The Science of Tutoring

LAURENCE HOLT

The Science of Tutoring

www.scienceoftutoring.com

ISBN: 979-8-9875380-0-5
Ebook ISBN: 979-8-9875380-1-2

Printed in the United States of America
Design by Liliana Guia

10 9 8 7 6 5 4 3 2 1

Version 1.0

The Science of Tutoring

Why deep tutoring is superior but rare

LAURENCE HOLT

CONTENTS

1 / INTRODUCTION 1

2 / EXPLANATION 8

3 / SCAFFOLDING 19

4 / SURFACE V DEEP 25

5 / FEEDBACK 32

6 / CHECKING FOR UNDERSTANDING 37

7 / FINDING BIGFOOT 43

8 / SILENCE AND CONFUSION 50

ENDNOTES 71

1 / INTRODUCTION

An often repeated but, if you stop to think about it, somewhat bizarre finding from research on tutoring is that the experience and training of the tutor don't make much difference to how well students learn.

That doesn't mean you can take someone off the street and feel confident they will be as successful as anyone else. Just as in any walk of life, some people really are better at tutoring than others. But neither their experience in the subject they are teaching (above a minimum level), their experience of tutoring, nor the training they received tell you whether a new tutor will be successful when faced with real students.

This seems like a critical gap in our understanding of tutoring, an approach that has become a significant element of K-12 education since the global pandemic. It is not the only gap, as we shall see. In fact, it often seems that the only things we know for sure about tutoring are that it works and it's expensive[1].

This odd result about tutor experience dates back to a 1982 meta-analysis by Peter Cohen of Dartmouth College. A meta-analysis assembles findings across a set of research studies and aggregates them. Cohen took studies in which tutors received training, for instance, and compared them with studies in which they did not. He found that, on average, the amount of students' learning between the two sets was the same. There does not appear to have been further investigation of the question in the 40 years that followed.

And yet, something must be making some tutors better than others, which, as anyone who has been tutored can tell you, they are. What is it? Can it be distilled, bottled, and delivered to newly recruited tutors for their

[1] References for claims throughout the book can be found in the end notes.

nourishment? If so, we could provide the elixir to any group of tutors, re-run Cohen, and prove him wrong. With this, we will have founded the science of tutoring.

This book explores that wild idea.

Tutoring is as old as the ancients: Aristotle himself tutored Alexander the Great. Even older, in fact: bees, ants, pied babblers, and cheetahs all tutor, in the sense that they *modify their behavior to help others learn without an immediate benefit to themselves*. Adult meerkats consume prey as soon as they catch it. But if their pups are watching, they disable the prey so that the youngsters can practice killing it—for instance, the adult might remove the sting from a scorpion. As the pups age, they are given increasingly intact prey. In the language of tutoring, meerkats "scaffold" learning and "fade" the scaffolding over time. If pups struggle, adults help more.

But tutoring is not universal. Our closest relatives, chimpanzees, don't appear to do it. And, according to Kevin Laland, an evolutionary biologist, "there is surprisingly little evidence of teaching among modern-day hunter-gatherers, in comparison to learning by imitation." But it is difficult to imagine modern human society as having arisen without tutoring and, later, school teaching. Schools, being the cheaper option, overtook tutors as the main source of education in England during the reign of Elizabeth I. But tutoring never went away and, indeed, is resurgent, driven by distrust toward schooling, an increasingly desperate need to keep kids on track academically, and a growing recognition of individual differences.

Despite its long history, tutoring is something of an enigma. There is strong evidence for its remarkable effects: when you add tutoring to a student's education mix, they gain a greater understanding, are more motivated, and work faster. But we don't know why or how. What are the active ingredients of tutoring? Without an answer, we know neither how to train new tutors most effectively nor what aspects of tutoring we can vary without doing violence to it. For example, why does tutoring work for a group of three tutees but not for thirteen, or thirty? Does tutoring work as well when the tutor is remote? Can computers be used to tutor or at least to aid human tutors? If so, in what ways?

As with many advances in our understanding of a subject, a crucial step is to generate higher-resolution data and look at it closely. Understanding disease and cell biology mostly happened through guesswork until the invention of the microscope. Understanding the movement of celestial bodies was mostly guesswork too until the invention of the telescope[2].

Tutoring finally got its invention-of-the-microscope moment in the 1990s. A few researchers began to record tutoring sessions, transcribe them, and look for common "moves" such as *tutor-gives-explanation*, *tutor-asks-question*, *student-answers-question*, and *tutor-gives-feedback*. Once you have painstakingly coded enough

[2] Since one of those instruments is essentially the other flipped around, it should perhaps not surprise us that both breakthroughs occurred around the same time, in the early 1600s.

transcripts, you can ask "Which moves predict student learning?" For instance, do tutoring sessions with a lot of *tutor-gives-explanation* moves result in more learning than those without?

Still, in a country with thousands of education researchers, it is curious that so few of them have turned their attention to tutoring. With the manual transcription and coding of tutoring sessions—and more recently, semi-automatic transcription and coding—studying the science of tutoring is becoming possible. What researchers are finding is not what you might have expected.

This is not a book about how to organize and run a tutoring program, any more than a book on aeronautics is about how to organize and run an airline. Certainly, if you are to run an airline, you will want to be sure that the people who designed your airplanes know a great deal about aeronautics. And the more your pilots know of it, the better. But you don't need to. So, this is a book for people who design tutoring programs and for the tutors themselves, people who need to know how and why tutoring works.

There are many excellent books and courses on the science of flight. But to my knowledge, there are none at all on the science of tutoring. Given the enormous surge in tutoring happening around the world, that is a problem—one tantamount to dozens of airlines starting up without anyone really knowing how to keep planes in the air. This book attempts to fill that gap—or at least to begin to fill it.

2 / EXPLANATION

Richard Feyman, a Nobel Prize-winning physicist, told the story of an explanation his father gave him when he was a boy.

> *One day, I was playing with what we call an express wagon, which is a little wagon with a railing around. I had a ball in it, and I noticed something about the way the ball moved. So, I went to my father, and I said, "Pop, I noticed that when I pull the wagon, the ball rolls to the back, and when I'm pulling along and I suddenly stop, the ball rolls to the front of the wagon. Why is that?"*
>
> *And he said, "Nobody knows."*
>
> *He said the general principle is that things that are moving try to keep on moving and things that are standing still tend to stand still unless you push*

on them. He said this tendency is called inertia, but nobody knows why it's true.

That's a deep understanding. He doesn't give me a name. He knew the difference between knowing the name of something and actually knowing something.

"If you look close," he said, "you'll find the ball does not rush to the back of the wagon, but it's the back of the wagon that you're pulling against the ball. The ball stands still."

So, I ran back to the little wagon and set the ball up again and pulled the wagon from under it and looked sideways and saw he was right. The ball never moved backwards in the wagon when I pulled the wagon forward. It moved backward relative to the wagon, but relative to the sidewalk, it stood still until the wagon caught up with it.

What more surprising and curiosity-inducing expression can a small boy hear from an expert than "nobody knows"? Having hooked young Feynman, his father gave him a brief "general principle" and an idea for an experiment to confirm it.

