Designing Instructional Materials for Coherence and Endurance

Why science units should be built to help students both build understanding and keep it

An overview for teachers, curriculum leaders, and instructional-materials designers

Victor Sampson · The University of Texas at Austin

SUGGESTED CITATION

Sampson, V. (2026). *Designing instructional materials for coherence and endurance: Why science units should be built to help students both build understanding and keep it* [White paper]. The University of Texas at Austin.

1. The premise

Every set of instructional materials we put in front of students carries, whether or not its designers make the commitment explicit, a theory of how people learn. A unit that opens with a puzzling phenomenon and asks students to explain it reflects one set of claims about learning; a unit that opens with a definition and a worked example reflects another. Because this is so, the most useful question we can ask of any curriculum is not whether a theory of learning is present, since the most thoughtfully designed materials plainly rest on one, but whether that theory draws on the *whole* of what we have learned about how people learn, or on only part of it.

This distinction matters because the science of learning has, over the past several decades, settled into two broad and complementary bodies of evidence, and a great deal of excellent contemporary practice draws deeply on one of them while drawing only lightly on the other. The first concerns **meaning-making and coherence**: how learners build new understanding from the ideas they already hold, how motivation and self-direction arise from the way a learning experience is arranged, how talk and collaboration function as engines of thinking, and how knowledge becomes usable when it is organized around a few powerful ideas. The second concerns not how understanding is first constructed but **whether it endures and proves usable over time**: the finding that memory is reconstructed rather than played back, so that spaced and effortful retrieval is what makes learning last; the finding that recall is cued and context-dependent, so that practice across varied conditions is what makes it reliable; and the recognition that transfer is a design problem rather than a hoped-for byproduct.

Both bodies of evidence are well established, and both are indispensable. A learner who never constructs meaning has nothing to hold onto later; a learner who constructs meaning but has no support to retain it is left, weeks afterward, with little to show for the work. The purpose of this paper is to introduce a model of unit design, the **Science of Learning Instructional Design (SOLID) framework**, that is built to honor both from the start, so that the considerable strengths teachers have already built into their practice are joined by a structure that carries the second body of evidence as well.

I want to be clear at the outset that naming this pattern is not a criticism of anyone's craft. That the most influential current approaches draw principally on the meaning-making evidence is a reasoned choice about what to privilege, not a failure to know the rest, and the tools that guide and evaluate

curriculum design have simply not yet asked designers to attend to endurance. SOLID is offered as a complement to strong current practice, not a replacement for it.

> *A unit can draw on the meaning-making evidence with great sophistication and still leave learning more fragile than its designers intended — not through any shortcoming of theirs, but because its structure was never asked to carry the second body of evidence.*

2. Two traditions, and the opportunity between them

The current landscape of science curriculum design is shaped by two mature bodies of design knowledge, and they divide along a consistent line. Approaches grounded in the vision of *A Framework for K–12 Science Education* (NRC, 2012) and realized through storyline units (Reiser et al., 2021), project-based instruction (Krajcik et al., 2008), and ambitious science teaching (Windschitl et al., 2018) draw with real care on the evidence concerning meaning-making and coherence. They begin from students' own questions about phenomena, they treat sensemaking as the central work of the classroom, and they organize instruction so that each step follows from the one before in a way that is coherent from the students' perspective. This is a genuine advance, and teachers who have invested in it have built something worth keeping.

A second and largely separate line of work, grounded in cognitive science, has focused on the conditions under which learning is retained and can be recalled and used after instruction ends. Its central findings concern a set of practice conditions that reliably improve long-term retention: retrieving information from memory rather than rereading it (the testing effect), distributing practice across time rather than massing it (the spacing effect), interleaving related skills rather than blocking them, and studying worked examples. These findings are largely content-general, which is the source of both their power and their limits: they describe properties of human memory that hold across subjects, and they specify the conditions under which learning endures without specifying what, in a given discipline, is most worth making endure.

The two traditions are not in conflict. They are answers to different questions — how understanding is built, and whether it lasts — and a unit that draws on only one of them is answering only one of the two questions the design of instruction must address. What the field has had few occasions to pursue is the design of units that draw on both from the outset, so that the same unit that helps students build understanding also helps that understanding last. That is the opportunity the SOLID framework takes up.

