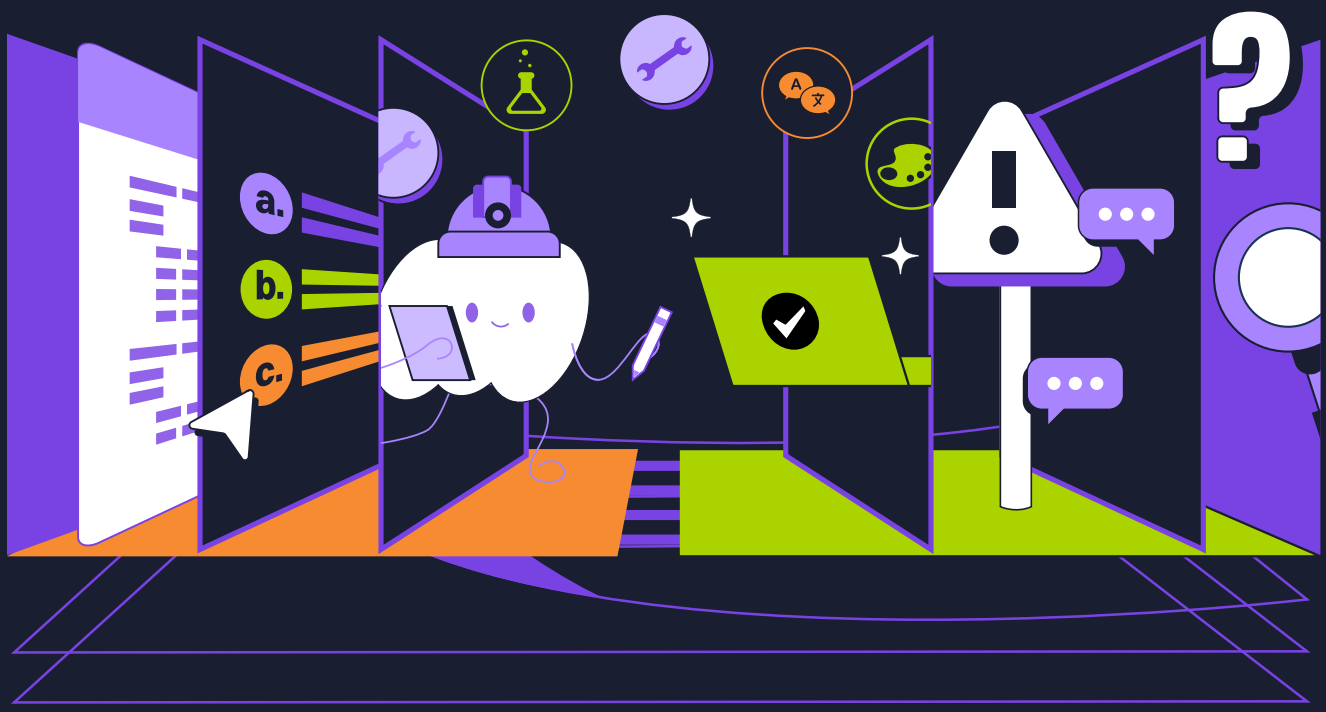
Gradeo

Making the Shift

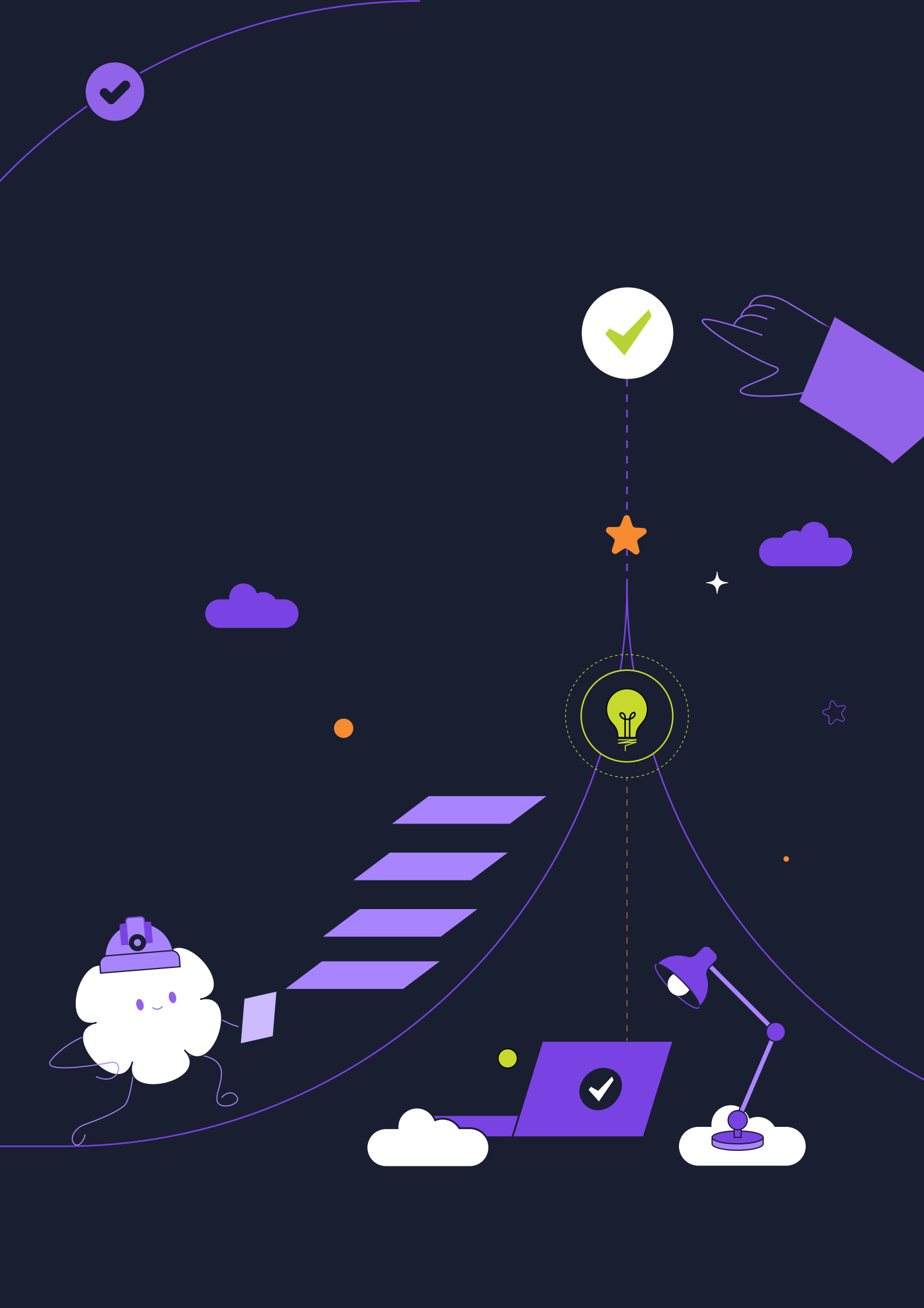
How online exams and AI are changing assessment



The CSSA Online Trial HSC Examinations 2026

CSSA

Catholic Schools NSW



Contents

Executive Summary	3
Background	7
Methodology	9
Highlights at a glance	10
Preparation and readiness	12
Exam experience	16
Marking and reporting	23
Reviews and integrity	26
The Future of Exams	28
AI and the Evidence of Learning	31
Implications for schools and assessment authorities	48
Conclusion	49
Acknowledgements	50
References	51

Executive Summary

As artificial intelligence (AI) continues to improve rapidly, the question of how human capabilities, knowledge and skills might fit into a world of pervasive AI, has become a matter of intense scrutiny.

If we believe that the future is one where human capability, knowledge and skills do, *or should*, continue to matter, then how we measure those attributes - and the development of those attributes - matters too.

How can we continue to fairly, authentically, validly and reliably measure student learning when AI can perform or convincingly imitate the traditional outputs of that learning without the "desirable difficulties" or "friction" of learning? That is the central and abiding challenge for educational assessment and measurement in the AI era.

This report examines that challenge, while evaluating the second year of large-scale online trial examinations delivered by the CSSA through Gradeo.

Drawing on surveys of over 650 students and 160 teachers from across 200 schools in New South Wales, focus-group discussions, and 8.4 million platform interaction events, it follows the examination journey from preparation and secure delivery through to marking, reporting and review. It then asks what students and teachers now consider credible evidence of learning in the age of AI.

The two strands meet in a common question: how can technology be used to enhance, enrich and diversify, rather than mask or narrow, human learning and assessment?

Take-home assessment has lost the confidence of the classroom.

Some 89% of teachers say AI has made it harder to trust that take-home work is a student's own, while 84% of teachers and 68% of students say genuine knowledge and capability have become harder to judge. Asked which assessment formats are fairest in the age of AI, only one teacher or leader out of 167 selected take-home written or research tasks.

Students and teachers instead placed supervised handwritten examinations, lockdown digital examinations and in-class written tasks at the top. The message is clear that a finished assignment can no longer be assumed to provide sufficient evidence of independent capability.

Students describe widespread AI use, including where it is prohibited.

Nearly two thirds of students (63%) say students in their year commonly use AI to complete homework or take-home assessment. Their comments range from generating practice questions and seeking explanations to submitting AI-produced work despite explicit school rules. One student described friends receiving full marks on take-home work and then failing the next task. These accounts show why the challenge extends beyond writing a clearer policy: schools need assessment conditions that make each student's understanding visible.

High-achieving students feel the pressure of work that is easier to imitate.

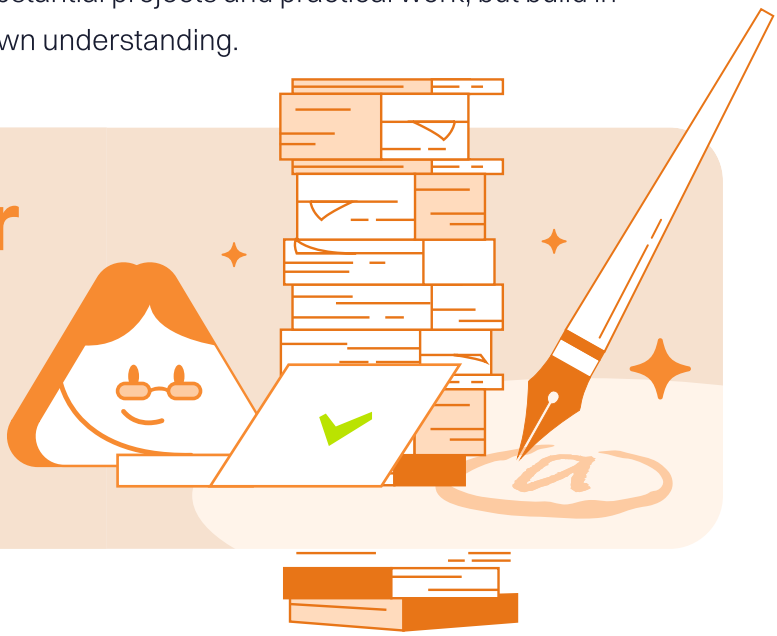
More than half of students say AI has made it harder for their work to stand out. That rises to 61% among students who describe themselves as high performing or above average, compared with 47% of students describing themselves as average. Students who feel their work is harder to distinguish are also more than twice as likely to support a temporary ban on take-home assessment while schools adapt. The concern is not confined to catching misconduct; it is also about whether knowledge, effort and achievement remain visible and fairly recognised.

Teachers want assessment redesigned rather than policed by better detection.

Some 91% of teachers and leaders agree that schools need to redesign assessment rather than rely mainly on detection or enforcement, with agreement outnumbering disagreement by 38 to one. Better plagiarism or AI detection tools ranked last among nine priorities. Teachers instead favoured explicit teaching of acceptable AI use, clearer policies and changes to assessment itself. At the same time, 86% say AI concerns are narrowing the range of tasks schools can confidently use. Their comments make the design challenge clear: preserve substantial projects and practical work, but build in manageable ways to establish the student's own understanding.

Just 1 teacher in 167

identified **take-home written** or **research tasks** as among the fairest assessment formats in the age of AI.



The second year of online examinations delivered a stronger student experience.

Students rated every part of the examination experience more positively than in 2025, covering login, the interface, Safe Exam Browser, navigation and answering questions. Every answer type repeated from last year also received stronger ease-of-use ratings, including substantial gains for Python and spreadsheets. Teachers reported easier setup, login and monitoring. Together, the results show a program becoming more familiar and more usable as schools and Gradeo build on the first year of delivery.

Preparation gave students confidence before the examination began.

Most students who sat an online examination (86%) recalled completing a familiarisation activity. The proportion feeling “very prepared” more than doubled, from 9% in 2025 to 22% in 2026. Almost ten students said the trial helped them prepare for the online Higher School Certificate (HSC) for every one who said it did not. Their leading request for the next examination period was more practice, particularly past papers and familiarisation with the specialist answer types they would encounter.

Students want to type, while subject suitability still shapes the transition.

Three in four students who sat an online examination would prefer to type their HSC examinations rather than handwrite them, broadly consistent with last year. Half supported more HSC subjects moving online; for every two students opposed, nearly five supported expansion.

Technologies and English attracted the strongest support, while Mathematics remained the clearest preference for paper. The pattern supports a subject-by-subject transition built around what students need to demonstrate and how naturally the available tools allow them to do it.

Online marking is returning time to teachers.

