ThinkActuary

RESEARCH WHITEPAPER / OCTOBER 2026

Beyond the Mark

From Automated Feedback to Evidence-Driven Learning in Professional Examinations

How learning science and AI can turn assessed performance into measurable, durable professional capability.

A research-led agenda for better exam feedback

Contents

From marking scripts to developing transferable actuarial capability.

Executive Summary . . . 3

1. The Unfinished Business of Assessment . . . 4

2. What Learning Science Tells Us About Effective Feedback . . . 5

3. Defining State-of-the-Art Feedback . . . 8

4. The Role of Artificial Intelligence . . . 10

5. A Practical Roadmap: From Text Feedback to Adaptive Learning . . . 11

6. Measuring What Matters: From Feedback Quality to Learning Impact . . . 14

7. Why Actuarial Education Is an Important Test Case . . . 16

8. Building a Learning Evidence Infrastructure . . . 17

9. A New Opportunity for Professional Examining Bodies . . . 18

10. The ThinkActuary Research Agenda . . . 19

11. Conclusion: The Future of Feedback Is Learning . . . 19

References . . . 21

About ThinkActuary . . . 21

© 2026 ThinkActuary

Executive Summary

THE CENTRAL THESIS

Feedback is not the end of assessment. It is the beginning of the next learning decision.

For decades, assessment in professional education has followed a familiar sequence: a student attempts an examination question, an examiner evaluates the response, and the student receives a mark accompanied by feedback.

The process is effective at measuring performance. It is less effective at ensuring that performance improves.

A student may understand precisely why they lost marks yet make the same mistakes on their next attempt. Feedback is provided, but the learning process is not necessarily completed.

Advances in artificial intelligence create an opportunity to fundamentally rethink this model. Yet the most important innovation is not the ability to produce longer, faster or more personalised explanations. It is the ability to connect feedback to subsequent learning activities and objectively measure whether those activities produce improvement.

Learning science offers a foundation for this transformation. Decades of research suggest that effective feedback is specific, informative and actionable; that learning requires students to actively engage with information; and that retrieval practice, appropriate challenge and repeated assessment can contribute to durable learning.

This paper proposes a new paradigm for professional exam preparation: **evidence-driven feedback systems that continuously diagnose performance, guide deliberate improvement and verify the transfer of learning to unfamiliar problems.**

The proposed progression is deliberate. Rather than attempting to build an advanced AI tutor immediately, educational platforms should begin by evaluating their existing feedback, testing simple improvements and accumulating evidence of what works.

Over time, these capabilities can evolve into personalised, adaptive learning systems.

For actuarial education, where success demands not only technical knowledge but also the ability to apply that knowledge, communicate reasoning and exercise professional judgement, the opportunity is particularly compelling.

Our central proposition is simple:

© 2026 ThinkActuary

> **The measure of effective feedback is not the quality of the explanation a student receives. It is the improvement the student can subsequently demonstrate.**

1. The Unfinished Business of Assessment

Assessment measures performance. Learning requires something more.

Professional examinations represent an important quality-control mechanism. They establish standards, evaluate knowledge and provide evidence that candidates possess essential technical and professional capabilities.

However, assessment and learning serve different functions.

An examination measures what a candidate can demonstrate at a particular moment. Learning involves developing knowledge, understanding and skills that remain available and can be applied in the future.

The distinction matters because an accurate mark does not automatically produce learning.

Consider an actuarial student who receives 45% for a question. The examiner identifies several deficiencies: insufficient explanation, failure to apply relevant information from the scenario and omission of important financial consequences.

The student reads the comments and understands them.

What happens next?

In a traditional feedback process, the responsibility for translating those observations into improved performance falls almost entirely on the student. They must interpret the feedback, identify the underlying weaknesses, select relevant study activities and determine whether those activities have worked.

This is a demanding process, particularly for students who have not yet developed strong self-regulated learning skills.

The result is a structural gap between **identifying a performance deficiency** and **correcting the underlying capability**.

A learning-science perspective

Black and Wiliam's influential work on formative assessment established the importance of using assessment evidence to improve subsequent teaching and learning [1].

© 2026 ThinkActuary

Hattie and Timperley subsequently framed effective feedback around three questions [2]:

- Where am I going?
- How am I going?
- Where to next?

