1

Learning Is Misunderstood

EARLY IN HIS CAREER as a pilot, Matt Brown was flying a twin-engine Cessna northeast out of Harlingen, Texas, when he noticed a drop in oil pressure in his right engine. He was alone, flying through the night at eleven thousand feet, making a hotshot freight run to a plant in Kentucky that had shut down its manufacturing line awaiting product parts for assembly.

He reduced altitude and kept an eye on the oil gauge, hoping to fly as far as a planned fuel stop in Louisiana, where he could service the plane, but the pressure kept falling. Matt has been messing around with piston engines since he was old enough to hold a wrench, and he knew he had a problem. He ran a mental checklist, figuring his options. If he let the oil pressure get too low he risked the engine's seizing up. How much further could he fly before shutting it down? What would happen when he did? He'd lose lift on the right side, but could he stay aloft? He reviewed the tolerances he'd memorized for the Cessna 401. Loaded, the best you could do on one engine was slow your descent. But he had a light load, and he'd burned through most of his fuel. So he shut down the ailing right engine, feathered the prop to reduce drag, increased power on the left, flew with opposite rudder, and limped another ten miles toward his intended stop. There, he made his approach in a wide left-hand turn, for the simple but critical reason that without power on his right side it was only from a left-hand turn that he still had the lift needed to level out for a touchdown.

While we don't need to understand each of the actions Matt took, *he* certainly needed to, and his ability to work himself out of a jam illustrates what we mean in this book when we talk about learning: we mean acquiring knowledge and skills and having them readily available from memory so you can make sense of future problems and opportunities.

There are some immutable aspects of learning that we can probably all agree on:

First, to be useful, learning requires memory, so what we've learned is still there later when we need it.

Second, we need to keep learning and remembering all our lives. We can't advance through middle school without some mastery of language arts, math, science, and social studies. Getting ahead at work takes mastery of job skills and difficult colleagues. In retirement, we pick up new interests. In our dotage, we move into simpler housing while we're still able to adapt. If you're good at learning, you have an advantage in life.

Third, learning is an acquired skill, and the most effective strategies are often counterintuitive.

Claims We Make in This Book

You may not agree with the last point, but we hope to persuade you of it. Here, more or less unadorned in list form, are some of the principal claims we make in support of our argument. We set them forth more fully in the chapters that follow.

Learning is deeper and more durable when it's *effortful*. Learning that's easy is like writing in sand, here today and gone tomorrow.

We are *poor judges* of when we are learning well and when we're not. When the going is harder and slower and it doesn't feel productive, we are drawn to strategies that feel more fruitful, unaware that the gains from these strategies are often temporary.

Rereading text and *massed practice* of a skill or new knowledge are by far the preferred study strategies of learners of all stripes, but they're also among the *least productive*. By massed practice we mean the single-minded, rapid-fire repetition of something you're trying to burn into memory, the "practice-practice-practice" of conventional wisdom. Cramming for exams is an example. Rereading and massed practice give rise to feelings of fluency that are taken to be signs of mastery, but for true mastery or durability these strategies are largely a waste of time.

Retrieval practice—recalling facts or concepts or events from memory—is a more effective learning strategy than review by rereading. Flashcards are a simple example. Retrieval strengthens the memory and interrupts forgetting. A single, simple quiz after reading a text or hearing a lecture produces better learning and remembering than rereading the text or reviewing lecture notes. While the brain is not a muscle that gets stronger with exercise, the neural pathways that make up a body of learning do get stronger, when the memory is retrieved and the learning is practiced. Periodic practice arrests forgetting, strengthens retrieval routes, and is essential for hanging onto the knowledge you want to gain.

When you *space out practice* at a task and get a little rusty between sessions, or you interleave the practice of two or more subjects, retrieval is harder and feels less productive, but the effort produces longer lasting learning and enables more versatile application of it in later settings.

Trying to solve a problem *before being taught the solution* leads to better learning, even when errors are made in the attempt.

The popular notion that you learn better when you receive instruction in a form consistent with your preferred *learning style*, for example as an auditory or visual learner, is *not supported by the empirical research*. People do have multiple forms of intelligence to bring to bear on learning, and you learn better when you "go wide," drawing on all of your aptitudes and resourcefulness, than when you limit instruction or experience to the style you find most amenable.

When you're adept at extracting the *underlying principles or "rules"* that differentiate types of problems, you're more successful at picking the right solutions in unfamiliar situations. This skill is better acquired through *interleaved and varied practice* than massed practice. For instance, interleaving practice at computing the volumes of different kinds of geometric solids makes you more skilled at picking the right solution when a later test presents a random solid. Interleaving the identification of bird types or the works of oil painters improves your ability both to learn the unifying attributes within a type and to differentiate between types, improving your skill at categorizing new specimens you encounter later.

We're all *susceptible to illusions* that can hijack our judgment of what we know and can do. Testing helps calibrate

our judgments of what we've learned. A pilot who is responding to a failure of hydraulic systems in a flight simulator discovers quickly whether he's on top of the corrective procedures or not. In virtually all areas of learning, you build better mastery when you use testing as a tool to identify and bring up your areas of weakness.

All new learning requires a *foundation of prior knowledge*. You need to know how to land a twin engine plane on two engines before you can learn to land it on one. To learn trigonometry, you need to remember your algebra and geometry. To learn cabinetmaking, you need to have mastered the properties of wood and composite materials, how to join boards, cut rabbets, rout edges, and miter corners.

In a cartoon by the *Far Side* cartoonist Gary Larson, a bug-eyed school kid asks his teacher, "Mr. Osborne, can I be excused? My brain is full!" If you're just engaging in mechanical repetition, it's true, you quickly hit the limit of what you can keep in mind. However, if you practice *elaboration*, there's no known limit to how much you can learn. Elaboration is the process of giving new material meaning by expressing it in your own words and connecting it with what you already know. The more you can explain about the way your new learning relates to your prior knowledge, the stronger your grasp of the new learning will be, and the more connections you create that will help you remember it later. Warm air can hold more moisture than cold air; to know that this is true in your own experience, you can think of the drip of water from the back of an air conditioner or the way a stifling summer day turns cooler out the back side of a sudden thunderstorm. Evaporation has a cooling effect: you know this because a humid day at your uncle's in Atlanta feels hotter than a dry one at your cousin's in Phoenix, where your sweat disappears even before your skin feels damp. When you study the principles of heat transfer, you understand conduction from warming your hands around a hot cup of cocoa; radiation from the way the sun pools in the den on a wintry day; convection from the life-saving blast of A/C as your uncle squires you slowly through his favorite back alley haunts of Atlanta.

Putting new knowledge into a *larger context* helps learning. For example, the more of the unfolding story of history you know, the more of it you can learn. And the more ways you give that story meaning, say by connecting it to your understanding of human ambition and the untidiness of fate, the better the story stays with you. Likewise, if you're trying to learn an abstraction, like the principle of angular momentum, it's easier when you ground it in something concrete that you already know, like the way a figure skater's rotation speeds up as she draws her arms to her chest.

