



CHAPTER XI

Stories and Scoreboards

“Whatever interest attaches to your subject-matter must be evoked here and now; whatever powers you are strengthening in the pupil, must be exercised here and now; whatever possibilities of mental life your teaching should impart, must be exhibited here and now.”

Alfred North Whitehead, “The Aims of Education”

🍷 Day 412

At 11:58 on a Tuesday night, in the back seat of the car on the way home from her aunt's, a twelve-year-old named Iris opens her language app and taps through a lesson. She doesn't look at the words much. She's watching the little flame in the corner, which turns over from 411 to 412 with two minutes to spare. She has studied Spanish every single day for more than a year.

Three weeks later, on a family trip, a waiter asks her something in Spanish. She freezes, smiles, and points at the menu. Her younger brother orders for both of them, badly, with his hands.

That same month, at Sunday dinner, Iris's grandfather mentions that the Titanic carried lifeboats for only about half the people on board, and that this was perfectly legal. Iris asks why it was legal. Nobody at the table knows.

By the next Sunday she knows. The British rules in force had been written in 1894, when the largest ships were a fraction of the Titanic's size, and they set lifeboat requirements by a ship's tonnage, topping out at sixteen boats for anything over ten thousand tons. The Titanic measured more than forty-six thousand gross tons and carried twenty boats, with room for about 1,178 people. There were about 2,200 aboard. Iris has done the subtraction, drawn the boat deck from a photograph, and worked out which boats left half empty and why. She has a new question now, about the radio operator on the nearest ship, and she's reading about that.

Both of these were designed. The flame was built by a company of engineers who test every color and sound on millions of users. The question was built, more or less by accident, by a grandfather who knows how to drop a fact at dinner. The flame held Iris for 412 days and left her pointing at a menu. The question held her for a weekend and became a year.

It would be easy to conclude that stories are good and scoreboards are bad. That's the wrong lesson. The flame could have counted conversations instead of taps. The story could have been a cartoon about a ship with no subtraction in it. Both are tools, and the designer decides where they point.

This chapter is about the outside drivers that a designer builds: **stories** (narrative, curiosity, play, entertainment) and **scoreboards** (rewards, grades, money, progress, streaks, leaderboards, deadlines, games). Chapter 9 built the tests for judging them, and Chapter 10 took the drivers that come from other people.¹ The book's hypothesis, which this chapter tests driver by driver, is that stories are best at opening Inquiry and scoreboards are best at carrying Practice.

Part One: Stories

Stories are the drivers that work through meaning: a plot, a gap, a role, a delight. They're the oldest teaching tools there are, older than writing.

Narrative

A textbook can feel like chewing sand. A thriller will keep you up until three in the morning on a work night.

Why Stories Stick

Raymond Mar, Jingyuan Li, Anh Nguyen, and Cindy Ta (2021) pooled more than 75 samples and 33,000 participants and found that narrative texts were more easily understood and better recalled than expository ones, and the result held across studies and conditions. The cognitive scientist Daniel Willingham (2004) calls stories “psychologically privileged”: the mind treats them differently from other kinds of material.

Willingham points to four features that make a story a story, and all four do work for memory. **Causality**: one event leads to another, and causal chains are far easier to remember than lists. **Conflict**: someone wants something and something stands in the way. **Complications**: the obstacles multiply, and each one raises a new problem. **Character**: the events happen to someone you can follow. A good story makes the listener do a little inference at every step (why did she do that? what will happen now?), and inference is the mental work that builds memory. A list of facts asks for none.

Iris's lifeboat question had all four. The cause: rules written for smaller ships. The conflict: the rule said one thing, and the sea said another. The complications: half-empty boats, a radio operator asleep on the ship that might have come. The characters: 2,200 of them.

A Story Inside the Material

The question for a designer is where the story sits. A story about the object carries the learner into the material: the dig team that has to measure a femur

to learn how big the dinosaur was. A story laid on top of the material carries the learner away from it: the cartoon between drills, where the math is the toll you pay to reach the story.

The best stories inside the material are usually true ones, because the material has a history, and the history is full of people who wanted something and hit obstacles. Take half-life, which is normally taught as a formula. Marie Curie's laboratory notebooks, more than a century after she wrote them, are still radioactive enough that they're kept in lead-lined boxes, and readers who want to see them sign a waiver. The radium that contaminated them has a half-life of about 1,600 years. How radioactive will the notebooks be when your great-great-grandchildren could visit them? What fraction will remain in the year 3500? That's the formula, inside a story, with a character, a conflict, and a reason to calculate.

The real history of chemistry, with its eccentrics poisoning themselves and its discoveries missed by a whisker, beats any invented space alien. So does the real history of almost everything. Kieran Egan's tools for building the story into the material, the heroes, mysteries, binary opposites, and extremes, are collected in Chapters 8 and 14.²

Game designers have found the same thing. In gamification research, adding **game fiction**, a story the learner plays inside, was among the features that most improved learners' behavior (Michael Sailer and Lisa Homner, 2020).

Story is the strongest tool of early childhood, and the advantage of narrative holds for adults as well. Every good lecture you remember was probably a story.