Teachers estimated an average saving of 27 minutes per student through online marking and/or AI-generated feedback, compared with 18 minutes reported by AI-feedback users last year. Uptake of AI feedback approximately doubled, while teachers retained responsibility for reviewing the comments and awarding marks. Since the 2025 report, session reporting has gone live and AI feedback has expanded to coding, spreadsheet and diagram responses. Teachers’ next requests are practical refinements for work at scale: faster marking pages, persistent criteria and sample answers, smoother scrolling and bulk script downloads.

Security preserved the examination conditions while schools asked for a simpler workflow.

Students sat in supervised rooms using timed access, individual passwords and Safe Exam Browser. The lockdown restricted outside applications, task-manager access and unauthorised exit while retaining the tools required by each examination. Approximately 8.4 million interaction events were recorded across the examination fortnight, supporting saved responses, event histories and session replay. Schools’ feedback identified the next priority: simplifying the secure-browser and onboarding workflow across different device environments and examination programs.

The findings point to a new balance between learning and proof.

AI can help students practise, seek explanations and approach difficult work. That benefit does not restore the evidentiary value of a finished take-home response. Schools and assessment authorities now need a broader program in which supervised examinations, observed practical work, oral explanation and well-designed checkpoints each provide credible evidence of different capabilities. The immediate task is to redesign assessment so that useful assistance can coexist with results that still represent the student.

The finished assignment is no longer enough.

Assessment must establish what the student knows.



Background

Technology is reshaping both the practice and meaning of educational assessment. As formal examinations move digital, schools are gaining new opportunities for secure delivery, more flexible answer tools, faster marking and richer feedback. This report first evaluates the second year of large-scale online Trial HSC examinations delivered through Gradeo. It then turns to a broader question raised by generative AI: when a tool can perform the reasoning, drafting or problem-solving itself, what does a finished piece of unsupervised work establish about the student?

For almost 60 years, CSSA has delivered Trial HSC examinations that help schools assess student performance and prepare students for the HSC. In 2025, CSSA and Gradeo delivered Australia's first large-scale online Trial HSC examinations. The program expanded substantially in 2026 with thousands of students completing synchronous examinations in Software Engineering, Enterprise Computing, and Science Extension.

More than 85% of NSW students in subjects already examined online for the HSC undertook their digital trial examinations and/or preparation on Gradeo. In 2027, English Extension 1 and English Extension 2 will also move online for the HSC. The 2026 program therefore provides evidence from schools already managing a transition that will soon reach a much larger group of students and teachers.

While generative AI has elevated the importance of supervised forms of assessment, for unsupervised alternatives, it has changed what a finished piece of work can prove. Within three years, these tools have reached an estimated 53% of the global population, while frontier models now meet or exceed human performance on several established benchmarks.[1] Students have ready access to systems capable of producing essays, explanations, code and other work at or above the standard many assessment tasks were designed to elicit.

That work cannot be reliably authenticated by software. Regulators warn that AI detectors produce false positives and false negatives and cannot establish misconduct on their own.[2] Teacher familiarity remains valuable, but it is not proof of authorship. Knowing a student can help identify an anomaly, but it cannot make a conventional take-home task AI-proof.

The concern extends beyond malpractice. The OECD describes a "mirage of false mastery", in which generative AI improves performance without producing equivalent learning gains,[3] while Pope Leo XIV identifies the corresponding human risk: "The speed and ease with which answers or summaries can be obtained risk extinguishing the desire to ask questions."[4]

OECD:
the "mirage of false mastery"

Better performance with generative AI does not necessarily mean stronger learning.[3]

Pope Leo XIV:
protecting the desire to learn

Easy access to answers can diminish the questioning and intellectual effort through which understanding develops.[4]

Assessment authorities are responding. In August 2026, NESA announced that Stage 6 school-based assessment programs could include no more than one take-home task, weighted at no more than 15%. The limit takes effect for HSC courses in Term 4 2026 and Preliminary courses in Term 1 2027, with exemptions for specified courses and major projects.[5]

Education systems are reconsidering the conditions in which students use AI.

The policy response extends beyond Australia. New York City and Los Angeles have restricted student access to generative AI in their public school systems, Norway has tightened examination conditions, and United Kingdom assessment guidance emphasises authenticating students' own work. These approaches differ; restrictions on classroom access address a different question from the conditions under which a result is awarded, but together, they place this evaluation within a wider discussion about how schools can use new tools while protecting learning and confidence in assessment. [6-9]

Jurisdiction	Development
New York City	For 2026-27, student-facing generative AI is subject to a moratorium in grades 2K-8, with limited use in high schools. This is a city school-system policy, not a ban across New York State. [6]
Los Angeles	LAUSD confirmed restrictions on student access to generative AI platforms for 2026-27 while it reviews instructional use and safeguards. [7]
Norway	For 2026 written examinations, general internet access is blocked, only a small set of online aids is available, the preparation component has been removed and supervision strengthened. [8]
United Kingdom	JCQ guidance requires sufficient supervision to authenticate candidates' work. Work outside school may remain possible where the centre is confident it is the candidate's own and subject requirements permit it. [9]

Building on Taking the Leap, Gradeo's 2025 evaluation, this report considers the complete examination experience—from preparation and delivery to integrity, marking and reporting—alongside the wider question now confronting schools: when AI can improve the work a student submits, and what evidence best demonstrates what that student knows and can do?

Methodology

Surveys

Over 160 teachers/ school leaders and 650 students completed a voluntary, anonymous survey, covering exam setup, delivery, usability, integrity, marking and reporting, and the impact of AI on assessment. This compared with 181 students and 72 teachers surveyed last year.

167 teachers/ school leaders

653 students

Focus Groups

Four hour-long sessions were held with teachers, students and school leaders to understand key survey trends. Discussions were reviewed for recurrence, but quotes illustrate perspective not prevalence.

4x 1hr sessions

Platform telemetry

Over 18.8 million timestamped interaction events (e.g. navigation, input, saves) were recorded by the platform from May to August 2026, including ~8.4 million across the examination fortnight. Support tickets and help centre traffic was separately analysed to identify common themes and trends.

18.8m
timestamped interaction events

8.4m
across examination fortnight

Highlights at a glance

More than 34,000 hours recorded in four months.

Between May and September, heavy platform adoption generated over 88,000 individual sessions averaging 23 minutes and 27 seconds. Across both trial and non-trial delivery, users logged more than 34,700 hours of cumulative activity—representing over 1,440 days of total system engagement by candidates and staff. Usage surged throughout August as schools conducted trial sittings and familiarisation sessions, highlighting the expansion of digital delivery during peak assessment periods.

Only one teacher in 167 chose take-home assessment as one of the fairest formats.

Only one teacher or leader selected take-home written or research tasks among the fairest assessment formats in the age of AI. Supervised handwritten exams, lockdown digital exams and in-class written tasks led the responses.

Read more: AI and the Evidence of Learning p. 36

Teachers and students agree: take-home work is becoming harder to trust.

84% of teachers and 68% of students say AI makes genuine knowledge and capability harder to judge from take-home work.

Read more: AI and the Evidence of Learning p. 31

Teachers back assessment redesign by 38 to one.

Teachers' agreement that some assessments need redesign outnumbers disagreement by 38 to one. Better detection tools ranked last among nine school priorities.

Read more: AI and the Evidence of Learning p. 41

AI is making it harder for high-achieving students to stand out.

61% of self-described high-performing or above-average students say AI makes it harder for their work to stand out, compared with 47% of students describing themselves as average.

Read more: AI and the Evidence of Learning p. 34

Nearly two in three students say AI use is common in their year group.

63% say students in their year commonly use AI to complete homework or take-home assessment. Their comments describe both valuable assistance and use that bypasses school rules.

Read more: AI and the Evidence of Learning p. 30

The student exam experience improved across every measure.

Student ratings improved across all five examination-experience measures and all eleven answer types repeated from 2025. Teachers also reported easier setup, login and monitoring.

Read more: Exam Experience p. 16

Teachers estimate online marking saves 27 minutes per student.

Teachers estimated saving nearly half an hour through online marking and/or AI feedback. Last year's report recorded 18 minutes among AI-feedback users. Teachers continued to review feedback and award marks.

Read more: Marking and reporting p. 23

AI feedback use doubled and expanded to complex answer types.

AI-feedback uptake approximately doubled among exam teachers. Feedback now supports diagrams, spreadsheets and coding, extending a benefit teachers valued last year.

Read more: Marking and reporting p. 24

For every student who did not feel better prepared for the HSC, almost ten did.

Almost ten students said the online trial helped them prepare for the HSC for every one who said it did not.

Read more: Preparation and Readiness p. 14

Three in four would rather type

75% of students who sat an online exam would prefer to type their HSC examinations, broadly consistent with last year.

Read more: The Future of Exams p. 28

Preparation and readiness

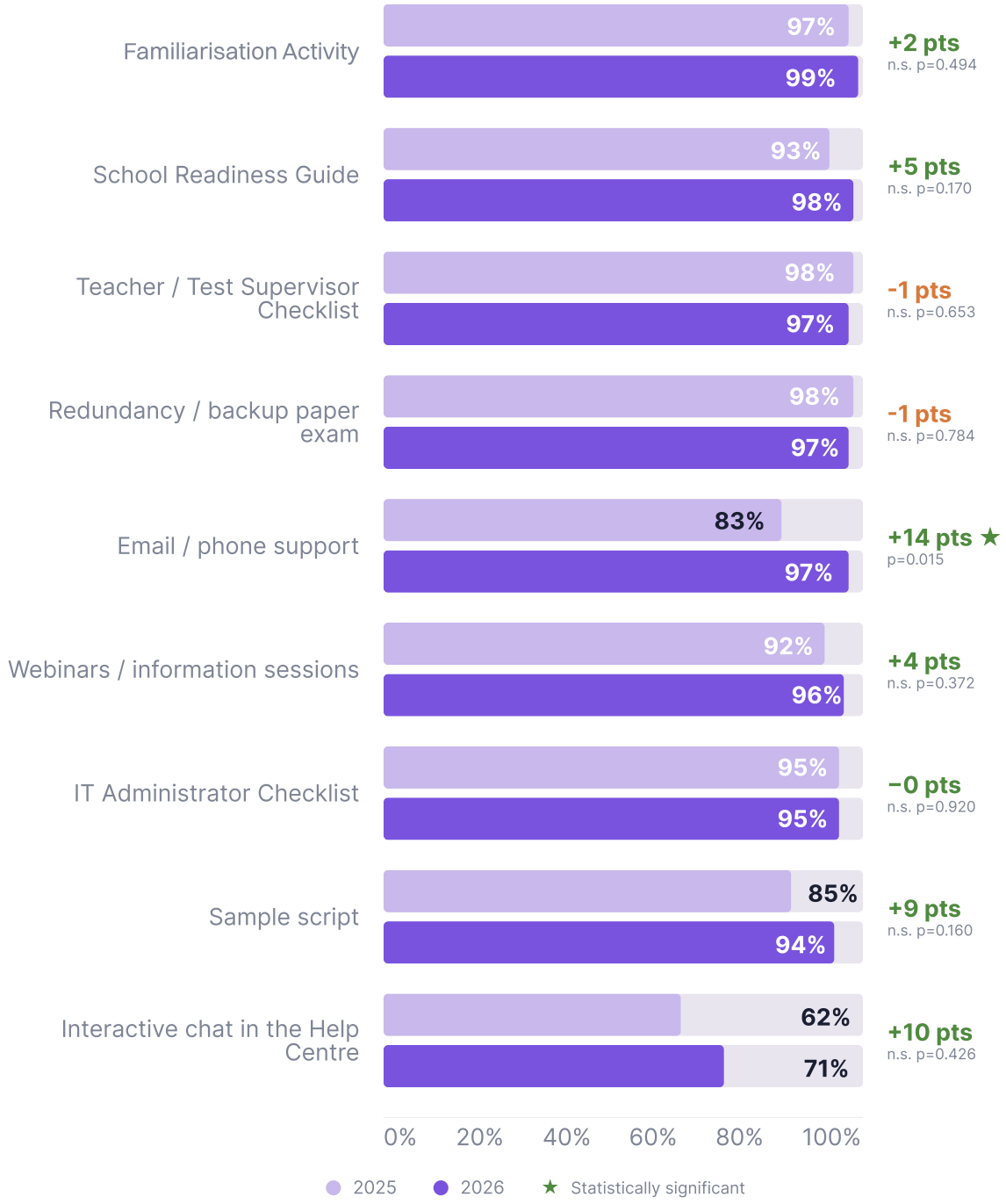
Teachers placed greater value on every preparation resource repeated from last year.

Last year's evaluation showed the importance of giving teachers, supervisors, IT staff and students a common starting point. The School Readiness Guide continued to bring preparation materials together in 2026, supported by checklists, webinars, familiarisation activities and a dedicated ITE Resource Hub to clearly differentiate between exam licence workflows. These resources helped schools coordinate the work that precedes an examination: preparing devices, understanding secure access and giving students a chance to use the answer tools.

Teachers rated that preparation more highly this year. The share selecting "very useful" increased for every resource repeated from 2025. Among staff who used them, 99% found familiarisation useful, 98% the School Readiness Guide and 97% the Teacher/Test Supervisor Checklist. Their comments recognised a smoother process as schools built experience.