People who learn to *extract the key ideas from new material and organize them into a mental model* and connect that model to prior knowledge show an advantage in learning complex mastery. A mental model is a mental representation of some external reality.[1] Think of a baseball batter waiting for a pitch. He has less than an instant to decipher whether it's a curveball, a changeup, or something else. How does he do it? There are a few subtle signals that help: the way the pitcher winds up, the way he throws, the spin of the ball's seams. A great batter winnows out all the extraneous perceptual distractions, seeing only these variations in pitches, and through practice he forms distinct mental models based on a different set of cues for each kind of pitch. He connects these models to what he knows about batting stance, strike zone, and swinging so as to stay on top of the ball. These he connects to mental models of player positions: if he's got guys on first and second, maybe he'll sacrifice to move the runners ahead. If he's got men on first and third and there is one out, he's got to keep from hitting into a double play while still hitting to score the runner. His mental models of player positions connect to his models of the opposition (are they playing deep or shallow?) and to the signals flying around from the dugout to the base coaches to him. In a great at-bat, all these pieces come together seamlessly: the batter connects with the ball and drives it through a hole in the outfield, buying the time to get on first and advance his men. Because he has culled out all but the most important elements for identifying and responding to each kind of pitch, constructed mental models out of that learning, and connected those models to his mastery of the other essential elements of this complex game, an expert player has a better chance of scoring runs than a less experienced one who cannot make sense of the vast and changeable information he faces every time he steps up to the plate.

Many people believe that their intellectual ability is hardwired from birth, and that failure to meet a learning challenge is an indictment of their native ability. But every time you learn something new, *you change the brain*—the residue of your experiences is stored. It's true that we start life with the gift of our genes, but it's also true that we become capable through the learning and development of mental models that enable us to reason, solve, and create. In other words, the elements that shape your intellectual abilities lie to a surprising extent within your own control. Understanding that this is so enables you to see failure as a badge of effort and a source of useful information—the need to dig deeper or to try a different strategy. The need to understand that when learning is hard, you're doing important work. To understand that striving and setbacks, as in any action video game or new BMX bike stunt, are essential if you are to surpass your current level of performance toward true expertise. Making mistakes and correcting them builds the bridges to advanced learning.

Empirical Evidence versus Theory, Lore, and Intuition

Much of how we structure training and schooling is based on learning theories that have been handed down to us, and these are shaped by our own sense of what works, a sensibility drawn from our personal experiences as teachers, coaches, students, and mere humans at large on the earth. How we teach and study is largely a mix of theory, lore, and intuition. But over the last forty years and more, cognitive psychologists have been working to build a body of evidence to clarify what works and to discover the strategies that get results.

Cognitive psychology is the basic science of understanding how the mind works, conducting empirical research into how people perceive, remember, and think. Many others have their hands in the puzzle of learning as well. Developmental and educational psychologists are concerned with theories of human development and how they can be used to shape the tools of education—such as testing regimes, instructional organizers (for example topic outlines and schematic illustrations), and resources for special groups like those in remedial and gifted education. Neuroscientists, using new imaging techniques and other tools, are advancing our understanding of brain mechanisms that underlie learning, but we're still a very long way from knowing what neuroscience will tell us about how to improve education.

How is one to know whose advice to take on how best to go about learning?

It's wise to be skeptical. Advice is easy to find, only a few mouse-clicks away. Yet not all advice is grounded in research—far from it. Nor does all that passes as research meet the standards of science, such as having appropriate control conditions to assure that the results of an investigation are objective and generalizable. The best empirical studies are experimental in nature: the researcher develops a hypothesis and then tests it through a set of experiments that must meet rigorous criteria for design and objectivity. In the chapters that follow, we have distilled the findings of a large body of such studies that have stood up under review by the scientific community before being published in professional journals. We are collaborators in some of these studies, but not the lion's share. Where we're offering theory rather than scientifically validated results, we say so. To make our points we use, in addition to tested science, anecdotes from people like Matt Brown whose work requires mastery of complex knowledge and skills, stories that illustrate the underlying principles of how we learn and remember. Discussion of the research studies themselves is kept to a minimum, but you will find many of them cited in the notes at the end of the book if you care to dig further.

People Misunderstand Learning

It turns out that much of what we've been doing as teachers and students isn't serving us well, but some comparatively simple changes could make a big difference. People commonly believe that if you expose yourself to something enough times—say, a textbook passage or a set of terms from an eighth grade biology class—you can burn it into memory. Not so. Many teachers believe that if they can make learning easier and faster, the learning will be better. Much research turns this belief on its head: when learning is harder, it's stronger and lasts longer. It's widely believed by teachers, trainers, and coaches that the most effective way to master a new skill is to give it dogged, single-minded focus, practicing over and over until you've got it down. Our faith in this runs deep, because most of us see fast gains during the learning phase of massed practice. What's apparent from the research is that gains achieved during massed practice are transitory and melt away quickly.

The finding that rereading textbooks is often labor in vain ought to send a chill up the spines of educators and learners, because it's the number one study strategy of most people—including more than 80 percent of college students in some surveys—and is central in what we tell ourselves to do during the hours we dedicate to learning. Rereading has three strikes against it. It is time consuming. It doesn't result in durable memory. And it often involves a kind of unwitting self-deception, as growing familiarity with the text comes to feel like mastery of the content. The hours immersed in rereading can seem like due diligence, but the amount of study time is no measure of mastery.[2]

You needn't look far to find training systems that lean heavily on the conviction that mere exposure leads to learning. Consider Matt Brown, the pilot. When Matt was ready to advance from piston planes, he had a whole new body of knowledge to master in order to get certified for the business jet he was hired to pilot. We asked him to describe this process. His employer sent him to eighteen days of training, ten hours a day, in what Matt called the "fire hose" method of instruction. The first seven days straight were spent in the classroom being instructed in all the plane's systems: electrical, fuel, pneumatics, and so on, how these systems operated and interacted, and all their fail-safe tolerances like pressures, weights, temperatures, and speeds. Matt is required to have at his immediate command about eighty different "memory action items"—actions to take without hesitation or thought in order to stabilize the plane the moment any one of a dozen or so unexpected events occur. It might be a sudden decompression, a thrust reverser coming unlocked in flight, an engine failure, an electrical fire.

Matt and his fellow pilots gazed for hours at mind-numbing PowerPoint illustrations of their airplane's principal systems. Then something interesting happened.

"About the middle of day five," Matt said, "they flash a schematic of the fuel system on the screen, with its pressure sensors, shutoff valves, ejector pumps, bypass lines, and on and on, and you're struggling to stay focused. Then this one instructor asks us, 'Has anybody here had the fuel filter bypass light go on in flight?' This pilot across the room raises his hand. So the instructor says, 'Tell us what happened,' and suddenly you're thinking, Whoa, what if that was me?

"So, this guy was at 33,000 feet or something and he's about to lose both engines because he got fuel without antifreeze in it and his filters are clogging with ice. You hear that story and, believe me, that schematic comes to life and sticks with you. Jet fuel can commonly have a little water in it, and when it gets cold at high altitude, the water will condense out, and it can freeze and block the line. So whenever you refuel, you make good and sure to look for a sign on the fuel truck saying the fuel has Prist in it, which is an antifreeze. And if you ever see that light go on in flight, you're going to get yourself down to some warmer air in a hurry."[3] Learning is stronger when it matters, when the abstract is made concrete and personal.