The third question is especially important.

A mark tells the student how they performed. A marking schedule explains what was expected. But neither necessarily establishes what the student should do next to improve.

For feedback to become genuinely formative, information about performance must influence subsequent learning.

This leads to the first principle of a next-generation feedback system:

Feedback should be designed as the beginning of an improvement process, not the end of an assessment process.

2. What Learning Science Tells Us About Effective Feedback

The educational research literature provides several important lessons for the design of feedback systems.

2.1 The information contained in feedback matters more than its presentation

Feedback is not a single educational intervention.

A grade, a statement identifying an incorrect answer, a worked example and an explanation of an underlying misconception all provide information. However, they do not necessarily generate equivalent learning outcomes.

Wisniewski, Zierer and Hattie examined 435 studies involving more than 61,000 participants. Their 2020 meta-analysis identified an overall moderate positive effect of feedback on learning, but substantial variation between different feedback approaches [3].

The authors emphasised the importance of the information conveyed by feedback.

Similarly, a meta-analysis of computer-based feedback by Van der Kleij and colleagues found that elaborated feedback, which provides explanations, was generally more effective than simply stating whether an answer was correct, particularly for higher-order learning outcomes [4].

The implication for examination preparation is significant.

A student who receives the comment:

© 2026 ThinkActuary

“You did not adequately explain the impact on profitability.”

has learned something about the deficiency in their answer.

A student who receives:

“You identified the increase in claims costs but did not explain its financial consequence. On similar questions, connect the change in expected claims to premium adequacy, the loss ratio or underwriting profitability.”

receives a more transferable explanation of how to improve.

The second response does more than describe the deficiency. It introduces a reusable approach.

The aim should therefore be to optimise the **instructional value of feedback**, not its length, detail or apparent sophistication.

2.2 Feedback must be acted upon

Carless and Boud's framework of student feedback literacy shifts attention from the information delivered by an educator to the learner's capacity to interpret and use that information [5].

Their work highlights four interconnected capabilities: appreciating feedback, making judgements, managing affect and taking action.

This has an important consequence for educational technology.

Even excellent feedback may accomplish little if students do not understand it, do not know how to act on it or never have an opportunity to use it.

A next-generation feedback system must therefore go beyond delivering recommendations.

It should create a practical opportunity to apply those recommendations.

A suggestion to improve causal explanations becomes more valuable when the student can immediately practise writing a causal explanation and receive evidence of whether the attempt was successful.

In this model, feedback is not merely information delivered to a learner. It is part of an active learning process.

2.3 Retrieval and practice contribute to durable learning

Research on retrieval practice has demonstrated that attempting to retrieve knowledge from memory can improve subsequent retention compared with additional study alone.

Roediger and Karpicke's experimental work showed that repeated testing can improve delayed retention even when additional study produces stronger immediate performance [6].

For professional examination preparation, this points towards an important design principle:

Students should not only review their mistakes. They should attempt new problems that require them to overcome those mistakes.

© 2026 ThinkActuary

However, repeated practice must be designed carefully.

Simply repeating the original question may improve familiarity with that particular answer without establishing that the underlying skill has improved.

A stronger approach tests the same capability in a new context.

For example, if a student fails to explain how increasing claims severity affects general insurance profitability, a subsequent exercise might ask them to explain how increasing longevity affects annuity liabilities.

The scenarios differ, but the required reasoning skill is related.

Research on the transfer of test-enhanced learning suggests that retrieval can support transfer beyond the original learning material, although the extent of transfer depends on the task and how practice is designed [7].

The implication is that the ultimate objective should be **transferable capability**, not memorisation of model answers.

2.4 Learning must persist beyond the immediate correction

There is an important distinction between performing better immediately after receiving help and developing a lasting capability.

Soderstrom and Bjork's review distinguishes observable performance during instruction from the more durable changes that constitute learning [8].

A student may answer a question correctly immediately after seeing a worked example but fail to apply the principle independently several weeks later.

The immediate result can create a misleading impression of mastery.

Research on distributed practice also suggests that appropriately spaced learning episodes can support longer-term retention [9].

An effective feedback system should therefore assess improvement at different points:

- Immediately after feedback, to establish whether the correction has been understood.
- On a new question, to establish whether the skill transfers.
- After an appropriate delay, to establish whether the learning is retained.