> **Conducting the examination was so much better than last year. Smoother process. This did relieve the stress involved considerably.**
>
> — Teacher

Staff who rated each preparation resource moderately or very useful (%)

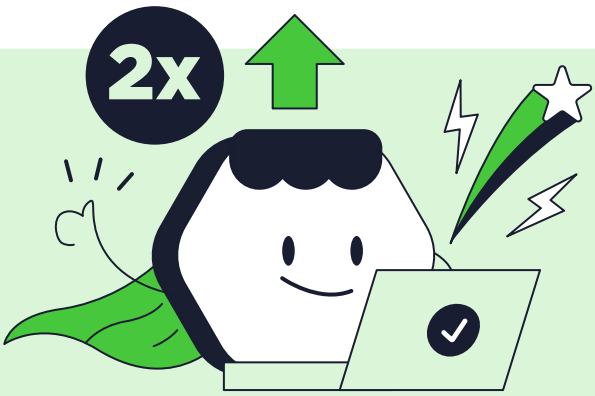


More than twice the proportion of students arrived feeling very prepared.

Most students who sat an online exam (86%) recalled completing a familiarisation activity. Two thirds felt prepared to use the system, broadly consistent with last year, but the share feeling “very prepared” rose from 9% to 22%. Students who completed familiarisation were also more likely to feel prepared, at 69% compared with 54% of those who did not. The association reinforces a practical lesson from 2025: a rehearsal gives students an opportunity to learn the examination workflow before they need to use it under timed conditions.

More than twice the proportion of students felt “very prepared”

22%, compared with **9%** in 2025.

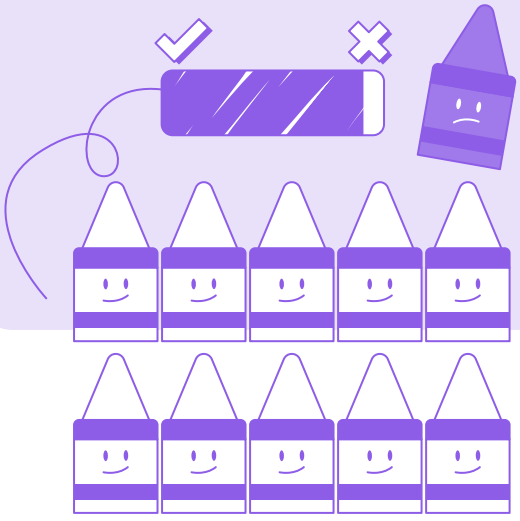


Students’ leading request was more practice.

Asked what would help them feel more confident, students most often selected more past papers or practice exams (61%), followed by more practice using the system (39%) and answering questions online (23%). Their comments made the request specific: practice should cover the coding, drawing and other answer types they will encounter in the paper. This brings subject preparation and platform familiarisation together, allowing students to concentrate on the question rather than working out how to enter their response.

> **Wider spread** of question types — practice did not include coding.
>
> — Year 12 student

Approx 10 students said the online trial helped them prepare for the HSC for every 1 who said it did not.



The online trial gave students a useful rehearsal for their final examinations.

Students continued to value experiencing the format before their final examinations. Almost ten said the online trial helped them prepare for every one who said it did not, closely echoing last year’s finding. Familiarisation introduces students to the system, whereas a full trial examination adds the demands of timing, moving between questions and producing answers under supervision. The preparation finding therefore concerns more than access to software: it reflects the value students place on knowing what an online examination will feel like before the final sitting.

Responsive support helped schools prepare.

School preparation continued into the examination period. Schools' support requests covered secure-browser setup, licensing, content and platform questions. Email and phone support were rated useful by 97% of staff who used them. We received 428 support tickets relating to exams, with a large majority of tickets relating to the lockdown browser, licence provisioning, and exam content. Less than 1 in 3 tickets was related to a platform query. Teachers' comments described the value of being able to reach someone who understood their setup and could help them resolve a query, an important part of running examinations across different school environments.

> I found [the team] super **helpful and responsive to queries** and have always been pleased with the outcomes of contacts.
>
> — Teacher

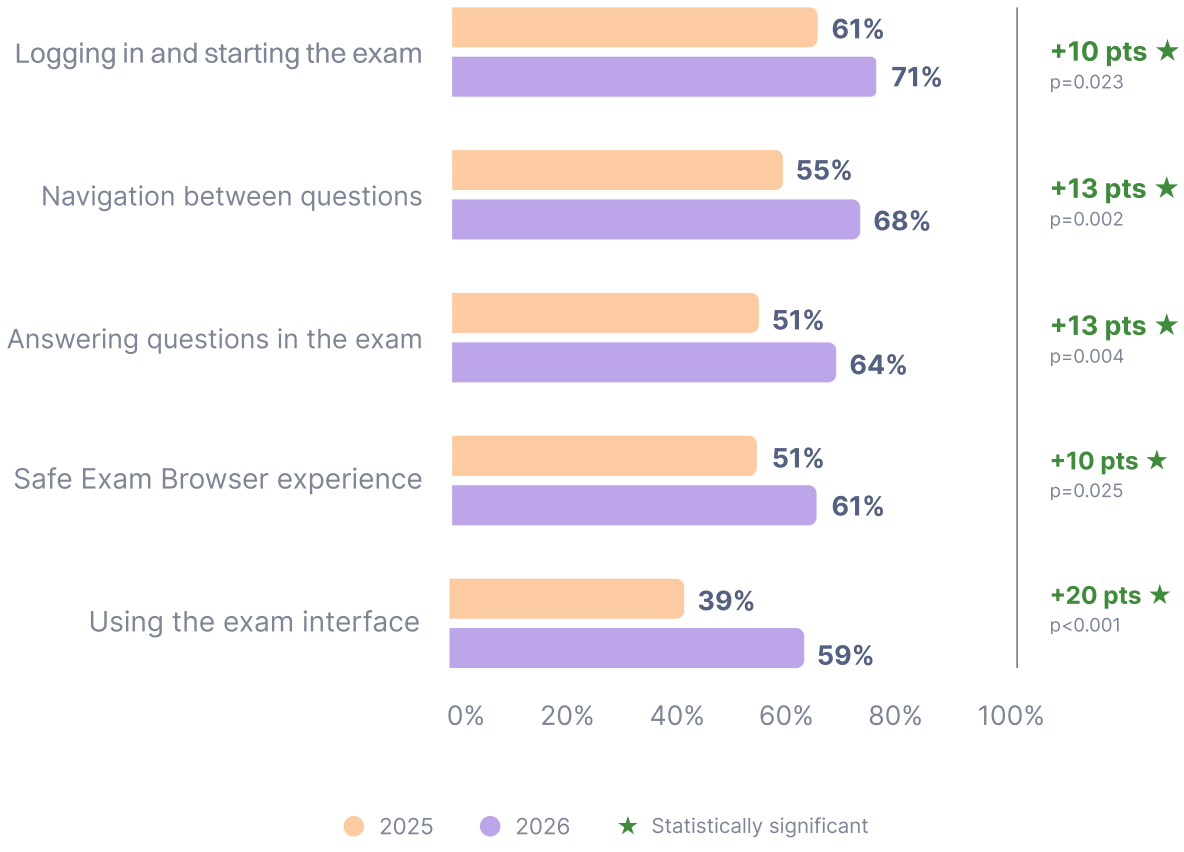
Exam Experience

Students rated every part of the online examination experience more positively.

Students rated all five experience measures more positively than in 2025, covering login, the interface, Safe Exam Browser, navigation and answering questions. The largest gain was in the interface, where favourable ratings rose by 20 percentage points. Improvements therefore extended across the workflow students used during the exam, rather than being confined to one tool.

Students rating each part of the online exam good or excellent

2025 and 2026 (%)



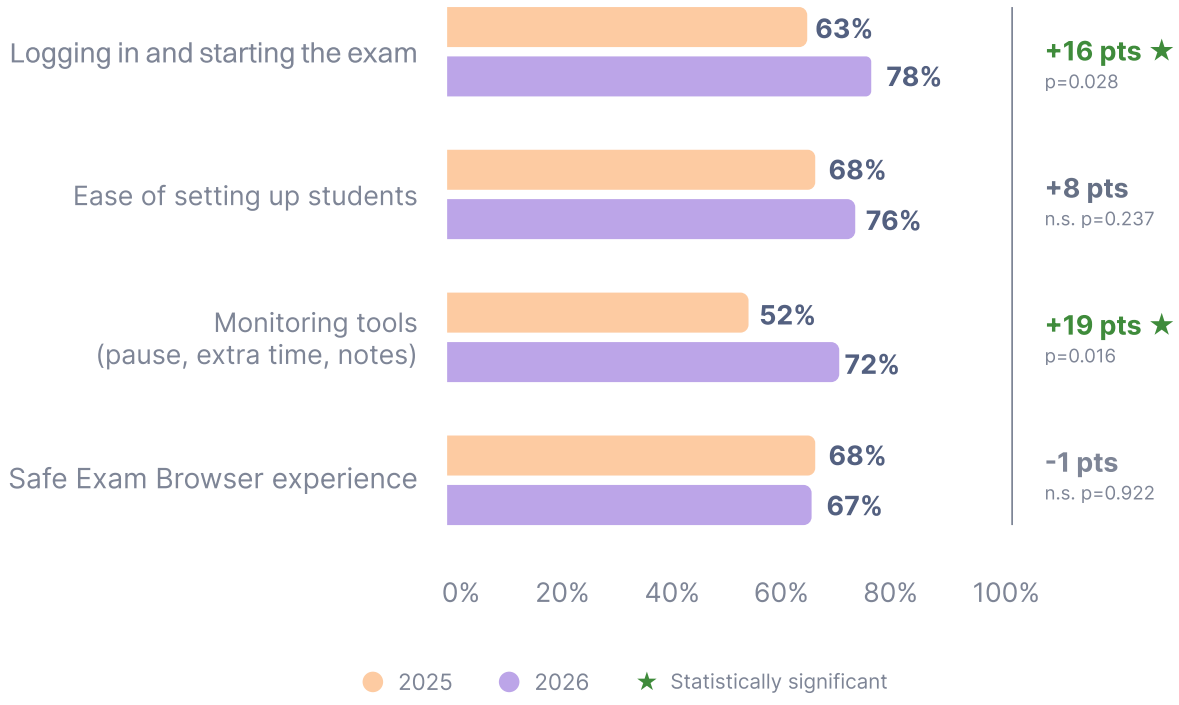
Teachers found setup, login and monitoring easier.

Teacher ratings improved for setting up students, starting the exam and monitoring progress. Monitoring showed the largest gain, with 72% rating it good or excellent, up from 52%. Teachers' Safe Exam Browser ratings were broadly unchanged, reinforcing the priority of simplifying secure access. Feedback described how pause and extra-time controls helped supervisors respond to individual circumstances.

> **One student's computer crashed** and had to use the backup laptop, nothing to do with Gradeo, but the **pause function** was very handy.
>
> — Teacher

Staff rating each part of exam delivery good or excellent

2025 and 2026 (%)



Every answer type received stronger ease-of-use ratings.

All eleven answer types received more favourable student ratings than in 2025. Some of the largest gains were in specialist tools: Python increased from 45% to 62% easy or somewhat easy to use, and spreadsheets from 39% to 58%. These tools matter because they allow students to demonstrate skills through code, data and diagrams within the examination itself. Teachers described the ability to write and test code as a benefit that a handwritten paper could not reproduce.

> The online exam allowed students to **write Python code and test it.** This can't be done on a paper exam.
>
> — Teacher

Students finding each answer type easy or somewhat easy to use

2025 and 2026 (%)

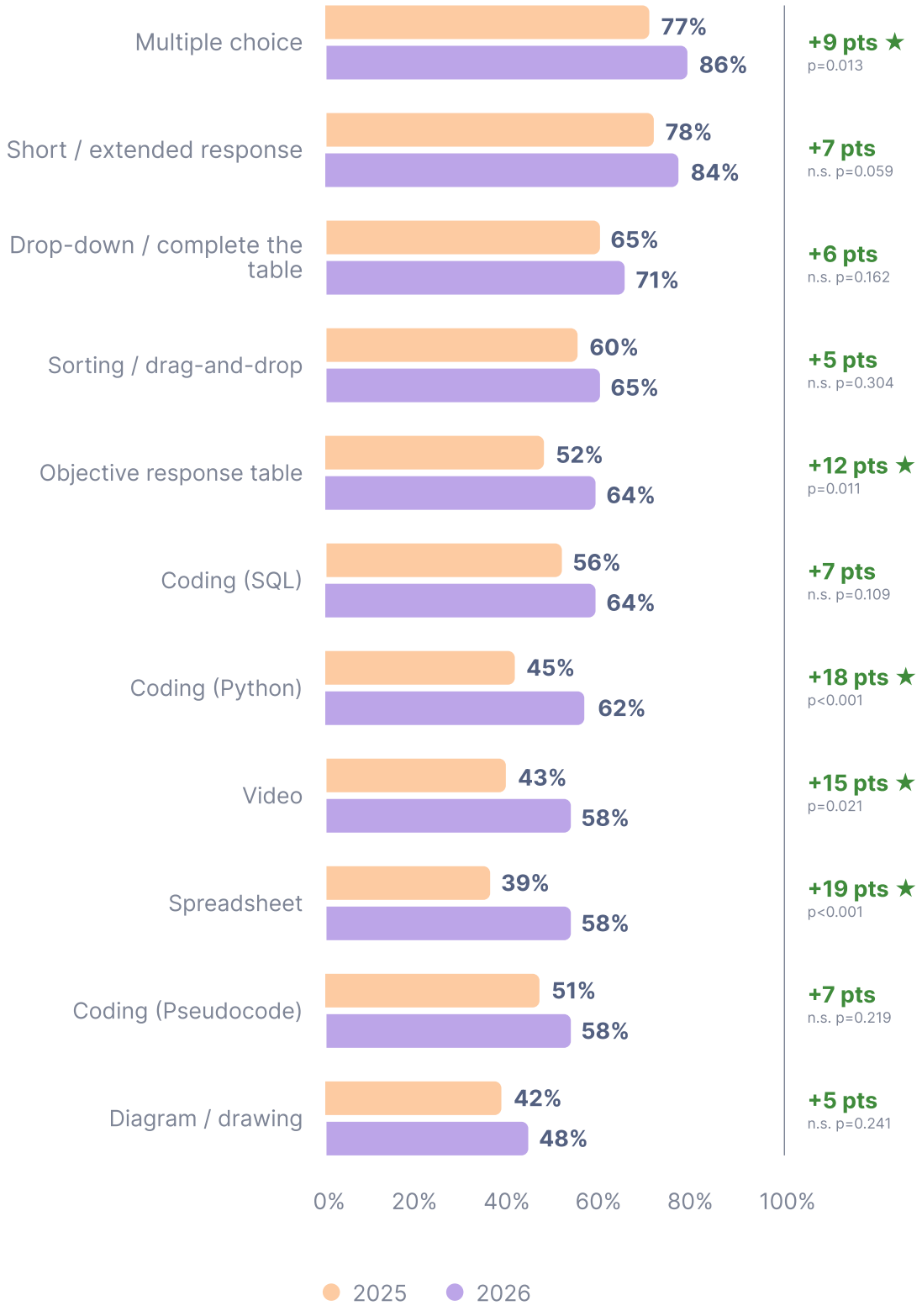




Photo Credits: Students sitting Gradeo examinations at St Joseph's International Catholic College, Papua New Guinea.