Then the nature of Matt's instruction shifted. The next eleven days were spent in a mix of classroom and flight simulator training. Here, Matt described the kind of active engagement that leads to durable learning, as the pilots had to grapple with their aircraft to demonstrate mastery of standard operating procedures, respond to unexpected situations, and drill on the rhythm and physical memory of the movements that are required in the cockpit for dealing with them. A flight simulator provides retrieval practice, and the practice is spaced, interleaved, and varied and involves as far as possible the same mental processes Matt will invoke when he's at altitude. In a simulator, the abstract is made concrete and personal. A simulator is also a series of tests, in that it helps Matt and his instructors calibrate their judgment of where he needs to focus to bring up his mastery.

In some places, like Matt Brown's flight simulator, teachers and trainers have found their way to highly effective learning techniques, yet in virtually any field, these techniques tend to be the exception, and "fire hose" lectures (or their equivalent) are too often the norm.

In fact, what students are advised to do is often plain wrong. For instance, study tips published on a website at George Mason University include this advice: "The key to learning something well is repetition; the more times you go over the material the better chance you have of storing it permanently."[4] Another, from a Dartmouth College website, suggests: "If you intend to remember something, you probably will."[5] A public service piece that runs occasionally in the *St. Louis Post-Dispatch* offering study advice shows a kid with his nose buried in a book. "Concentrate," the caption reads. "Focus on one thing and one thing only. Repeat, repeat, repeat! Repeating what you have to remember can help burn it into your memory."[6] Belief in the power of rereading, intentionality, and repetition is pervasive, but the truth is you usually can't embed something in memory simply by repeating it over and over. This tactic might work when looking up a phone number and holding it in your mind while punching it into your phone, but it doesn't work for durable learning.

A simple example, reproduced on the Internet (search "penny memory test"), presents a dozen different images of a common penny, only one of which is correct. As many times as you've seen a penny, you're hard pressed to say with confidence which one it is. Similarly, a recent study asked faculty and students who worked in the Psychology Building at UCLA to identify the fire extinguisher closest to their office. Most failed the test. One professor, who had been at UCLA for twenty-five years, left his safety class and decided to look for the fire extinguisher closest to his office. He discovered that it was actually right next to his office door, just inches from the doorknob he turned every time he went into his office. Thus, in this case, even years of repetitive exposure did not result in his learning where to grab the closest extinguisher if his wastebasket caught fire.[7]

Early Evidence

The fallacy in thinking that repetitive exposure builds memory has been well established through a series of investigations going back to the mid-1960s, when the psychologist Endel Tulving at the University of Toronto began testing people on their ability to remember lists of common English nouns. In a first phase of the experiment, the participants simply read a list of paired items six times (for example, a pair on the list might be "chair—9"); they did not expect a memory test. The first item in each pair was always a noun. After reading the listed pairs six times, participants were then told that they would be getting a list of nouns that they would be asked to remember. For one group of people, the nouns were the same ones they had just read six times in the prior reading phase; for another group, the nouns to be learned were different from those they had previously read. Remarkably, Tulving found that the two groups' learning of the nouns did not differ—the learning curves were statistically indistinguishable. Intuition would suggest otherwise, but prior exposure did not aid later recall. Mere repetition did not enhance learning. Subsequent studies by many researchers have pressed further into questions of whether repeated exposure or longer periods of holding an idea in mind contribute to later recall, and these studies have confirmed and elaborated on the findings that repetition by itself does not lead to good long-term memory.[8]

These results led researchers to investigate the benefits of rereading texts. In a 2008 article in *Contemporary Educational Psychology*, Washington University scientists reported on a series of studies they conducted at their own school and at the University of New Mexico to shed light on rereading as a strategy to improve understanding and memory of prose. Like most research, these studies stood on the shoulders of earlier work by others; some showed that when the same text is read multiple times the same inferences are made and the same connections between topics are formed, and others suggested modest benefits from rereading. These benefits had been found in two different situations. In the first, some students read and immediately reread study material, whereas other students read the material only once. Both groups took an immediate test after reading, and the group who had read twice performed a bit better than the group who had read once. However, on a delayed test the benefit of immediate rereading had worn off, and the rereaders performed at the same level as the one-time readers. In the other situation, students read the material the first time and then waited some days before they reread it. This group, having done spaced readings of the text, performed better on the test than the group who did not reread the material.[9]

Subsequent experiments at Washington University, aimed at teasing apart some of the questions the earlier studies had raised, assessed the benefits of rereading among students of

differing abilities, in a learning situation paralleling that faced by students in classes. A total of 148 students read five different passages taken from textbooks and *Scientific American*. The students were at two different universities; some were high-ability readers, and others were low-ability; some students read the material only once, and others read it twice in succession. Then all of them responded to questions to demonstrate what they had learned and remembered.

In these experiments, multiple readings in close succession did not prove to be a potent study method for either group, at either school, in any of the conditions tested. In fact, the researchers found no rereading benefit at all under these conditions.

What's the conclusion? It makes sense to reread a text once if there's been a meaningful lapse of time since the first reading, but doing multiple readings in close succession is a time-consuming study strategy that yields negligible benefits at the expense of much more effective strategies that take less time. Yet surveys of college students confirm what professors have long known: highlighting, underlining, and sustained poring over notes and texts are the most-used study strategies, by far.[10]

Illusions of Knowing

If rereading is largely ineffective, why do students favor it? One reason may be that they're getting bad study advice. But there's another, subtler way they're pushed toward this method of review, the phenomenon mentioned earlier: rising familiarity with a text and fluency in reading it can create an illusion of mastery. As any professor will attest, students work hard to capture the precise wording of phrases they hear in class lectures, laboring under the misapprehension that the essence of the subject lies in the syntax in which it's described. Mastering the lecture or the text is not the same as mastering the ideas behind them. However, repeated reading provides the illusion of mastery of the underlying ideas. Don't let yourself be fooled. The fact that you can repeat the phrases in a text or your lecture notes is no indication that you understand the significance of the precepts they describe, their application, or how they relate to what you already know about the subject.

Too common is the experience of a college professor answering a knock on her office door only to find a first-year student in distress, asking to discuss his low grade on the first test in introductory psychology. How is it possible? He attended all the lectures and took diligent notes on them. He read the text and highlighted the critical passages.

How did he study for the test? she asks.

Well, he'd gone back and highlighted his notes, and then reviewed the highlighted notes and his highlighted text material several times until he felt he was thoroughly familiar with all of it. How could it be that he had pulled a D on the exam?

Had he used the set of key concepts in the back of each chapter to test himself? Could he look at a concept like "conditioned stimulus," define it, and use it in a paragraph? While he was reading, had he thought of converting the main points of the text into a series of questions and then later tried to answer them while he was studying? Had he at least rephrased the main ideas in his own words as he read? Had he tried to relate them to what he already knew? Had he looked for examples outside the text? The answer was no in every case.

He sees himself as the model student, diligent to a fault, but the truth is he doesn't know how to study effectively.

