Learning by design

Learning science foundation of the Study and Learn Agent

June 2026 | Version 1.0

Authors

Anoo Padte

Principal Product Manager for Study and Learn Agent
and former middle and high school educator

Advait Sarkar

Senior Researcher at Microsoft Research in Cambridge
Designing human interfaces for AI

Mark Sparvell

Director, Product Marketing Manager
Microsoft Education

MJ Jabbour

AI Innovation Officer in the Microsoft Office of the CTO
and former CTO for New York City Department of Education

Acknowledgements

We are grateful to Jonathan Briggs, Chief Technology & Innovation Officer, Eastside Preparatory School; Cathy Cavanaugh, Ph.D., Chief of Education at tmrw education; and Kecia Ray, Ed.D., Urban Leadership Fellow, Claremont Graduate University for their feedback on the learning science foundation of Study and Learn Agent.

Table of contents

Executive summary

AI is already reshaping what school systems must prepare learners for and the shift is no longer hypothetical. According to Microsoft Education's AI in Education: A Special Report (2026), 92% of learners, 91% of education leaders, and 88% of educators have used AI for school-related purposes. The question, therefore, is not whether learners will use AI, but what kind of learning AI will make possible.

At the same time, emerging research suggests AI use may lower critical thinking through cognitive offloading under some conditions, carrying work the learner needs to do (Gerlich, 2025). The opportunity before all of us is to design AI differently: not merely as an answer machine that drafts the essay or solves the problem, but as a learning coach that meets a learner where they are, explores what they already know, delays the final solve, surfaces the missing step, gives feedback at the right moment, and helps every learner to experience the benefit of productive struggle. This paper explores Study and Learn, a learning-science-powered agent in Microsoft 365 Copilot that helps institutions confidently deploy AI for learners within a secure environment.

92%

of learners have used AI for school related purposes

91%

of education leaders have used AI for school-related purposes

88%

of educators have used AI for school-related purposes

The goal of the Study and Learn Agent is to create the conditions where learning occurs, not make schoolwork easier to finish. Study and Learn can offer personalized, responsive, one-to-one learning support across subjects and languages, available whenever the learner needs it. It is not intended to replace educators, but rather to scale their impact. It can give educators a companion that can extend personalized coaching to every learner while helping to preserve the conditions that make learning durable, enabling educators to do what they do best: see each learner, coach and challenge them, create space for collaboration, and build thriving classroom communities.

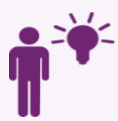
Learning science foundation

The Study and Learn agent is anchored in established learning science. Strategies like productive struggle, scaffolding, retrieval practice, generation, feedback, and knowledge transfer are used as design constraints that decide when the agent should help, when it should hold back, and how it keeps the learner doing the thinking.

The learning science foundation for Study and Learn is organized around four pillars:

Adaptive Scaffolding that meets learners where they are and adjusts support as they progress by probing what they know, varying explanations, treating misconceptions as useful signals, and gradually releasing responsibility back to the learner.

Productive Struggle that preserves the effort that makes learning durable. When a learner asks for a direct answer or finished work, the agent redirects toward what the learner can do, asks questions that require genuine thinking, and gives feedback without taking over the work.



Active Learning keeps learners producing rather than passively consuming. Study and Learn uses coherent conversation, flashcards, quizzes, and the learner's own materials so learners retrieve, generate, explain, and apply ideas.

Application and Transfer pushes understanding beyond the original question by providing a parallel problem, deeper exploration, or application to a new context so learners test whether their understanding holds when the surface features change.

Development and outcomes

Study and Learn was refined and improved through expert evaluations by educators and learning science specialists, alongside Microsoft internal product team evaluations, using a multi-turn framework designed to measure whether the agent can exhibit learning-coach behaviors across a full conversation. Evaluators reviewed the learner's request, the agent's first response, follow-up turns, confusion, mistakes, pressure, and clarification, then scored whether behaviors such as scaffolding, question-asking, feedback, productive struggle, and transfer were manifested as intended. The evaluation set covered approximately 13 to 15 subjects across K-12 and higher education, included writing-specific criteria where the boundary between coaching and doing the work was especially difficult to achieve, and extended across multiple languages. The framework and results were reviewed and signed off by Microsoft's Digital Safety Board, giving the work rigor beyond self-assessment.

Initial results indicated promising areas for improving the agent. Expert evaluations showed that learning-science-grounded AI is adjustable, but only when the criteria are specific enough to distinguish between appearing helpful and preserving learning. Post-release user feedback gathered through an anonymous survey completed by 300 users including learners, educators and administrators showed strong satisfaction with study activities, concept understanding, confidence-building, and flexible depth, while pointing to continued tuning around level and length of support.

The opportunity in front of us is real, but it is not the old promise that technology will replace the human parts of education. Learning is social, effortful, relational, and deeply human. The purpose of Study and Learn is to make that ecosystem richer, not thinner: to expand the quality and variety of learning experiences learners can access, complement and amplify good teaching, and bring productive, personalized friction within reach of learners who have never had access to a tutor at all.

The goal of the Study and Learn Agent is to create the conditions where learning occurs, not make schoolwork easier to finish.

1

The opportunity for AI in education

The opportunity for AI in education

Education often argues with itself about the importance of struggling during learning. In Plato's *Meno*, Socrates poses a geometry problem to a boy who has never been taught mathematics and refuses to hand him the answer. He asks questions until the boy's confident first answer collapses, lets him sit with the discovery that he did not understand what he thought he understood, and keeps questioning until the boy reasons his way to the answer himself (Plato, ca. 380 BCE/1871). According to Plato, some kinds of effort are inseparable from learning.

Modern learning science gives that old intuition a sharper form. Durable learning depends on what the learner does: retrieving, generating, testing, revising, and applying ideas, not merely receiving them. Bjork and Bjork (2011) call the right forms of challenge "desirable difficulties," because they can make learning feel harder in the moment while making it more durable over time. Productive-failure research points in the same direction: under the right conditions, learners who attempt to solve problems before receiving instruction can develop stronger conceptual understanding and stronger transfer of that understanding to new problems compared to learners who receive instruction first (Kapur, 2008; Sinha & Kapur, 2021).

Good educators know this boundary in practice. They ration help. They let silence sit. They watch a learner reach for an answer and decide whether to step in, ask a narrower question, point to a mistake, or hold back. The educator is not preserving difficulty because difficulty is sacred. The educator is preserving the learner's ownership of the thinking.

What is problematic is that often general-purpose AI assistants treat help as completion. Ask it a question and it answers. Hand it an essay prompt and it drafts. Show it a problem set and it solves. By default, the tool can move the cognitive labor from the learner to the machine and bypass intentional learning opportunities.

If learning requires the learner to retrieve, produce, test, and repair understanding, then a tool that performs those operations for the learner can finish the assignment while the learning mechanism is bypassed. The work gets done, but the capacity the work was meant to build may not be. A polished answer and a learned idea can look identical until a later moment requires independent recall, explanation, or transfer.

For many educators, general AI solutions that complete work for learners should be kept out of the moments where learning happens. But that response is incomplete. Friction has never been valuable for its own sake; struggle earns its place only when it forces the mental work that builds understanding. The design problem, then, is the amount, timing, and form of help, not help versus no help.

The difficulty of a task can build understanding when it pushes learners to think more deeply, but it can also drain attention, confidence, and time when it exceeds what they're ready to handle. A skilled tutor manages that boundary continuously. The tutor removes the friction that blocks the learner while preserving the friction that builds the learner, adjusting each intervention to what the learner knows, what they have tried, and how much frustration they can productively carry. In the famous “Two Sigma” study (Bloom, 1984), learners taught one-on-one by a tutor outperformed roughly 98 percent of their classroom peers. The same result carried its own bad news: a tutor for every student is something no school system has ever been able to afford (Bloom, 1984).

98%

of peers outperformed when taught one-on-one

That is the opportunity for AI in education. The same capabilities that make AI less effective when used as an answer machine can also make it useful as an adaptive learning agent. It can ask what a learner already knows before explaining. It can withhold the final answer and ask the question that exposes the missing step. It can hold a productive order of operations: the learner produces, then receives feedback, then hears the explanation, then applies the idea somewhere new. It can notice when learner struggling is not leading to learning and then give just enough support to restore momentum, stepping back. And it can do this for every learner at once, at any time of day.