Online delivery connected students across locations.

Teachers working across locations described reduced printing, distribution and collection, with students able to sit the same paper at their local school. Students across metropolitan, regional and rural New South Wales, Singapore, Malaysia, Papua New Guinea and Hong Kong sat exams on Gradeo.

> **Students right across the region** could sit the same paper at their own local school, with no papers posted out and back. No printing, packing or posting, and **responses were ready as soon as the exam closed.**
>
> — Teacher

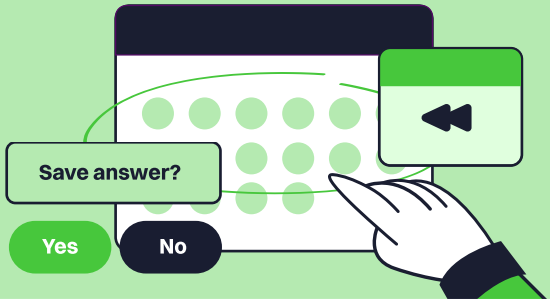
Ten times faster, at four times the scale

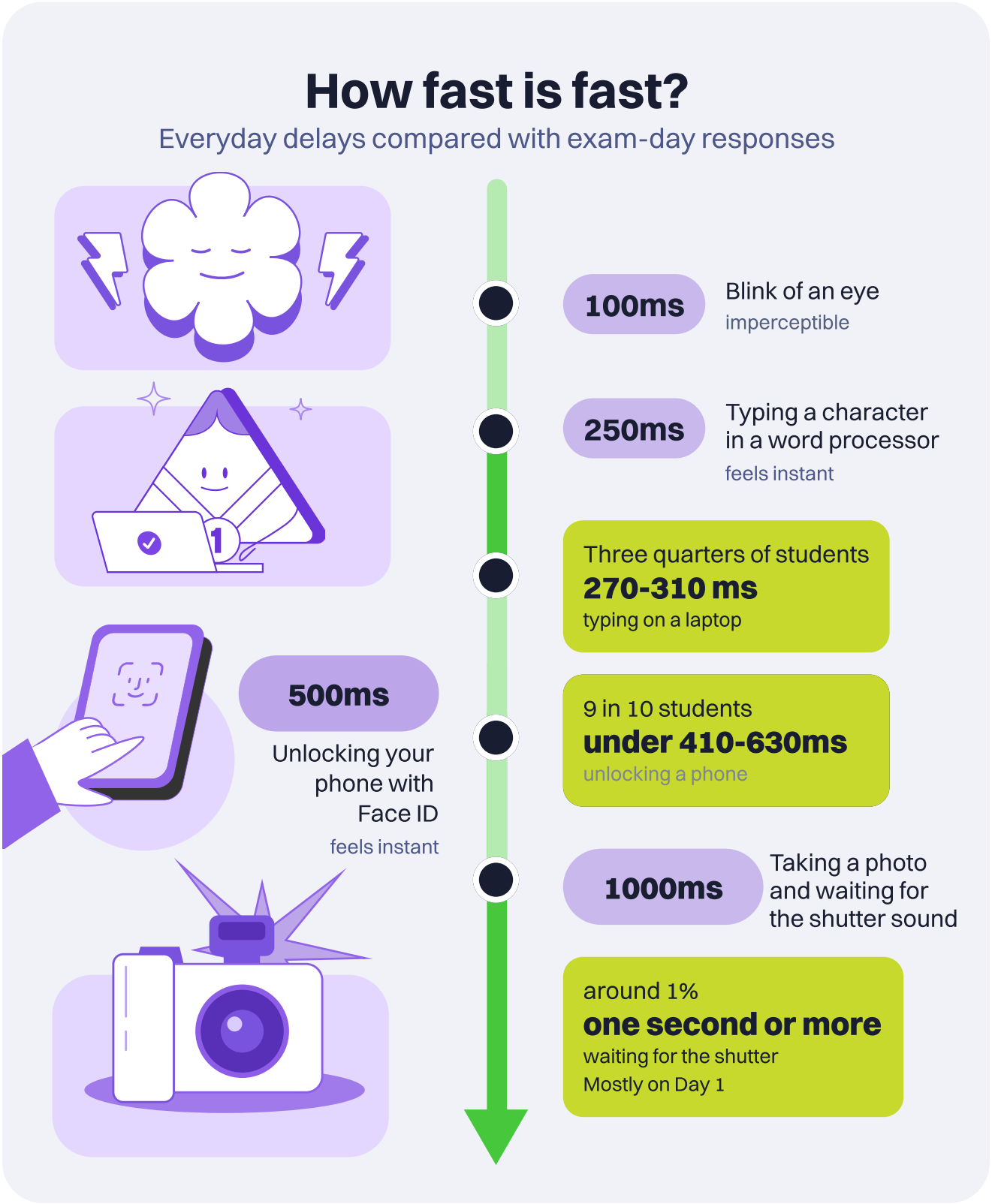
Gradeo went into the 2026 exam window having load-tested the platform well beyond the 2025 peak. Across the exam period the platform recorded about 18.8 million events between May and the end of August (page views, clicks, answer saves, logins and logouts), peaking at 3.9–4.5 million events a week in the first fortnight of August and 8.4 million across the two exam weeks.

Through that peak, average server response time was 27–28 ms on the exam days of 10 and 11 August, roughly ten times faster than the 2025 exam days (300 ms and 929 ms). A period of higher-than-anticipated demand during the Software Engineering sitting on 7 August caused delays for some students and manifested in an error message warning. Server and database capacity was increased during the sitting, with the additional capacity retained for subsequent exams.

Schools from across borders participated

spanning Australia, Hong Kong, Malaysia, Papua New Guinea and Singapore





Schools' feedback identifies practical refinements for the next examination period.

Students asked for more intuitive drawing interactions and practice with subject-specific symbols. Staff called for a simpler secure-browser workflow and clearer connection and save information when a sitting is interrupted. These requests are informing Gradeo's priorities for onboarding, answer tools and continuity through connection interruptions.

School feedback	Improvement priority
Make setup easier for teachers and test supervisors	Simplify the SEB workflow and onboarding across ITE, CSSA and NSW Department of Education arrangements, and the technical check for devices.
Make drawing and diagram tools more intuitive	Feedback on these interaction types will be provided to NESA. Continue to refine common interactions and provide practice covering the symbols and shape types students need.
Keep work safe and make save status clear	Strengthen transparency around resilience through interruptions and give supervisors clearer connection and save information to bolster confidence of test supervisors and students.
Make paper and online content consistent	Strengthen quality assurance checks of question content, stimulus, reading time and working time before release, with an additional review layer to be added to the process.

Delivery at scale

approximately

8.4 million interaction events

were recorded across the **two examination weeks.** These records include navigation, input, saving and submission activity; they are not a count of examinations.



Marking and Reporting

Online marking is returning nearly half an hour per student to teachers.

Marking was one of the clearest and material benefits reported by teachers. Typed responses removed the need to decipher handwriting, multiple-choice questions were marked automatically, and teachers could double or triple mark, hide marker feedback or scores, anonymise marking, split-marking and move through the same question across a cohort without sorting or sharing paper booklets. Their comments connected these changes with quicker marking and more thorough individual feedback.

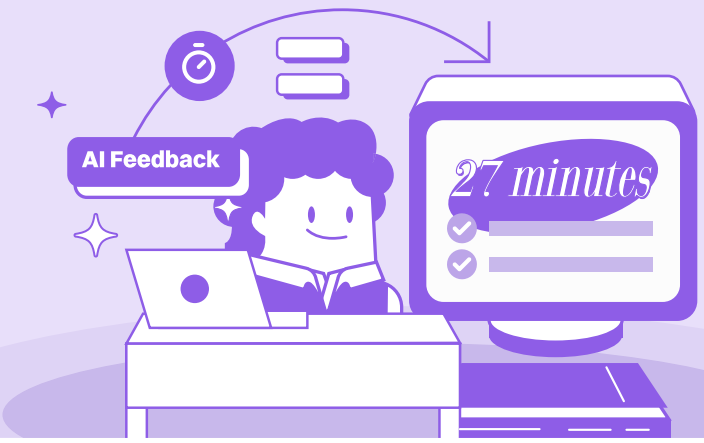
Teachers estimated an average saving of 27 minutes per student through online marking and/or AI-generated feedback. Last year's report recorded 18 minutes among AI-feedback users. The estimates point to a substantial practical benefit in a task teachers repeat across an entire class: time saved moving through scripts can be spent reviewing the response and giving students useful feedback.

> **Marking online** allowed for quicker and more thorough **individual feedback**
>
> — Head Teacher

> The **biggest advantage** of running exams online was that it saved time and made it **easier to generate feedback.** It also allowed teachers to **quickly access and print copies of students' responses** when required.
>
> — Teacher

27 minutes per student

enough time to **watch an episode of Seinfeld,** saved through online marking and/or AI feedback.



AI-feedback uptake approximately doubled, with teachers continuing to award the marks.

Compared with 2025, approximately twice the proportion of exam teachers used AI-generated feedback. Gradeo's workflow retains the same division of responsibility: AI drafted comments using marking criteria and supporting guidance; teachers reviewed the feedback, edited it as needed and determined the mark. Teacher comments are informing further refinements to the length and tone of feedback.

Feedback now supports coding, spreadsheets and diagrams

Last year, teachers asked for AI feedback to extend beyond written responses to the complex answer types used in computing and other practical subjects. That capability has now been delivered. This year, teachers were able to draw on drafted feedback for code, spreadsheet and diagram responses, while retaining responsibility for reviewing the comments and awarding marks. This extended the assistance to work where a student's method, structure and working can be as important as the final answer.

You said → We did

AI feedback now covers coding, spreadsheet and diagram responses.

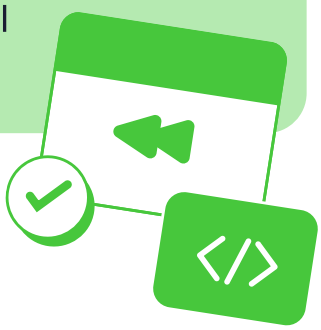


Session reporting gives teachers more to work with after the exam.

An examination produces more information than a final mark. Since last year's report, session reporting has gone live, allowing teachers to review question-level performance, time spent on questions and response word counts alongside students' answers. This gives teachers a starting point for examining how students approached the paper, identifying questions for class discussion and planning feedback. Read with the responses themselves, the information can support the next lesson as well as the record of achievement.

You said → We did

Session reporting shows how your class performed in an exam relative to everyone else in the state at the question, outcome, and topic level



> I'd like to be able to **produce a summary for feedback** as a class and individually, ready for next lesson back in class.
>
> — Teacher, Government school

Bulk script download is the next step in reducing paper administration.

Printable student scripts and CSV results downloads addressed two requests from 2025. Teachers could retain assessment records and access responses without exchanging exam booklets. This year's feedback asks for completed scripts to be downloaded together, saving the repeated steps of opening and saving each student's PDF.

> One of the benefits is being **able to mark independently** and not having to share the exam booklets.
>
> — Teacher

Small marking refinements matter across a large cohort.

Teachers asked for criteria to stay open as they move between responses, smoother scrolling and easier access to the question and supporting information. These requests are shaping minor usability refinements to the marking screen, alongside a focus on speed for very large cohorts.

> If I **open the marking guide,** can that please stay open until I close it, as opposed to disappearing each time I open a question?
>
> — Head Teacher

Reviews and Integrity

Secure conditions preserved access to examination tools while restricting outside assistance.

Students sat in supervised rooms at their schools on school-provided or personal devices. Safe Exam Browser, an examination launcher and one-time-password access worked alongside invigilation to restrict access to the examination and the resources available during the sitting. Students could use the tools provided within the paper, including coding, spreadsheets and diagrams.