Explanations are the mainstay of school instruction, accounting for the vast majority of what teachers do in class each day. But research has shown repeatedly that they don't reliably lead to learning. (They are, in fact, very effective for the giver but seldom for the receiver.)

Why not? Nobody knows. But researchers have uncovered some general principles.

Two expert cardiologists based in Chicago, Joel Michael and Allen Rovick, analyzed transcripts of tutoring sessions in which they were the tutors. Their approach was to show tutees scenarios in which a person's blood pressure was perturbed and to ask them to predict cardiovascular responses. Michael or Rovick then began a dialog with the tutee about the accuracy of their predictions. The researchers were well aware of the limited value of explanations. But when they examined transcripts of their own tutoring, they were shocked to discover that, at the first sign of an error in tutee thinking, they flipped into full-on explanation mode and never flipped back. They had unwittingly demonstrated that, like a trout and a dragonfly lure, not even an experienced tutor can resist the draw of explaining.

Most pernicious are unsolicited explanations in which tutors launch into a lecture, often at the start of a session. As well as being ineffective, they waste valuable tutoring time.

But there is "a time for telling," as the title of a famous paper by education researchers Dan Schwartz and John Bransford has it. They showed that there are moments when an explanation is not only okay but that failing to provide one would be a missed opportunity.

For example, a student might have the insight that when comparing similar right triangles—triangles that are the same except for their size—the two sides next to the right angle are always in the same ratio to each other. But they will not figure out for themselves that the ratio is called the tangent and that someone back in antiquity made a handy lookup table of that ratio for every triangle. That is a time for telling.

Schwartz and Bransford call this the "detective story" approach. The explanation is the solution to the story. It comes at the end. Instead of starting with an explanation of trigonometry and then giving examples to illustrate it, the student divines something in the examples first and the denouement comes with the tutor's explanation.

In other words, an explanation is the right move when the tutee is ready for it, not when the tutor is. This is the first of many examples we are going to encounter showing that *effective tutors are following the student's lead*, even though it is almost universally assumed by tutors that it should be the other way around.

The research literature is unequivocal on one way through which most tutors can instantly improve their effectiveness: talk less. "Most tutors just won't shut up," says Micki Chi, a professor in the Institute for the Science of Teaching and Learning at Arizona State University. She once ran an experiment in which she attempted to train tutors to stop making long explanations. ("How long is long?" I asked. "You know it when you hear it," she replied.) Chi trained the tutors to replace explanations with "content-free" prompts such as these:

What's going on here?

You look like you are thinking about something. What is it?

Could you put that in your own words?

Do you have any ideas about why that might be the case?

Could you connect what you just read with something you read before?

Anything else?

But the training didn't take. Tutors found it difficult to keep to purely content-free prompts. Instead, they often substituted content-specific prompts such as "Which of Newton's laws would help here?" or "How could you get the *x* on its own?" The problem with these prompts is that they reflect the pathway the tutor has in mind, which may not match the pathway the student has in mind. This approach to tutoring isn't totally ineffective, though it can turn sessions into a game of "guess what the tutor wants me to do next." Students can get quite good at this game, but it leaves them unprepared when they face similar problems in the future and the tutor isn't around.

Although Chi's tutors didn't keep to content-free prompts, they were successful in reducing long explanations. What effect did that have on student learning? Surprisingly, learning without explanations was as good as learning with explanations, even though students who got the long explanations heard a lot more information than those who heard only the prompts. Without explanations, the tutor–tutee conversation became much more interactive. Students initiated more of the dialog and did proportionately more of the work and that led them to learn more.

Research has uncovered some principles for how, as well as when, to give an explanation. The first is that the explanation must be at the learner's knowledge level. We'll call that a *considerate explanation*. When technical support agents help customers fix technology problems, they very often bamboozle the caller. That's if they give an explanation at all. They may simply tell you a sequence of mysterious steps. As a result, when the problem reoccurs, you are no better able to solve it. But when agents were given some information about a caller's IT expertise—callers answered questions such as "Do you know what FTP is used for?"—the agents were able to adapt their explanations in a way that made them far more effective for the callers. The way they did this was to give more definitions to less-skilled customers while talking more about processes and events (that rely on an understanding of definitions) to more skilled customers. They recalibrated their mental model of whom they were talking to: instead of picturing a generic layperson, they pictured a specific individual with competencies and confusions.

Tutors, likewise, must encourage tutees to make their thinking visible, like video gamers who narrate their gameplay on YouTube. That allows the tutor to pitch their explanations at the level of the tutee, or a bit above. This is akin to adjusting your language when trying to be comprehensible to a non-native speaker: slowing down and using higher-frequency vocabulary (rather than just talking more loudly). Hence, the term "considerate explanation".

The second general principle, which was noted over 30 years ago by Paul Vedder, a Dutch teacher, is a wonderfully simple technique for supercharging explanations. We'll call it *active explanation*. Here is an example from vocabulary instruction:

S: What does "heathen" mean?

T: Sort of uncivilized. It's not usually meant kindly.

(Pause)

So, would you call the King of England a heathen, do you think?

S: No.

T: Neither would I. Who would you call a heathen?

S: (Thinks) You?

T: What! Why?

S: You eat with your fork and knife in the wrong hands.

Vedder suggested explanations that include an opportunity for the learner to immediately put new knowledge (such as the meaning of "heathen") to work are more effective. An easy way to do this is to come up with an example or counterexample and ask the learner to say which it is—for example, is the King of England a heathen? This technique is effective for two reasons. It gives the learner an opportunity to use their new knowledge, which helps memory. And it uncovers situations where the explanation didn't click.

I once ran a workshop with a group of business executives, and after 20 minutes of explaining, I finally gave them an exercise to check how much they had absorbed. The answer was next to nothing. I was shocked, though every teacher I've described this to has a similar story. The solution, of course, is not to wait 20 minutes.

The revelation that they didn't understand can also come as a surprise to the learner, who frequently has the illusion of understanding to match the teacher's illusion of having explained. Active explanations rapidly disabuse them both of the notion.

Richard Feyman's father intuitively followed these prescriptions: he pitched his description of the physics at the level of a boy's understanding (considerate explanation), gave young Feyman a follow-up to pursue (active explanation), and provided the explanation in response to a specific question (time for telling). Perhaps it should not be surprising that Feyman himself went on to become one of the world's great explainers.

3 / SCAFFOLDING

I first met Micki Chi at a conference lunch, where seven researchers and I were seated at a large, circular table. The standard protocol at these events is to begin by going around the table giving introductions. As Chi introduced herself in a quiet voice, eyes fixed on the centerpiece, the young woman next to me became unable to sit still. Finally, she interrupted, “Wait, you’re Chi, 1989?”

“Yes,” said Chi, softly.

“Whoa,” said the young researcher, “I have cited you so many times!”

In the world of peer-reviewed journals and researchers whose lives turn to the rhythm of publication, “Chi, 1989” is how you refer to ideas. That paper, one of the most cited in all of cognitive science, tells of the discovery that having learners pause and explain to themselves what they just read increases how much of it they can remember later. Chi called it *self-explanation*. The marvel is that a technique so simple and so powerful has not found its way into every classroom in the country[3].