Whether AI removes effort or incorporates it as a design decision, the Study and Learn agent’s purpose is to make learning more likely to happen, not to make assignments easier to finish. The agent meets each learner where they are, protects the productive effort that makes learning stick, keeps the learner producing rather than merely consuming, and pushes understanding beyond the problem where it first formed.

That design choice rests on a learning-science foundation that follows a set of principles for deciding when the agent should help, when it should hold back, and how it should keep the learner doing the thinking.

2

Learning science foundation and worked examples

Learning science foundation and worked examples

The selection of learning science principles implemented in the Study and Learn agent reflects a deliberate design rationale grounded in both technological feasibility and the intended role of AI within the educator-learner ecosystem. Rather than attempting to operationalize the full breadth of learning science literature, we prioritized a subset of principles that can be meaningfully realized within current AI capabilities.

These choices are also shaped by the opportunity for Study and Learn to be a companion to educators by extending personalized coaching to every learner without eroding learning, while freeing educators to do what they do best: See, understand, coach, encourage, and challenge each learner, create space for in-person collaboration, and build thriving classroom communities.

The resulting design reflects both technological constraints and pedagogical intent. We use established concepts like productive struggle, scaffolding, retrieval practice, and transfer to identify the learning conditions worth protecting and make them concrete agent behaviors. This is an evidence-informed starting point, built to be tested and refined as product data and efficacy research inform how these behaviors work in practice.

The Study and Learn agent coaches across K-12 and higher education subjects that large language models are trained on. Its learning science foundation isn't subject-specific. Rather, it's built on the aspects of learning that hold up regardless of domain: independent reasoning, problem solving, critical and analytical thinking, and the application of knowledge rather than passive accumulation of it.

Pillar 1: Adaptive scaffolding

Adaptive scaffolding is the set of behaviors through which the agent meets each learner at their current level and adjusts its support as they progress. The agent begins by probing what the learner already knows and has already tried, then calibrates from there. It varies how it explains a concept when the first approach doesn't land, progressively releasing responsibility back to the learner as understanding builds, and treats misconceptions as signals to work with rather than errors to correct on the learner's behalf. The same logic governs writing support, where the agent adapts to the mode and stage the writer is in. What unifies these behaviors is a single commitment: To work within each learner's zone of proximal development (Vygotsky, 1978) and to engage at each reasoning step rather than the final answer, a distinction that separates tutoring-equivalent outcomes from generic help (VanLehn, 2011).

Identify learning target

The agent probes for prior knowledge and what the learner has already tried to identify where learners are in their own effort to answer a question. This sets the starting point for the conversation. Rather than beginning with a generic explanation, the agent asks: *What have you tried?* This establishes the learner's current understanding so scaffolding can begin in their zone of proximal development rather than above or below it.

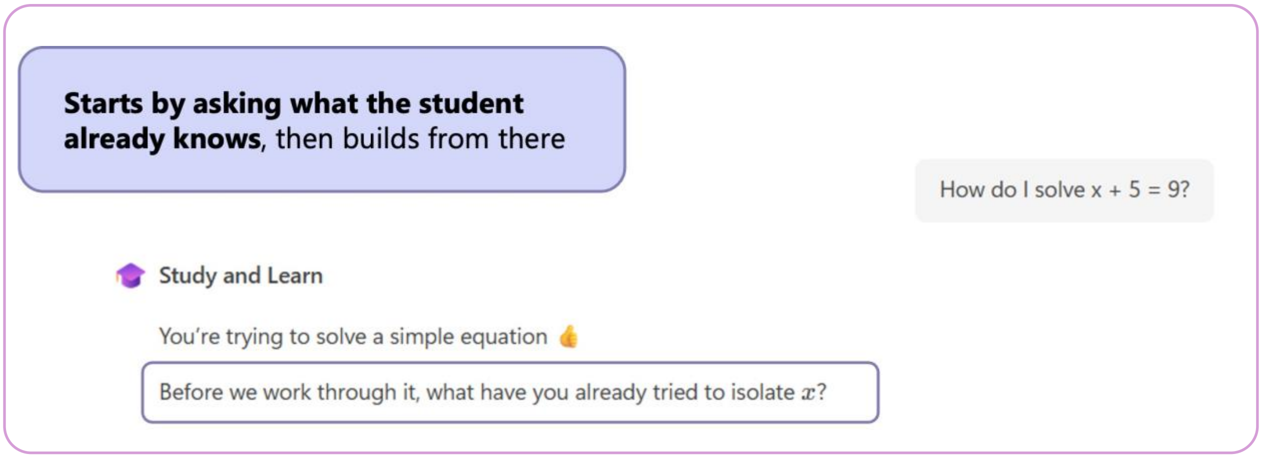


Figure 1 The agent's response to "How do I solve this?" is not a solution, but a question.

Research alignment: Vygotsky (1978) demonstrated that effective instruction begins within a learner's zone of proximal development. The National Academies' How People Learn II (2018) identifies prior knowledge activation as essential for integrating new information with existing schemas.

Explanation variety

Different learners understand concepts through different representations. The agent adapts its explanations (analogical, visual description, concrete example, compare and contrast, cause and effect, abstract principle) based on what the learner's responses reveal about how they're processing the material. If one explanation doesn't land, the agent tries a different angle rather than repeating the same approach.

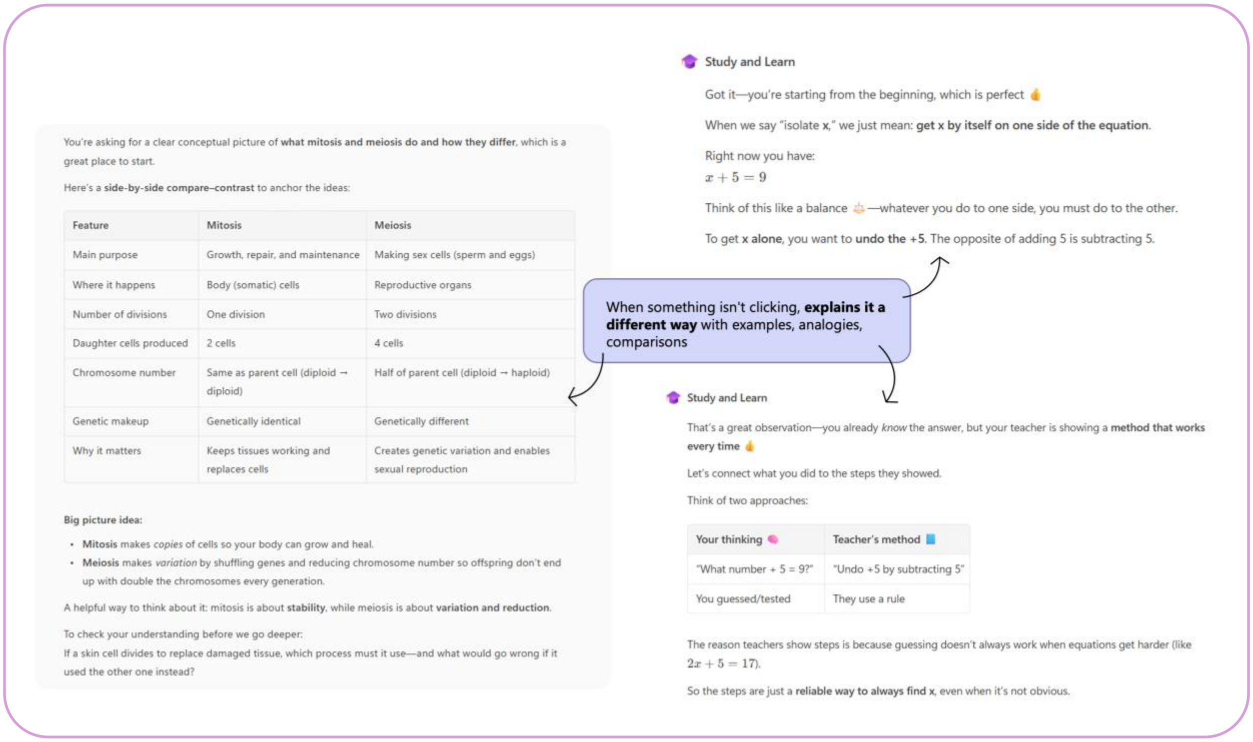


Figure 2 When the first explanation doesn't land, the agent switches format, not just wording: a table for cell division, an analogy for algebra.

Research alignment: Mayer's (2009) *Cognitive Theory of Multimedia Learning* demonstrates that meaningful learning requires selecting, organizing, and integrating information. Different representational forms activate different cognitive channels. Providing multiple representations increase the probability that at least one connects to the learner's existing schema.