These assessments answer different questions and should not be treated as interchangeable.

2.5 Feedback should manage cognitive load

Actuarial examination questions are often cognitively demanding. They require candidates to integrate technical knowledge, interpret scenarios, construct explanations and manage time.

Providing a lengthy list of every possible improvement can increase the demands placed on the learner.

© 2026 ThinkActuary

Cognitive load theory highlights the limitations of working memory and the importance of structuring instruction so that learners can devote sufficient mental resources to learning [10].

A practical implication is to prioritise a small number of important improvements rather than overwhelming students with an exhaustive report.

This does not mean that three recommendations are universally optimal. The appropriate number should be established through testing.

It means that feedback should be organised around the learner's capacity to interpret and apply it.

The guiding question becomes:

What is the smallest amount of feedback that will produce the greatest useful change in future performance?

2.6 Feedback should progressively develop self-regulation

At the beginning of a learning journey, students may need explicit guidance.

They may benefit from structured explanations, model answers and simple frameworks such as:

Point → Why → Consequence

As they become more capable, the support should progressively shift towards independent judgement.

Instead of always telling the student what is missing, the system might ask:

“Which part of your explanation would an examiner find insufficient?”

Eventually, students should become increasingly capable of evaluating and improving their own responses.

This connects feedback to metacognition and self-regulated learning: the capacity to plan, monitor and evaluate one's own learning.

The long-term objective is not to create dependence on an AI tutor.

It is to develop independent, capable candidates who require progressively less external guidance.

3. Defining State-of-the-Art Feedback

These principles suggest a different conception of the future of educational feedback.

We define state-of-the-art feedback as:

© 2026 ThinkActuary

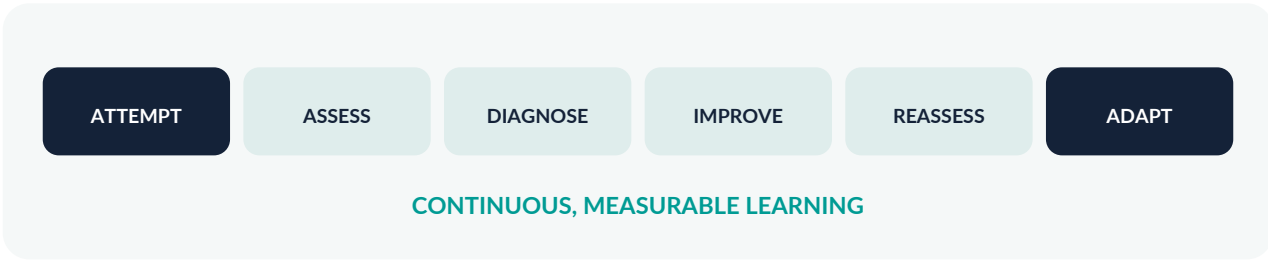
> **A continuously improving learning system that uses assessment evidence to identify individual development needs, selects appropriate educational interventions and verifies whether those interventions result in durable, transferable improvements.**

Such a system requires six capabilities.

Capability	Purpose
Accurate diagnosis	Identify the underlying reasons for lost marks
Intelligent prioritisation	Determine which deficiencies matter most
Actionable coaching	Explain how the student should improve
Targeted practice	Create opportunities to apply the correction
Independent reassessment	Verify whether the capability has improved
Longitudinal adaptation	Remember progress and select future interventions

Together, these capabilities create a closed learning loop:

Attempt → Assess → Diagnose → Improve → Reassess → Adapt



This represents a conceptual shift from feedback as a document to feedback as an ongoing educational process.

An illustrative student journey

Consider an actuarial student who repeatedly loses marks in questions requiring explanations of financial consequences.

An advanced feedback system identifies that the student understands the underlying actuarial concepts but frequently stops after identifying relevant risks.

© 2026 ThinkActuary

The system prioritises this weakness because it recurs across multiple questions and may be limiting performance.

It introduces a simple technique:

Point → Why → Consequence

The student then completes a five-minute exercise requiring two short causal explanations.

The response is assessed against the targeted skill. If the student can identify the point and explain the mechanism but omits the consequence, the next intervention focuses specifically on that missing step.