At the time I met her, Chi had just published another paper, "Learning from Human Tutoring," perhaps the most in-depth study of tutoring ever conducted. For the study, Chi and her team compiled transcripts of tutoring sessions on the human circulatory system. Students learned about the passage of blood through the body, lungs, and heart and why the system is arranged the way it is.

Chi coded each move that a tutor—or a tutee—made. She then calculated which moves correlated with learning gains. She found that explanations result in little learning, as we have already seen. According to Chi, that's because garden-variety explanations don't cause a student to do anything. In Chi's view, it's what a student *does* that

[3] That is the way of theory and practice in education. As one wag said, "… theory and practice are very similar, in theory, but in practice, they turn out to be very different indeed."

drives learning. Explanations mostly elicit "uh huhs." But a different move used by some of the tutors in the study was more promising: *scaffolding*.

"Scaffolding," says Chi, "is a kind of guided prompting that pushes the student a little further along the same line of thinking." It is a coordinated performance by the tutor and tutee. The tutor sets up a part for the tutee to play, one that requires them to do an increasing share of the cognitive work. The student is in the spotlight; the tutor is the director.

Chi provides a list of 15 types of scaffolding:

1. Pumping for more ("What else?")
2. Hinting ("So, it's kind of leaving out the lungs there?")
3. Making fill-in-the-blank requests ("Okay, and it goes from the atrium to the …")
4. Highlighting critical features of the task, calling attention to discrepancies between the current state and the goal ("This is good, but what about that?")
5. Decomposing the task, reducing complexity by allowing the student to focus on one part at a time

6. Executing parts of the skill, allowing the student to execute others
7. Providing physical props or cue cards
8. Describing the problem in a way that orients the student to the important features
9. Comparing the current problem with a previously solved problem
10. Maintaining focus on the goal by recognizing progress toward it
11. Initiating a reasoning step or a task ("Name every part on there.")
12. Completing the student's reasoning step, jumping in and providing the correct answer when the student makes an error, without acknowledging that an error has been made
13. Asking a leading question ("And where do you think it goes?")
14. Redirecting the student
15. Providing an example.

Scaffolding isn't just asking a question; it's nudging the student along a path toward a solution or a new understanding. The form of scaffolding—whether it's a question, statement, command, or something else—doesn't matter. It's the intent of the move itself that matters.

With scaffolding, the intent is for the student to get further than they would without it. Carpenters lay plywood over joists to give them a safe place to stand while they build the next level. Picture a student standing on the plywood, wobbling. Perhaps one of the planks is cracked, or has fallen. The tutor's role is to help them get stable and begin building the next level.

Russian psychologist Lev Vygotsky called this being in the *zone of proximal development* or ZPD. It's the zone in which the scaffolded student finds they are able to achieve something they may not have achieved alone.

Once the student is stable, the scaffolding can be gradually removed, leaving them with the ability to perform the new skill unaided. Educators, with an awkward mixing of metaphors, call this *fading*. The Romans built arches this way: they used wooden scaffolding to hold the arch up until the final piece, the keystone, was put in place. When they removed the scaffolding, the arch stood on its own. Some of them still do. Something similarly satisfying can happen for students: the last piece clicks into place, and their mental model achieves an integrity, a wholeness that allows it to stand on its own.

In Chi's study of biology tutors, scaffolding was quite rare: around 6% of tutor utterances were scaffolds. Explanations were nearly 10 times as frequent. Students were much likelier to respond to a scaffold with something meaningful, not just "uh huh" but another brick in the arch. You could often see them doing real cognitive work, a strong indicator of learning.

Did students whose tutors provided more scaffolding learn more? Well, there is good news and bad. The good news is that scaffolding did indeed correlate with learning. The bad news is that the learning was shallow.

4 / SURFACE V DEEP

In the 1970s, Ference Marton and Roger Säljö of the University of Gothenburg in Sweden noticed that students took different approaches to learning. Some students focused on remembering information. Others focused on understanding it: connecting it to other information, figuring out its structure, when it might be useful, making predictions based on it, and so on. Marton and Säljö christened the former *surface* or *shallow learning* and the latter *deep learning*.

Reading the last paragraph, you may have already formed the opinion that surface learning is bad and deep learning is good. But that is not always true. Some essential knowledge just doesn't have much depth to go after. For instance, the letter *m* makes the "mmm" sound. There simply is no conceptually deep understanding of this fact to be had. And almost all deep learning relies on knowledge of surface details. You can't construct an argument integrating multiple causes of World War II if you can't recall any of them.

Still, moving from being a beginner to an expert in any topic requires deep learning, so students who default to a surface approach will, sooner or later, have to be induced to go deeper. The task is made more difficult whenever we tell students they will be tested. Since most assessments operate at a surface level, the prospect of being tested can act as a signal to students that they need only memorize, not understand.

How do you push students to take a deep approach to a learning task? An effective method is to give them a goal that requires deeper understanding. For example, instead of the goal being to pass an assessment, help the student identify a project in which they are invested and which requires depth to complete.

Another approach is to lead the student gently into deeper waters. This is where tutoring comes in.

Take this example from the circulatory system study. A student and a tutor have just read the sentence, "If a substance can pass through a membrane, the membrane is permeable to it."

S: *So, it explains itself. If something is permeable to something, then that thing can pass through the other thing.*

T: *So, how would the …*

S: *And if it's impermeable, it can't.*

T: *And how does that relate back to the capillary walls?*

S: *Well, the capillary walls …*

T: *Can you explain?*

S: *Well, this is how I learned it.*

T: *Uh huh.*

S: *In the cell, it's made up of these things, and then, it has these protein things, like this* (draws a protein lying across a cell wall like a channel through the wall). *They're really, really big. And then, there's a little substance* *like oxygen, and it can just go through here* (pointing to the wall). *But a big substance like sugar, which is tons of letters, has to go through the protein first.*

T: *And how does, how does that relate to the cell membrane or the capillary?*

S: *Well, if it's too big—if something's too big—to go into the capillary through the capillary membrane, it can't, but then maybe, if it has protein, it can. Okay.*

T: *Okay.*

S: *Alright.*

Twice the tutor asked "How does that relate …"—a move that is specifically designed to prompt a deep response (Chi's scaffolding type #13). The tutor could have, instead, asked "What is passing through the cell membrane here [in the lungs]?" But even if the student had answered, correctly, "oxygen and carbon dioxide," they would be retrieving information they had read earlier, a surface response.

This is precisely the difference between the "content-free" prompts Chi tried to get tutors to use in the experiment we described earlier versus the "content-specific" prompts tutors used instead. The former push students towards deeper responses and so deeper learning.

Constructing deep prompts like this does not come naturally to tutors, as Chi found. Closed yes-or-no questions won't do. Deep prompts typically begin with "how," "why," or "what if." But the critical feature is *it should not be possible to answer just by repeating something you've learned*. A deep prompt presses the student to produce the meaning or implication of knowledge. For example, in the dialog above, the tutor asks the student to take the idea of permeability and apply it to the specific case of what is going on in the capillaries.

As it does here, a deep prompt will compel a student to do some cognitive heavy lifting. Often, that means fitting together multiple pieces of information. Some of those pieces the student will only have encountered recently, so the tutor can help by more scaffolding:

serving up missing pieces of information but pressing the student to fit them together. The process of a student constructing or generating something they didn't know before—or only sort of knew—has been shown to be a highly effective way of learning. But if the tutor does the fitting together on behalf of the student, the student will gain little. Whoever does the work does the learning.