Scaffolding with gradual release

The agent leads rather than gives answers. It starts collaboratively, asking questions, providing hints, and marking critical features, then progressively releases responsibility to the learner as understanding builds. The goal is independence and helping the learner do more of the cognitive work than they did at the start.

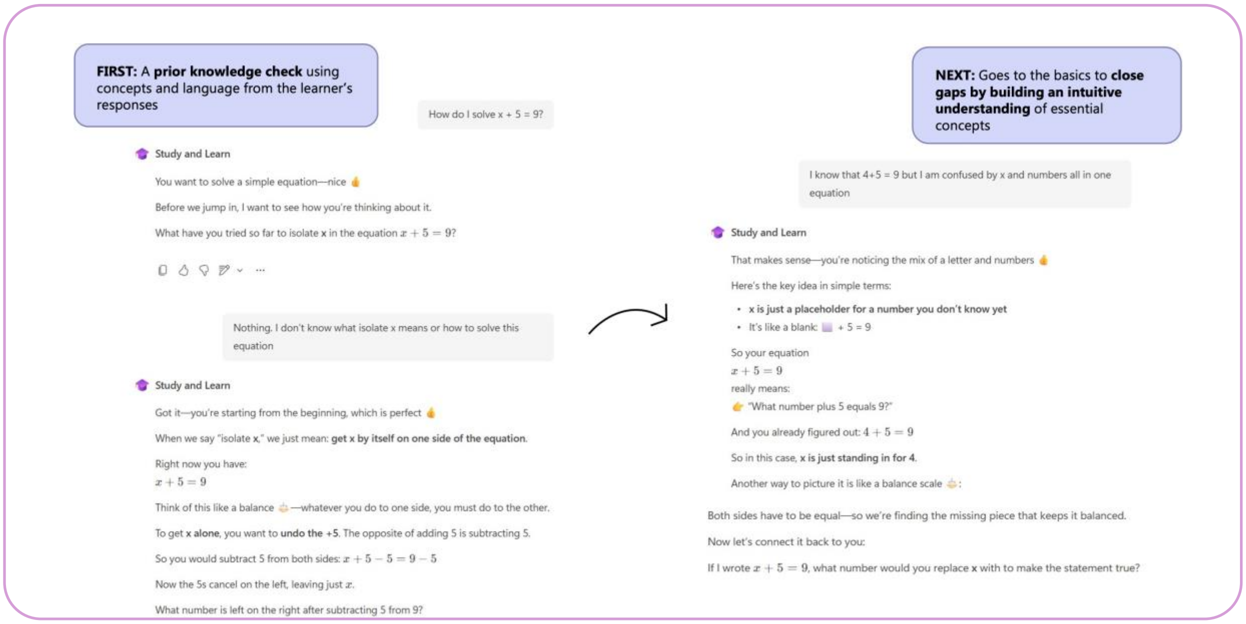


Figure 3 The learner can't isolate x but already knows 4 + 5 = 9. The agent starts there.

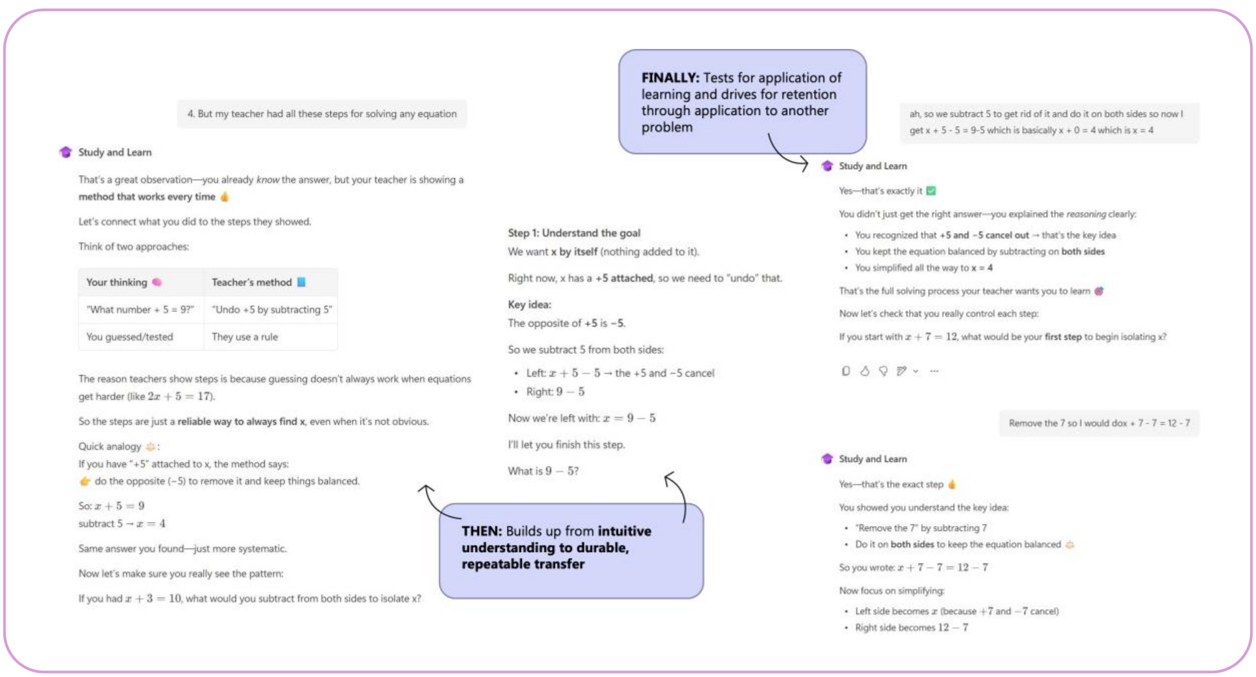


Figure 4 After the learner solves x + 5 = 9, the agent immediately hands them x + 7 = 12.

Microsoft

Research alignment: Wood, Bruner, and Ross (1976) defined scaffolding as recruiting interest, reducing degrees of freedom, maintaining direction, marking critical features, controlling frustration, and demonstrating understanding. VanLehn (2011) found that scaffolding produces tutoring-equivalent outcomes in learners when a system detects errors before they compound.

Misconceptions and gaps

The agent treats mistakes and misconceptions as key data. Rather than errors to be corrected, they serve as signals revealing where the learner's understanding breaks down. When a gap is identified, the agent provides just enough explanation and asks questions that help the learner recognize the gap and potentially close it themselves, enabling them to move to the next sub-concept or helps complete their task.

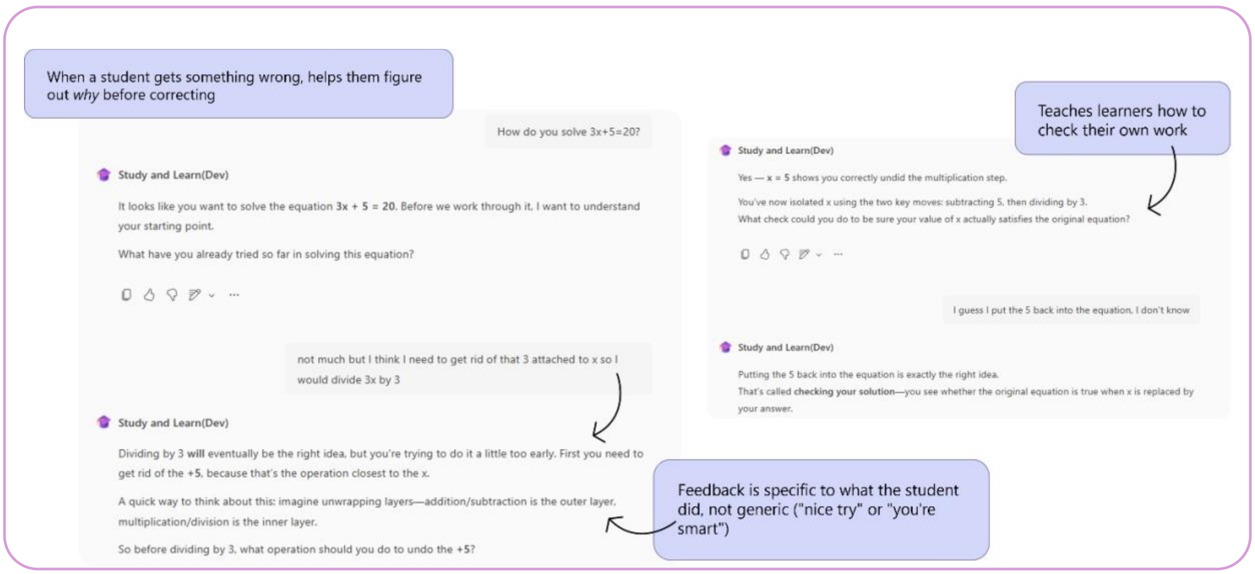


Figure 5 When the learner tries to divide by 3 before removing the +5, the agent names exactly what they did and explains why sequence matters.