A subsequent question uses a different actuarial scenario.

The student attempts it without the original scaffolding.

Several days later, the system introduces another opportunity to demonstrate the skill.

Over time, the system develops evidence about whether the deficiency is persistent, improving or resolved.

The experience becomes personalised not because the AI knows the student's preferred learning style, but because it has evidence of their actual performance and learning needs.

That is a more defensible basis for adaptation.

4. The Role of Artificial Intelligence

Artificial intelligence changes the economics and feasibility of this model.

Historically, providing detailed individual feedback, identifying longitudinal patterns and designing personalised remedial exercises required substantial educator time.

Modern AI systems make it possible to perform many of these tasks at scale.

Potential applications include:

- Analysing written responses against marking criteria.
- Identifying recurring patterns in student reasoning.
- Generating clear explanations of mistakes.
- Recommending or creating targeted practice activities.
- Evaluating subsequent attempts.
- Identifying which interventions appear to produce improvement.

However, AI capability should not be confused with educational effectiveness.

A language model may produce convincing feedback that is factually incorrect, poorly aligned with a marking schedule or educationally unhelpful.

© 2026 ThinkActuary

Research on AI-generated feedback has found promising potential alongside variation in effectiveness, transparency and the quality of human versus automated feedback [11, 12].

The correct strategy is therefore not to automate educational judgement indiscriminately.

It is to use AI within an evidence-based, auditable system.

Human expertise remains essential

For actuarial education, the distinction between marking and coaching is particularly important.

The marking process must remain anchored to defensible marking criteria, acceptable alternative answers and professional standards.

The feedback process should use the marking evidence without silently changing the awarded marks.

Human subject-matter experts should play a central role in developing evaluation benchmarks, reviewing ambiguous cases, identifying unsupported diagnoses and evaluating whether proposed feedback strategies are educationally sound.

AI provides scale and responsiveness.

Human expertise establishes standards, validates difficult judgements and governs the system.

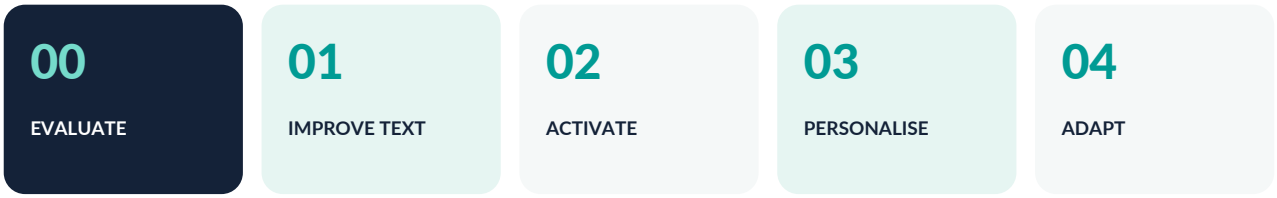
The opportunity is not merely more intelligence

Access to increasingly capable AI models will become widespread.

A potentially more durable source of advantage will be the ability to determine **which educational interventions actually work**.

That requires something models alone cannot provide: reliable longitudinal evidence connecting feedback, student action and subsequent performance.

5. A Practical Roadmap: From Text Feedback to Adaptive Learning



Advance on evidence of learning benefit, not feature complexity.

The desired end-state is ambitious.

© 2026 ThinkActuary

Attempting to construct the entire system at once would introduce unnecessary technical, educational and operational risk.

A more effective approach is to develop the system through five progressive milestones.

Milestone 0: Establish a scientific baseline

Objective: Learn how to measure feedback quality before attempting to improve it.

The starting point is a systematic evaluation of existing feedback.

A representative sample of marked responses should be assembled, including the original questions, student answers, marking schedules, awarded marks and generated feedback.

Expert reviewers should assess the feedback against clearly defined dimensions:

1. Accuracy and alignment with marking evidence.
2. Quality of diagnosis.
3. Specificity and actionability.
4. Transferability to future questions.
5. Clarity and appropriateness for the student.

Accuracy should function as an essential requirement, not simply another dimension that can be offset by good presentation.

The resulting benchmark becomes a repeatable evaluation instrument.

It allows alternative feedback approaches to be compared against a common standard.