The illusion of mastery is an example of poor metacognition: what we know about what we know. Being accurate in your judgment of what you know and don't know is critical

for decision making. The problem was famously (and prophetically) summed up by Secretary of State Donald Rumsfeld in a 2002 press briefing about US intelligence on Iraq's possible possession of weapons of mass destruction: "There are known knowns; there are things we know that we know. There are known unknowns; that is to say, there are things that we now know we don't know. But there are also unknown unknowns—*there are things we do not know we don't know.*"

The emphasis here is ours. We make it to drive home the point that students who don't quiz themselves (and most do not) tend to overestimate how well they have mastered class material. Why? When they hear a lecture or read a text that is a paragon of clarity, the ease with which they follow the argument gives them the feeling that they already know it and don't need to study it. In other words, they tend not to know what they don't know; when put to the test, they find they cannot recall the critical ideas or apply them in a new context. Likewise, when they've reread their lecture notes and texts to the point of fluency, their fluency gives them the false sense that they're in possession of the underlying content, principles, and implications that constitute real learning, confident that they can recall them at a moment's notice. The upshot is that even the most diligent students are often hobbled by two liabilities: a failure to know the areas where their learning is weak—that is, where they need to do more work to bring up their knowledge—and a preference for study methods that create a false sense of mastery.[11]

Knowledge: Not Sufficient, but Necessary

Albert Einstein declared "creativity is more important than knowledge," and the sentiment appears to be widely shared by college students, if their choice in t-shirt proclamations is any indication. And why wouldn't they seize on the sentiment? It embodies an obvious and profound truth, for without creativity where would our scientific, social, or economic breakthroughs come from? Besides which, accumulating knowledge can feel like a grind, while creativity sounds like a lot more fun. But of course the dichotomy is false. You wouldn't want to see that t-shirt on your neurosurgeon or on the captain who's flying your plane across the Pacific. But the sentiment has gained some currency as a reaction to standardized testing, fearing that this kind of testing leads to an emphasis on memorization at the expense of high-level skills. Notwithstanding the pitfalls of standardized testing, what we really ought to ask is how to do better at building knowledge *and* creativity, for without knowledge you don't have the foundation for the higher-level skills of analysis, synthesis, and creative problem solving. As the psychologist Robert Sternberg and two colleagues put it, "one cannot apply what one knows in a practical manner if one does not know anything to apply."[12]

Mastery in any field, from cooking to chess to brain surgery, is a gradual accretion of knowledge, conceptual understanding, judgment, and skill. These are the fruits of variety in the practice of new skills, and of striving, reflection, and mental rehearsal. Memorizing facts is like stocking a construction site with the supplies to put up a house. Building the house requires not only knowledge of countless different fittings and materials but conceptual understanding, too, of aspects like the load-bearing properties of a header or roof truss system, or the principles of energy transfer and conservation that will keep the house warm but the roof deck cold so the owner doesn't call six months later with ice dam problems. Mastery requires both the possession of ready knowledge and the conceptual understanding of how to use it.

When Matt Brown had to decide whether or not to kill his right engine he was problem solving, and he needed to know from memory the procedures for flying with a dead engine and the tolerances of his plane in order to predict whether he would fall out of the air or be unable to straighten up for landing. The would-be neurosurgeon in her first year of med school has to memorize the whole nervous system, the whole skeletal system, the whole muscular system, the humeral system. If she can't, she's not going to be a neurosurgeon. Her success will depend on diligence, of course, but also on finding study strategies that will enable her to learn the sheer volume of material required in the limited hours available.

Testing: Dipstick versus Learning Tool

There are few surer ways to raise the hackles of many students and educators than talking about testing. The growing focus over recent years on standardized assessment, in particular, has turned testing into a lightning rod for frustration over how to achieve the country's education goals. Online forums and news articles are besieged by readers who charge that emphasis on testing favors memorization at the expense of a larger grasp of context or creative ability; that testing creates extra stress for students and gives a false measure of ability; and so on. But if we stop thinking of testing as a dipstick to measure learning—if we think of it as practicing retrieval of learning from memory rather than "testing," we open ourselves to another possibility: *the use of testing as a tool for learning.*

One of the most striking research findings is the power of active retrieval—testing—to strengthen memory, and that the more effortful the retrieval, the stronger the benefit. Think flight simulator versus PowerPoint lecture. Think quiz versus rereading. The act of retrieving learning from memory has two profound benefits. One, it tells you what you know and don't know, and therefore where to focus further study to improve the areas where you're weak. Two, recalling what you have learned causes your brain to reconsolidate the memory, which strengthens its connections to what you already know and makes it easier for you to recall in the future. In effect, retrieval—testing—interrupts forgetting. Consider an eighth grade science class. For the class in question, at a middle school in Columbia, Illinois, researchers arranged for part of the material covered during the course to be the subject of low-stakes quizzing (with feedback) at three points in the semester. Another part of the material was never quizzed but was studied three times in review. In a test a month later, which material was better recalled? The students averaged A- on the material that was quizzed and C+ on the material that was not quizzed but reviewed.[13]

In Matt Brown's case, even after ten years piloting the same business jet, his employer reinforces his mastery every six months in a battery of tests and flight simulations that require him to retrieve the information and maneuvers that are essential to stay in control of his plane. As Matt points out, you hardly ever have an emergency, so if you don't practice what to do, there's no way to keep it fresh.

Both of these cases—the research in the classroom and the experience of Matt Brown in updating his knowledge—point to the critical role of retrieval practice in keeping our knowledge accessible to us when we need it. The power of active retrieval is the topic of Chapter 2.[14]

The Takeaway

For the most part, we are going about learning in the wrong ways, and we are giving poor advice to those who are coming up behind us. A great deal of what we think we know about how to learn is taken on faith and based on intuition but does not hold up under empirical research. Persistent illusions of knowing lead us to labor at unproductive strategies; as recounted in Chapter 3, this is true even of people who have participated in empirical studies and seen the evidence for themselves, firsthand. Illusions are potent persuaders. One of the best habits a learner can instill in herself is regular self-quizzing to recalibrate her understanding of what she does and does not know. Second Lieutenant Kiley Hunkler, a 2013 graduate of West Point and winner of a Rhodes Scholarship, whom we write about in Chapter 8, uses the phrase "shooting an azimuth" to describe how she takes practice tests to help refocus her studying. In overland navigation, shooting an azimuth means climbing to a height, sighting an object on the horizon in the direction you're traveling, and adjusting your compass heading to make sure you're still gaining on your objective as you beat through the forest below.

The good news is that we now know of simple and practical strategies that anybody can use, at any point in life, to learn better and remember longer: various forms of retrieval practice, such as low-stakes quizzing and self-testing, spacing out practice, interleaving the practice of different but related topics or skills, trying to solve a problem before being taught the solution, distilling the underlying principles or rules that differentiate types of problems, and so on. In the chapters that follow we describe these in depth. And because learning is an iterative process that requires that you revisit what you have learned earlier and continually update it and connect it with new knowledge, we circle through these topics several times along the way. At the end, in Chapter 8, we pull it all together with specific tips and examples for putting these tools to work.

2

To Learn, Retrieve

MIKE EBERSOLD GOT CALLED into a hospital emergency room one afternoon late in 2011 to examine a Wisconsin deer hunter who'd been found lying unconscious in a cornfield. The man had blood at the back of his head, and the men who'd found and brought him in supposed he'd maybe stumbled and cracked his skull on something.

Ebersold is a neurosurgeon. The injury had brain protruding, and he recognized it as a gunshot wound. The hunter regained consciousness in the ER, but when asked how he'd hurt himself, he had no idea.