So, one way to evaluate a tutoring session is to ask whether it pushed the tutee toward deep learning or whether it remained skating at the surface. Micki Chi's study of circulatory system tutoring asked exactly that. When Chi looked at tutor moves and which resulted in surface or deep learning, there were three big takeaways.

First, tutor explanations led, at best, to surface responses from the student. In fact, most of the time, they didn't lead to content-ful responses at all—just an "uh huh."

Second, scaffolding mostly led to surface responses too and deep responses only occasionally. But, as we just saw, that is most likely due to the kind of scaffolding prompts that tutors used rather than being a problem intrinsic to scaffolding. In other words, scaffolding can lead to deep responses, but few of the tutors in Chi's study knew how to do so.

Third, there was one move that was more likely than any other to lead to deep responses. We will meet it soon.

5 / FEEDBACK

In Western Australia, in the summer of 2006, John Hattie, an education researcher, had a brainwave. What if, he asked, you were to take the entire literature of educational methods—tens of thousands of papers—and run one enormous bake-off? Might you identify the handful of interventions that schools everywhere should implement?

It was a colossal task, but when Hattie published the list—and each time he has updated it since—the same method appeared at or near the top: feedback. By one estimate, feedback can double a student's rate of learning.

On the face of it, there is not much to say about feedback. It is a powerful inducer of learning. And studies have shown that tutors do it liberally and instinctively. If a child points at a pig and says "sheep," any adult within earshot will find it difficult to resist giving feedback[4].

But, as ever, there is more to it than that. To get the full benefit of feedback—doubling the learning rate—the literature calls for feedback to be "well implemented." What does that mean?

It does not mean non-stop. "Too much feedback, too often and too soon ... can turn students into feedback junkies, too reliant on advice and therefore less likely to do well when left to their own devices," says Hattie.

So, what's the right way to do feedback? At its core, it is *to help the student improve by revealing the gap between what they know and what they could know*—call that the gap between A and B. This suggests two strategies: sometimes, the tutor can simply point out B; other

[4] It doesn't always work. My four-year-old daughter, visiting a zoo, pointed at a fish and said, "Flounder." I read the display label and gave her some feedback. "Actually, it's a ray." "Ray-flounder," said my daughter. She's been negotiating ever since.

times, the tutor can point out the existence of a gap and let the student figure out what it is and how to close it. In one case, you reveal the specific A-B gap, in the other you reveal only that there is a gap in need of closing.

The first strategy fits best with surface learning—perhaps during an earlier phase of learning in which students are collecting facts and not yet doing anything particularly interesting with them. The second strategy, letting them bridge the gap themselves, fits best with the deeper learning phase, when students are trying to fit facts together into something new.

For instance, if a student can't recall Newton's Third Law, tell them what it is ("Every force has an equal and opposite force"). If, instead, they can recite those words but are stuck when asked to draw the forces keeping a book sitting on a table, your feedback might remind them that the law exists and tempt them to use it, like seed laid down for a bird.

This means that, in deciding what feedback to give and when, tutors have to consider what type of learning they want to induce: surface or deep. In general, for any new concept, students progress from surface to deep. The tutor can close gaps of the former variety but simply highlight and allow the tutee to close those of the latter. The mix is key. A tutor who relies too heavily on the one will impart only surface learning and on the other risks imparting only frustration, as the lesson slows to a crawl and the arc of it is lost.

A powerful way to highlight the A-to-B gap is to use what James Nottingham, a teacher coach, calls a WAGOLL: What A Good One Looks Like. Instead of saying "Your diagram is missing a force over here," say "A good force diagram will have arrows for all the forces, and every arrow will have an equal and opposite arrow." The beauty of this approach is that it doesn't divert the student from the work; it makes clear, succinctly, what the work is.

It is surprising how often the WAGOLL is news to the student. There is a story of a boy who read slowly whenever his teachers tested him. Let's call him Sal. One day, a famous researcher visited the classroom and was asked to sit with Sal. The researcher listened to the boy read a passage accurately but plodding slow and said, “Now, Sal, would you read the passage again, but this time, the goal is to read it quickly, like this …” and read the first sentence at target pace—a WAGOLL. “Oh, okay,” Sal said and proceeded to read perfectly on grade level. Until that moment, he had thought the goal was to read each word carefully and with precision, just like he had heard his teacher do when he was first learning to read.

6 / CHECKING FOR UNDERSTANDING

For many years, it was thought that the answer to the mystery of how the earth came to have a moon was that the moon was wandering by and the earth captured it. Subsequent investigation, though, showed that the laws of physics would not allow such a thing. The moon would have been slung out into space, or into the sun[5]. Very gradually, the prevailing assumption changed.

Similarly, for many years, the prevailing assumption as to how tutors operate—and the answer to the mystery of why tutoring is so effective—was that, faced with a single student rather than an entire class, the tutor could

[5] In fact, the moon is a chunk of earth that split off in a huge cosmic collision. When astronauts brought back moon rock, we found that it was strangely familiar.

diagnose their weak spots and address each of them, aligning tutor and tutee orbits so that they would revolve together for a while.

For that to be possible, tutors must continually check a student's understanding so they can adapt their next move accordingly. But multiple studies have shown that tutors are really not very good at checking for understanding. The prevailing view of how tutoring works is going to have to change.

In particular, tutors do not reliably pick up on a student's alternative—and perhaps flawed—thinking. Instead, they do what might more accurately be called a "check for correct": does the student's answer match that of the tutor? If not, they set about adjusting the student's approach to match their own. They assume the student's thinking is uninteresting and discard it.

Again, there is a crucial difference between surface and deep. For example, if my tutor asks me to add 317 and 45 and I write down the numbers left aligned instead of right aligned, a surface assessment would conclude that I have forgotten how to line up numbers when doing addition. A deep assessment might conclude that I don't understand place value—that addition only works when I add the ones to the ones and the tens to the tens. A tutor making a surface assessment would tell me no more than how to align the numbers. A tutor making a deep assessment would dig further and may wind up changing their plan for the session to teach, instead, a mini lesson on place value.

True checks for understanding are very human acts, like the bus driver who sees you may be having a bad day and holds the doors open versus the robot that closes them in order to stay on schedule. A surface assessment is simple and swift. A deep assessment requires divining the peculiar circumstance that led the student to their answer. It is not just a 2D snapshot of what is presented but adds the dimension of time: how did we get here? Perhaps you don't understand place value. Or perhaps you do, but you were distracted by thinking about the argument you had with your sister this morning. One is a fundamental misconception that needs to be addressed; the other is just a slip.

Supporting deep learning depends on acts of *diagnosis* like this. The two diagnoses above call for two very different responses from the tutor, and we don't know which diagnosis is correct. A follow-up question may be needed to find out. For instance, "In the number forty-five, what does the four stand for?"

Why do tutors so often skip the diagnosis stage? One reason is that it is difficult to do. The core characteristic of deep learning is that it requires *connections between elements of knowledge*[6]. Diagnosing a breakdown in deep learning, then, requires the tutor to ask a question that can only be answered by connecting knowledge.