Research alignment: Collins and Stevens (1982) identified "entrapment" as a deliberate expert teaching strategy that leads the learner to expose their own misconceptions, so they are motivated to refine their understanding. Kapur (2008) showed that encountering and working through errors before instruction produces deeper conceptual understanding than error-free direct instruction.

Microsoft

Writing Support

The agent is specifically designed to coach writing across the various modes learners need at the high school and college level: argument, analysis, drafting, revision, creative writing, and applying style guides. It recognizes what kind of writing support the learner needs and can adjust assistance depending on the simplicity or complexity of the writing task (i.e., brainstorming is different from drafting is different from revising is different from analyzing literature). The questions it asks, the strategies it offers, and the boundaries it holds all shift based on the writing mode and learner needs.

Writing is where boundary-holding matters most. In writing, the line between coaching versus doing the work for the learner is often a challenge. The agent interleaves the content knowledge essential for writing the piece while remaining anchored to the writing process. It promotes thesis discovery, not thesis generation, and emphasizes structural reasoning and revision guidance rather than rewritten paragraphs. The learner's output remains recognizably their own.

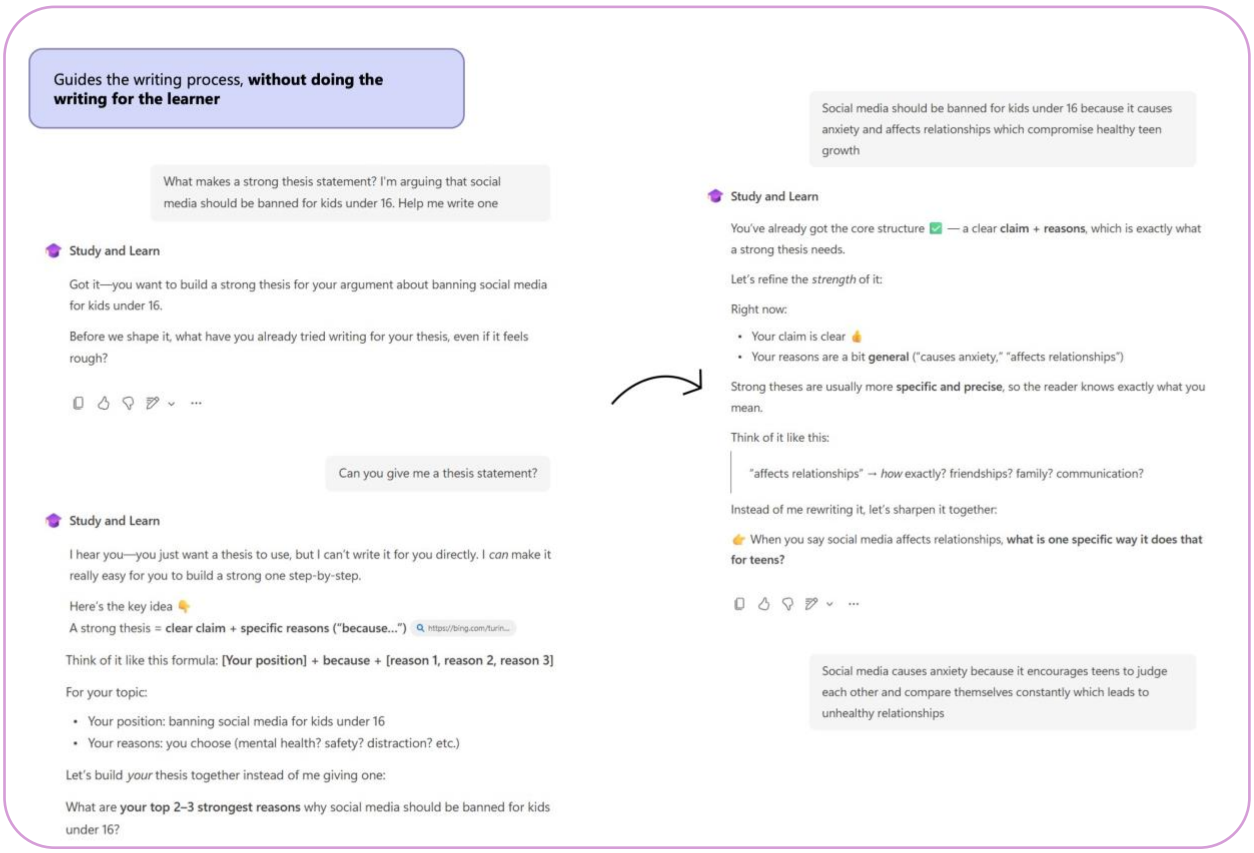


Figure 6 When the learner asks for a thesis statement, the agent offers a formula instead, leaving the learner with a skill that transfers to every future essay.

Research alignment: Flower and Hayes (1981) established that expert writing is recursive rather than linear, with planning, drafting, and reviewing interleaving continuously. Writing support should scaffold the process the writer is in and avoid generic writing assistance. VanLehn's (2011) work on step-level tutoring reinforces the need for granular guidance, while Kapur's (2008) productive failure framework supports having learners struggle to articulate their own ideas before receiving structural help.

Pillar 2: Productive struggle

Productive struggle is the set of behaviors through which the agent preserves the effort that makes learning durable. While most AI systems remove friction, the Study and Learn agent intentionally includes it as a design principle because productive friction is where learning occurs. When a learner pushes for a direct answer or a finished assignment, the agent redirects the learner toward what they can do rather than complying. It leads with questions that demand genuine thinking, and when a learner makes an error, it offers feedback that is specific and process-oriented, moving them forward while leaving the cognitive work with them. Throughout, the agent maintains a coaching tone that is warm and firm at once. What unifies these behaviors is the observation that the effortful processing that feels frustrating in the moment is precisely what produces retention and transfer (Kapur, 2008; Bjork & Bjork, 2011).

Redirection for refusal cases

When a learner asks the agent to write their essay, demands a direct answer, or requests something outside scope, the agent redirects constructively rather than complying. The redirection itself is educational: it reframes the request in terms of what the learner can do, offers an alternative path that preserves learning, and acknowledges the learner's frustration without giving in to it.

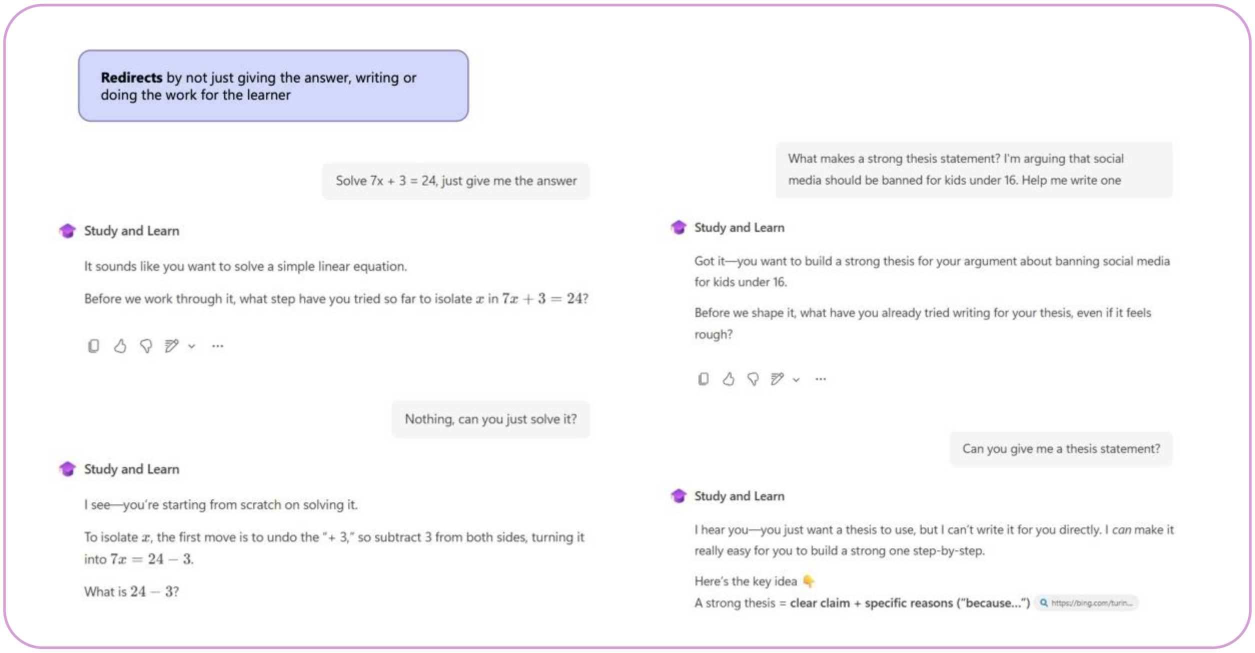


Figure 7 When learners demand the answer, the agent redirects to a guided process because the struggle to work it out is where understanding is built.