Recounting the incident later, Ebersold said, "Somebody from some distance away must have fired what appeared to be a 12-gauge shotgun, which arced over God only knows what distance, hit this guy in the back of his head, fractured his skull, and lodged into the brain about an inch. It must have been pretty much spent, or it would have gone deeper."[1]

Ebersold is tall, slender, and counts among his forebears the Dakota chiefs named Wapasha and the French fur traders named Rocque who populated this part of the Mississippi River Valley where the Mayo brothers would later found their famous clinic. Ebersold's formal training included four years of college, four years of medical school, and seven years of neurosurgery training—building a foundation of knowledge and skills that has been broadened and deepened through continuing medical education classes, consultations with his colleagues, and his practice at the Mayo Clinic and elsewhere. He carries himself with a midwestern modesty that belies a career that counts a long list of high-profile patients who have sought out his services. When President Ronald Reagan needed treatment for injuries after a fall from his horse, Ebersold participated in the surgery and postsurgical care. When Sheikh Zayed bin Sultan Al Nahyan, president of the United Arab Emirates, needed delicate spinal repair, he and what seemed like half the nation's ministry and security forces settled in Rochester while Mike Ebersold made the repair and oversaw Zayed's recovery. Following a long career at Mayo, Mike had returned to help out at the clinic in Wisconsin, feeling indebted to it for his early medical training. The hunter whose bad luck put him in the way of an errant 12-gauge slug was luckier than he likely knows that Mike was on the job that day.

The bullet had entered an area of the skull beneath which there is a large venous sinus, a soft-tissue channel that drains the brain cavity. As he examined the hunter, Ebersold knew from experience that when he opened up the wound, there was a high probability he would find this vein was torn. As he described it,

> You say to yourself, "This patient is going to need surgery. There's brain coming out of the wound. We have to clean this up and repair this as best we can, but in so doing we may get into this big vein and that could be very, very serious." So you go through the checklist. You say, "I might need a blood transfusion for this patient," so you set up some blood. You review the steps, A, B, C, and D. You set up the operating room, telling them ahead of time what you might be encountering. All of this is sort of protocol, pretty much like a cop getting ready to pull over a car, you know what the book says, you've gone through all these steps.
>
> Then you get to the operating room, and now you're still in this mode where you have time to think through it. You say, "Gee, I don't want to just go and pull that bullet out if there might be major bleeding. What I'll try to do is I'll work around the edges and get things freed up so I'm ready for what could go wrong, and then I'll pull it out."

It turned out that the bullet and bone were lodged in the vein, serving as plugs, another lucky turn for the hunter. If the wound hadn't corked itself in the field, he would not have lived for more than two or three minutes. When Ebersold removed the bullet, the fractured bone chips fell away, and the vein let loose in a torrent. "Within five minutes, you've lost two or so units of blood and now you sort of transfer out of the mode where you're thinking through this, going through the options. Now it becomes reflex, mechanical. You know it's going to bleed very, very much, so you have a very short time. You're just thinking, 'I have to get a suture around this structure, and I know from previous experience I have to do it in this particular way.' "

The vein in question, which is about the size of an adult's small finger, was torn in several places over a distance of about an inch and a half. It needed to be tied off above and below the rupture, but it's a flat structure that he knows well: you can't just put a stitch around it, because when you tighten it, the tissue tears, and the ligature leaks. Working urgently and mechanically, he fell back on a technique he'd developed out of necessity in past surgeries involving this vein. He cut two little pieces of muscle, from where the patient's skin had been opened up in surgery, and imported them to the site and stitched the ends of the torn vein to them. These plugs of muscle served to close the vein without deflecting its natural shape or tearing its tissue. It's a solution Mike has taught himself—one he says you won't find written anywhere, but handy in the moment, to say the least. In the sixty or so seconds it took to do, the patient lost another two hundred cubic centimeters of blood, but once the plugs were in place, the bleeding stopped. "Some people can't tolerate this sinus vein being closed off. They get increased brain pressure because the blood doesn't drain properly. But this patient was one of the fortunate who can." The hunter left the hospital a week later. He was minus some peripheral vision but otherwise remarkably unscathed from a very close brush with mortality.

Reflection Is a Form of Practice

What inferences can we draw from this story about how we learn and remember? In neurosurgery (and, arguably, in all aspects of life from the moment you leave the womb), there's an essential kind of learning that comes from reflection on personal experience. Ebersold described it this way:

> A lot of times something would come up in surgery that I had difficulty with, and then I'd go home that night thinking about what happened and what could I do, for example, to improve the way a suturing went. How can I take a bigger bite with my needle, or a smaller bite, or should the stitches be closer together? What if I modified it this way or that way? Then the next day back, I'd try that and see if it worked better. Or even if it wasn't the next day, at least I've thought through this, and in so doing I've not only revisited things that I learned from lectures or from watching others performing surgery but also I've complemented that by adding something of my own to it that I missed during the teaching process.

Reflection can involve several cognitive activities that lead to stronger learning: retrieving knowledge and earlier training from memory, connecting these to new experiences, and visualizing and mentally rehearsing what you might do differently next time.

It was this kind of reflection that originally had led Ebersold to try a new technique for repairing the sinus vein at the back of the head, a technique he practiced in his mind and in the operating room until it became the kind of reflexive maneuver you can depend on when your patient is spouting blood at two hundred cubic centimeters a minute.

To make sure the new learning is available when it's needed, Ebersold points out, "you memorize the list of things that you need to worry about in a given situation: steps A, B, C, and D," and you drill on them. Then there comes a time when you get into a tight situation and it's no longer a matter of thinking through the steps, it's a matter of reflexively taking the correct action. "Unless you keep recalling this maneuver, it will not become a reflex. Like a race car driver in a tight situation or a quarterback dodging a tackle, you've got to act out of reflex before you've even had time to think. Recalling it over and over, practicing it over and over. That's just so important."

The Testing Effect

A child stringing cranberries on a thread goes to hang them on the tree, only to find they've slipped off the other end. Without the knot, there's no making a string. Without the knot there's no necklace, there's no beaded purse, no magnificent tapestry. Retrieval ties the knot for memory. Repeated retrieval snugs it up and adds a loop to make it fast.

Since as far back as 1885, psychologists have been plotting "forgetting curves" that illustrate just how fast our cranberries slip off the string. In very short order we lose something like 70 percent of what we've just heard or read. After that, forgetting begins to slow, and the last 30 percent or so falls away more slowly, but the lesson is clear: a central challenge to improving the way we learn is finding a way to interrupt the process of forgetting.[2]

The power of retrieval as a learning tool is known among psychologists as the testing effect. In its most common form, testing is used to measure learning and assign grades in school, but we've long known that the act of retrieving knowledge from memory has the effect of making that knowledge easier to call up again in the future. In his essay on memory, Aristotle wrote: "exercise in repeatedly recalling a thing strengthens the memory." Francis Bacon wrote about this phenomenon, as did the psychologist William James. Today, we know from empirical research that practicing retrieval makes learning stick far better than reexposure to the original material does. This is the testing effect, also known as the retrieval-practice effect.[3]

To be most effective, retrieval must be repeated again and again, in spaced out sessions so that the recall, rather than becoming a mindless recitation, requires some cognitive effort. Repeated recall appears to help memory consolidate into a cohesive representation in the brain and to strengthen and

multiply the neural routes by which the knowledge can later be retrieved. In recent decades, studies have confirmed what Mike Ebersold and every seasoned quarterback, jet pilot, and teenaged texter knows from experience—that repeated retrieval can so embed knowledge and skills that they become reflexive: the brain acts before the mind has time to think.