---

[6] At least, that's the characteristic we're focusing on here. In practice, people use "deep" to mean "anything that isn't surface," where "surface" means "repeating information you heard." So, "deep" has also been taken to mean mapping the structure of the knowledge, using it to make predictions, knowing which situations to apply it in, and understanding the intention behind it. All of these rely on knowledge connections, hence our shorthand.

Why is X true? Well, if you put X together with Y, you can see that X has to be true. Why did Shakespeare open Hamlet with Barnardo asking the guard "Who goes there?" instead of the other way around? Well, it shows how messed up everything is, even from the first line of the play. Why is your left ventricle more muscular than your right? Well, it has to pump blood all around your body, not just to your lungs. How are fractions like whole numbers? Well, you can add them; in fact, you can do all the same things to fractions that you can do to whole numbers. Which other president is most like Andrew Jackson?

Another reason why tutors don't do much deep diagnosis is that they typically come to a tutoring session with a plan, some form of "script" such as a series of problems to work through. The very existence of a plan creates pressure to follow it. Stopping to uncover the root cause of a bug in a student's thinking might feel to the tutor like taking the session off track. If there are multiple students in the session, diagnosing each of their understandings can take a lot of time. Instead of diagnosing, tutors repeatedly cajole students toward their (the tutor's) own solution pathway for a problem. It's as if the tutor wants to pick the tutee up and place them at the start of the pathway and then, every time the tutee takes a step off the path, nudge them back onto it.

But this is missing the true value of tutoring. The goal is not *covering*—moving diligently through a set course of material—but *uncovering*—creating moments that reveal a student's thinking and where it can be advanced. Such moments are extremely difficult to pull off in a lesson taught to a whole class. For a student to have taken a leap, stumbled, and betrayed that to their tutor is a rare opportunity indeed, an open-handed gift. Take it. Throw out the script—or at least set it aside for a bit. Show me. Show me how you got to this place in your thinking. Because I don't care whether you got to the right place this time or not. I care that you learn how to navigate—to get yourself to wherever it is you will need to go.

This moment is something to get excited about.

7 / FINDING BIGFOOT

We have now completed our tour of the major tutor moves: *explanation* (which is not very effective, except as the coda to a learning episode), *scaffolding* (effective for surface learning but difficult for deep), *feedback* (ditto), and *checks for understanding* (easy for surface understanding but harder for deep diagnoses).

The clear pattern is that surface learning is quite easy to obtain—indeed, it is difficult to avoid. That perhaps accounts for the observation we started with: the general effectiveness of tutoring however and by whomever it is delivered. But deep learning is much harder to achieve, requiring a great and targeted effort by tutors. Even then, some of a student's misconceptions—such as a "folk" understanding of the physics of motion that objects slow down and stop unless a force is pushing them—may have been held for so long that a single session of tutoring is simply not enough to dislodge them.

Or, to restate the previous paragraph for the optimistic reader, there is a colossal opportunity to improve what tutors and tutoring can achieve. We can imagine two kinds of tutors: those who produce perfectly respectable surface learning easily and declare victory; and those who make a determined effort to engender deep learning—succeeding, even so, only part of the time. What follows is intended for the second group.

Let us first repeat that surface learning is not without value. Some content isn't sufficiently meaningful to allow anything else: Knowing your multiplication facts does not require a great conceptual leap. Neither does decoding written text in the early grades, nor punctuating writing in later elementary. Most content, though, can be treated at both a surface and deep level. Indeed, this might be the definition of good art.

"The world breaks everyone and afterward many are strong at the broken places."

“At the still point, there the dance is.”

“The pieces I am, she gather them and gave them back to me in all the right order.”

In math and science, students may find themselves being given a procedural, surface-level understanding—tips and tricks such as dividing fractions by flipping and multiplying or reversing the sign when you move a term from one side of an equation to the other. Instruction designed to impart a deeper, conceptual understanding is harder to serve up, and so, it is often simply left out for several grades. This is why, when you arrive at calculus, the wheels come off.

Whatever the topic, traveling beyond beginner levels and into the realm of experts requires depth. Perhaps this is why schools have evolved to turn out beginners in a wide range of topics and experts in none.

So deep learning, sooner or later, is a cognitive necessity. What remains is for us to identify the tutor moves that lead reliably to it. Unfortunately, detailed studies of tutoring sessions have not been able to find them. Micki Chi analyzed the rare occasions when deep learning was glimpsed in tutoring sessions in order to answer the question: which tutor moves predicted it? She couldn't find any. Nothing tutors did in the sessions she studied reliably led to deep learning.

This points to the strong possibility that, like Bigfoot and UFOs, deep-learning tutor moves are difficult to find for the simple reason that they don't exist. Kurt VanLehn, a computer scientist and education researcher at Arizona State University, writes, rather shockingly, of "the intriguing possibility that the content of the tutor's comments may not matter much." He doesn't mean a tutor is unnecessary. Rather, he means that tutors' key contribution to learning may not be the content they share but their ability to orchestrate the session so that the student engages with the topic in a sustained way. Indeed, it is a striking fact of tutoring that tutees are so engaged. They almost never fail to respond when the tutor asks a question. (If that doesn't seem odd, contrast it with the classroom, where most students don't respond to most questions, or to online learning, where student attention wanders all over the place.) Perhaps the secret sauce of tutoring is simply making it difficult for a tired or bored student to hide.

There is something in that, but Chi the indefatigable—who, just to make things interesting, is VanLehn's spouse—wasn't content with it. Spooling compulsively through the rows of numbers in her statistics app, she finally noticed one thing that actually did predict deep learning. It was the moment of a student reflecting on her own progress—for instance, saying something like "Hmm, I understand most of this, but not all of it." Students who did that demonstrated more deep learning later on. Perhaps, thought Chi, we have been looking in the wrong place all along: it isn't what the tutor does that matters; it's what the *student* does.

Once we switch focus from tutor-moves to student-moves, a whole raft of potential signs of deep learning suddenly comes into view: forming hypotheses ("Maybe, germs can get in through your skin, through a cut"), extrapolating to other situations ("There's fighting going on between the good and the bad germs ... a whole new world inside the body"), coming up with predictions ("What if there was a super fighter germ?"), forming analogies ("So, the septum is like a wall in your heart"), generating justifications, generating critiques, and revising existing knowledge to deal with conflicting information.

What everything in that list has in common is, once again, *connecting knowledge*, fitting newly learned information together with something that was already there, like rain falling on a pond. For instance, to reflect is to ask whether the new information fits. Hypotheses, extrapolations, predictions, and analogies require you to begin with the new information and take it somewhere. Generating justifications and critiques and resolving conflicts require you to bring existing knowledge to bear on the new information.

If we can get a student to make any of these connecting moves with new knowledge they just encountered—or even better, more than one connecting move—deep learning is more likely to follow. Are there tutor moves that prompt students to make such a connection? And if so, why didn't those moves show up in the research? Didn't we just conclude that, maybe, they don't exist?