Design moves. Build the story from the material's own history: who discovered it, what they were trying to solve, what went wrong. Cast the learner in a role inside the problem. Let the story end in the material, so that the last step of the plot is the calculation, the experiment, or the argument.

❧ Curiosity and Mystery

The Itch

George Loewenstein (1994) proposed that curiosity arises when people notice a gap between what they know and what they want to know. The gap feels like an itch, and people will work to scratch it. Daniel Berlyne, decades earlier (1960), had catalogued what produces the itch: novelty, complexity, surprise, and uncertainty. Each one says the same thing to the mind: there's something here you don't understand yet.

Curiosity helps memory. Matthias Gruber, Bernard Gelman, and Charan Ranganath (2014) had people read trivia questions in a brain scanner and rate how curious they were about each answer. People remembered the answers they'd been curious about better, both immediately and a day later, and curiosity recruited the same brain reward circuits that respond to tangible rewards (see also Min Jeong Kang and colleagues, 2009). Closing a gap you care about registers, in the brain, something like being paid. The same study found better memory even for unrelated faces shown during high-curiosity moments, though later studies suggest that spillover is weaker and less consistent. Celeste Kidd and Benjamin Hayden (2015), reviewing the field, describe curiosity as a drive to gather information, as basic as the drives for food and water.

Curiosity Peaks at Partial Knowledge

In Kang's study, curiosity was highest when people were moderately confident of an answer. Knowing nothing gives the itch nothing to grip. Knowing everything leaves no itch. Knowing a little makes you want the rest.

That makes curiosity a matter of fit, the dial from Chapter 5.³ A question at the edge of what a learner knows itches. A question far past the edge is noise, and a question behind it is a quiz. Iris knew the Titanic sank and roughly how. She didn't know about the rules. That's the edge.

Surprise Teaches

Even infants learn hardest where the world surprises them. Aimee Stahl and Lisa Feigenson (2015) showed eleven-month-olds objects that behaved impossibly: a ball that seemed to roll through a solid wall, or a toy car that rolled off a ledge and hovered in midair. Afterward, the infants learned new things about the surprising object better than about an ordinary one, and they explored it as if testing a hypothesis. They banged the ball that had gone through the wall, as if checking its solidity. They dropped the car that had floated, as if checking whether it would fall.

Older learners can be surprised on purpose. Garvin Brod, Marcus Hasselhorn, and Silvia Bunge (2018) found that asking learners to predict an outcome before seeing it improved their learning, and the benefit ran through surprise: predictions that turned out wrong produced the biggest gains. A prediction makes a gap visible. Ask a ten-year-old whether a bowling ball or a tennis ball will hit the ground first when dropped together, and get her to commit. Then drop them. The gasp is the itch.

School Quiets Questions

Barbara Tizard and Martin Hughes (1984) recorded 30 four-year-old girls at home and at nursery school. At home they asked about 26 questions an hour, often in long chains on a single topic: why, and then why that, and then what about this. At nursery school they asked about 2 an hour, and that kind of sustained exchange almost never happened. The children were the same children. The setting changed what they voiced.

Children don't stop wondering when they start school. The capacity is there at every age. What changes is whether anything in the room leaves a gap open long enough to be felt, and whether anyone has time for the chain.

Capture Wearing Curiosity's Clothes

A gap inside the object leads into the object: why does the maple turn red weeks before the oak? A gap built to be clicked leads nowhere. The headline that ends in a cliffhanger, the video that promises the answer "at the end," the

notification that says *you won't believe what happened*: these are information gaps engineered to produce the itch with nothing behind it. They pass the attention test in the first second and fail every other test in Chapter 9. You know the feeling of finishing one: a small, empty sense of having been had.

Design moves. Start with a question the learner half knows. Ask for a prediction before the answer. Leave the gap open long enough to be felt. Answer a question with the next question. Make room for chains of questions on one topic, and keep a place, a wall or a notebook, where unanswered questions wait.

🐭 Play and Make-Believe

A Head Taller

“In play a child always behaves beyond his average age, above his daily behavior,” wrote Lev Vygotsky; “in play it is as though he were a head taller than himself.”

Psychologists in his tradition tested the idea. In a study from the 1940s, Z. V. Manuilenko asked preschoolers to stand still for as long as they could. Asked plainly, they managed only briefly. When they played sentry, guarding a post in a pretend game, they stood still far longer, holding themselves to the role. The study is old and small, but it captures something any parent has seen: a child who can't sit still for a story will sit motionless for twenty minutes as a cat waiting for a mouse.

The Batman effect is the modern version. Rachel White and her colleagues (2017) gave four- and six-year-olds a dull, repetitive computer task for ten minutes, with an attractive video game on a tablet beside them that they could switch to whenever they liked. Some children were asked to think about how they were doing in their own voice (“Am I working hard?”). Some were asked to refer to themselves in the third person (“Is Jamie working hard?”). Some were given a costume prop and asked to pretend to be a hard-working character such as Batman (“Is Batman working hard?”). The pretending children spent the most time working, more than the third-person group, and

far more than the children who stayed in their own perspective. A child may want the tablet. Batman has a job to do. Play can carry Practice, and it connects to the self-control research in Chapter 8.⁴ (If you ever need to get through your taxes, consider a cape.)