Yet despite what research and personal experience tell us about the power of testing as a learning tool, teachers and students in traditional educational settings rarely use it as such, and the technique remains little understood or utilized by teachers or students as a learning tool in traditional educational settings. Far from it.

In 2010 the *New York Times* reported on a scientific study that showed that students who read a passage of text and then took a test asking them to recall what they had read retained an astonishing 50 percent more of the information a week later than students who had not been tested. This would seem like good news, but here's how it was greeted in many online comments:

> "Once again, another author confuses learning with recalling information."

> "I personally would like to avoid as many tests as possible, especially with my grade on the line. Trying to learn in a stressful environment is no way to help retain information."

> "Nobody should care whether memorization is enhanced by practice testing or not. Our children cannot *do* much of anything anymore."[4]

Forget memorization, many commenters argued; education should be about high-order skills. Hmmm. If memorization is irrelevant to complex problem solving, don't tell your neurosurgeon. The frustration many people feel toward standardized, "dipstick" tests given for the sole purpose of measuring learning is understandable, but it steers us away from appreciating one of the most potent learning tools available to us. Pitting the learning of basic knowledge against the development of creative thinking is a false choice. Both need to be cultivated. The stronger one's knowledge about the subject at hand, the more nuanced one's creativity can be in addressing a new problem. Just as knowledge amounts to little without the exercise of ingenuity and imagination, creativity absent a sturdy foundation of knowledge builds a shaky house.

Studying the Testing Effect in the Lab

The testing effect has a solid pedigree in empirical research. The first large-scale investigation was published in 1917. Children in grades 3, 5, 6, and 8 studied brief biographies from *Who's Who in America*. Some of them were directed to spend varying lengths of the study time looking up from the material and silently reciting to themselves what it contained. Those who did not do so simply continued to reread the material. At the end of the period, all the children were asked to write down what they could remember. The recall test was repeated three to four hours later. All the groups who had engaged in the recitation showed better retention than those who had not done so but had merely continued to review the material. The best results were from those spending about 60 percent of the study time in recitation.

A second landmark study, published in 1939, tested over three thousand sixth graders across Iowa. The kids studied six-hundred-word articles and then took tests at various times before a final test two months later. The experiment showed a couple of interesting results: the longer the first test was delayed, the greater the forgetting, and second, once a student

had taken a test, the forgetting nearly stopped, and the student's score on subsequent tests dropped very little.[5]

Around 1940, interest turned to the study of forgetting, and investigating the potential of testing as a form of retrieval practice and as a learning tool fell out of favor. So did the use of testing as a research tool: since testing interrupts forgetting, you can't use it to measure forgetting because that "contaminates" the subject.

Interest in the testing effect resurfaced in 1967 with the publication of a study showing that research subjects who were presented with lists of thirty-six words learned as much from repeated testing after initial exposure to the words as they did from repeated studying. These results—that testing led to as much learning as studying did—challenged the received wisdom, turned researchers' attention back to the potential of testing as a learning tool, and stimulated a boomlet in testing research.

In 1978, researchers found that massed studying (cramming) leads to higher scores on an immediate test but results in faster forgetting compared to practicing retrieval. In a second test two days after an initial test, the crammers had forgotten 50 percent of what they had been able to recall on the initial test, while those who had spent the same period practicing retrieval instead of studying had forgotten only 13 percent of the information recalled initially.

A subsequent study was aimed at understanding what effect taking multiple tests would have on subjects' long-term retention. Students heard a story that named sixty concrete objects. Those students who were tested immediately after exposure recalled 53 percent of the objects on this initial test but only 39 percent a week later. On the other hand, a group of students who learned the same material but were not tested at all until a week later recalled 28 percent. Thus, taking a single test boosted performance by 11 percent after a week.

But what effect would three immediate tests have relative to one? Another group of students were tested three times after initial exposure and a week later they were able to recall 53 percent of the objects—the same as on the initial test for the group receiving one test. In effect, the group that received three tests had been "immunized" against forgetting, compared to the one-test group, and the one-test group remembered more than those who had received no test immediately following exposure. Thus, and in agreement with later research, multiple sessions of retrieval practice are generally better than one, especially if the test sessions are spaced out.[6]

In another study, researchers showed that simply asking a subject to fill in a word's missing letters resulted in better memory of the word. Consider a list of word pairs. For a pair like *foot-shoe*, those who studied the pair intact had lower subsequent recall than those who studied the pair from a clue as obvious as *foot-s_ _e*. This experiment was a demonstration of what researchers call the "generation effect." The modest effort required to generate the cued answer while studying the pairs strengthened memory of the target word tested later (shoe). Interestingly, this study found that the ability to recall the word pair on later tests was greater if the practice retrieval was delayed by twenty intervening word pairs than when it came immediately after first studying the pair.[7] Why would that be? One argument suggested that the greater effort required by the delayed recall solidified the memory better. Researchers began to ask whether the schedule of testing mattered.

The answer is yes. When retrieval practice is spaced, allowing some forgetting to occur between tests, it leads to stronger long-term retention than when it is massed.

Researchers began looking for opportunities to take their inquiries out of the lab and into the classroom, using the kinds of materials students are required to learn in school.

Studying the Testing Effect "In the Wild"

In 2005, we and our colleagues approached Roger Chamberlain, the principal of a middle school in nearby Columbia, Illinois, with a proposition. The positive effects of retrieval practice had been demonstrated many times in controlled laboratory settings but rarely in a regular classroom setting. Would the principal, teachers, kids, and parents of Columbia Middle School be willing subjects in a study to see how the testing effect would work "in the wild"?

Chamberlain had concerns. If this was just about memorization, he wasn't especially interested. His aim is to raise the school's students to higher forms of learning—analysis, synthesis, and application, as he put it. And he was concerned about his teachers, an energetic faculty with curricula and varied instructional methods he was loath to disrupt. On the other hand, the study's results could be instructive, and participation would bring enticements in the form of smart boards and "clickers"—automated response systems—for the classrooms of participating teachers. Money for new technology is famously tight.

A sixth grade social studies teacher, Patrice Bain, was eager to give it a try. For the researchers, a chance to work in the classroom was compelling, and the school's terms were accepted: the study would be minimally intrusive by fitting within existing curricula, lesson plans, test formats, and teaching methods. The same textbooks would be used. The only difference in the class would be the introduction of occasional short quizzes. The study would run for three semesters (a year and a half), through several chapters of the social studies textbook, covering topics such as ancient Egypt, Mesopotamia, India, and China. The project was launched in 2006. It would prove to be a good decision.

For the six social studies classes a research assistant, Pooja Agarwal, designed a series of quizzes that would test students on roughly one-third of the material covered by the teacher. These quizzes were for "no stakes," meaning that scores were not counted toward a grade. The teacher excused herself from the classroom for each quiz so as to remain unaware of which material was being tested. One quiz was given at the start of class, on material from assigned reading that hadn't yet been discussed. A second was given at the end of class after the teacher had covered the material for the day's lesson. And a review quiz was given twenty-four hours before each unit exam.