Success criterion: The organisation can reliably distinguish stronger from weaker feedback and identify recurring defects in the existing system.

This milestone does not require a sophisticated new interface or tutoring architecture.

Milestone 1: Produce demonstrably better text feedback

Objective: Improve the existing experience using controlled experimentation.

The next step is to develop alternative text-feedback approaches using the same marking evidence.

For example, one approach may emphasise detailed explanations of missing marks.

Another may prioritise transferable improvement rules.

A third may combine a concise diagnosis with a worked example.

These alternatives should be evaluated using the established quality benchmark and, wherever practical, randomised experiments with students.

The central research question becomes:

© 2026 ThinkActuary

Does changing the structure and content of feedback improve subsequent performance?

A useful experimental comparison might test existing feedback against feedback that explicitly identifies the deficiency, explains the underlying mechanism and provides an actionable rule for future questions.

Initially, expert quality ratings and student engagement may provide useful signals.

However, improved performance on new, comparable questions remains the stronger indicator of educational value.

Success criterion: At least one new approach demonstrates reliable quality improvements and credible evidence of educational benefit.

Milestone 2: Convert feedback into action

Objective: Connect each important feedback recommendation to relevant practice.

Once useful feedback strategies have been identified, the platform can begin linking them to learning activities.

These activities need not initially be complex.

They might include an existing past-paper question, a short structured response or a focused five-minute retry.

The essential requirement is that the practice addresses the diagnosed weakness.

A student struggling with explaining consequences should not simply receive another randomly selected question.

They should receive an exercise that creates an appropriate opportunity to practise that skill.

This milestone introduces the first genuine closed-loop capability.

The system can now observe not only the feedback delivered, but also the student's response to that feedback.

Success criterion: Targeted practice produces measurable improvement compared with feedback alone.

Milestone 3: Introduce longitudinal personalisation

Objective: Build an evidence-based understanding of each student's development.

At this stage, the platform begins tracking important capabilities across multiple attempts.

These might include technical knowledge, scenario application, causal explanation, breadth of answer, numerical execution and professional judgement.

A student's profile should represent demonstrated performance rather than speculative personality or learning-style classifications.

© 2026 ThinkActuary

The system can then identify recurring weaknesses, distinguish persistent from isolated mistakes and recommend activities based on accumulated evidence.

For example, one student may require technical revision, while another requires practice communicating the implications of technical knowledge.

The appropriate intervention depends on the nature of the deficiency.

Success criterion: Personalised recommendations outperform non-personalised recommendations on relevant learning outcomes.

Milestone 4: Develop an adaptive learning system

Objective: Continuously improve the selection and sequencing of educational interventions.

The mature system can use longitudinal evidence to choose among different feedback and practice strategies.

It may determine that a particular student benefits from a worked example before attempting a new problem, while another benefits from immediate independent practice.

It may adjust the level of scaffolding, select appropriate challenge levels and schedule reassessment.

Such decisions should remain governed by clear educational objectives, quality controls and human oversight.

Importantly, the system must distinguish observed associations from demonstrated causal effects.

An intervention associated with improvement is not necessarily the cause of that improvement.

Controlled experiments remain essential.

Success criterion: The adaptive system demonstrates sustained improvement in independent performance compared with simpler approaches, without sacrificing accuracy, fairness or transparency.

6. Measuring What Matters: From Feedback Quality to Learning Impact

01	02	03
QUALITY	**ACTION**	**IMPACT**
Is it accurate and useful?	Does the student use it?	Can they now perform better?

A major risk in educational AI is optimising the wrong outcome.

© 2026 ThinkActuary

Platforms can readily measure how much feedback was generated, how quickly it was delivered and whether students reported finding it useful.

These measures are valuable operational signals.

They are not sufficient evidence of learning.

A scientifically grounded feedback strategy should distinguish three measurement levels.

Level 1: Feedback quality

Does the feedback accurately identify the deficiency? Is it aligned with marking evidence? Is it understandable, specific and actionable?

These questions establish whether the intervention is educationally credible.

Level 2: Student engagement and behaviour

Does the student review the feedback, understand the recommendation and attempt an appropriate learning activity?

These measures establish whether the feedback influences behaviour.

Level 3: Learning outcomes

Can the student demonstrate the relevant capability on a new question?