Control	What it meant during the exam
Leaving the lockdown browser	Students could not quit the secure session without the specific exit or quit password held by the supervisor.
External applications and system access	Access to other applications and the task manager was restricted; Ctrl-Alt-Delete access was disallowed in the configured examination workflow.
External displays	The default configuration disallowed external monitors. A backup launcher was issued to schools to accommodate desktop setups that relied on a separate display.
Permitted examination tools	Students used the question content and answer tools provided within the examination, including subject-specific coding, spreadsheet and diagram tools.
Local supervision	School supervisors monitored the room and managed authorised interruptions, adjustments and exit from the secure session.

The secure-browser workflow will be simpler across different school setups.

The default configuration disallowed external displays, which also affected desktop computers relying on separate monitors. Gradeo supplied two launcher files for each examination to accommodate these setups. The extra configuration preserved secure delivery but added administration for schools. Simplifying this workflow, alongside onboarding across ITE, CSSA and NSW Department of Education arrangements, is a focus for the next program.

Session replay gives schools evidence for reviewing a concern.

For schools reviewing a student's concern, the final script is only part of the available evidence. Saved responses, event logs and session replays showed the sequence of navigation, input, saving and submission. A review can examine what happened before a response changed or a session ended, rather than relying only on recollection. Teachers described using replays to see activity within a session; they also asked for clearer monitoring descriptions so supervisors can understand what triggered a flag and respond appropriately.

> **Session replay** shows the mouse movements as the **student draws the diagrams** for those questions
>
> — Examination Coordinator

The Future of Exams

Three in four students would rather type their HSC exams.

Among students who sat an online exam in 2026, 75% preferred typing their HSC exams to handwriting, broadly consistent with 78% last year. Students described the physical demands of extended handwriting, while teachers recognised legibility and the opportunity to use digital tools to demonstrate subject knowledge.



> **For English**, it'd be excellent if we could **type all of our essays**, because you don't get that cramp in your hand when you've been writing for two hours in Paper 2.
>
> — Year 12 student

Three in four students

who **sat an online exam** would prefer to **type their HSC exams.**



Teacher support for more online examinations increased.

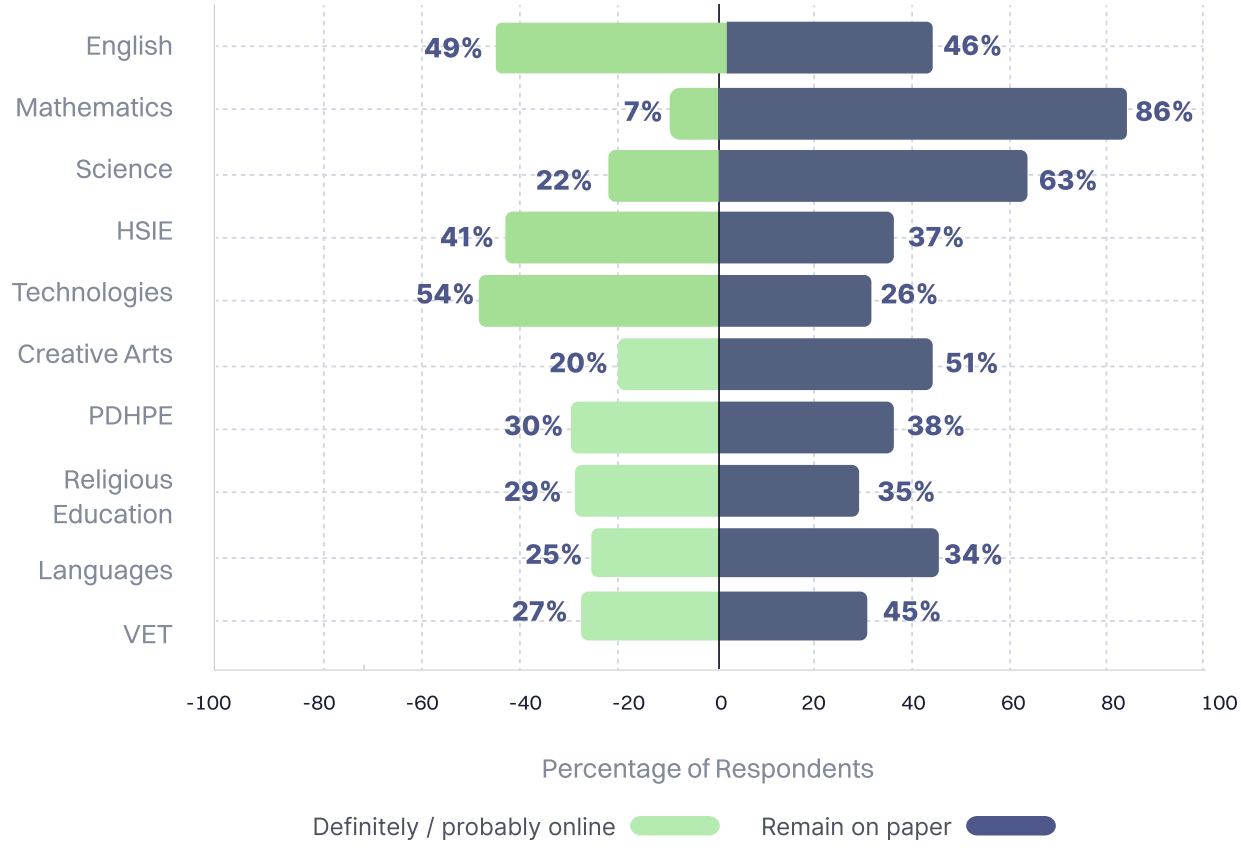
Among staff who administered an examination, support for moving more exams online rose from 38% in 2025 to 49% in 2026. Among student sitters, 50% supported expansion, with supporters outnumbering opponents by nearly five to two. Among students, Technologies and English attracted the strongest student support for moving more HSC examinations online, continuing last year's pattern.

Mathematics remained the clearest preference for paper, with 86% favouring handwriting. Students' comments explain the distinction: typing can make an extended response easier to produce, while mathematical working needs a quick and natural way to enter symbols, equations and intermediate steps. These preferences support a subject-by-subject transition, with NESA placing English Extension 1 and 2 online from 2027. [10]

> **Maths** kind of only works on paper, because you can't really do an online maths exam. **Writing out all of your working** takes too long, rather than just drawing it.
>
> — Year 12 student

> I think the fairest way is a balance. There's not **one catch-all assessment task type.** It's finding a balance that allows everyone to have a chance to try their best in their best method.
>
> — Year 12 student

Student exam sitters' preferences for each subject's HSC examination format (%)
(with +/- net support)



AI and the Evidence of Learning

For generations, the take-home assessment has been a central pillar of schooling. Take-home tasks have been like rites of passage from the book review poster to the research essay to the diorama; the reflection journal to math worksheet. But in a world of pervasive AI tools that can do or assist many of these in seconds, the foundational premise of unsupervised assessment risks permanent disruption.

Where take-home assignments have traditionally served two purposes – an opportunity for learning and evidence of that learning – the findings in this section examine what happens when those purposes pull apart. Students and teachers describe growing difficulty connecting a finished response to the knowledge of the student who submitted it.

This section moves from how students use AI to what that use means for trust, fairness and recognition. It then examines the assessment formats respondents favour and the changes teachers want schools to prioritise. The central question is whether take-home work can still carry the same weight as evidence of independent capability.

AI use is common, including where school rules prohibit it.

Nearly two thirds of students (63%) say students in their year commonly use AI to complete homework or take-home assessment. Written and spoken accounts describe assistance ranging from generating practice questions to producing assessment content. Some students openly described using AI despite school restrictions, showing why a stated rule and classroom practice are different kinds of evidence.

> For our school, we have a **blanket ban on all AI use in assessment tasks** [...] even though it says you can't use AI, everyone, including me, used AI to generate ideas for their project."
>
> — Year 12 student

63% of students

say students in their year commonly **use AI to complete homework** or take-home assessment.



Teachers and students both question what take-home work now proves.

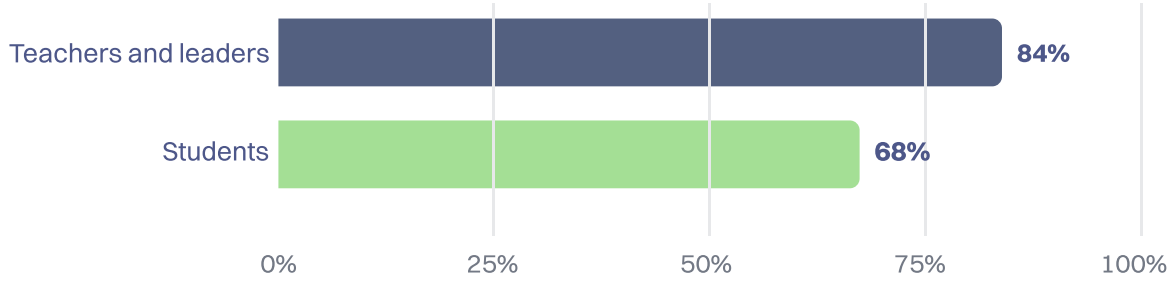
Some 89% of teachers say AI makes it harder to trust that take-home work is a student's own. Asked about judging genuine knowledge and capability from that work, 84% of teachers and 68% of students agree it has become harder. Students describe the problem in recognisable terms: an accomplished assignment followed by difficulty demonstrating the same understanding in another task. Such a contrast does not establish misconduct on its own, but it explains why the finished product may leave teachers with unanswered questions about learning.

> I've had friends who, on a **take-home task**, score **100 out of 100**, and then the next task right after, score really low. I've had some friends fail right after.
>
> — Year 12 student

> **AI tools** have meant that the average quality of work across a cohort has increased. The **downside** is that less of this work can actually be attributed to the student...
>
> — Year 12 student

> The **skill floor has lowered**, and it is harder for examiners to tell where there are **gaps in knowledge.**
>
> — Year 12 student

Students and staff agreeing that AI makes genuine capability harder to judge from take-home work (%)



Research insight

The illusion of learning: stronger performance can mask weaker understanding.

In a field experiment involving nearly 1,000 secondary mathematics students, Bastani and colleagues found that a general-purpose GPT-4 interface improved grades during assisted practice by 48% relative to a no-AI control. In a later unaided assessment, that group performed 17% worse than the control. A tutoring interface with learning safeguards largely mitigated the negative effect. This is external experimental evidence, separate from the Gradeo survey. [11]

The design of assistance matters. In a separate randomised trial involving 194 university physics students, Kestin and colleagues found that a purpose-built AI tutor produced stronger immediate learning gains in less time than an active-learning class. The studies concern different learners and interventions; together they show why "using AI" is too broad a description to predict its effect on learning. [12]

External mathematics study: relative change in grades during GPT-4-assisted practice and a later unaided assessment

+48%

Performance during **AI-Assisted Practice**

-17%

Performance later, without **AI access**

Research insight

Why the effort between question and answer matters.

Research on “desirable difficulties” distinguishes immediate performance from durable learning. Retrieving knowledge, generating a solution and revisiting an idea can strengthen retention even when practice feels harder. The relevant question is which learning process assistance supports or removes. [13]

Students and teachers describe a ‘illusion of false mastery’ or confidence.

Student and teachers described a tension between self-confidence in being able to complete a task and developing the knowledge to do it independently. Teachers recognised the value of quick explanations while worrying about the loss of productive struggle. Students described reduced motivation, the ease of taking shortcuts and pressure on those completing work without AI.

“**Lost:** sound understanding of **fundamental concepts**, depth of knowledge, problem-solving skills, memorising capabilities.

Gained: confidence in giving things a go, whether it is coding or an essay; **false belief** that they now know things.”

— Teacher

“**Faster access** to explanation and feedback - a student stuck at 9pm can get unstuck instead of stopping.”

— Teacher

"The **productive struggle.** Working through confusion is where a lot of the learning happens, and that step is easy to skip now."

— Teacher

"Students have **lost their ability to perform long-horizon tasks such as major projects** [...] It is sad to see such a crazy tool smoothing brains rather than developing them, in my opinion."

— Year 12 student

"**Students have lost a lot of motivation** to actually complete homework themselves, as the entire time there is the **subconscious thought that I could just use AI** and do something else."

— Year 10 student

> I experienced this multiple times, spending **long periods of time completing work** while others completed it within minutes, where the teacher kept giving more work thinking the class was finished, causing my own workload to increase as I was completing the work **without AI assistance.**"
>
> — Year 12 student

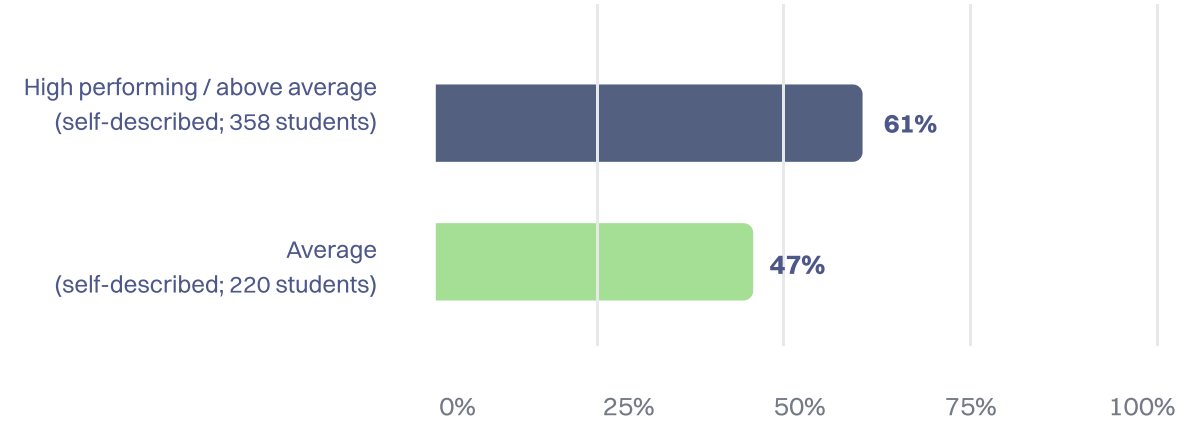
Higher-achieving students are particularly concerned about recognition.