What the Curricula Show

Tools of the Mind, a preschool curriculum built on structured make-believe, improved executive function in early trials (Adele Diamond and colleagues, 2007) and in a kindergarten trial with 759 children (Clancy Blair and Cybele Raver, 2014). A larger trial with 877 preschoolers in 60 schools found no benefits and possible harms (Dale Farran and Sandra Wilson, 2014). Positive effects appeared most often in high-poverty settings.

A major review found the evidence insufficient to show that pretend play is crucial to development; it may be one route among many (Angeline Lillard and colleagues, 2013). Play works as a driver. Claims that it's necessary for development remain unproven. Guided play, where the child leads inside a frame an adult has set, has the strongest evidence of the forms of play in learning, and it's covered in Chapter 8.⁵

Play Inside the Object

Play inside the object carries the object: a pretend shop where making change takes real arithmetic, a pretend hospital where the patient's temperature has to be read off a real thermometer. Play beside it, a game of tag as a reward for finishing the worksheet, leaves the object behind.

With age, play changes shape without disappearing. Make-believe is strongest in early childhood. Games with rules take over in middle childhood. In adolescence and adulthood, play becomes simulation and role: the mock trial, the model legislature, the flight simulator, the war game. These are make-believe with high fidelity, the dial from Chapter 6, and they're among the most powerful learning environments adults have built.⁶

Design moves. Give learners roles that require the skill. Use pretend to carry dull reps. Keep real materials and real numbers inside the play. For older learners, build simulations where the role demands the competence.

❧ Entertainment

Seductive Details

Afraid of boring an audience, most of us decorate: a joke about the cat, a funny image, a story from the weekend. Narmada Sundararajan and Olusola Adesope (2020) pooled the research on interesting but irrelevant material and found a small-to-moderate negative effect on recall and transfer (see also Günter Daniel Rey, 2012). Working memory is small. Spend some of it on the cat, and on working out what the cat has to do with chemical bonds, and there's less left for the bonds. You've probably sat through a charming, funny presentation and walked out unable to say what the new policy was.

Game designers have a name for the worst version: chocolate-covered broccoli. The learning is the broccoli, the entertainment is the chocolate, and children can taste the difference immediately.

Entertainment Built Around the Material

Sesame Street is the counterexample, and it was built on purpose. When the Children's Television Workshop launched the show in 1969, it put researchers and producers in the same building. The researchers built a device they called the Distractor: a slide projector beside the television that showed a new image every seven and a half seconds, while an observer recorded whether the children were watching the show or the slides. Segments that lost the children to the slideshow got rewritten. That was a measure of capture, second by second. The researchers also measured learning: independent evaluations in the show's first years found that children who watched more learned more of the letters, numbers, and skills the show taught (Samuel Ball and Gerry Ann Bogatz, 1970).

And the learning lasted. Children who had access to Sesame Street when it first aired, because of where they happened to live, were more likely to be at the right grade level for their age later on, especially boys (Melissa Kearney and Phillip Levine, 2019). The entertainment was the letters. The Muppets sang about the letter J.

Humor

Reviews of classroom humor find its effects depend on the kind. Humor about the material tends to help; humor unrelated to it, or at students' expense, doesn't (John Banas and colleagues, 2011).

The child from Chapter 4, shouting with glee as her rocket hits a record height, is entertained by the object itself.⁷ Entertainment about the object is engagement. Entertainment about something else is capture.

Design moves. Make the material itself the source of delight. Cut decorative jokes and images from hard lessons. If you measure whether learners are watching, also measure whether they're learning. Save pure entertainment for rest.

Part Two: Scoreboards

Scoreboards are the machinery of school and work: rewards, grades, money, progress bars, streaks, rankings, deadlines. They're the easiest drivers to install and the easiest to count, which is why they're everywhere. They're also the drivers most likely to fail the removal test, because they're the easiest to remove.

Rewards

Ask a friend to help you move a couch on a Saturday and they'll grumble and come. Buy them a pizza afterward and the favor is sealed. Offer them \$4.50 for the afternoon and they'll be insulted. The money turns a favor into a job, and a badly paid one.

What the Research Shows

A meta-analysis of 128 experiments found that expected, tangible rewards reduced free-choice intrinsic motivation, while informational positive feedback increased it. Unexpected rewards, given after the fact, didn't reduce motivation (Edward Deci, Richard Koestner, and Richard Ryan, 1999). The pizza after the couch is safe.

The size of the effect has been argued over for decades. Judy Cameron and David Pierce (1994) read the same literature and concluded that rewards do little overall harm, except in the specific case of expected, tangible rewards given simply for doing a task. That exception is exactly the sticker chart. The two camps disagree about how often the harm occurs; they largely agree about where it does.

A Little Money Is Worse Than None

Uri Gneezy and Aldo Rustichini (2000) sent high school students door to door collecting donations for charity. Some were paid nothing, some a small commission, some a larger one. The students paid a small commission collected less than the students paid nothing. A larger commission brought effort back up to about where the unpaid students had been. The title of their paper was its conclusion: "Pay enough or don't pay at all." A small payment told the students the work was a job, and not much of one.