There was concern that if students tested better in the final exam on material that had been quizzed than on material not quizzed, it could be argued that the simple act of reexposing them to the material in the quizzes was responsible for the superior learning, not the retrieval practice. To counter this possibility, some of the nonquizzed material was interspersed with the quiz material, provided as simple review statements, like "The Nile River has two major tributaries: the White Nile and the Blue Nile," with no retrieval required. The facts were quizzed for some classes but just restudied for others.

The quizzes took only a few minutes of classroom time. After the teacher stepped out of the room, Agarwal projected a series of slides onto the board at the front of the room and read them to the students. Each slide presented either a multiple choice question or a statement of fact. When the slide contained a question, students used clickers (handheld, cellphone-like remotes) to indicate their answer choice: A, B, C, or D. When all had responded, the correct answer was revealed, so as to provide feedback and correct errors. (Although teachers were not present for these quizzes, under normal circumstances,

with teachers administering quizzes, they would see immediately how well students are tracking the study material and use the results to guide further discussion or study.)

Unit exams were the normal pencil-and-paper tests given by the teacher. Exams were also given at the end of the semester and at the end of the year. Students had been exposed to all of the material tested in these exams through the teacher's normal classroom lessons, homework, worksheets, and so on, but they had also been quizzed three times on one-third of the material, and they had seen another third presented for additional study three times. The balance of the material was neither quizzed nor additionally reviewed in class beyond the initial lesson and whatever reading a student may have done.

The results were compelling: The kids scored a full grade level higher on the material that had been quizzed than on the material that had not been quizzed. Moreover, test results for the material that had been reviewed as statements of fact but not quizzed were no better than those for the nonreviewed material. Again, mere rereading does not much help.

In 2007, the research was extended to eighth grade science classes, covering genetics, evolution, and anatomy. The regimen was the same, and the results equally impressive. At the end of three semesters, the eighth graders averaged 79 percent (C+) on the science material that had not been quizzed, compared to 92 percent (A–) on the material that had been quizzed.

The testing effect persisted eight months later at the end-of-year exams, confirming what many laboratory studies have shown about the long-term benefits of retrieval practice. The effect doubtless would have been greater if the retrieval practice had continued and occurred once a month, say, in the intervening months.[8]

The lesson from these studies has been taken to heart by many of the teachers at Columbia Middle School. Long after concluding their participation in the research studies, Patrice Bain's sixth grade social studies classes continue today to follow a schedule of quizzes before lessons, quizzes after lessons, and then a review quiz prior to the chapter test. Jon Wehrenberg, an eighth grade history teacher who was not part of the research, has knitted retrieval practice into his classroom in many different forms, including quizzing, and he provides additional online tools at his website, like flashcards and games. After reading passages on the history of slavery, for example, his students are asked to write down ten facts about slavery they hadn't known before reading the passages. You don't need electronic gadgetry to practice retrieval.

Seven sixth and seventh graders needing to improve their reading and comprehension skills sat in Michelle Spivey's English classroom one period recently with their reading books open to an amusing story. Each student was invited to read a paragraph aloud. Where a student stumbled, Miss Spivey had him try again. When he'd gotten it right, she probed the class to explain the meaning of the passage and what might have been going on in the characters' minds. Retrieval and elaboration; again, no technology required.

Quizzes at Columbia Middle School are not onerous events. Following completion of the research studies, students' views were surveyed on this question. Sixty-four percent said the quizzing reduced their anxiety over unit exams, and 89 percent felt it increased learning. The kids expressed disappointment on days when clickers were not used, because the activity broke up the teacher's lecture and proved enjoyable.

Principal Chamberlain, when asked what he thought the study results indicated, replied simply: "Retrieval practice has a significant impact on kids' learning. This is telling us that

it's valuable, and that teachers are well advised to incorporate it into their instructional technique."[9]

Are similar effects found at a later age?

Andrew Sobel teaches a class in international political economics at Washington University in St. Louis, a lecture course populated by 160–170 students, mostly freshmen and sophomores. Over a period of several years he noticed a growing problem with attendance. On any given day by midsemester, 25–35 percent of the class would be absent, compared to earlier in the semester when maybe 10 percent would be absent. The problem wasn't unique to his class, he says. A lot of professors give students their PowerPoint slides, so the students just stop coming to class. Sobel fought back by withholding his slides, but by the end of the semester, many students stopped showing up anyway. The class syllabus included two big tests, a midterm and a final. Looking for some way to leverage attendance, Sobel replaced the big tests with nine pop quizzes. Because the quizzes would determine the course grade and would be unannounced, students would be well advised to show up for class.

The results were distressing. Over the semester, a third or more of the students bailed out. "I really got hammered in the teaching reviews," Sobel told us. "The kids hated it. If they didn't do well on a quiz they dropped the course rather than get a bad grade in it. Of those who stayed, I got this bifurcation between those who actually showed up and did the work, and those who didn't. I found myself handing out A-plusses, which I'd never given before, and more Cs than I'd ever given."[10]

With so much pushback, he had little choice but to drop the experiment and reinstate the old format, lectures with a midterm and final. A couple of years later, however, after hearing a presentation about the learning benefits of testing, he added a third major test during the semester to see what effect it might have on his students' learning. They did better, but not by as much as he'd hoped, and the attendance problems persisted.

He scratched his head and changed the syllabus once again. This time he announced that there would be nine quizzes during the semester, and he was explicit about when they would be. No surprises, and no midterm or final exams, because he didn't want to give up that much of his lecture time.

Despite fears that enrollments would plummet again, they actually increased by a handful. "Unlike the pop quizzes, which kids hate, these were all on the syllabus. If they missed one it was their own fault. It wasn't because I surprised them or was being pernicious. They were comfortable with that." Sobel took satisfaction in seeing attendance improve as well. "They would skip some classes on the days they didn't have a quiz, particularly the spring semester, but they showed up for the quizzes."

Like the course, the quizzes were cumulative, and the questions were similar to those on the exams he used to give, but the quality of the answers he was getting by midsemester was much better than he was accustomed to seeing on the midterms. Five years into this new format, he's sold on it. "The quality of discussions in class has gone way up. I see that big a difference in their written work, just by going from three exams to nine quizzes." By the end of the semester he has them writing paragraphs on the concepts covered in class, sometimes a full-page essay, and the quality is comparable to what he's seeing in his upper division classes.

"Anybody can design this structure. But I also realize that, Oh, god, if I'd done this years ago I would have taught them that much more stuff. The interesting thing about adopting this strategy is I now recognize that as good a teacher as I

might think I am, my teaching is only a component of their learning, and how I structure it has a lot to do with it, maybe even more." Meanwhile, the course enrollment has grown to 185 and counting.

Exploring Nuances

Andy Sobel's example is anecdotal and likely reflects a variety of beneficial influences, not least being the cumulative learning effects that accrue like compounded interest when course material is carried forward in a regime of quizzes across an entire semester. Nonetheless, his experience squares with empirical research designed to tease apart the effects and nuances of testing.