More than half of students (54%) say AI has made it harder for their work to stand out and show what they can do. Among those describing themselves as high performing or above average, 61% agree, compared with 47% of students describing themselves as average. The comments reveal a concern about recognition as well as fairness: students want the knowledge and effort behind their work to be visible when others can produce polished responses quickly.

> Since **many students use AI** to rapidly complete assignments with general, wide information, it means that **students who would rather not use AI a**re sometimes forced to use AI to keep up with these other students.
>
> — Year 9 student

> **AI has offered fast clarity** for unsure or new concepts; however, its overuse has led to assignments being a less accurate depiction of work ethic and skill, as authentic work is rewarded less than AI.
>
> — Year 12 student

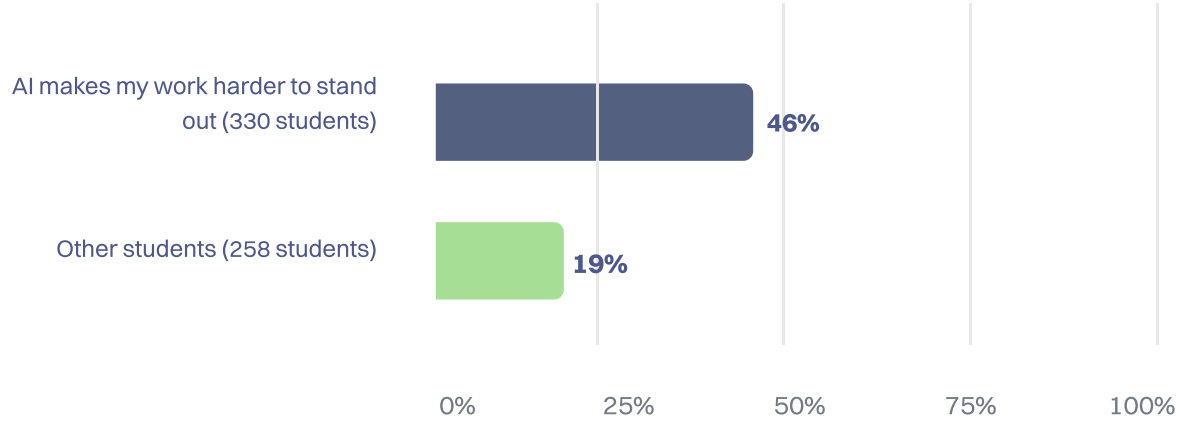
Students saying AI makes their work harder to distinguish, by self-described academic performance (%)



Concern about recognition is associated with stronger support for a temporary ban.

Students who say AI makes their work harder to stand out are more than twice as likely to support a temporary ban on take-home assessment while schools adapt: 46%, compared with 19% of other students. This comparison spans achievement groups. It shows that support for restrictions is associated with students' concerns about recognition, adding a student perspective to a debate often framed around the preferences of teachers and assessment authorities.

Student support for a temporary take-home assessment ban, by concern about work standing out (%)



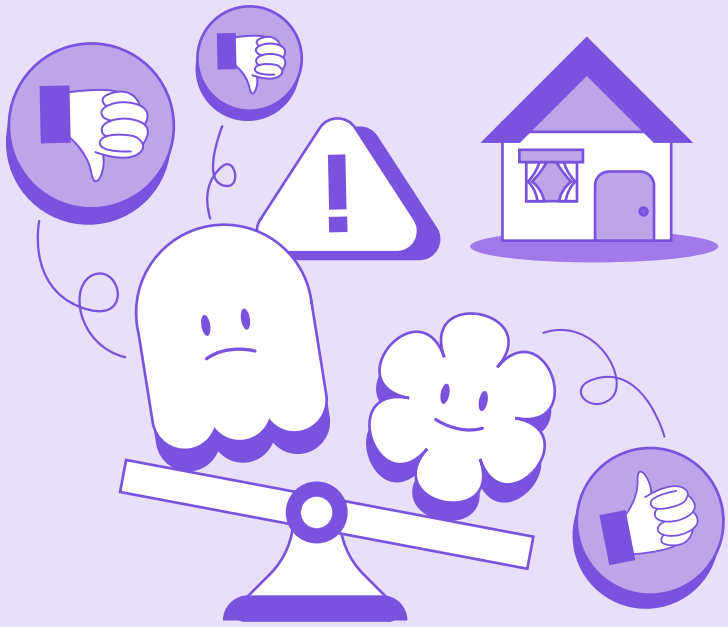
Students are concerned that take-home work is no longer 'fair'.

Since AI tools became common, 40% of students say take-home work feels less fair, compared with 20% who say it feels fairer. Their comments connect fairness to whether assessed work reflects the student's effort and understanding, particularly when some complete tasks independently and others use AI to produce answers.

> The effort required to **cheat in homework and take-home assessment** has been dramatically reduced at the same time that detecting it has become **significantly more difficult.** It is a demotivating factor for students who do not wish to cheat.
>
> — Teacher

Students reporting take-home work feels less fair

outnumber those saying it feels fairer by **two to one.**



> Assessment needs to change, but "we can't make **every assessment task** an exam." — Teacher

Research insight

What should the mark measure?

The Standards for Educational and Psychological Testing define validity in terms of the evidence supporting the interpretation of results for their intended use. Applied to AI-assisted assessment, the issue is whether a mark supports a conclusion about the student's intended knowledge or skill. A polished product can be insufficient evidence if the capability being assessed has not been demonstrated by the learner. [14]

Validity concerns the interpretation a result supports; reliability concerns consistency across tasks, occasions or markers. An assessment can also omit important capabilities, or be influenced by factors outside the intended skill. These distinctions help explain why securing a task and measuring learning are related, but separate, design requirements. [14]

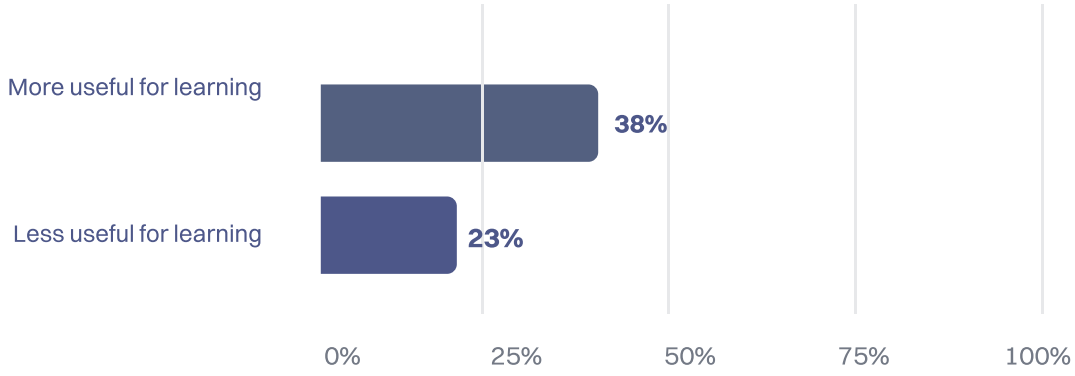
Learning value and **assessment value** can diverge.

Students express a dualistic view of the challenge and opportunity of AI. While more than one in three students (38%) say take-home work has become more useful for learning since AI became common, compared with 23% who say it has become less useful, 73% of those who say it is more useful, still agree that AI makes genuine capability harder to judge. These are the same students recognising useful learning support and uncertainty about what a completed task proves.

> Used in the right way it's **very helpful for preparation,** but not in any assessment task and not really doing the work for me. It's often just building upon and refining what I've made.
>
> — Year 12 student

> Don't get it to give you the answer, get it **to guide you to an answer**, get it to explain why an answer is better than another one.
>
> — Year 12 student

Students reporting take-home work as more or less useful for learning since AI became common (%)



Teachers described students approaching coding and extended writing with more confidence, and gaining access to help when a teacher was unavailable. Students also described using AI to generate practice questions and build subject knowledge. These accounts identify specific learning uses: explanation, guided practice and assistance that helps a student continue.

> The students at the lower end of our classes, they can now access coding. They're not scared of coding. Even if it's an essay, they're not scared of the question anymore.
>
> — Teacher

> Previously, I was essentially the bottleneck for that skill. So if they had trouble with something, then I would build a partial solution for them to get them over that hurdle so they could continue with the rest of their project.
>
> — Teacher

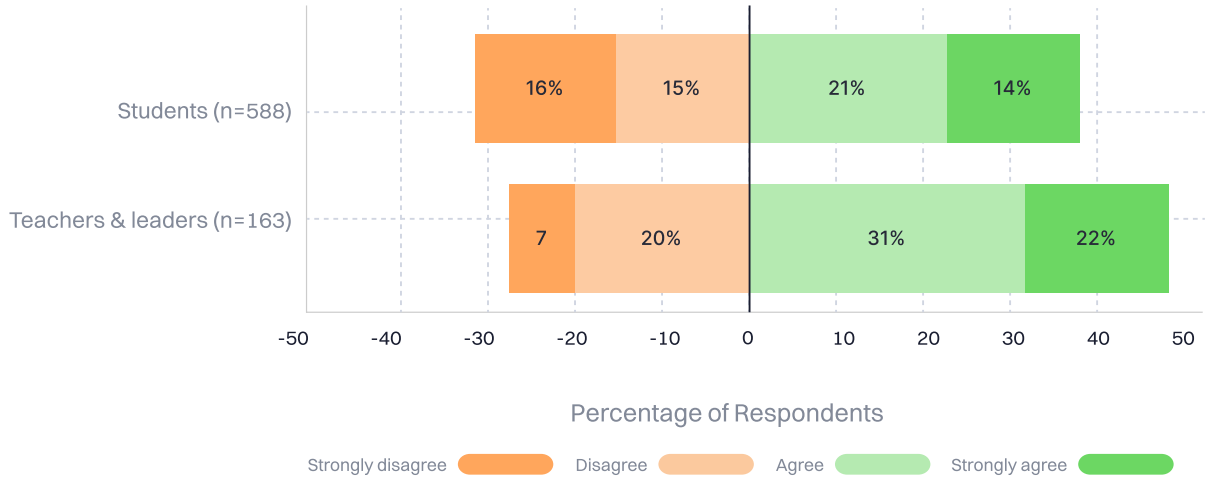
> We were actually told to use AI, because we don't have enough contact with our teacher to be able to ask them about every single problem.
> — Year 12 student

> “At the same time, students have gained new ways to practise for upcoming assessments. Many students are starting to use AI tools to generate practice questions in order to build their fundamental skills in maths and practise content recall in humanities-based subjects.”
> — Year 12 student

Teachers are divided about how to respond to these challenges.

Some 91% of teachers and leaders agree that schools need to redesign some assessments rather than rely mainly on detection or enforcement. Agreement outnumbers disagreement by 38 to one. Even among teachers who oppose a temporary take-home assessment ban, 42 of 45 support redesign. Restrictions divide opinion, but the need to reconsider assessment commands much wider agreement. Teachers’ comments focus on tasks that anticipate AI access and provide better evidence of what students can demonstrate themselves.

A temporary ban on take-home assessment while schools adapt: to what extent do you agree? Students and teachers.

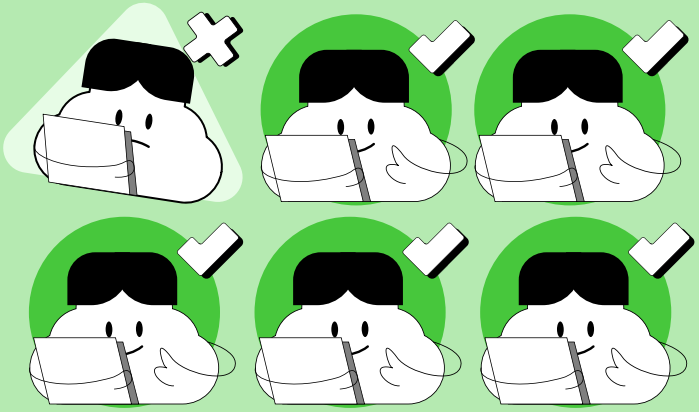


> Redesign all assessments under the assumption that students **WILL use AI if available**. Whether that be moving to in-class tasks, in-class project-based learning, in-class performance, or merely redesigning tasks so that **AI will not necessarily assist.**
>
> — Teacher