What Rewards Buy

Incentives predict the quantity of performance better than its quality (Christopher Cerasoli and colleagues, 2014), and acting for rewards isn't associated with persistence (Joshua Howard and colleagues, 2021). Rewards buy more of the work. They rarely buy better work, or work that lasts after the payments stop. In the same 1999 meta-analysis, the undermining effect of tangible rewards was larger for children than for college students.

Rewards point at the object when they carry information about the work, which is why the swim coach's second speech in Chapter 9 worked.⁸ A prize for doing the task points at the prize.

Design moves. Save rewards for tasks with no interest to lose, like the dishes, and keep them away from work the learner already enjoys. Make rewards unexpected and informational. If you pay, pay enough to be respectful, or don't pay. Phase rewards out as the work takes over, using the fades from Chapter 9.

🍷 Grades

A Recent Invention

Grades feel like a law of nature, but they're young. In 1785, Ezra Stiles, the president of Yale, recorded the results of an examination of his seniors in four Latin categories, from *Optimi*, the best, down to *Peiores*, the worst. Letter grades spread through American schools and colleges only in the late nineteenth and early twentieth centuries, as schooling grew and institutions needed a compact way to pass judgments about students to each other. That last function explains a great deal. Grades persist because they're a currency between institutions, which is a different job from helping a learner learn.

Comments Beat Grades

Ruth Butler (1988) randomly assigned twelve fifth- and sixth-grade classes to receive individual comments, numerical grades, or both after interesting tasks. Interest and performance were highest after comments alone, for high and low achievers alike, and a grade given alongside a comment wiped out the comment's benefit. Later studies haven't always found grades this damaging, and the effects vary widely.

Hand a student a paper with three paragraphs of careful comments and a B+ at the top, and watch where the eyes go. (To the B+, and then to the neighbor's paper.)

Grades also change what students aim for. Caroline Pulfrey, Céline Buchs, and Fabrizio Butera (2011) found that students who expected a grade for their work adopted more **performance-avoidance** goals, trying above all not to do worse than others, than students who expected only comments, and the shift

ran through a loss of autonomous motivation. A grade turns a task into a test of standing, and the safest strategy in a test of standing is not to lose.

Removing the Rank

When Mayo Medical School moved its first-year courses to pass/fail, you might have expected the future doctors to coast. Students reported less stress and greater group cohesion, and their scores on the national board exam didn't change (Rohe and colleagues, 2006). A review of such changes found improvements in well-being across studies (Spring and colleagues, 2011).

A comment points at the work: what to change next. A grade points at a rank. Given both, learners look at the grade. Grades are the outside driver the classroom form runs on, which is part of why that form struggles to produce ownership, and through Girard's lens they're the purest metaphysical object in school: a symbol of standing among peers.⁹

Grades arrive with schooling and gain weight every year, peaking in the competition for college.

Design moves. Comments first. Grades rarely, and never alongside comments. Pass/fail or mastery standards wherever ranking serves no purpose. Where grades must exist, for transcripts and admissions, keep them as far as possible from the daily work.

Money

Paying for Inputs and Outputs

Roland Fryer (2011) ran randomized trials in more than 200 urban schools. Students in Dallas were paid to read books (\$2 a book, confirmed by a short quiz), in New York for test performance, and in Chicago for grades. Effects on state test scores were statistically zero in each city. The only significant effect was for English-speaking students in Dallas, the group paid to read. Fryer's interviews suggested why: students were excited by the money and didn't know how to turn effort into higher scores.

Offer me a million dollars to play a Mozart concerto tomorrow and you'll get my full attention and no Mozart. I can't play the piano. Motivation without capacity produces a lot of staring at the keys.

Immediacy Is Everything

In field experiments with thousands of students around Chicago, rewards announced just before a low-stakes test raised scores, and the effect vanished entirely when the reward was delayed. Rewards framed as losses (here's the money; you keep it if you improve) did somewhat better. Trophies worked about as well as cash for younger students and less well for older ones (Steven Levitt, John List, Susanne Neckermann, and Sally Sadoff, 2016).

Paying for an act that is itself contact with the object, reading a book, can help. Paying for an outcome points at the payout. The test-day rewards changed effort on the test and left learning where it was. They raised the measure, which is what capture does.

Design moves. If you pay at all, pay for contact with the object, and pay at once. Never pay for outcomes learners don't know how to produce.

Progress and Feedback

Progress in Meaningful Work

Teresa Amabile and Steven Kramer (2011) collected nearly 12,000 daily diary entries from 238 professionals on 26 project teams in seven companies. The single strongest booster of motivation and positive emotion at work was making progress in meaningful work, even small progress.

Progress also pulls harder as the goal gets closer. In the 1930s, Clark Hull found that rats in a maze ran faster the nearer they got to the food. Ran Kivetz, Oleg Urminsky, and Yuhuang Zheng (2006) found the same in people with coffee-shop loyalty cards: customers bought coffee more often as they approached the free one. It's called the **goal gradient**, and it's why a progress bar at 90 percent is so hard to walk away from. Chapter 8's car-wash cards, with two free stamps to start, put it to work.¹⁰

Feedback That Helps, and Feedback That Hurts

Avraham Kluger and Angelo DeNisi (1996) pooled 607 effects and found that feedback improved performance on average ($d = 0.41$). More than a third of feedback interventions made performance worse. Feedback helped less the more it drew attention toward the self and away from the task.