We did. Though perhaps they do exist, just not where we are looking. The corpus of tutoring sessions that researchers sift through may not be rich enough to allow them to be found. Sessions in that corpus are typically led by inexperienced tutors—frequently college students, a ubiquitous and affordable resource right outside every research lab's door—and focus on surface topics, such as solving simple physics problems. It's like trying to discover new species of whales by studying the lakes of England. They aren't there. You won't find deep souls in shallow water.

8 / SILENCE AND CONFUSION

The first clue to the secret of deep tutoring came, almost by accident, in a 1998 experiment. Schwartz and Bransford were working on what became "A Time for Telling," the ground-breaking paper we encountered when discussing explanations. The researchers did something curious. Instead of having students listen to a lecture and then attempt to solve some associated problems, they switched the order. They had students attempt problems, such as measuring the density of clowns packed in a bus, without having been introduced to the concept of density. Students listened to the lecture only after working the problems.

On the face of it, this seems like a recipe for frustration. How can you solve problems before you have been given the tools to do so? Indeed, the problem-first students did poorly at solving the problems before hearing the lecture. But after the lecture, they performed as well as lecture-first students on surface-learning questions and much better on deep-learning questions. Schwartz and Bransford had discovered a way to induce deep learning.

Why might this topsy-turvy approach work? Schwartz and Bransford suggest that putting the problem first gives students a chance to develop the prior knowledge they need to get the most out of the lecture. For example, when students analyzed pictures of clowns in different numbers on buses of different sizes—lots of clowns on a small bus versus a few clowns on a large bus—and tried to think of ways to summarize the differences, they may have begun to develop a proto-concept: call it the “crammed-ness” of clowns on a bus. “Crammed-ness” then provides a seed for the formal notion of density presented in the lecture. Attempting to solve the problems first made them better prepared to learn. This is akin to the idea we encountered when considering explanations: later is better.

Manu Kapur, a learning scientist based at ETH, a public university in the heart of Zurich, Switzerland, took this idea and ran with it. He systematically showed across dozens of studies that the cognitive free-for-all induced by problem-first instruction is actually beneficial for learning. His name for this phenomenon was *productive failure*, though widespread usage has morphed the term into the slightly less abrasive *productive struggle*.

Kapur even showed that the lecture-first approach may be damaging. When asked to invent ways of summarizing the variation in a basketball player's point scores over 20 games, students quickly came up with five or six. When asked to do the same thing after having been told the "official" answer—standard deviation—students couldn't come up with any; they just kept repeating the solution they'd been given. "Students may infer from instruction by a knowledgeable adult," says Kapur, "that all the relevant knowledge and procedures that they need to learn have already been taught." Problem-first students have no such apprehension.

According to Kapur, problem-first instruction works not just because it seeds a framework of proto-concepts such as "crammed-ness" but also because it improves motivation. If you struggle with a problem for a few minutes without much success, you are considerably more interested in hearing the solution than if you are just presented with it cold. In one study, Kapur compared watching other students struggle with struggling yourself and found, perhaps unsurprisingly, that there is nothing like your own failure to incentivize you to hear the canonical solution. The enhanced motivation may enable instructors to push beyond surface learning into more cognitively demanding deep learning.

One of the most counter-intuitive findings in the science of learning featured in a 1992 paper by Robert Bjork and Richard Schmidt, professors at the University of California, Los Angeles. The paper systematically establishes that our intuitions about what it looks like when a student is learning successfully, or what it feels like when *we* are learning successfully, are way off.

Imagine, for example, that you are learning French vocabulary. You decide to cram all your practice into one afternoon, and when you do, it feels like you have made great progress. Your friend decides instead to space practice over several days. That feels like much slower progress, and she is frustrated by how much she keeps forgetting. But if you both take the same test a week later, you may be surprised to discover that she has retained far more than you.

Or perhaps, to study a textbook chapter, you decide to re-read it. Your friend instead turns to the test at the end of the chapter and takes it. Your approach gives you a growing feeling of familiarity. Hers is painful. It is also much more effective.

Or take the order of learning: interleaving three topics—ABCABCABC—is more work than blocking—AAABBBCCC—but it reliably results in more learning.

The problem, Bjork points out, is that near-term performance feels good, but *performance* is not the same as *learning*. The sense that learning is hard may even be a good shortcut indicator of when it is actually happening.[7] Of course, no tutor likes to watch a student struggling, becoming more and more confused and frustrated. It is almost impossible to stop yourself jumping in and relieving the frustration by giving the answer, or at least a big hint. This entirely human impulse is, though, robbing the student of the very cognitive exertion that leaves learning as its residue. The harder you work, the more you will retain.

As a result, tutoring sometimes becomes a game of how much cognitive work you can get a student to do before they become demotivated. If the content is genuinely interesting to the student, or can be made so by relating it to their interests in an authentic way, or if they are motivated by some longer-term goal, the tutor's job is easier. In other situations, the tutor may need to interleave tougher content with breaks or game-like interludes.

[7] That is not to say that learning has to be hard to be effective. We are learning all the time with no outward effort at all.

Where a tutor–tutee arrangement is expected to persist for several weeks or months, it makes sense for the tutor to invest time in the relationship itself: building trust, taking an interest in the tutee's interests, helping them navigate ups and downs, and so on. There do not appear to have been long-term studies of tutor–tutee relationships to guide us here, but just knowing that solid learning demands hard work is enough to tell us that tutors need to find ways to increase student stamina for it. Plus, it is of course more pleasant for both parties to spend time with someone they like. And, perhaps most importantly, a tutee who trusts their tutor is more likely to take leaps into the unknown and risk falling, which is exactly what reveals how to make a better leap next time.

Bjork and Schmidt's finding that performance is not a good indicator of learning was focused on surface knowledge. Is the same true of deep learning? Sidney D'Mello, an expert on confusion at the University of Colorado, thinks it is. "One important form of deep learning," says D'Mello, "occurs when there is a discrepancy in the information stream and the discrepancy is identified and corrected." If there is no discrepancy, "there is no learning, at least from a perspective of conceptual change."

When researchers tracked tutee emotional states during tutoring, the most common emotion displayed was confusion. "Confusion reigns supreme during deep learning activities," says D'Mello, reporting on studies in which students worked on complex scientific concepts. And confusion was the only emotion that significantly predicted learning. Not even tutee engagement could match it.

D'Mello concludes that there is value in running lessons that "intentionally perplex learners." Of course, confusion has to be resolved for it to lead to learning. The trick is to help the student uncover the resolution for themselves, not to serve it up to them. According to D'Mello, "confusion resolution requires the individual to stop, think, engage in careful deliberation, problem solve, and revise their existing mental models." These are all examples of what Bjork calls *desirable difficulties*.

There are a wide range of ways to induce confusion. D'Mello lists "obstacles to goals, interruptions of organized action sequences, impasses, contradictions, anomalous events, dissonance, unexpected feedback, exposure of misconceptions, and general deviations from norms and expectations." In short, bumps in the road. Kurt Van Lehn calls them *impasses* that "motivate a student to take an active role in constructing a better understanding." He points out that impasses aren't only about getting stuck but include whenever a student "does an action correctly but expresses uncertainty about it."

That last point is fascinating: asking a student "Are you sure?" might create a learning moment, even if their answer was correct. Contrast that with the common practice of moving on quickly once a correct answer is arrived at. The great expert in math for little kids, Herb Ginsburg, has a wonderful technique for getting students to think harder: after every answer they give, right or wrong, simply ask them "How did you know that?"