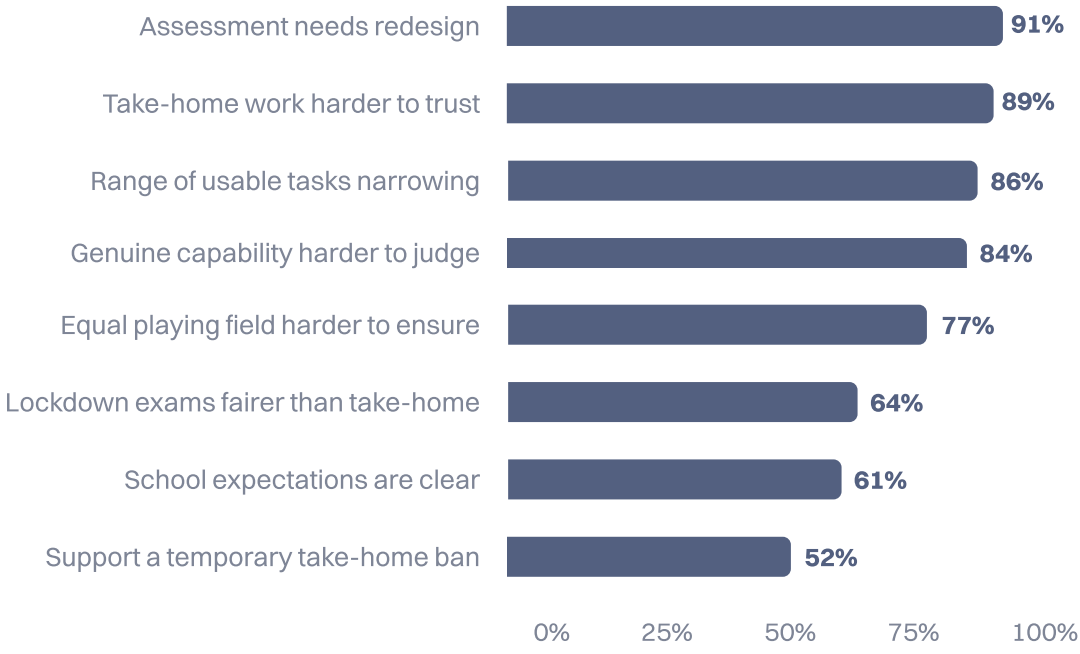
38:1

For every teacher who disagrees, thirty-eight agree

that assessment needs redesign



Staff agreeing with statements about AI and assessment (%)

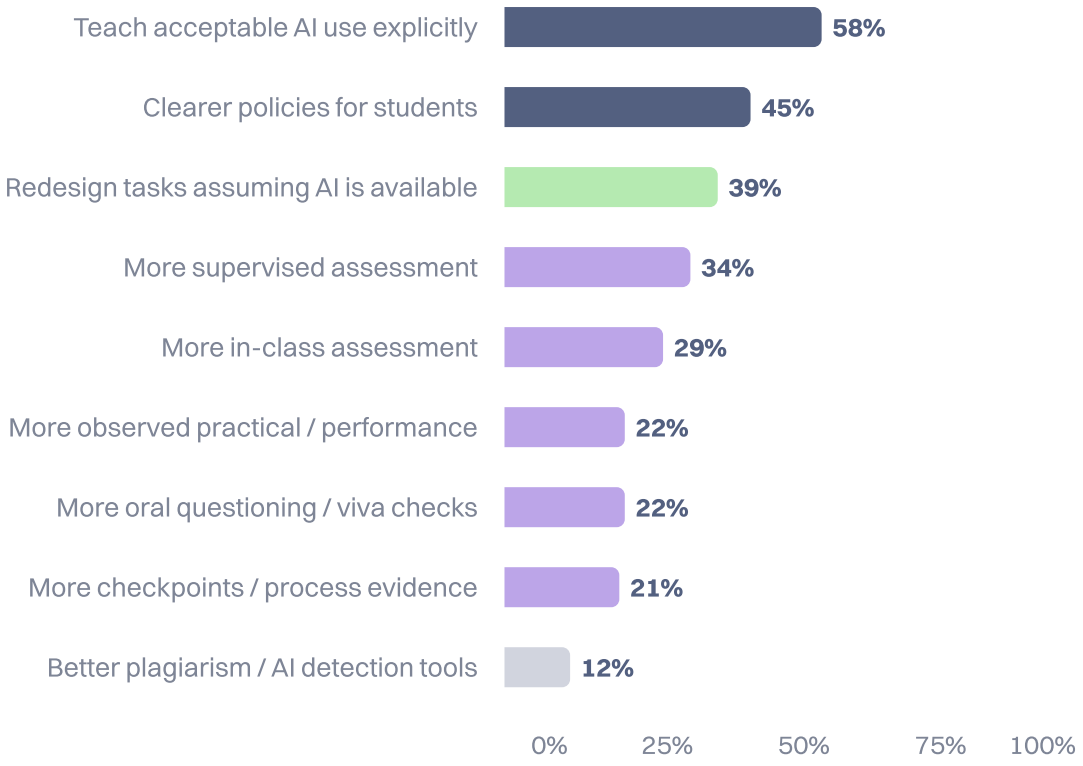


Teaching acceptable AI use leads teachers' priorities; better detection ranks last.

Asked what schools should prioritise over the next two to three years, teachers most often chose explicit teaching of acceptable AI use (58%), followed by clearer policies (45%). Better plagiarism or AI detection tools ranked last among the nine options, at 12%. The pattern gives practical substance to the support for redesign: teachers are looking for changes to expectations, teaching and the assessment task itself. Detection is a much less prominent priority in their responses.

> Schools should place **greater emphasis on supervised, process-based assessment**, where students provide evidence of how they developed their work, rather than **relying only** on the **final submitted product.**
>
> — Teacher

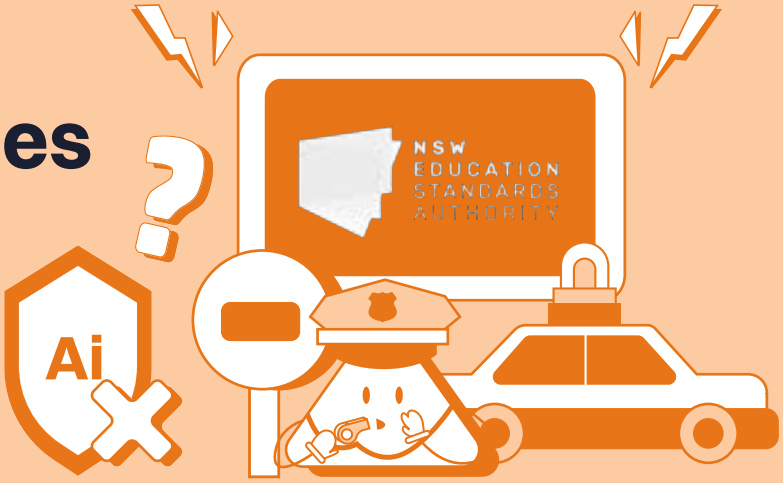
School responses to AI nominated as priorities by teachers and leaders (%)



Policy Context

NESA advises schools

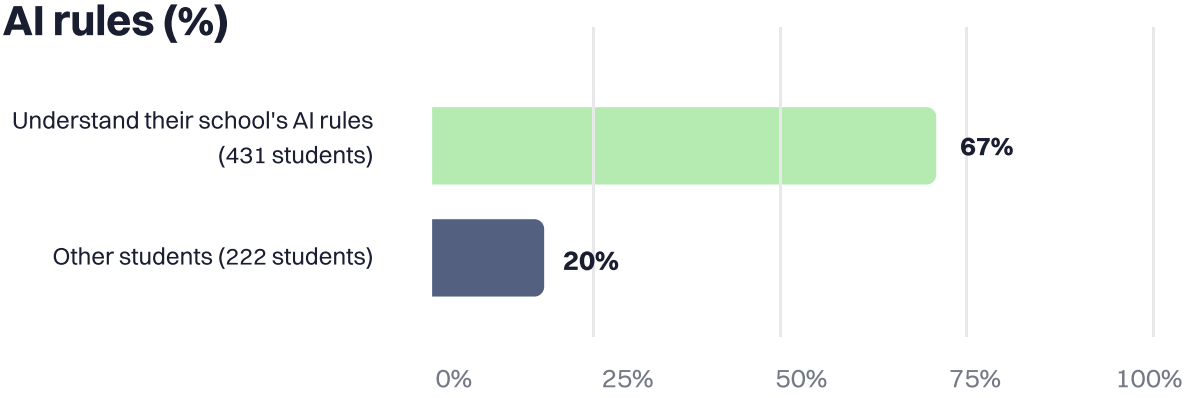
not to rely on AI detection tools as their main safeguard against malpractice, or treat detection results as sole proof of misuse. [2]



Students who understand the rules have greater confidence in assessment.

Among students who understand their school's AI rules, 67% agree that assessments give them a fair opportunity to show their learning, compared with 20% of other students. The difference is substantial. Students' understanding of expectations is closely associated with confidence that their work will be assessed fairly, giving schools a practical reason to attend to how policies are explained and experienced. The survey cannot isolate the effect of rules from other features of those students' school experience.

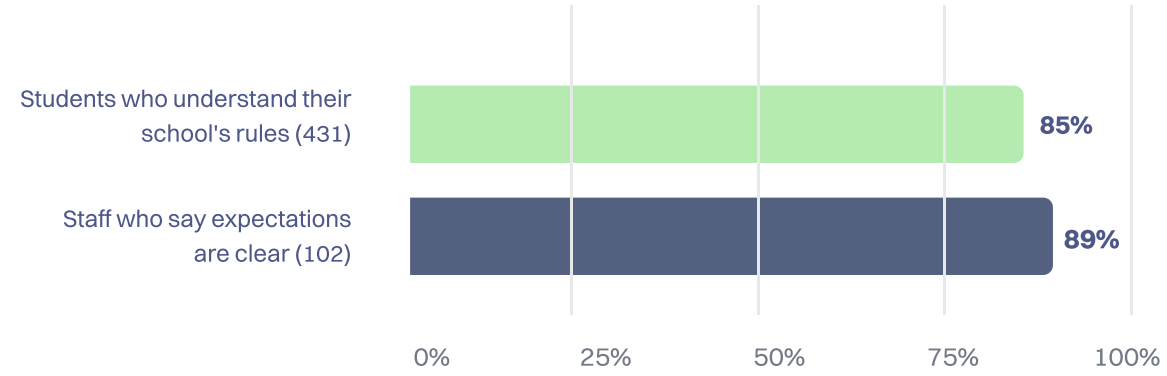
Students agreeing assessment gives them a fair opportunity to show learning, by understanding of school AI rules (%)



Clear rules do not remove the difficulty of judging independent capability.

Even among students who understand their school's AI rules, 85% say AI makes genuine capability harder to judge from take-home work. The corresponding figure is 89% among staff who describe their school's expectations as clear and consistently understood. Knowing what is permitted does not necessarily make a finished response more informative about the learner. This finding helps explain why teachers want clear policies to sit alongside changes to the way assessment is designed.

Students and staff who report clear AI rules but still say genuine capability is harder to judge (%)



"It definitely narrows the field, for the **simple reason** that you don't get those **rich assessment tasks."**

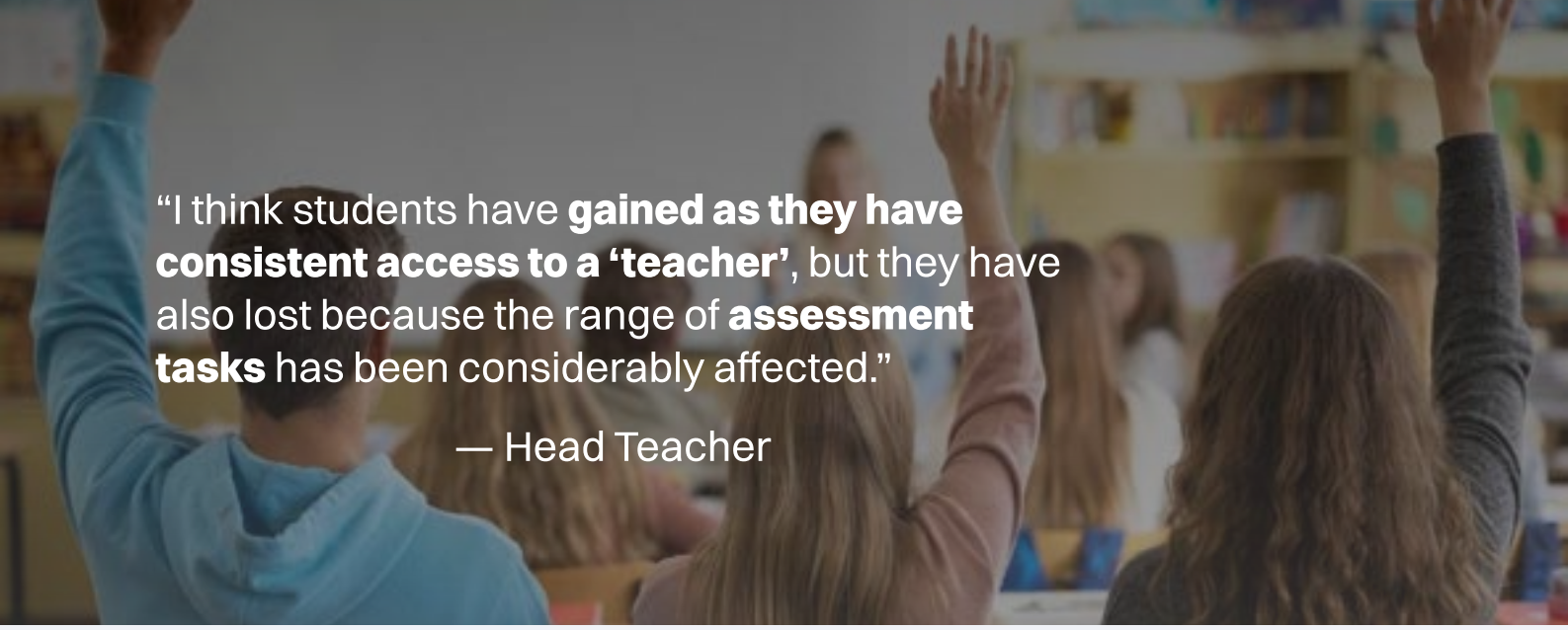
— Teacher

"The **limited time available** in the HSC to teach the required content leaves very little time for students to complete **in-class assessment tasks**, which may result in a significant increase in the reliance on summative tests."

