Pedagogical Changes

We also acknowledge that the ways that people have been teaching with (and sometimes against) AI have developed in the past 18 months in unexpected ways. Universities such as the University of Sydney have invested enormous time and resources in the [creation of open access LLM-based applications such as Cogniti](#), which can be trained on the corpus of an instructor’s materials to create a custom chatbot similar to a GPT or a Claude.

Projects leveraging these tools offer incredible opportunities for educators. The original AIAS was seen by many as a tacit permission for educators to use and experiment with artificial intelligence in the classroom, and we have carried that forward into this new version. Again, the only level which rules out artificial intelligence use is Level One: No AI. We now provide more clarity on why educators might choose that level.

To avoid a shopping list of AI tasks, we retain some of the distinction between Levels 2 and 3, articulating Level 2 as assessments where students may use AI as part of the planning, initial composition, research, or in later years of university courses (once they have established foundational skills) through AI-powered research tools. At Level 3, the focus is instead on evaluation, feedback and the use of AI chatbots, such as custom designed course tutors

provided by the education institute or the educator. Students can also use AI at Level 3 for the development of their writing, as long as they are still applying a critical eye to the output.

Whilst some approaches like University of Sydney's [two-lane framework](#) are deliberately broad to acknowledge the reality of education and educators' understanding of these technologies, we still believe that people need more support in exploring different applications of generative AI in their contexts.

Assessment Security

One of the most significant changes in our revised AI Assessment Scale is our perspective on assessment security and assessment validity. This has been shaped in large part by discussions in Australia through webinars and materials provided by [TEQSA](#) and [Deakin University's CRADLE](#), and in particular, Phillip Dawson, Margaret Bearman, Mollie Dollinger, and David Boud in their recent publication [Validity matters more than cheating](#).

It is a potentially disruptive but absolutely necessary perspective which understands that permitting *any* use of AI effectively permits *all* use of AI, and since it is undetectable and sophisticated across domains, the distinction between previous Levels 2, 3, 4, and even 5 is somewhat arbitrary.

In the revised version, we take the stance that assessment security, in the traditional sense, is only possible at Level 1, but that assessment validity is possible throughout. We will discuss this further in a forthcoming journal article.

Stylistic Changes

A surprising amount of time and attention goes into the branding of academic frameworks for both pragmatic and theoretical reasons: research needs to be communicated clearly, made accessible, and be practical.

The first version of the AIAS was successful in part due to its simplicity and the bold use of the red to green gradient. The AIAS at times has been conflated with traffic light approaches suggesting “No AI” means stop and Green Level Five assessments “Full AI” means go. As the authors have pointed out in numerous publications, including the peer-reviewed JUTLP, the stop-slow-go of the traffic light was never our intention.

Still, the red to green colour scheme of the AI Assessment Scale, whilst being fundamental to its adoption, has troubled us for a while. [In a recent LinkedIn article, Danny Liu of University of Sydney criticised traffic light approaches](#), including ones with “yellow-green, or yellow-red in the mix as well”.

We agree.

So, in the new version, we’ve adopted more neutral colours, still easily identifiable (e.g., the “pink level”, the “blue level”), but without the symbolic connotations of stopping or going. We also chose these colours for accessibility purposes. We rejected the idea of a gradient, since it can be difficult for partially sighted people to discern between levels of the gradient if the contrast between levels is not distinct. We further tested our ideas using online accessibility tools and talking to people with colour blindness and low sight to arrive at the five new distinct colours.

Accessibility and inclusion will be another strand of our forthcoming paper.

Although we have kept the table in the same style as the original AIAS, we also acknowledge that the five levels suggest a hierarchy to some people, perhaps in the sense that “more AI is better”, or suggesting that there is a percentage of allowable AI use at a given level (e.g., 25% at level two, 30% at level three and so on).

We understand, however, these are arbitrary distinctions. Depending on the context and the discipline and the type of assessment, different levels of AI use may be appropriate at different times. We have therefore designed a new circular version of the Scale graphic, which suggests that all levels may be treated equally in terms of their merits for assessment design.

Mike, Jasper, and Jason will have their own nuanced perspectives and positions on many of these decisions, and we'll draw those together in the full article.

Here is the revised AI Assessment Scale:

The AI Assessment Scale

1	NO AI	The assessment is completed entirely without AI assistance in a controlled environment, ensuring that students rely solely on their existing knowledge, understanding, and skills You must not use AI at any point during the assessment. You must demonstrate your core skills and knowledge.
2	AI PLANNING	AI may be used for pre-task activities such as brainstorming, outlining and initial research. This level focuses on the effective use of AI for planning, synthesis, and ideation, but assessments should emphasise the ability to develop and refine these ideas independently. You may use AI for planning, idea development, and research. Your final submission should show how you have developed and refined these ideas.
3	AI COLLABORATION	AI may be used to help complete the task, including idea generation, drafting, feedback, and refinement. Students should critically evaluate and modify the AI suggested outputs, demonstrating their understanding. You may use AI to assist with specific tasks such as drafting text, refining and evaluating your work. You must critically evaluate and modify any AI-generated content you use.
4	FULL AI	AI may be used to complete any elements of the task, with students directing AI to achieve the assessment goals. Assessments at this level may also require engagement with AI to achieve goals and solve problems. You may use AI extensively throughout your work either as you wish, or as specifically directed in your assessment. Focus on directing AI to achieve your goals while demonstrating your critical thinking.
5	AI EXPLORATION	AI is used creatively to enhance problem-solving, generate novel insights, or develop innovative solutions to solve problems. Students and educators co-design assessments to explore unique AI applications within the field of study. You should use AI creatively to solve the task, potentially co-designing new approaches with your instructor.