John Hattie and Helen Timperley (2007) sorted feedback into four levels. Feedback about the **task** says whether the work is right. Feedback about the **process** says how the work was done and how to do it better. Feedback about **self-regulation** helps the learner check and steer her own work. Feedback about the **self** (“good girl,” “you’re so clever”) says something about the person. The first three help, with process and self-regulation feedback the most powerful for deep learning. The fourth helps least, because it carries no information about the work.

Picture a tennis coach. “Tilt your wrist down a little as you swing” gives the player something to do on the next ball. “You’re just not very coordinated” gives her something to think about instead, and the thing is herself. So does “you’re in the bottom fifth of the group,” however gently it’s said. Kluger and DeNisi’s finding is Chapter 9’s test in its most general form.

Progress that reports movement on the object points at the object: a first loaf that rises, a passage that finally holds at tempo, a mastery chart where each filled square is a skill the learner can now perform. Points that climb whatever the learning does point at the points.

Young children need progress they can see and touch: the jar of marbles, the chart on the fridge, the height marks on the doorframe. Older learners can track progress in the abstract.

Design moves. Show progress on the object itself. Describe the next step, and leave the person out of it. Use the goal gradient honestly: break long work into stages a learner can see the end of. Keep progress private where comparison would hurt.

🔥 Points, Streaks, and Leaderboards

These three are the core of what's usually called gamification, and they share one mechanism. Each takes a running record of the learner's behavior and makes the record itself the thing to protect or improve.

Streaks

The little flame with a 45 beside it does get you to open the language app. Then you travel, miss a day, lose the 45, and don't open the app again for six months.

Across seven studies, highlighting an intact streak in a behavior log increased people's next engagement in the behavior, compared with highlighting a broken one. The effect depended only on how the log represented past behavior, and the behavior itself didn't matter, because people come to treat keeping the streak as a goal in its own right. The damage from a broken streak was larger when people blamed themselves for the break, and smaller when they could repair it (Joseph Silverman and Alixandra Barasch, 2023). Tap through one lesson at 11:58 p.m. to keep the flame alive, and Spanish has become incidental to the chain. That's Iris.

The habit is more forgiving than the streak. Phillippa Lally and her colleagues (2010) followed 96 people forming a new daily habit. The behavior took a median of 66 days to become as automatic as it would get (well past the 21 days of self-help lore), with a range of 18 to 254 days. Missing a single day didn't materially affect the process. Forty-four days out of forty-five is a fine record.

Leaderboards

Leaderboards work like goals. In an experiment, adding a leaderboard to a task raised performance about as much as assigning a difficult goal. People implicitly aimed near the top of the board, and the effect depended on how committed they were to that goal (Richard Landers, Kristina Bauer, and Rachel Callan, 2017).

Comparison has costs. In a virtual classroom experiment with 80 women taking a math test, leaderboards shifted academic identification and, in some conditions, lowered math performance. The harm ran through social comparison more than through stereotype threat (Katheryn Christy and Jesse Fox, 2014).

A leaderboard is other-referenced competition made permanent and public, and Chapter 10's N-effect applies to it directly.¹¹ On a board of thousands, almost nobody is near enough to anyone to care. A woman who jogs twice a week gains nothing from learning she's 10,412th in the app, somewhere behind an Olympian. Show her she's neck and neck with the neighbor who also jogs twice a week, and she laces up. The board works best for learners near the top, and for everyone else it works only when it's small.

The Bundle

Gamification helps on average. A meta-analysis of 40 experiments found that gamification improved cognitive ($g = 0.49$), motivational ($g = 0.36$), and behavioral ($g = 0.25$) learning outcomes. Game fiction, and competition combined with collaboration, were among the most effective elements (Sailer and Homner, 2020). Well-designed digital games beat non-game instruction ($g = 0.33$), and enhanced game designs added further gains (Douglas Clark, Emily Tanner-Smith, and Stephen Killingsworth, 2016).

But novelty fades. In one of the few semester-long studies, one course used leaderboards and badges and a matched course didn't. Over sixteen weeks, students in the gamified course grew less motivated, less satisfied, and less empowered, and their lower motivation carried through to lower final exam scores (Michael Hanus and Jesse Fox, 2015). The badges that sparkled in September were a chore by December.

Games and Gamification

Games and gamification are different things. In a good learning game, the mechanics are the object. There is a tablet game in which a child isolates a little box on one side of the screen by applying moves to both sides at once:

dividing everything, adding the same card to each side. A few levels in, the cards turn into numbers and letters, and the child discovers she has been solving equations all along. To win, she has to understand the thing. In most gamification, points and badges are laid on top of an unchanged task. The first points at the object. The second points at itself.

Design moves. Count the work in streaks, and leave logins out of it. Allow repair. Emphasize the long pattern over the unbroken chain. Replace public rankings with personal goals; if a board is used, show learners their near neighbors instead of the whole field, and change what it measures often. Prefer game fiction and team competition to points and badges. Build the mechanics from the material. Plan for the novelty to fade.

Time

You may be thinking that deadlines are the only reason anything gets done. Without one, the report waits and the evening goes to television.

Deadlines

In a classic experiment, people working under a deadline later showed less interest in the task than people working without one (Teresa Amabile, William DeJong, and Mark Lepper, 1976).