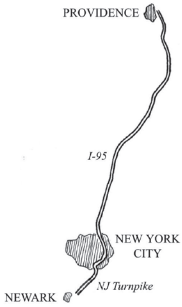
Here, then, is a way for any tutor to reliably induce deep learning. The research of Bjork, Kapur, Van Lehn, and others has given us a formula—one that will be familiar to every Hollywood screenplay writer.

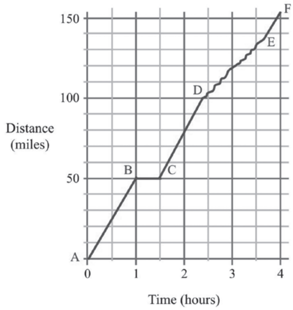
The secret to deep learning, like the secret to a good story, is (1) a conflict or impasse that leads to (2) a resolution. And in learning, as in a good story, you do not want to rush either of them. The impasse has to feel like a genuine impasse, which it won't if it is not properly established or if, from an abundance of eagerness, it is resolved too quickly. The resolution, when it comes, has to come from the actions of the main character. In learning, the main character is the student, not the teacher.

This is what makes the job of the tutor a rich and rewarding one. Crafting a problem to bring the student to an impasse, letting them dig in just deep enough, and then nudging them toward a resolution take skill and often art. When done well, it will feel to the student as if the tutor did nothing at all—except perhaps add some formalism, such as correct terminology, after the fact.

THE JOURNEY

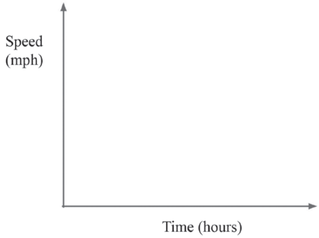
The map and graph below describe a car journey from Providence, RI to Newark, NJ using I-95 and the New Jersey Turnpike.





(i) Describe each stage of the journey, making use of the graph *and* the map. In particular describe and explain what is happening from A to B; B to C; C to D; D to E and E to F.

(ii) Using the information given above, sketch a graph to show how the speed of the car varies during the journey.



Here is an example of this impasse–resolution technique—call it *deep tutoring*—in action. A student is asked to read the first part of the task shown on the previous page (created by Malcolm Swan, a virtuoso designer of math problems at the University of Nottingham).

The student has established that segment AB represents the initial part of the journey, starting from Providence, RI.

T: *Okay, what's happening from* B *to* C*?*

S: *It shows a short distance traveled.*

T: *Okay. It looks like the line from* B *to* C *is flat. What do you think that means?*

S: *I think it means the road flattened out a bit; then, they went up another hill.*

T: *I see what you're thinking, but if* B *to* C *is flat, that means there is no distance being traveled.*

S: *Ah, okay, that makes sense.*

The student is reading the line segment BC as an illustration, not a graph. This is a common misread of graphs, especially distance–time or velocity–time graphs. In fact, it was precisely this misread that Swan was trying to uncover when he created the task. The tutor makes a very standard tutor move: they give feedback along with the correct interpretation of BC. What is the likely effect of choosing that move?

It's possible that it leads to surface learning: the student might remember "flat means no distance traveled." But it's doubtful that they have understood why (even if they think they have). And any deeper learning is unlikely to have happened here. What could the tutor have done instead?

You may want to look away from this page and come up with your own answer to that question based on the idea of impasse/resolution before reading on.

The first thing we need is an impasse. The student is reading BC as a line that depicts a flat road. But that isn't an impasse. They are not confused. They simply haven't glimpsed the power of a graph like this yet. To create an impasse for them, we need them to see an alternative to their answer without telling them it. One way to do that is to ask a content-free question such as "How do you know that?" Sometimes, that's enough for the student to realize there's a problem. In this case, though, we could prompt the student to see a contradiction.

S: *I think it means the road flattened out a bit; then, they went up another hill.*

T: *Hmm, okay. Can you tell me from the graph how far they are traveling between points* B *and* C*?*

S: *(Pause) Looks like a half.*

T: *A half what?*

S: *Er … half an hour? That can't be right.*

Bingo. Now, we wait.

T: *(Silence)*

S: *I'm not sure I get it.*

We could now jump in with an explanation. But it's worth checking that impulse and seeing if we can scaffold the student to the same conclusion, to do what they didn't do earlier: read the graph. It could be that they just don't know how to do that or that they do but they took a shortcut this time. The tutor may know which of those is true from earlier work with the student. But let's assume not.

T: Well, the y-axis tells distance.

S: Okay … So, from B *to* C *… that's zero distance.*

T: (Silence)

S: So … they didn't move.

T: For how long didn't they move?

S: For half an hour.

T: Right. So, what do you think is happening?

S: They stopped. Maybe they went to the bathroom?

T: Awesome! Who knew these charts could tell you about a bathroom visit? So, what does a flat line mean on this graph?

S: It means you are stopped.

T: Right. No distance traveled, but time passes. So, you must have stopped. Where do you think B *and* C *are on the map?*

S: Well, if he's stopped, B *and* C *must be in the same place. But I'm not sure where …*

What happened here is that the student began to produce the resolution to their own confusion. The tutor shifts to positive feedback to keep them moving. You can feel the new insight scratching like a pet at the door. The student opens it.

Do we have deep learning here? Not yet. The student doesn't truly understand what the graph is telling them about this journey in a way that would allow them to read other graphs. That's why the problem was designed with several more graph segments to make sense of. If the tutor had instead chosen a problem with just one line segment, we would lose the cumulative effect and the student's growing sense of "Oh, now I can read these graphs!"

Making sense of a distance–time graph—a crucial skill if and when the student encounters calculus, since many of the examples will be about motion—is only part of what this problem is designed to tackle. Perhaps the more important challenge is translating among three different representations of the journey: the graph, the map, and a verbal description. Let's rejoin the conversation a little later.

T: Okay, what's happening from D to E?

S: Well, the time is … an hour and a bit more. And they traveled … twenty-something miles.

T: Great. Can you figure out how that relates to the map?

S: You mean where they are?

T: Right.

S: Hmm … we don't have enough information.

T: Okay. Can you see anything different about the line on the graph from D to E?

S: Ah … it's bumpy.

T: (Silence)

S: Which doesn't mean the road is bumpy.

T: Nice.

S: Bumpy … oh, maybe, when the line is flat, they are stopped, like in B to C. And the bumpy line is flat, or nearly flat, some of the time. So, maybe, they stop and start.

T: You're on to something.

S: So, they're in New York!

T: How do you know that?

S: Because they keep stopping and starting in traffic.

This is real progress. Ultimately, we are trying to get to something along these lines:

> *They drove at 60 mph from Providence on* I-95 *for one hour, then stopped for half an hour, and then carried on for another hour before reaching New York, and so on.*

A lot has to come together to produce that description, and it is the *coming-together*—the connections—as we have seen, that characterize deep learning. Of course, that ability might not transfer—the student may stumble with the very next graph they see—but a new track has been etched in the learner's brain, ready to be deepened.