— Teacher

"There's got to be a realistic understanding that **AI is here and it's going to influence their work**, and trying to police it, I think, is a losing battle in the end."

— Teacher



"I think students have **gained as they have consistent access to a 'teacher'**, but they have also lost because the range of **assessment tasks** has been considerably affected."

— Head Teacher

"How can we possibly teach those skills and assess those skills **without it being a take-home task**, without giving students the time to actually deliver a project?"

— Year 12 student

"I **asked students** to explain sections of [their assessment], and it was very clear **who actually did the work** and **who didn't.**"

— Teacher

"If you want a big surprise and just say, 'I'll see you guys in seven weeks,' you're going to get an **AI surprise**. [In-class check-ins are] more work, but I think it's more authentic, because I can see the **product developing.**"

— Teacher

"Giving each student an hour, that means I'm doing **27 hours of [assessment during class hours]**."

— Teacher

Implications for schools & assessment authorities

Support assessment redesign beyond restrictions on take-home work.

The strong agreement on redesign, including among teachers who oppose a temporary ban, supports practical guidance on assessment design. Subject-specific examples could show how to preserve substantial projects and practical work while incorporating evidence of individual understanding. Teachers' concerns about narrowing the task range make that breadth important.

Make assessment literacy a professional-learning priority.

Teachers distinguish between formats they consider fair and those that best reveal knowledge and skills. That finding, alongside difficulty attributing take-home work to students, makes assessment literacy an urgent area for support. Professional learning should help teachers evaluate what a result can legitimately tell them about the learner, drawing on validity, reliability and the intended syllabus outcomes.[14]

Make expectations understandable to students and usable in tasks.

Students who understand school AI rules report greater confidence in assessment fairness, while explicit teaching of acceptable use is teachers' leading priority. Shared examples of permitted and prohibited assistance would help make policies concrete. Rules should be accompanied by assessment conditions that allow students to demonstrate the knowledge and skills being marked.

Keep verification proportionate to teaching time.

Teachers describe the workload of individual questioning and supervised checkpoints. Guidance should support proportionate checks, common criteria and moderation so that stronger assurance remains workable across whole cohorts. The comments also support preserving opportunities for extended learning rather than replacing every task with an examination.

Expand digital assessment where the format serves the subject.

Student preferences for typing and improved answer-tool ratings support continued, subject-specific expansion. Teachers' accounts of coding and other practical tools show where online formats add value. Simpler onboarding, familiarisation and accessible answer tools remain important to implementation across different school environments.

Conclusion

A stronger examination experience and a clearer question about learning.

The 2026 evaluation shows the value of building on schools' experience. Students rated the examination workflow and every answer type more positively than last year; teachers recognised improvements in preparation, delivery and marking. Requests from 2025 have become working features, including session reporting and AI feedback for complex answer types. The next refinements focus on simplifying secure access and improving the everyday workflow across large cohorts.

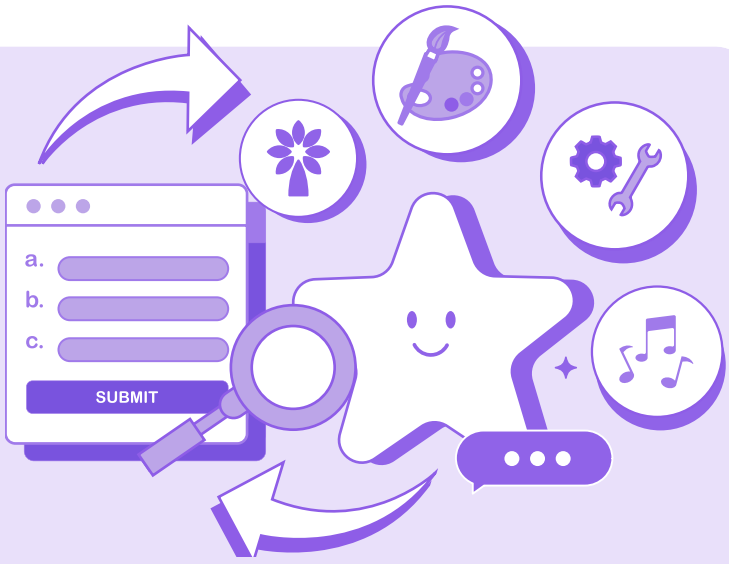
The wider findings make a case for assessment redesign that reaches beyond the debate about academic integrity. Teachers and students both question what take-home work now proves. Their responses challenge the assumption that a completed assignment can, by itself, provide dependable evidence of the student's capability.

Supervised assessment attracts greater confidence, while practical work and extended projects remain valuable when individual understanding can be established. Their strongest agreement is that assessment needs to change, with clearer evidence of individual understanding at its centre.

For schools and assessment authorities, the opportunity is to connect those aims. Secure digital examinations can provide supervised evidence while allowing students to use the tools their subjects require. Practical tasks, projects and well-designed checkpoints can reveal capabilities that a single examination may not capture. The test of a stronger assessment program is how confidently its results can be connected to what each student has learnt.

The finished assignment is no longer enough.

Assessment must establish what the **student knows.**



Acknowledgements

This report was made possible by the students, teachers, school leaders, examination coordinators and IT staff who shared their experiences and perspectives. Their survey responses, discussion contributions and practical feedback informed the evaluation and its improvement priorities.

We extend our sincere thanks to Catholic Schools NSW (CSNSW), CSSA, Marita Winters (Deputy CEO, CSNSW), Monica O'Brien (Head of Curriculum and Assessment, CSNSW), Dr Andrew Eaton (Senior Manager, CSSA Exams), Jill McGettigan, Klaudia Prasnicki, Benjamin Lum, Mora Soliman, Felicity Cassidy, Phil Oke and respective CSSA Exam Committee Members, Julia Klaryz (Designer), and to all of the teachers, students, and schools who participated.

We also acknowledge St Joseph's International Catholic College in Papua New Guinea for the examination photographs included in the report, and extend our thanks to teachers who shared insights and feedback in focus groups, including Matthew Topp (*St Marys Senior High School*), Ghamdan El-Miski (Amity College), Rebecca Dam (Sydney Boys High School), Kelvin Kaing (Burwood Girls High School), Nicholas Zupan (Lismore Online Education Centre), Tim Bailey (Redeemer Baptist School), Sarah Tor (Macarthur Anglican School), Jenny Lill (Henry Kendall High School), Raymond Montalban (Arthur Phillip High School), Rachael Jones (Inaburra School), Daniel Comben (Sydney Boys High School), Sherobhi Rajamantri (Barker College), Cody McDonald (St John's College (Dubbo)), Heath Phillips-Johnson (Bourke High School), Sharon Hibbert (All Saints' College Maitland), Ben Jones (Tempe High School) and Darren Chantler (Marist College Kogarah).

References

[1] Stanford Institute for Human-Centered Artificial Intelligence (2026). The 2026 AI Index Report.

[2] NSW Education Standards Authority (2026). Use of AI by students for the HSC.

[3] OECD (2026). OECD Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education.

[4] Pope Leo XIV (15 May 2026). Magnifica Humanitas, particularly paragraphs 139–141.

[5] NSW Education Standards Authority (31 August 2026). New limit on take-home assessment tasks from Term 4 2026.

[6] New York City Public Schools (2026). Guidance on Artificial Intelligence and Screen Time.

[7] Merod, A. (3 September 2026). LAUSD restricts all students from using AI tools. K-12 Dive.

[8] Norwegian Directorate for Education and Training (20 April 2026). Measures to prevent AI use in examinations.

[9] Joint Council for Qualifications. Instructions for conducting non-examination assessments: GCE and GCSE specifications.

[10] NSW Education Standards Authority (17 March 2025). Sample online exams for HSC English Extension 1 and 2 courses.

[11] Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö. and Mariman, R. (2025). Generative AI without guardrails can harm learning. Proceedings of the National Academy of Sciences.

[12] Kestin, G., Miller, K., Klales, A., Milbourne, T. and Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. Scientific Reports, 15, 17458.

[13] Bjork, E.L. and Bjork, R.A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning.

[14] American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2014). Standards for Educational and Psychological Testing.

About Gradeo

Gradeo delivers assessment infrastructure to help teachers to author, deliver, grade and report paper and digital formative and summative assessments.

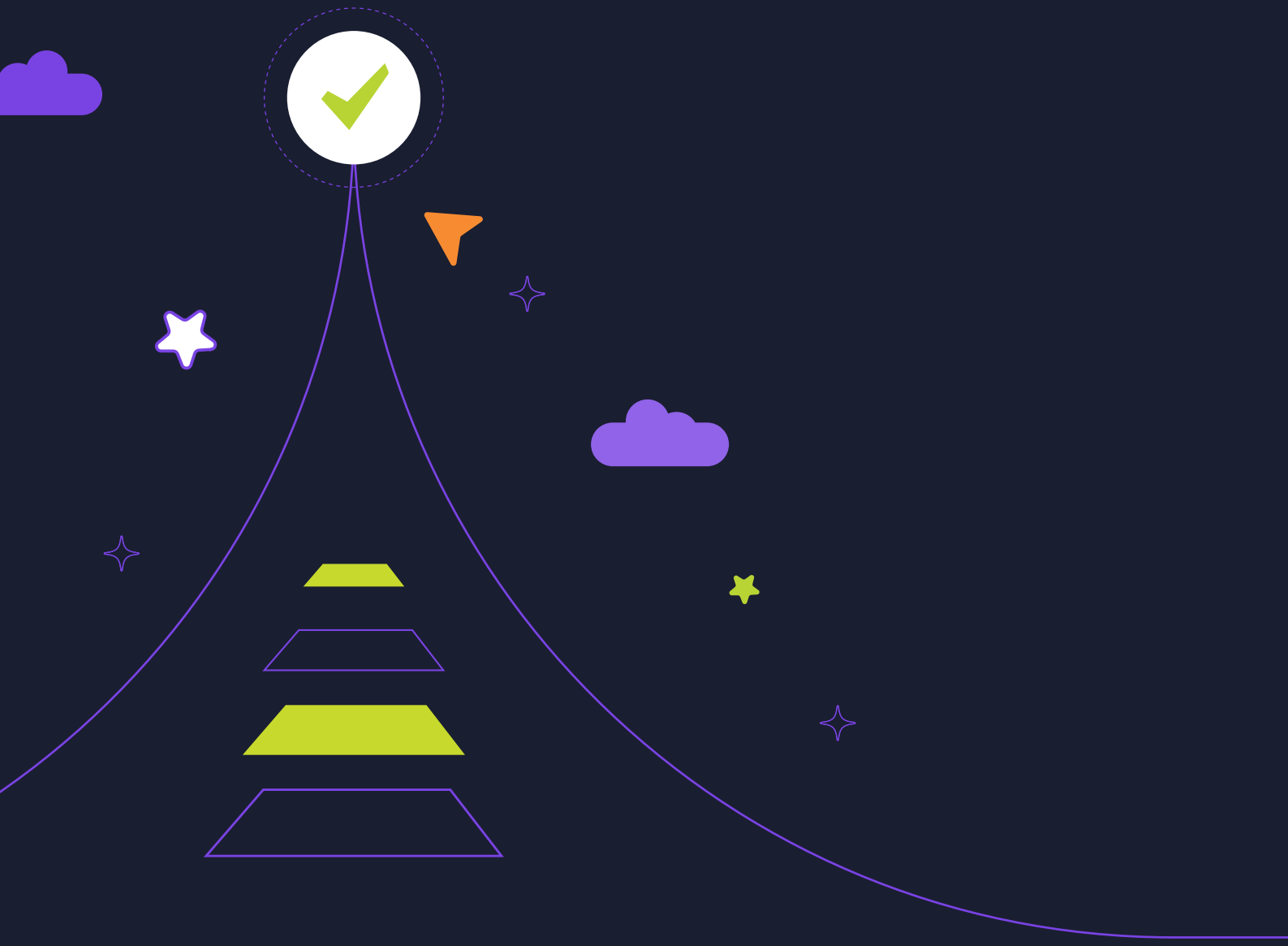
Schools use Gradeo to securely deliver trial exams, class assessments and quizzes in 55+ subjects (with over 1500 exams and 40,000 questions written by exam committees) across a wide-range of answer types — from multiple choice and extended response through to coding (Python/SQL), spreadsheets, drag-and-drop, diagram, drawing and objective response.

On Gradeo, teachers can author their own questions, create syllabus-aligned quizzes and exams, mark responses for an entire class, and generate detailed insights into student progress. Students gain access to past papers, can generate their own quizzes, and access valuable reporting tools that support their learning.

Since launching in March 2025, over 450 schools in NSW use Gradeo for online assessment. The platform operates under two models: subscriptions, which give schools year-round access to past papers, courses and question banks; and one-off exam licences for in-year CSSA and NSW Independent Trial Exams, which allow high-stakes delivery of restricted exams during a designated exam window.

Citation

Gradeo (2026), *Making the Shift: How online exams and AI are changing assessment*, 2026, Sydney



gradeo.com.au

Copyright © 2026 Gradeo Pty Ltd

This document is for informational purposes only and does not constitute professional advice. Any references to products, services, or prices have been prepared using publicly available information and best endeavours. While Gradeo Pty Ltd strives for accuracy, no representation or warranty is made as to the completeness or correctness of the information. Readers should seek their own professional advice before making decisions based on the content herein. Gradeo Pty Ltd assumes no liability for actions taken based on this document.