For example, in one experiment college students studied prose passages on various scientific topics like those taught in college and then either took an immediate recall test after the initial exposure or restudied the material. After a delay of two days, the students who took the initial test recalled more of the material than those who simply restudied it (68 v. 54 percent), and this advantage was sustained a week later (56 v. 42 percent). Another experiment found that after one week a study-only group showed the most forgetting of what they initially had been able to recall, forgetting 52 percent, compared to a repeated-testing group, who forgot only 10 percent.[11]

How does giving feedback on wrong answers to test questions affect learning? Studies show that giving feedback strengthens retention more than testing alone does, and, interestingly, some evidence shows that delaying the feedback briefly produces better long-term learning than immediate feedback. This finding is counterintuitive but is consistent with researchers' discoveries about how we learn motor tasks, like making lay-ups or driving a golf ball toward a distant green. In motor learning, trial and error with delayed feedback is a more awkward but effective way of acquiring a skill than trial and correction through immediate feedback; immediate feedback is like the training wheels on a bicycle: the learner quickly comes to depend on the continued presence of the correction.

In the case of learning motor skills, one theory holds that when there's immediate feedback it comes to be part of the task, so that later, in a real-world setting, its absence becomes a gap in the established pattern that disrupts performance. Another idea holds that frequent interruptions for feedback make the learning sessions too variable, preventing establishment of a stabilized pattern of performance.[12]

In the classroom, delayed feedback also yields better long-term learning than immediate feedback does. In the case of the students studying prose passages on science topics, some were shown the passage again even while they were asked to answer questions about it, in effect providing them with continuous feedback during the test, analogous to an open-book exam. The other group took the test without the study material at hand and only afterward were given the passage and instructed to look over their responses. Of course, the open-book group performed best on the immediate test, but those who got corrective feedback after completing the test retained the learning better on a later test. Delayed feedback on written tests may help because it gives the student practice that's spaced out in time; as discussed in the next chapter, spacing practice improves retention.[13]

Are some kinds of retrieval practice more effective for long-term learning than others? Tests that require the learner to

supply the answer, like an essay or short-answer test, or simply practice with flashcards, appear to be more effective than simple recognition tests like multiple choice or true/false tests. However, even multiple choice tests like those used at Columbia Middle School can yield strong benefits. While any kind of retrieval practice generally benefits learning, the implication seems to be that where more cognitive effort is required for retrieval, greater retention results. Retrieval practice has been studied extensively in recent years, and an analysis of these studies shows that even a single test in a class can produce a large improvement in final exam scores, and gains in learning continue to increase as the number of tests increases.[14]

Whichever theories science eventually tells us are correct about *how* repeated retrieval strengthens memory, empirical research shows us that the testing effect is real—that the act of retrieving a memory changes the memory, making it easier to retrieve again later.

How widely is retrieval practice used as a study technique? In one survey, college students were largely unaware of its effectiveness. In another survey, only 11 percent of college students said they use this study strategy. Even when they did report testing themselves, they mostly said they did it to discover what they didn't know, so they could study that material more. That's a perfectly valid use of testing, but few students realize that retrieval itself creates greater retention.[15]

Is repeated testing simply a way to expedite rote learning? In fact, research indicates that testing, compared to rereading, can facilitate better transfer of knowledge to new contexts and problems, and that it improves one's ability to retain and retrieve material that is related but not tested. Further research is needed on this point, but it seems that retrieval practice can make information more accessible when it is needed in various contexts.

Do students resist testing as a tool for learning? Students do generally dislike the idea of tests, and it's not hard to see why, in particular in the case of high-stakes tests like midterms and finals, where the score comes with significant consequences. Yet in all studies of testing that reported students' attitudes, the students who were tested frequently rated their classes more favorably at the end of the semester than those tested less frequently. Those who were frequently tested reached the end of the semester on top of the material and did not need to cram for exams.

How does taking a test affect subsequent studying? After a test, students spend more time restudying the material they missed, and they learn more from it than do their peers who restudy the material without having been tested. Students whose study strategies emphasize rereading but not self-testing show overconfidence in their mastery. Students who have been quizzed have a double advantage over those who have not: a more accurate sense of what they know and don't know, and the strengthening of learning that accrues from retrieval practice.[16]

Are there any further, indirect benefits of regular, low-stakes classroom testing? Besides strengthening learning and retention, a regime of this kind of testing improves student attendance. It increases studying before class (because students know they'll be quizzed), increases attentiveness during class if students are tested at the end of class, and enables students to better calibrate what they know and where they need to bone up. It's an antidote to mistaking fluency with the text, resulting from repeated readings, for mastery of the subject. Frequent low-stakes testing helps dial down test anxiety among students by diversifying the consequences over a much larger sample: no single test is a make-or-break event. And this kind of testing enables instructors to identify gaps in students' understanding and adapt their instruction to fill them. These benefits of low-stakes testing accrue whether instruction is delivered online or in the classroom.[17]

The Takeaway

Practice at retrieving new knowledge or skill from memory is a potent tool for learning and durable retention. This is true for anything the brain is asked to remember and call up again in the future—facts, complex concepts, problem-solving techniques, motor skills.

Effortful retrieval makes for stronger learning and retention. We're easily seduced into believing that learning is better when it's easier, but the research shows the opposite: when the mind has to work, learning sticks better. The greater the effort to retrieve learning, provided that you succeed, the more that learning is strengthened by retrieval. After an initial test, *delaying subsequent retrieval practice* is more potent for reinforcing retention than immediate practice, because delayed retrieval requires more effort.

Repeated retrieval not only makes memories more durable but produces knowledge that can be retrieved more readily, in more varied settings, and applied to a wider variety of problems.

While cramming can produce better scores on an immediate exam, the advantage quickly fades because there is much greater forgetting after rereading than after retrieval practice. The benefits of retrieval practice are *long-term.*

Simply including one test (retrieval practice) in a class yields a large improvement in final exam scores, and gains continue to increase as the frequency of classroom testing increases.

Testing doesn't need to be initiated by the instructor. Students can practice retrieval anywhere; no quizzes in the classroom are necessary. Think flashcards—the way second graders learn the multiplication tables can work just as well for learners at any age to quiz themselves on anatomy, mathematics, or law. Self-testing may be unappealing because it takes more effort than rereading, but as noted already, the greater the effort at retrieval, the more will be retained.

Students who take practice tests have *a better grasp of their progress* than those who simply reread the material. Similarly, such testing enables an instructor to *spot gaps and misconceptions* and adapt instruction to correct them.

Giving students *corrective feedback* after tests keeps them from incorrectly retaining material they have misunderstood and produces better learning of the correct answers.

Students in classes that incorporate low-stakes quizzing come to embrace the practice. Students who are tested frequently rate their classes more favorably.

What about Principal Roger Chamberlain's initial concerns about practice quizzing at Columbia Middle School—that it might be nothing more than a glorified path to rote learning? When we asked him this question after the study was completed, he paused for a moment to gather his thoughts. "What

I've really gained a comfort level with is this: for kids to be able to evaluate, synthesize, and apply a concept in different settings, they're going to be much more efficient at getting there when they have the base of knowledge and the retention, so they're not wasting time trying to go back and figure out what that word might mean or what that concept was about. It allows them to go to a higher level."