**

The practice of tutoring is inherently asymmetrical. The tutor knows something the tutee does not. That fact itself creates pressure for the tutor to tell, an osmosis in which information wants to flow from tutor to tutee. But the membrane between them is semi-permeable: surface-learning molecules get across easily, but bigger deep-learning molecules do not. The fix is not to push harder but to put the tutee in a place where they are primed to learn and then to wait. Wait for them to assemble the deep-learning molecules for themself, on their side of the membrane. Deep tutoring is a matter of getting comfortable with silence and confusion.

This is the answer to the mystery of why all tutors seem to get similar results. Aside from a few rare practitioners, they are all pushing surface-learning molecules. Figuring out how to get bigger knowledge molecules through the membrane is what makes tutoring a demanding and rewarding enterprise. But at least it is possible for an artful tutor focused on no more than a handful of tutees at one time. If there is a way the same can be achieved in a classroom of one teacher and thirty students, I have no idea what it is. Tutoring may be the only reliably effective mechanism available to us.

2 In fact, it often seems ... **it works**. White, S., Carey, M., O'Donnell, A., & Loeb, S. (2021). Early Lessons from Implementing High-Impact Tutoring at Scale. *National Student Support Accelerator* and Nickow, A., Oreopoulos, P., & Quan, V. (2020). The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence. Working Paper 27476. *National Bureau of Economic Research*.

2 **A 1982 meta-analysis**. Cohen, P. A., Kulik, J. A., & Kulik, C. L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American educational research journal, 19*(2), 237-248.

3 **Even older, in fact**. Hoppitt, W. J., Brown, G. R., Kendal, R., Rendell, L., Thornton, A., Webster, M. M., & Laland, K. N. (2008). Lessons from animal teaching. *Trends in ecology & evolution*, 23(9), 486-493.

3 **In the sense that they modify their behavior**. Caro, T. M., & Hauser, M. D. (1992). Is there teaching in nonhuman animals?. *The quarterly review of biology*, 67(2), 151-174.

3 **According to Kevin Laland**. Hoppitt, W. J., Brown, G. R., Kendal, R., Rendell, L., Thornton, A., Webster, M. M., & Laland, K. N. (2008). Lessons from animal teaching. *Trends in ecology & evolution, 23*(9), 486-493.

4 There is **strong evidence** for its remarkable effects. Slavin, R. E. (1987). Making Chapter 1 make a difference. *The Phi Delta Kappan, 69*(2), 110-119 and White, S., Carey, M., O'Donnell,

A., & Loeb, S. (2021). Early Lessons from Implementing High-Impact Tutoring at Scale. *National Student Support Accelerator*.

10 They **don't reliably lead to learning**. Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist, 43*(1), 49-64.

10 Joel Michael and Allen Rovick **analyzed transcripts**. Cho, B. I., Michael, J. A., Rovick, A. A., & Evens, M. W. (2000, June). An analysis of multiple tutoring protocols. In *International Conference on Intelligent Tutoring Systems* (pp. 212-221). Springer, Berlin, Heidelberg.

11 A **famous paper** by education researchers Dan Schwartz and John Bransford. Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and instruction, 16*(4), 475-5223.

12 She once ran **an experiment** … stop making long explanations. Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive science, 25*(4), 471-533.

15 **The explanation must be at the learner's knowledge level**. Nückles, M., Wittwer, J., & Renkl, A. (2005). Information about a layperson's knowledge supports experts in giving effective and efficient online advice to laypersons. *Journal of Experimental Psychology: Applied, 11*(4), 219.

16 **Noted** over 30 years ago by Paul Vedder. Vedder, P. (1985). *Cooperative learning: A study on processes and effects of cooperation between primary school children* (Vol. 1041). Taylor & Francis Group.

19 That **paper**, one of the most cited in all of cognitive science. Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser,

R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science, 13*(2), 145-182.

20 "**Learning from Human Tutoring**," perhaps the most in-depth study of tutoring ever conducted. Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive science, 25*(4), 471-533.

21 **Pumping** for more. Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied cognitive psychology, 9*(6), 495-522.

21 **Highlighting critical features** of the task. Stone, C. A. (1993). What is missing in the metaphor of scaffolding. *Contexts for learning: Sociocultural dynamics in children's development,* 169-183.

22 **Providing physical props or cue cards**. Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive science, 8*(2), 173-190.

22 **Describing the problem in a way that orients the student to the important features**. McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction, 7*(3), 197-244.

22 **Comparing the current problem with a previously solved problem**. McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction, 7*(3), 197-244.

22 **Maintaining focus on the goal**. Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Child Psychology & Psychiatry & Allied Disciplines*.

22 **Completing the student's reasoning step**. Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied cognitive psychology, 9*(6), 495-522.

25 Ference Marton and Roger Säljö … **noticed**. Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British journal of educational psychology, 46*(1), 4-11.

30 The process of a student constructing … has been **shown**. Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist, 49*(4), 219-243.

32 When Hattie published **the list**. Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of educational research, 66*(2), 99-136.

32 By **one estimate**, feedback can double a student's rate of learning. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112.

38 **Multiple studies** have shown that tutors are really not very good at checking for understanding. Putnam, R. T. (1987). Structuring and adjusting content for students: A study of live and simulated tutoring of addition. *American educational research journal, 24*(1), 13-48 and McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction, 7*(3), 197-244.

46 Kurt VanLehn … **writes**, rather shockingly. VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring?. *Cognition and Instruction, 21*(3), 209-249.

48 **Extrapolating** to other situations. Chan, C. K., Burtis, P. J., Scardamalia, M., & Bereiter, C. (1992). Constructive activity in learning from text. *American educational research journal, 29*(1), 97-118.

48 Coming up with **predictions**. Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction, 1*(2), 117-175.

48 **Forming analogies**. Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science, 18*(3), 439-477.

50 A **1998 experiment**. Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and instruction, 16*(4), 475-5223.

51 **Much better** on deep-learning questions. Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. J*ournal of educational psychology, 103*(4), 759.

52 Kapur even **showed** that the lecture-first approach may be damaging. Kapur, M. (2014). Productive failure in learning math. *Cognitive science, 38*(5), 1008-1022.

53 **Found**, perhaps unsurprisingly. Kapur, Manu. "Comparing learning from productive failure and vicarious failure." *Journal of the Learning Sciences* 23, no. 4 (2014): 651-677.

53 A **1992 paper** by Robert Byork and Richard Schmidt. Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological science, 3*(4), 207-218.

56 “One important form of deep learning,” says D’Mello. D’Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153-170.

57 When researchers tracked tutee emotional states during tutoring. Lehman, B., Matthews, M., D’Mello, S., & Person, N. (2008, June). What are you feeling? Investigating student affective states during expert human tutoring sessions. In *International conference on intelligent tutoring systems* (pp. 50-59). Springer, Berlin, Heidelberg.

57 Confusion was the only emotion that significantly predicted learning. D’Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153-170.

61 Malcolm Swan, a virtuoso designer of math problems. Swan, M. (1985). The Language of Functions and Graphs: An examination module for secondary schools. Shell Centre for Mathematical Education.

www.scienceoftutoring.com

$9.95
ISBN 979-8-9875380-0-5
50995>
9 798987 538005