Perkins, Furze, Roe & MacVaugh (2024). The AI Assessment Scale



Perkins, Furze, Roe & MacVaugh (2024). The AI Assessment Scale

Our forthcoming paper and the presentations at the UNESCO Digital Week will help to explain our choices further, and provide the theoretical and pedagogical grounding for the revisions. As always, we encourage K-12 and Higher Education educators to both adopt and adapt the AIAS as part of increasing transparency and clarity over the use of AI in learning and assessment.

Thank you to everyone across the world who has contributed to our ideas as we shape and refine the AIAS. We continue to seek feedback from all educators in K-12 and Higher Education!

Get in touch to discuss the AIAS, or anything related to Generative AI professional learning and consulting services:

If you'd like to discuss Generative AI consulting, advisory, and professional learning services or just have something to say about this post, get in touch via the form below:

Name (required)

Email (required)

Message

Contact Leon

Leon Furze

[August 28, 2024](#)

[AI](#)

[AI, AI Assessment Scale, AIAS, assessment](#)

15 responses to “Updating the AI Assessment Scale”

[The AI Assessment Scale: Version 1 – Leon Furze](#)
[August 28, 2024 at 1:39 pm](#)

[...] Updating the AI Assessment Scale [...]

Loading...

[Reply](#)

[Exploring New Frontiers: Level 5 of the AI Assessment Scale – Alchemy](#)
[August 31, 2024 at 3:14 am](#)

[...] updated AI Assessment Scale (AIAS) of 2024 by Perkins et al. presents an exciting opportunity for educators and students to [...]

Loading...

[Reply](#)

[Directing AI in Assessments: Level 4 of the AI Assessment Scale – Alchemy](#)
[August 31, 2024 at 3:29 am](#)

[...] AI continues to play a significant role in education, the AI Assessment Scale (AIAS) of 2024 by Perkins et al. provides a framework for integrating AI into academic tasks. This [...]

Loading...

[Reply](#)

[Collaborating with AI: Level 3 of The AI Assessment Scale – Alchemy](#)

[August 31, 2024 at 5:01 am](#)

[...] educators continue to consider the integration of AI into academic assessments, the 2024 AI Assessment Scale (AIAS) offers a structured approach to guide this process. This framework consists of five levels, [...]

Loading...

[Reply](#)

[AIAS: Why We've Driven Through the Traffic Lights – Leon Furze](#)

[September 2, 2024 at 8:30 am](#)

[...] week, I made a post about our updates to the AI Assessment Scale. These updates are the culmination of using the scale [...]

Loading...

[Reply](#)

[Sharing Diigo Links and Resources \(weekly\) | Another EducatorAI Blog](#)

[September 2, 2024 at 2:05 pm](#)

[...] Updating the AI Assessment Scale – Leon Furze [...]

Loading...

[Reply](#)

[AI Assessment Scale Level 1: No AI – Leon Furze](#)

[September 9, 2024 at 8:30 am](#)

[...] Updating the AI Assessment Scale [...]

Loading...

[Reply](#)

[Planning with AI: Level 2 of the AI Assessment Scale – Alchemy](#)
[September 11, 2024 at 2:17 am](#)

[...] 2024 AI Assessment Scale (AIAS) outlines how AI can support students in the planning and initial stages of their work while [...]

Loading...

[Reply](#)

[We Need to Talk About Deepfakes – Leon Furze](#)
[September 11, 2024 at 2:26 pm](#)

[...] context, I have worked with Assoc. Prof. Mike Perkins and Dr Jasper Roe (coauthors of the AI Assessment Scale) to articulate a Research Agenda for Deepfakes in Higher Education. You can read the full agenda in [...]

Loading...

[Reply](#)



[Joseph Thibault](#)
[September 13, 2024 at 6:56 am](#)

Leon, this is really wonderful. I'm a fan of the color scheme and non-hierarchical options for displaying this.

Thanks to you and your co-authors for providing such a great resource and for explaining all the updates and changes!

Loading...

[Reply](#)



[Leon Furze](#)

[September 13, 2024 at 7:31 am](#)

Thanks Joseph. Keep an eye out for the series over the next few weeks exploring each level of the scale in detail!

Loading...

[Reply](#)

[Upholding Academic Integrity – Level 1 of the AI Assessment Scale – Alchemy](#)
[September 14, 2024 at 3:40 am](#)

[...] for students to demonstrate their core knowledge and skills without external assistance. The AI Assessment Scale (AIAS) of 2024 by Perkins et al. provides a structured framework for this. To wrap up our 5-part [...]

Loading...

[Reply](#)

[Harnessing the Power of AI in Education: Alchemy's Commitment to Effective and Ethical AI Integration – Alchemy](#)
[September 18, 2024 at 3:44 am](#)

[...] to help educators determine the appropriate level of AI use in student assessments. This scale was revised in August 2024 to better reflect the rapidly changing AI landscape. The AIAS offers five levels of AI integration, [...]

Loading...

[Reply](#)

[Harnessing the Power of AI in Education: Commitment to Effective and Ethical AI Integration – Alchemy](#)
[September 18, 2024 at 3:45 am](#)

[...] to help educators determine the appropriate level of AI use in student assessments. This scale was revised in August 2024 to better reflect the rapidly changing AI landscape. The AIAS offers five levels of AI integration, [...]

Loading...

[Reply](#)

[The AI Assessment Scale – Pontydysgu EU](#)

[September 26, 2024 at 1:14 am](#)

[...] I don't know quite how I have managed to miss this up to now. The AI Assessment Scale (AIAS) has been around for over a year. On the occasion of updating to the latest version – see illustration above, Leon Furze, a Consultant, author and PhD candidate and one of the authors, said in his blog: [...]

Loading...

[Reply](#)

Leave a Reply

© Leon Furze 2022

Proudly Powered by [WordPress](#)