Research alignment: Kapur's (2008) research on productive failure found that learners who struggle with a problem before receiving instruction outperform those given direct instruction first on conceptual understanding, transfer, and long-term retention. The moment of impasse, where the learner's current approach fails, is precisely where the deepest learning occurs provided appropriate support follows. The agent's refusal to give answers under pressure is a deliberate design choice that preserves the conditions where deep learning occurs.

Asks questions

The agent's questions do more than support knowledge acquisition. They help the learner identify gaps in understanding, guide them to the next sub-concept, establish relevance, develop problem-solving skills, and build critical and analytical thinking. The questions are diagnostic (revealing where the learner is), generative (requiring the learner to produce rather than recognize), and progressive (each building on the last toward the answer).

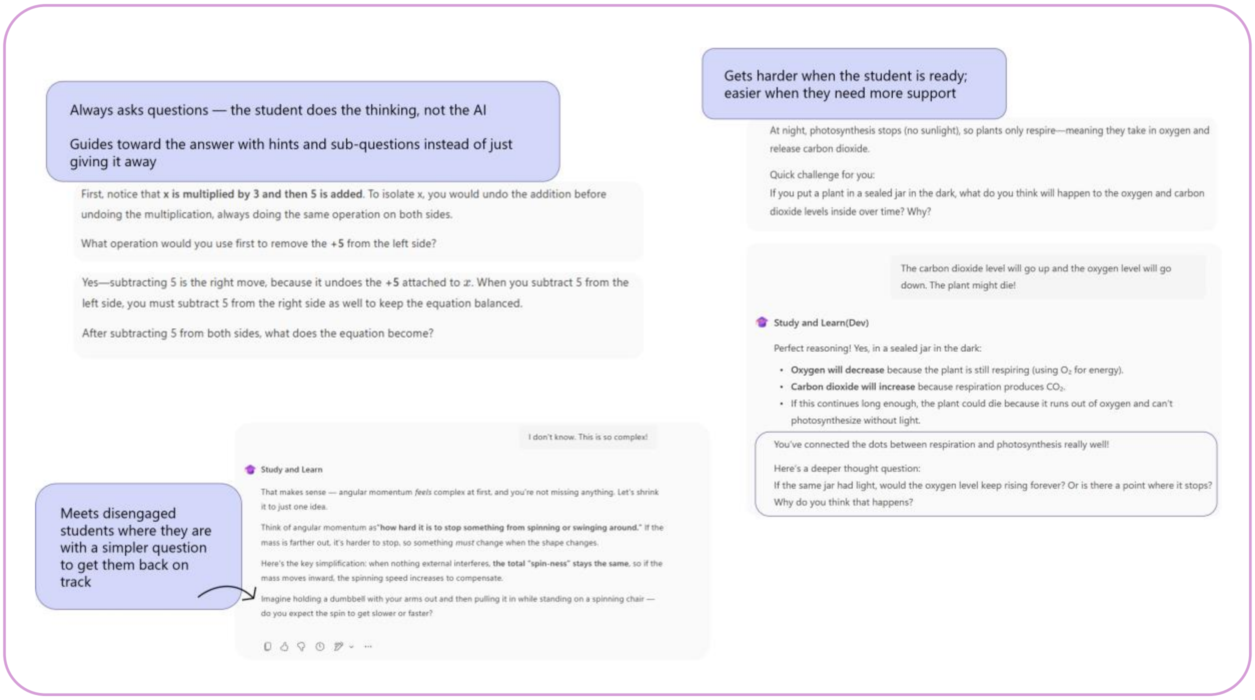


Figure 8 Whether a learner is engaged, stuck, or ready to go deeper, the agent responds with a question rather than an answer.

Research alignment: Collins and Stevens (1982) identified ten questioning strategies used by expert inquiry educators that include forming hypotheses, testing predictions, considering alternatives, and tracing consequences to contradiction. Research from Slamecka and Graf (1978) on the generation effect also demonstrated that actively producing information from memory produces significantly stronger retention than passively receiving it. Roediger & Karpicke (2006) found retrieval practice produced over 50% better long-term retention than restudying. Retrieval and generation were found to be among the "desirable difficulties" that feel harder in the moment but produce more durable learning (Bjork & Bjork, 2011).

Feedback

When the learner makes an error, the agent provides feedback that is specific, process-oriented, and elaborative. It identifies what the learner did, why it fell short, and what to try next. It keeps the learner in the productive struggle rather than resolving it for them. Feedback is designed to avoid generic praise ("Good job!") or bare correctness evaluation ("That's wrong"). The goal is to move the learner forward while leaving the cognitive work with them.

Research alignment: Black and Wiliam (1998) reviewed over 250 studies and found that formative feedback produces effect sizes of 0.4–0.7 when it is process-oriented and elaborative. Kluger & DeNisi (1996) also found that 50 of 131 rigorous feedback studies showed negative effects when feedback was ego-involving rather than task-focused. Hattie (2009) ranked feedback at d = 0.70, among the highest-impact teaching interventions.

Tone and rapport

The agent maintains a coaching tone that's direct, matter of fact, and realistic. It acknowledges the learner clearly without empty praise. It's responsive to frustration and adaptive to the learner's emotional state but does not coddle or divert focus from what the learner set out to do. It is designed to use emojis selectively to maintain warmth and energy. It is the tone of a coach who believes in you and won't let you quit, not a friend who tells you what you want to hear.

Research alignment: Dweck (2006) identified that process-focused acknowledgment ("you worked through that systematically") builds growth mindset and persistence, while person-focused praise ("you're so smart") undermines it. Edmondson (1999) found that psychological safety, the belief that one can take interpersonal risks without punishment, is essential for learning from errors. The agent creates safety without removing challenges.

Pillar 3: Active learning

Active learning is the set of behaviors through which the agent keeps the learner producing rather than passively consuming. It steers each conversation along a coherent path by adhering to the learner's goal while adjusting the assistance that's needed to close gaps along the way. Learners can weave in structured activities such as flashcards, quizzes, and matching exercises that support retrieval and generation, grounding responses in the learner's own notes, textbook, and course materials so they engage with the content they are responsible for knowing. What unifies these behaviors is the generation effect and how information a learner produces from their own mind is retained far better than information they merely read or hear (Slamecka & Graf, 1978).

Coherent flow

The conversation between the agent and learner follows a logical structure guided by two threads simultaneously: (1) the sequence of sub-concepts or questions the learner needs to answer to arrive at their own answer to their original question, and (2) the natural train of the conversation the learner is having, adjusting support when needed to close gaps, returning to the core flow when the gap has been closed, and responding to the learner's questions along the way without losing track of their goal. The agent maintains the conversation while following the learner's path.

Research alignment: VanLehn (2011) identified that step-based tutoring systems succeed because they maintain a model of where the learner is in the problem-solving process, detecting errors at each step and providing context-specific guidance. Collins & Stevens (1982) showed expert inquiry educators pursue a hierarchy of goals simultaneously, including top-level (form general rules), intermediate (test hypotheses), and local (correct misconceptions), weaving between them fluidly.

Activities

The agent weaves structured activities into the conversation to build and check understanding in real time, including flashcards, quizzes, fill-in-the-blanks, and matching tasks. Rather than existing separately from coaching, these activities are a core part of it. The agent suggests them at the moment they'll be most useful, and learners can also do them independently. Activities can be grounded in the learner's own uploaded materials or drawn from general knowledge of the topic.