Can they demonstrate it without assistance?

Does the improvement persist after a delay?

Does it generalise to comparable but unfamiliar problems?

These questions establish whether learning has occurred.

The most important outcome is improvement in independently demonstrated capability, not feedback consumption.

A proposed research methodology

ThinkActuary's initial experiments should randomly assign students to feedback treatments where feasible, while ensuring that all groups continue to receive essential marking information.

Experiments should predefine primary outcomes, observation windows and analytical methods.

A strong primary outcome would be performance on a previously unseen, appropriately calibrated assessment of the targeted skill.

Relevant considerations include baseline student ability, question difficulty, repeated attempts, differences between exam subjects and variation in the marking process.

© 2026 ThinkActuary

Where sample sizes are small, studies may be more useful for establishing feasibility and estimating likely effects than for making definitive claims.

A quality improvement identified by expert reviewers is a valuable milestone.

It should not be presented as proof that students subsequently earn more marks.

The research programme should progressively move from demonstrating feedback quality to demonstrating durable improvements in learning.

7. Why Actuarial Education Is an Important Test Case

Actuarial examinations provide an especially interesting environment for this approach.

Success requires more than recalling definitions and formulae.

Students must often interpret unfamiliar business circumstances, select appropriate methods, communicate technical reasoning, assess risks and explain implications.

These are precisely the types of capabilities for which merely seeing a correct answer may be insufficient.

Consider two students who both lose marks on an insurance pricing question.

The first lacks a fundamental understanding of claims inflation.

The second understands claims inflation but fails to explain how it affects expected claims costs and premium adequacy.

The numerical outcome may be similar.

The educational intervention should be different.

For the first student, additional conceptual instruction may be appropriate.

For the second, targeted practice constructing complete causal explanations may be more productive.

An advanced feedback system should distinguish these cases.

From examination performance to professional capability

This approach may ultimately have implications beyond examination preparation.

The ability to identify a weakness, apply feedback, reassess a decision and improve future work is central to professional development.

© 2026 ThinkActuary

The reasoning skills required in actuarial examinations frequently resemble those required in professional practice: explaining assumptions, interpreting results, considering financial consequences and communicating conclusions.

There is therefore a compelling hypothesis worth investigating:

A system that develops transferable actuarial reasoning through examination feedback may also help develop capabilities relevant to professional actuarial work.

This should be treated as a research proposition rather than an established outcome.

Demonstrating improvement in workplace judgement would require separate evaluation.

Nevertheless, the connection highlights the wider importance of moving beyond assessment as a one-off measurement event.

8. Building a Learning Evidence Infrastructure

The long-term significance of this approach may lie as much in its underlying data infrastructure as in the student-facing experience.

Every learning interaction can generate several important observations:

- What the student attempted.
- What the marking evidence established.
- Which weakness was identified.
- What feedback intervention was delivered.
- Whether the student acted on it.
- How the student performed subsequently.
- Whether improvement persisted.

When captured responsibly and analysed appropriately, these observations create an increasingly valuable body of educational evidence.

This infrastructure could support research into questions such as:

Which feedback strategies are most effective for particular types of actuarial reasoning?

When should students receive worked examples rather than independent practice?

Which weaknesses are most persistent across examination attempts?

How does the effectiveness of feedback vary with experience and baseline ability?

What intervention produces the greatest improvement per minute of student effort?

The answers should emerge from experimentation and observation, not from assumptions about what sophisticated AI ought to accomplish.

© 2026 ThinkActuary

Governance and professional trust

Educational evidence is only useful if it can be trusted.

A mature system therefore needs clear standards for data quality, privacy, access, retention and use.

Student responses and educational histories should be protected. AI-generated diagnoses should be traceable to the underlying evidence. Automated decisions should be reviewable, and consequential marking decisions should remain subject to appropriate oversight.

The system should also be evaluated for systematic differences in performance across relevant student groups and answer styles.

The goal is to create an infrastructure that supports learning without compromising assessment integrity or professional standards.

9. A New Opportunity for Professional Examining Bodies

Professional examining organisations have traditionally concentrated substantial effort on developing reliable assessments, maintaining standards and ensuring consistency in marking.

Those responsibilities remain essential.

However, AI-enabled feedback and learning systems create an opportunity to extend the role of assessment.