A caution about the evidence. The best-known study showing that deadlines improve performance (Dan Ariely and Klaus Wertenbroch, 2002) has been retracted, and a 2026 replication didn't reproduce its results.

Plans and Fresh Starts

Plans help more than pressure. A meta-analysis found that forming implementation intentions, specific if-then plans for when and where to act, had a medium-to-large effect on reaching goals (Peter Gollwitzer and Paschal Sheeran, 2006). “Finish the report by Friday” is a deadline. “Tuesday at eight, coffee poured, I sit down and write the first five hundred words” is a plan, and it tells you what to do when Tuesday at eight arrives.

The calendar has its own pull. Hengchen Dai, Katherine Milkman, and Jason Riis (2014) found that gym visits rose at the start of each week, month, and year, and after birthdays; so did searches for the word “diet.” They called it the **fresh start effect**. Temporal landmarks let people draw a line under past failures and begin again. A learner who lost her streak in March can start a new one on the first of April, and the date helps.

A deadline tied to a real event points at the object: opening night, the science fair, the morning the family leaves on the trip. An arbitrary Friday at five points at the calendar.

Design moves. Let learners set dates, with an adult checking in. Tie deadlines to real events. Pair them with specific plans. Use fresh starts on purpose, especially after a broken streak.

🐉 The Machine Built for Capture

Put every driver in these three chapters on a single screen, point each one at itself, and optimize the result against time on the app. You get the feed.

It has a variable reward: you never know whether the next swipe will bring something wonderful, and B. F. Skinner found that pigeons rewarded on unpredictable schedules pecked longer and more persistently than pigeons rewarded every time. It has no stopping cue: the page never ends, and the designer credited with inventing infinite scroll has since spoken publicly about regretting what it became. It has streaks and counts, curiosity gaps built to be clicked, and stories that resolve into the next story. And it has Chapter 10's room full of peers, ranked by likes, views, and followers, with no closing bell.¹²

Every driver in it points at itself, on purpose. That's what it was built to do, and it does it superbly. It passes the first two rows of Chapter 9's table, time and effort, better than anything humans have ever made. It is designed to fail the removal test, because a product that passes it loses its users.

The tools themselves are neutral. The same variable reward, pointed at the object, is the joy of a question whose answer you can't predict. The same streak, counting nights in a book, makes a reader. The same near-neighbor

board, in a running club of twelve, gets people to the track. The machine shows what these drivers can do when nobody points them at anything worth learning.

 The Whole Map

The two families in this chapter, and the social family in Chapter 10, fit the two systems differently. What follows is the book's hypothesis. The evidence is consistent with it, and none of it was designed to test it.

Structure suits Practice. Practice is repetition toward a standard, and repetition is where attention flags: the fortieth scale, the hundredth free throw, the vocabulary list on the fridge. Streaks that count the work, visible progress, and competition against your own record help carry the reps. The Batman effect was measured on exactly this kind of task: dull, repetitive, and worth finishing.

Story suits Inquiry. Inquiry runs on a question the learner owns, and curiosity is already its own reward. Story and mystery can open a question: the femur on the dig, the lifeboats on the Titanic. Rewards and rankings sit poorly with it, because the undermining research finds the strongest damage where the activity was interesting to begin with. Pay a child to solve a mystery she's already desperate to solve, and the next mystery may need paying for too.

Other people serve both. Cooperation, an audience, and admired models can carry a question or a rep.

	FITS WELL	USE WITH CARE
Practice	Streaks that count the work, visible progress, competing with your own record, rivals near mastery, play and pretend, teams	Arbitrary deadlines, public rankings, payment for outcomes
Inquiry	Story, mystery, prediction and surprise, partial knowledge, an audience that needs the answer, admired models	Expected rewards, grades, leaderboards
Both	Cooperation, teaching others, belonging to a group defined by the work, a teacher who cares about the work, informational feedback	Praise of the person, decorative entertainment

Most real learning runs both systems, often in the same hour, so most real settings mix the drivers. Iris's lifeboat question was Inquiry while the question was open, and Practice when she sat down to draw the boat deck to scale. The driver can change when the system does.

For reference, here is every outside driver in the three chapters, with the condition that turns it toward the object or away.

DRIVER	FAMILY	POINTS AT THE OBJECT WHEN	POINTS AT ITSELF WHEN	BEST FIT
Models	Social	The model visibly wants the work, a few steps ahead	The model is admired for status	Both
Competition	Social	Against a standard, a past self, or another team	Against the classmate at the next desk	Practice near mastery
Cooperation	Social	The shared goal is the work; each member accountable	The group itself becomes the goal	Both
Audience and teaching	Social	The audience needs the work	The audience judges the performer	Both; private for fragile learning
Belonging	Social	The group is defined by the work	Belonging depends on standing	Both
Praise	Social	It describes the work	It labels the person or ranks them	Both
The admired adult	Social	She cares visibly about the work	She cares mainly about scores	Both
Narrative	Story	The story is the material's own	The story is laid on top	Inquiry
Curiosity and mystery	Story	The gap is inside the object	The gap is built to be clicked	Inquiry
Play and make-believe	Story	The role requires the skill	The play is a reward beside the work	Practice and Inquiry