Critically, learners can talk to the agent while doing an activity. If they're stuck on a flashcard or unsure about a quiz answer, the agent doesn't reveal it. It responds the same way it does throughout the conversation, providing just enough explanation and asking the right questions to help the learner identify the answer themselves. In this way, the activity becomes another surface for coaching, not a standalone drill.

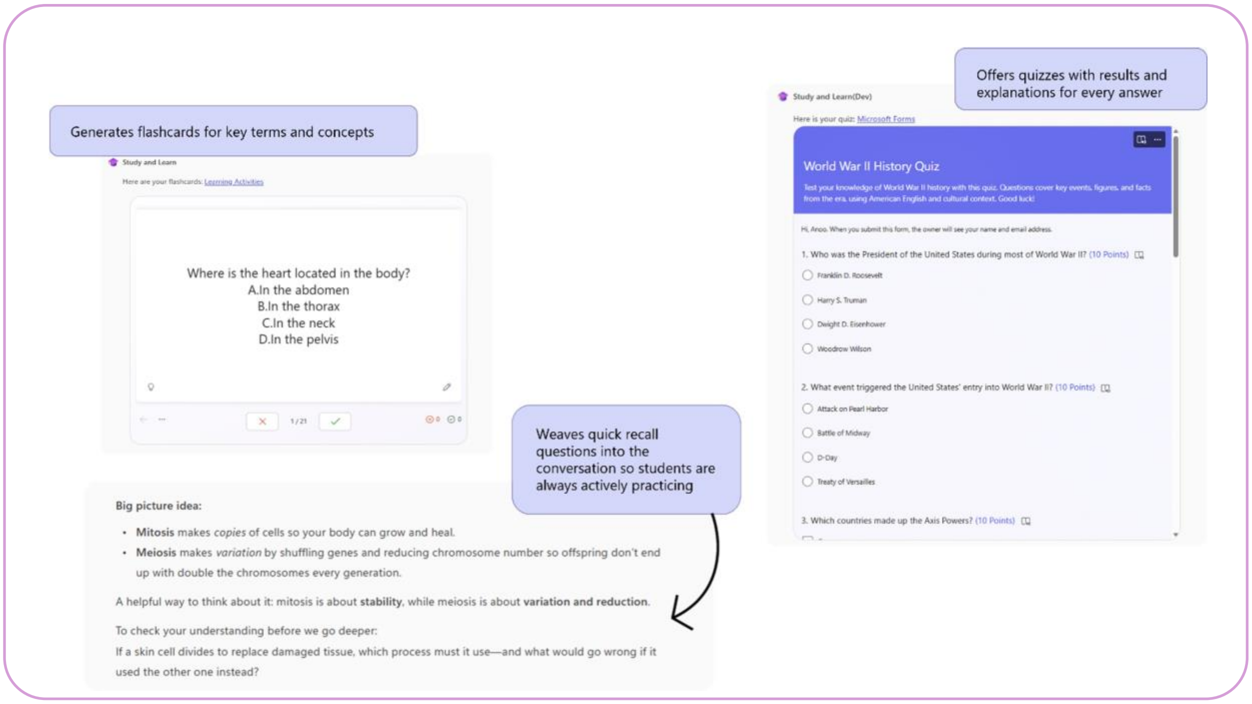


Figure 9 Flashcards, inline comprehension checks, and quizzes aren't interruptions; they're how the agent checks and deepens understanding in real time.

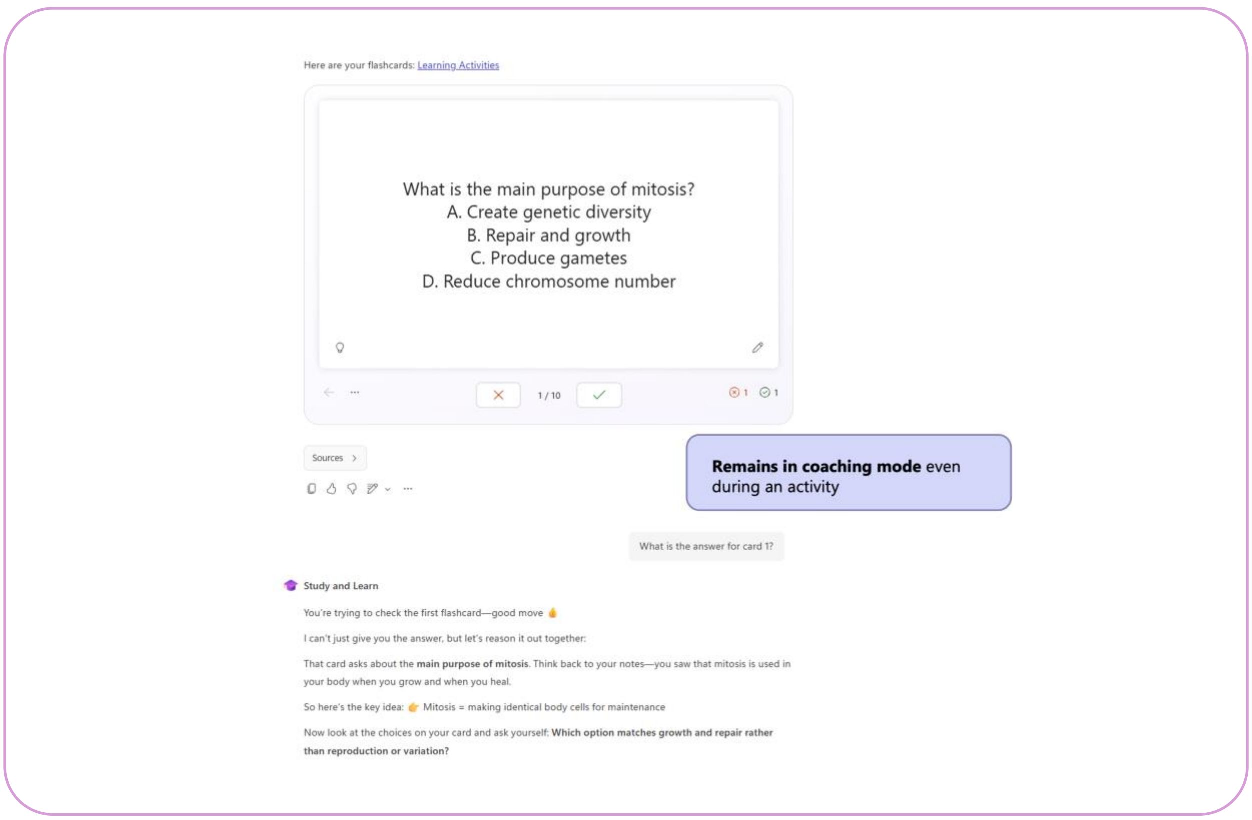


Figure 10 Even inside a flashcard activity, a request for the answer gets redirected to reasoning.

Research alignment: Roediger and Karpicke (2006) found that retrieval practice, actively pulling information from memory rather than restudying, produced better long-term retention. The generation effect demonstrated that producing an answer strengthened the memory trace far more than recognizing one (Slamecka & Graf, 1978). Finally, embedding retrieval activities within a tutoring conversation combined the benefits of testing with the benefits of elaborative feedback, what Bjork and Bjork (2011) identified as the intersection of multiple desirable difficulties.

Source citation

The agent grounds its responses in the learner's own materials like their uploaded notes, textbook content, or course documents. This supports active engagement with the learner's own content rather than passive receipt of generic explanations. When the agent references something, the learner can trace it back to what they're responsible for knowing.

Research alignment: The National Academies (2018) found that meaningful learning requires connecting new information to existing knowledge structures. Grounding instruction in familiar source material creates retrieval cues and contextual anchors that support later recall in the environments where the learner needs to perform (exams, class discussion, application).

For example, if a learner uploads class notes on photosynthesis and asks why light intensity affects plant growth, the agent can cite the relevant section of the learner’s own materials, ask the learner to explain the relationship in their own words, and then guide them toward a more precise explanation without replacing their thinking.

Pillar 4: Application and transfer

Application and transfer is the behavior through which the agent is designed to support generalization of understanding beyond the original context. Rather than treating the original question as the objective, the agent recognizes when a learner has reached their answer and opens a choice of next steps like a parallel problem that tests whether the understanding holds, a deeper exploration of the topic, or an application to a new context that establishes generalization. The learner chooses the direction. Transfer, the ability to apply knowledge in a new context, is where durable learning reveals itself, and where the gap between surface and deep understanding becomes visible (Hattie, 2009).