Rather than viewing examinations solely as a mechanism for determining whether candidates have met a standard, examining bodies could increasingly use assessment evidence to support the development of future competence.

This does not require weakening standards or compromising the independence of examinations.

On the contrary, it creates a stronger distinction between two complementary functions:

Summative assessment: Determine whether the required standard has been achieved.

Formative learning: Use assessment evidence to help candidates develop the capabilities necessary to achieve that standard.

The infrastructure supporting these functions can share educational knowledge while maintaining appropriate separation, permissions and governance.

For professional bodies, the opportunity is not simply to automate an existing process.

It is to contribute to a more evidence-driven model of professional education, in which the effectiveness of learning interventions can be evaluated systematically.

© 2026 ThinkActuary

10. The ThinkActuary Research Agenda

ThinkActuary's proposed approach begins with a deliberately modest premise.

We currently provide text-based feedback on marked exam responses.

Rather than assuming that a more sophisticated interface, more elaborate AI model or fully autonomous tutor will improve educational outcomes, we intend to test that proposition through a progressive research and development programme.

The first phase is to understand the strengths and weaknesses of existing feedback.

The second is to compare alternative feedback strategies.

The third is to connect feedback to targeted opportunities for practice.

The fourth is to track individual development across multiple attempts.

The final ambition is to use accumulated evidence to improve the selection, timing and delivery of educational interventions.

At every stage, the primary question remains unchanged:

Does this capability help students perform better on a subsequent independent attempt?

This approach has three advantages.

First, it reduces the risk of investing heavily in sophisticated features without evidence of educational value.

Second, it builds the evaluation capabilities and datasets required for increasingly intelligent learning systems.

Third, it makes the educational research programme itself a source of potential differentiation.

Over time, the platform's strength may depend less on the sophistication of any individual AI model and more on the quality of the evidence connecting assessment, intervention and learning.

11. Conclusion: The Future of Feedback Is Learning

The next generation of educational feedback systems will not be defined simply by their ability to produce detailed explanations.

That capability is becoming increasingly accessible.

The more important challenge is to design systems that use evidence about performance to help students develop skills, provide opportunities to apply those skills and verify that improvement persists.

Learning science provides a credible direction.

© 2026 ThinkActuary

Artificial intelligence provides new capabilities for implementing that direction at scale.

But neither removes the need for careful instructional design, high-quality assessment, human judgement or rigorous experimentation.

The journey from simple text-based feedback to intelligent adaptive learning should therefore proceed through evidence-based milestones rather than technological ambition alone.

Start by measuring feedback.

Improve the content.

Create opportunities for action.

Test transfer and retention.

Use the resulting evidence to personalise learning.

Only then move towards increasingly autonomous adaptation.

The underlying educational principle is enduring:

> **Assessment reveals what a learner can do today. Feedback should help them do something better tomorrow.**

The opportunity is to make that improvement visible, measurable and progressively more achievable for every student.

© 2026 ThinkActuary

References

1. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. https://doi.org/10.1080/0969595980050102

2. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. https://doi.org/10.3102/003465430298487

3. Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10, 3087. https://doi.org/10.3389/fpsyg.2019.03087

4. Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, 85(4), 475–511. https://doi.org/10.3102/0034654314564881

5. 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

6. Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

7. Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, 144(7), 710–756. https://doi.org/10.1037/bul0000151

8. Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2), 176–199. https://doi.org/10.1177/1745691615569000

9. Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

10. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

11. *Unraveling the mechanisms and effectiveness of AI-assisted feedback in education: A systematic literature review.* (2025). *Computers and Education Open*, 9, 100284. https://doi.org/10.1016/j.caeo.2025.100284

12. *Comparing the quality of human and ChatGPT feedback of students' writing.* (2024). *Learning and Instruction*. https://doi.org/10.1016/j.learninstruc.2024.101894

About ThinkActuary

ThinkActuary is an actuarial education technology platform focused on helping students prepare for professional actuarial examinations through practice, marking and feedback.

This whitepaper outlines a proposed research and product development direction. It draws on established educational literature but does not claim that the proposed system has yet demonstrated the learning outcomes described. These outcomes form the subject of the proposed experimental programme.

© 2026 ThinkActuary. Thought leadership and research perspective.