DRIVER	FAMILY	POINTS AT THE OBJECT WHEN	POINTS AT ITSELF WHEN	BEST FIT
Entertainment	Story	The delight is in the material	The decoration competes with it	Rest, and both with care
Rewards	Structural	They carry information, or come unexpected	They're expected and tangible	Tasks with no interest to lose
Grades	Structural	Rarely; mastery standards instead	They rank	Neither, where avoidable
Money	Structural	It pays for contact with the object, at once	It pays for outcomes	Practice inputs, sparingly
Progress and feedback	Structural	It reports movement on the object	It reports on the self	Both
Streaks	Structural	They count the work and allow repair	They count check-ins	Practice
Leaderboards	Structural	Small, near neighbors, changing measures	Large, permanent, public	Practice, sparingly
Deadlines	Structural	Tied to a real event, with a plan	Arbitrary	Both
Games	Structural	The mechanics are the material	Points and badges sit on top	Practice, and Inquiry in simulations

The design consequence. Reach for structure to carry Practice, for story and mystery to open Inquiry, and for other people to carry both. Be most careful with rewards around Inquiry. And run every driver through Chapter 9’s tests.

13

 Back to Day 412

Run Iris’s year again.

The app stays, but the streak counts something else. Her aunt has a friend in Bogotá with a daughter Iris’s age, and on Sunday afternoons the two girls talk on video for twenty minutes, in Spanish, about whatever comes up. The streak now counts Sundays, and a missed Sunday can be made up on Monday.

The app's daily lessons become what they always should have been: practice for Sunday. When Iris can't find the word for *lifeboat*, she looks it up, and uses it the next week. By the family trip she can order for both of them, and her brother is annoyed.

The lifeboat question gets a little structure too. Her grandfather asks whether she'd explain what she found to his friends at the Thursday coffee group, an audience of six retired men who would like to know. Iris makes a poster of the boat deck, to scale, which takes three evenings of careful measuring that she would never have done for a grade. She practices the talk twice. On Thursday one of the men asks a question she can't answer, about why the nearest ship's radio was switched off, and she goes home and finds out.

A flame and a question, both designed, both pointing at something worth learning.

• Open Questions

- Do story drivers help Inquiry at every age, or mainly while story is the learner's strongest tool?
- Can the fresh start effect be used to rescue learners after a broken streak, and does it work for children?
- How large does a reward have to be to escape Gneezy and Rustichini's trap of paying too little, and does the threshold differ for children?
- Can gamification be designed so its novelty hands over to the material before it fades?
- Can the machinery of the feed be pointed at the object at scale, and would anyone build it if its success meant users who eventually leave?

If you've watched a story open a question that lasted, or a scoreboard carry someone through the reps to mastery, I'd like to hear about it.

📌 Notes

- 1. See Chapter 9; Chapter 10.
- 2. See Chapter 8, “Seeding Ownership”; Chapter 14.
- 3. See Chapter 5.
- 4. See Chapter 8, “Passivity Is the Default.”
- 5. See Chapter 8, “Follow the Pull, Supply the Structure.”
- 6. See Chapter 6.
- 7. See Chapter 4.
- 8. See Chapter 9, “How a Driver Moves a Learner.”
- 9. See Chapter 2, “The Form Is the Cause”; Chapter 10, “The School Through Girard’s Lens.”
- 10. See Chapter 8, “Seeding Ownership.”
- 11. See Chapter 10, “Rivals.”
- 12. See Chapter 10, “The School Through Girard’s Lens.”
- 13. See Chapter 9, “Working Tests.”

📌 Sources

SOURCE	USE
Raymond Mar, Jingyuan Li, Anh Nguyen, and Cindy Ta, <i>Psychonomic Bulletin & Review</i> 28 (2021)	Narrative versus expository texts
Daniel Willingham, “The Privileged Status of Story,” <i>American Educator</i> (2004)	Why stories are remembered
Michael Sailer and Lisa Homner, <i>Educational Psychology Review</i> 32 (2020)	Gamification meta-analysis; game fiction
George Loewenstein, <i>Psychological Bulletin</i> 116 (1994)	The information-gap theory of curiosity
Daniel Berlyne, <i>Conflict, Arousal, and Curiosity</i> (1960)	Novelty, complexity, surprise, and uncertainty
Matthias Gruber, Bernard Gelman, and Charan Ranganath, <i>Neuron</i> 84 (2014); Min Jeong Kang and colleagues, <i>Psychological Science</i> 20 (2009)	Curiosity and memory