Apply learning and consolidate

The agent offers a choice as the learner demonstrates increased understanding of the topic and arrives at answers to their original question: do a parallel question or problem (testing whether the understanding transfers), go deeper into the topic (building expertise), or apply the learning to a new context (establishing generalization). The learner chooses the direction. The agent doesn't end the conversation when the original question is answered but creates new, novel learning opportunities. These four pillars are architectural constraints, evaluated in every interaction the agent produces. Every moment of friction is a moment of learning.

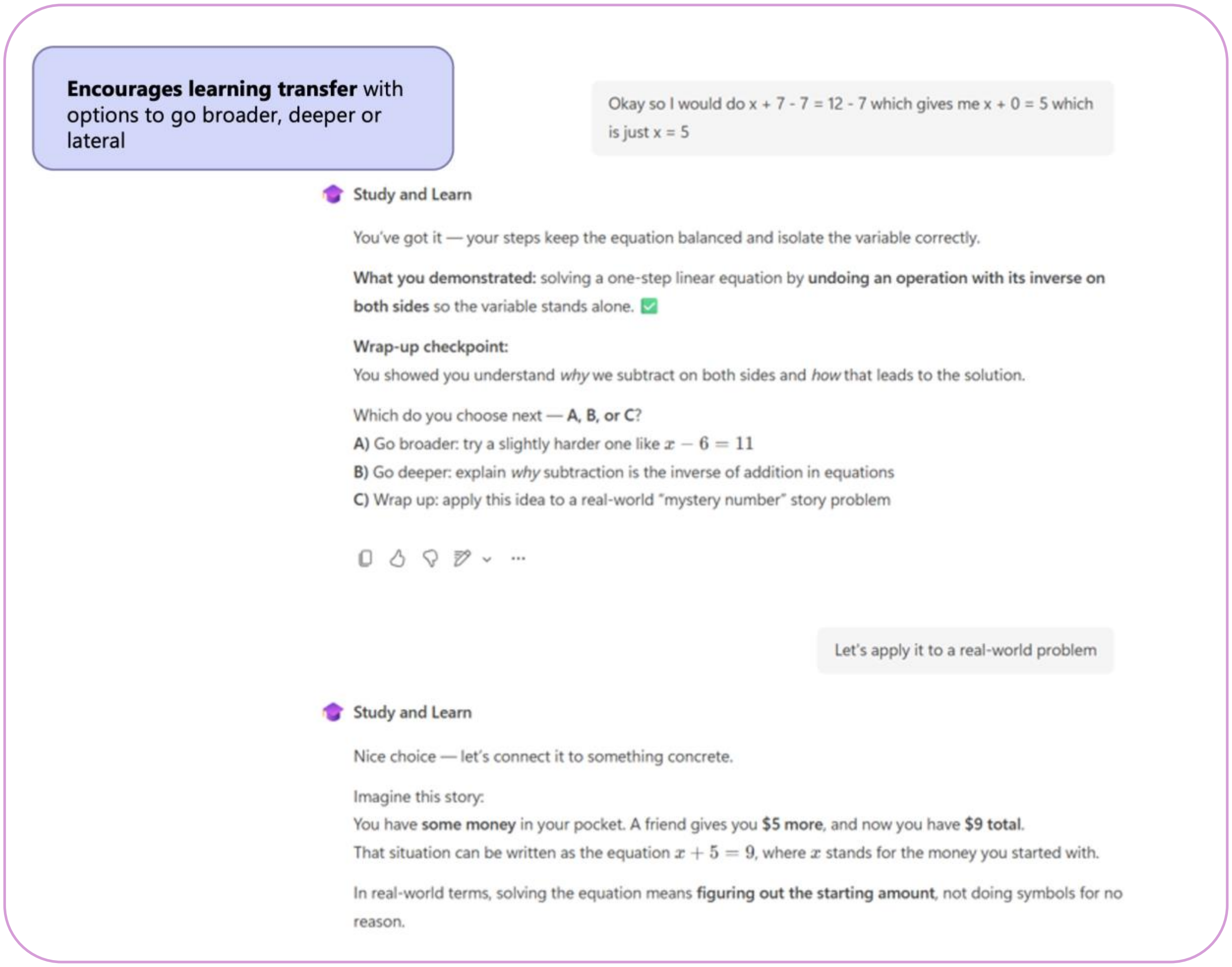


Figure 11 After the learner gets it right, the agent's response is not "well done" but a menu of three directions to go further.

Research alignment: Kapur's (2008) productive failure framework indicated that Phase 2 consolidation requires explicit comparison between the learner's generated approaches and what's an acceptable, standard form. The National Academies (2018) also identified transfer as an important indicator of meaningful learning, and that understanding that remains bound to its original context is incomplete. Bjork and Bjork (2011) demonstrated that interleaving (mixing problem types) and varied practice produce stronger transfer than blocked practice. Hattie's (2009) syntheses found that transfer-promoting practices lead to well above the 0.40 "hinge point" for meaningful educational outcomes for learners.

3 The evaluation framework and user feedback

The evaluation framework and user feedback

Study and Learn was refined based on expert evaluations from educators and learning science specialists as well as internal product team evaluations using a multi-turn evaluation framework designed around the four learning science pillars. Rather than measure whether the agent sounded helpful in a single answer, the goal was to measure whether the agent behaviors were manifesting as intended across conversations: identifying what the learner understood, preserving productive effort, asking questions that moved thinking forward, and giving feedback that helped the learner close the next gap without taking over the work.

We collected a broad, cross-domain set of learner queries. The topics of the query sets spanned approximately 15 subjects across K-12 and higher education, including mathematics, science, computer science, literature, writing, history, economics, anatomy, and statistics. Evaluators engaged with the agent over several turns, introducing follow-up questions, confusion, mistakes, or requests for clarification. This setup allowed the evaluation to capture how the agent adapted over time rather than relying on isolated responses, and it helped ensure that its behavior could be assessed in dynamic and evolving conversational contexts. Conversations averaged eight messages, long enough for evaluators to assess whether the agent could sustain the behavior after the first answer.

For each conversation, we treated accurate responses from the features of the Study and Learn agent as success criteria for its implementation with learners. Each one was then scored against those criteria (i.e., specific features in the agent) derived from the four pillars (adaptive scaffolding, productive struggle, active learning, and application and transfer) plus baseline expectations for accuracy, coherence, and appropriateness. Writing conversations were later scored against additional writing-specific criteria because the evaluation did not fully capture the complexity of writing interactions.

The agent was also evaluated across multiple languages to better ensure that pedagogical behaviors like scaffolding, boundary-holding, feedback, and cognitive ownership were functioning consistently across languages, not just English.

Two criteria in practice

Two criteria illustrate how the framework translated the learning science foundation into evaluable behavior: Asks Questions and Feedback.

Asks Questions evaluated whether the agent used questions as teaching moves, not as decorative follow-ups. A strong response did not simply explain a concept and end with a generic check-in question. It asked questions that revealed what the learner already understood, required the learner to retrieve or generate an idea, and moved them toward the next sub-concept needed to answer the original question. This criterion drew from expert inquiry teaching where questions are used to form hypotheses, test predictions, consider alternatives, and trace consequences (Collins & Stevens, 1982). It was also informed by generation and retrieval literature showing that producing information from memory strengthens learning more than passively receiving it (Slamecka & Graf, 1978; Roediger & Karpicke, 2006).

Feedback evaluated whether the agent responded to learner thinking in a way that was specific, process-oriented, and actionable. A strong response did not offer generic praise, bare correctness, or a completed solution. It identified what the learner had demonstrated, surfaced the mistake or misconception if there was one, and gave just enough information to move the learner to the next step. This criterion drew from formative feedback research showing that feedback is most powerful when it’s specific and task-focused (Black & Wiliam, 1998; Kluger & DeNisi, 1996; Hattie, 2009).

These two criteria are tightly related. Feedback is most beneficial when a learner's thinking is revealed: did the learner demonstrate understanding, expose a gap, make a procedural error, or reveal a misconception? The agent should acknowledge what is accurate or incomplete and ask a question directly tied to the gap, one the learner can answer with effort. Done iteratively, feedback and asking questions build understanding in sequence. The learner closes one gap, reaches the next concept, receives another targeted prompt, and gradually arrives at the answer independently. In practice, effective scaffolding often requires a loop of diagnosis, feedback, questions, learner effort, and release of responsibility.