SOURCE	USE
Celeste Kidd and Benjamin Hayden, <i>Neuron</i> 88 (2015)	The psychology and neuroscience of curiosity
Aimee Stahl and Lisa Feigenson, <i>Science</i> 348 (2015)	Infants' learning from violated expectations
Garvin Brod, Marcus Hasselhorn, and Silvia Bunge, <i>Learning and Instruction</i> 55 (2018)	Prediction, surprise, and learning
Barbara Tizard and Martin Hughes, <i>Young Children Learning</i> (1984)	Children's questions at home and at school
Lev Vygotsky, "Play and Its Role in the Mental Development of the Child," <i>Soviet Psychology</i> 5 (1967)	Play and the child "a head taller"
Z. V. Manuilenko (1948), as reported in Daniil Elkonin's work on play	The sentry study
Rachel White and colleagues, <i>Child Development</i> 88 (2017)	The Batman effect
Adele Diamond, Steven Barnett, Jessica Thomas, and Sarah Munro, <i>Science</i> 318 (2007); Clancy Blair and Cybele Raver, <i>PLOS ONE</i> (2014); Dale Farran and Sandra Wilson (2014)	Tools of the Mind trials
Angeline Lillard and colleagues, <i>Psychological Bulletin</i> 139 (2013)	Evidence on pretend play
Narmada Sundararajan and Olusola Adesope, <i>Educational Psychology Review</i> (2020); Günter Daniel Rey, <i>Educational Research Review</i> (2012)	Seductive details
Samuel Ball and Gerry Ann Bogatz, <i>The First Year of Sesame Street: An Evaluation</i> (1970)	Early evaluation of Sesame Street
Melissa Kearney and Phillip Levine, <i>American Economic Journal: Applied Economics</i> 11 (2019)	Sesame Street and school outcomes
John Banas, Norah Dunbar, Dariela Rodriguez, and Shr-Jie Liu, <i>Communication Education</i> 60 (2011)	Humor in educational settings
Edward Deci, Richard Koestner, and Richard Ryan, <i>Psychological Bulletin</i> 125 (1999)	Rewards and intrinsic motivation
Judy Cameron and W. David Pierce, <i>Review of Educational Research</i> 64 (1994)	Dissenting meta-analysis on rewards
Uri Gneezy and Aldo Rustichini, <i>Quarterly Journal of Economics</i> 115 (2000)	"Pay Enough or Don't Pay at All"
Christopher Cerasoli, Jessica Nicklin, and Michael Ford, <i>Psychological Bulletin</i> 140 (2014); Joshua Howard and colleagues, <i>Perspectives on Psychological Science</i> 16 (2021)	Incentives, quality, quantity, and persistence

SOURCE	USE
Ruth Butler, <i>British Journal of Educational Psychology</i> 58 (1988)	Comments, grades, and intrinsic motivation
Caroline Pulfrey, Céline Buchs, and Fabrizio Butera, <i>Journal of Educational Psychology</i> 103 (2011)	Grades and performance-avoidance goals
Rohe and colleagues, <i>Mayo Clinic Proceedings</i> 81 (2006); Spring and colleagues, <i>Medical Education</i> (2011)	Pass/fail grading
Roland Fryer, <i>Quarterly Journal of Economics</i> 126 (2011)	Paying students for inputs and outputs
Steven Levitt, John List, Susanne Neckermann, and Sally Sadoff, <i>American Economic Journal: Economic Policy</i> 8 (2016)	Immediate, delayed, and loss-framed rewards
Teresa Amabile and Steven Kramer, <i>The Progress Principle</i> (2011)	Progress in meaningful work
Clark Hull, <i>Psychological Review</i> 39 (1932); Ran Kivetz, Oleg Urminsky, and Yuhuang Zheng, <i>Journal of Marketing Research</i> 43 (2006)	The goal gradient
Avraham Kluger and Angelo DeNisi, <i>Psychological Bulletin</i> 119 (1996)	Feedback interventions and the self
John Hattie and Helen Timperley, <i>Review of Educational Research</i> 77 (2007)	Four levels of feedback
Joseph Silverman and Alixandra Barasch, <i>Journal of Consumer Research</i> 49 (2023)	Streaks
Phillippa Lally, Cornelia van Jaarsveld, Henry Potts, and Jane Wardle, <i>European Journal of Social Psychology</i> 40 (2010)	Habit formation
Richard Landers, Kristina Bauer, and Rachel Callan, <i>Computers in Human Behavior</i> 71 (2017)	Leaderboards as goals
Katheryn Christy and Jesse Fox, <i>Computers & Education</i> 78 (2014)	Leaderboards and social comparison
Douglas Clark, Emily Tanner-Smith, and Stephen Killingsworth, <i>Review of Educational Research</i> 86 (2016)	Digital games and learning
Michael Hanus and Jesse Fox, <i>Computers & Education</i> 80 (2015)	Gamification over a semester
Teresa Amabile, William DeJong, and Mark Lepper, <i>Journal of Personality and Social Psychology</i> 34 (1976)	Deadlines and intrinsic motivation
Dan Ariely and Klaus Wertenbroch, <i>Psychological Science</i> 13 (2002, retracted); Hyndman and Bisin, <i>Psychological Science</i> (2026)	Deadlines evidence and its retraction

SOURCE	USE
Peter Gollwitzer and Paschal Sheeran, <i>Advances in Experimental Social Psychology</i> 38 (2006)	Implementation intentions
Hengchen Dai, Katherine Milkman, and Jason Riis, <i>Management Science</i> 60 (2014)	The fresh start effect
B. F. Skinner and Charles Ferster, <i>Schedules of Reinforcement</i> (1957)	Variable reward schedules
Alfred North Whitehead, “The Aims of Education,” in <i>The Aims of Education and Other Essays</i> (1929)	Epigraph

From Inquiry and Practice, by Anatoly Leczben. Working draft, October 2026.