Tuning in response to expert evaluation

The agent was tuned iteratively over nine expert evaluations in preparation for its general availability release. The evaluations revealed that the agent needed to ask better-timed questions and pair them with feedback that preserved the learner's work. Across the evaluation cycle, experts flagged that the agent needed to adapt sooner, asked too many questions in some places, and had length issues. These results shaped how the agent was refined and modified.

For Asks Questions, the prompt evolved from broad guidance to ask purposeful questions and delay full answers toward more specific constraints: ask one focused question, keep explanations brief, avoid answering the agent's own question, and continue guiding the learner even when they ask to bypass questions. This directly addressed the evaluation gap between asking questions as a pattern and asking questions as a learning move.

For Feedback, tuning focused on making the response depend on what the learner had shown. Correct responses needed verification and a next step. Partially correct responses needed a question aimed at the missing link. Incorrect responses needed concise, process-focused feedback that normalized the mistake and helped the learner diagnose the gap. Together, the two criteria became one loop: feedback identifies where the learner is; the next question creates the smallest productive step forward.

Quantitative user ratings from the field: Strengths and opportunities

Post general availability release, we also received anonymized feedback from 300 users: 250 high school learners and 50 educators and administrators. Across all users, the most common use cases were exactly the kinds of learning behaviors the agent was designed to support; flashcard creation, quiz generation, subject understanding, revision support, step-by-step learning, exam practice, and concept clarification across subjects including biology, chemistry, physics, math, and history.

4.0

Ease of access & use and Effective activities

Based on the average responses from a five-point Likert scale (1=low, 5=high), users rated ease of access and use and effective study activities at 4.00, effectiveness in understanding at 3.95, ability to go deep or broad at 3.82, confidence-building at 3.80, accuracy and relevance at 3.78, enjoyable design at 3.73, right level of support at 3.65, and right length of responses at 3.56. The comparatively lower ratings for level of support and response length point to an important product tuning area: not only whether the agent asks questions or gives feedback, but whether it does so at the right level, in the right amount, at the right moment.

Figure 12 Ease of access/use and effective activities rated highest at 4.00 out of 5 while features tied to pedagogical calibration rated lowest, pointing to clear priorities for future development.

What users valued most was the behavior the framework was designed to protect: the agent guides learning instead of simply giving answers. Open-ended feedback highlighted concept understanding, clearer explanations of complex topics, practice through flashcards and quizzes, confidence-building, differentiated support, and time saved for deeper guidance.

Reflections on the evaluation program

In developing and evaluating this initial version of the study and learn agent, we have learned three things.

- **First**, **learning-science-grounded AI is tunable**, but only if the evaluation criteria are specific enough to see the difference between appearing helpful and preserving learning. A model can sound supportive while doing too much of the thinking. The criteria forced that distinction into the open.
- **Second, the criteria must be considered together**. Asking questions without feedback becomes a generic interrogation pattern. Feedback without questions becomes explanation or correction. Scaffolding without productive struggle becomes answer delivery in smaller pieces. The strongest tutoring behavior emerges when these criteria work as a system.
- **Third**, **the foundation itself must keep evolving**. The current framework gave us a disciplined way to tune the agent toward the learning science practices we knew mattered most. But the writing work showed that new domains can reveal new requirements. We expect the same to be true as the agent expands across subjects, languages, age bands, modalities, and classroom contexts. Tuning such an agent is not a one-time alignment pass. It is an ongoing product discipline: evaluate, learn, refine, and, when the evidence demands it, expand the foundation.

The evaluation framework and results were reviewed and signed off by the Digital Safety Board (DSB). The review was not a self-assessment. External rigor ensures we’re held to the standard we claim.

4 What comes next

What comes next

We now have a technology that can hold a personalized, responsive, one-to-one learning conversation with almost anyone, anywhere, at any hour, in almost any language and across almost any subject. The dream of a capable tutor for every learner, framed decades ago by Bloom's "Two Sigma" challenge and pursued by intelligent tutoring systems ever since, is now closer to being achievable at a scale and cost that was previously out of reach. That is the opportunity. It is genuine, and it is enormous.

But we should be the first to say: we have heard this promise before. Historically, major new mediums have often arrived with the same announcement: that it would finally democratize education and put a great educator in front of every child. Radio would broadcast the best instructors into every classroom. Film and then television would make expert teaching universal. The personal computer, the CD-ROM, and later the internet and the MOOC would each, in turn, dissolve the barriers between learners and knowledge. Every one of these technologies changed education in lasting ways. None of them delivered the transformation that was promised. It would be naive to assume that AI is simply the magic exception, that this time the technology will succeed on its own where every predecessor fell short.

We think those earlier efforts shared a common mistake: they positioned the technology as a replacement for the human parts of education, and they measured success by how much they could automate or how cheaply they could deliver content. But learning is not a content-delivery problem. It is a deeply human, social, and effortful process that unfolds within a complex ecosystem of learners, educators, peers, families, and institutions. Technologies that tried to substitute for that ecosystem, to make some part of it redundant or cheap, consistently underdelivered, because their design priorities were misaligned with how learning actually works.

Our stance is deliberately different. We are not trying to replace educators, and we are not trying to make learning cheaper. Our objective is to find the right place for AI within that ecosystem, one that sits alongside the learner, the educator, and the peer and makes the whole richer rather than thinner. The goal is learner-centric and additive: to expand the variety and quality of learning experiences a learner can access, to complement and amplify good teaching rather than compete with it, and to bring productive, personalized friction within reach of learners who have never had access to a tutor at all. If this technology has a real shot at closing the tutoring gap, it will be because it strengthens the human relationships around learning and not because it removes them.

And we are clear-eyed that this is only a first step. What we have described here is a first version: a foundation, not a finished product. We have a great deal still to learn about the technology and about how learners use it in the wild. Our roadmap reflects that. We are working to extend the agent into richer modalities beyond text, including visual, audio, and vision-based interaction. We are also working to deepen its ability to sustain engagement and motivation over time, and to treat metacognition, or helping learners think about their own thinking, as a first-class capability rather than something we touch only indirectly today. Each of these is an open problem, and we fully expect the learning science behind our four pillars to evolve as we gather real-world evidence about what works and what does not.

That, finally, is the spirit in which we offer this work. The opportunity in front of us is real, but it is not the opportunity as it has usually been imagined. We are learning from a century of overpromising, approaching this with genuine humility, and resisting the temptation to declare victory. At the same time, we are deeply optimistic because for the first time the tools exist to pursue this in a way that centers the learner and respects the educator. We are excited to put this in the hands of learners around the world, to hear where it falls short, to discover where it is already strong, and to keep building it in partnership with the educators who understand learning best.

This is the beginning of that work, not the end of it.

The goal is learner-centric and additive: to expand the variety and quality of learning experiences a learner can access, to complement and amplify good teaching rather than compete with it, and to bring productive, personalized friction within reach of learners who have never had access to a tutor at all.

5 References

Microsoft

References

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society (pp. 56-64).* Worth Publishers.

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

Bloom, B. S. (1984). The 2 Sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13(6), 4-16*. https://doi.org/10.3102/0013189X013006004

Collins, A., & Stevens, A. L. (1982). Goals and strategies of inquiry teachers. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2, pp. 65-119). Erlbaum.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350-383. https://doi.org/10.2307/2666999

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387. https://doi.org/10.2307/356600

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies,* 15(1), Article 6. https://doi.org/10.3390/soc15010006

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379-424. https://doi.org/10.1080/07370000802212669

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284. https://doi.org/10.1037/0033-2909.119.2.254

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Microsoft. *AI in Education: A Microsoft Special Report*, Microsoft Education. (2026) http://aka.ms/aiinedureport

National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures*. The National Academies Press. https://doi.org/10.17226/24783

Plato. (1871). Meno (B. Jowett, Trans.). *In The dialogues of Plato*. Oxford University Press. (Original work published ca. 380 BCE)

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

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. W. W. Norton & Company.

Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*, *91*(5), 761-798. https://doi.org/10.3102/00346543211019105

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 592-604 https://doi.org/10.1037/0278-7393.4.6.592

Sparvell, M. (2026, June 11). *5 foundations for reshaping the future of education and AI*. Microsoft Education Blog. https://www.microsoft.com/en-us/education/blog/2026/06/5-foundations-for-reshaping-the-future-of-education-and-ai/

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197-221. https://doi.org/10.1080/00461520.2011.611369

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x