It is worth naming who bears the cost when a unit leaves endurance to chance. A student with well-resourced support outside of school — with the time, the help, and the materials to reconstruct a fragile idea after the unit has moved on — can absorb a curriculum that does not design for endurance. A student without those resources cannot, and is left to depend on whatever the classroom itself made durable. Designing for endurance is, in this sense, also designing for fairness, a point I return to near the close.

3. The ten principles, in plain language

Researchers describe how people learn in many idioms, but for the purpose of designing and evaluating a unit, ten principles capture what matters, and they fall naturally into the two aspects introduced above. I draw them from *How People Learn II: Learners, Contexts, and Cultures* (NASEM, 2018), and I do so for a specific reason. *How People Learn II* is the successor to the report that helped motivate the *Framework* itself, prepared by the same authoritative body, and it is notable precisely because it carries both aspects of the evidence with equal standing. It gives the findings about memory, retrieval, and transfer the same authoritative treatment it gives the findings about prior knowledge, motivation, and social interaction. The warrant for attending to both, in other words, is not one designer's construction but the field's own consensus.

Table 1 sets the ten principles out in plain language, with each principle paired with the design implication it carries, so that a principle named in the abstract is joined, in the same line, to the concrete decision it asks a designer to make. The first five concern how understanding is constructed; the second five concern whether it endures. This division reflects a real distinction in the underlying findings rather than a rhetorical convenience, and it organizes everything that follows.

Table 1. *The ten principles of learning and what each asks of a unit.*

Aspect of learning	Principle	What it asks of a unit
Meaning-Making and Coherence-Building	1. People interpret new information through the ideas they already hold.	Elicit students' current ideas first, then design new learning forward from them.
	2. Culture shapes what a community treats as useful or acceptable.	Help students grow into the standards of the scientific community rather than imposing them from outside.
	3. People reason through interaction with others, with ideas, and with the representations they build together.	Make talk and the collaborative refinement of representations the engine of every learning experience.
	4. Motivation is a condition a learning experience must create, not a trait a student simply brings.	Give students a genuine reason to take up a new idea.
	5. Self-regulation brings coherence to a developing body of ideas.	Build in occasions for students to monitor their own understanding and set goals from it.
Consolidation and Stabilization	6. Feedback from others, ideas, and artifacts is how a learner monitors an evolving understanding.	Embed assessment for learning inside the learning experience.
	7. The ideas a learner retains are those that prove useful.	Give students occasions to apply an idea beyond the setting in which it was first learned.
	8. Memory is reconstructed, not played back.	Space the reconstruction of core ideas across occasions separated in time.
	9. Memory is context-dependent and shaped by the cues present at recall.	Practice reconstruction across varied contexts and cues.
	10. Ideas are easier to use and remember when organized around a few core concepts.	Have students build representations of organized, connected knowledge.

The consolidation-and-stabilization principles are not an external supplement grafted onto the learning sciences to make a case. They are an equal part of the same report that grounds the meaning-making principles the field already designs so well. That is what licenses the claim at the center of this paper: a unit drawing on only the first five principles is drawing on one part of a broader account that the field's own consensus synthesis presents as a whole.

4. What the principles require when the content is science

The ten principles describe how human beings learn anything, and for that reason they cannot, on their own, tell a designer how to build a *science* unit. They establish that learning must be built from prior knowledge and made to endure through spaced reconstruction, yet they are silent on which ideas in science a student should hold prior knowledge about, or on what a scientific explanation must contain to count as evidence of understanding. To design for science, SOLID holds the ten principles fixed and asks what they require when the content is science, and it answers through four consensus reports that form a deliberate chain.

- *A Framework for K–12 Science Education* (NRC, 2012) supplies the **vision**: proficiency in science is three-dimensional, integrating disciplinary core ideas, science and engineering practices, and crosscutting concepts, and it is a goal for every student — which makes equity a design requirement built into the vision itself.
- The *Next Generation Science Standards* (NGSS Lead States, 2013) supply the **targets**: performance expectations that give concrete, assessable shape to three-dimensional proficiency.
- *Developing Assessments for the Next Generation Science Standards* (NRC, 2014) supplies the **logic of evidence**: measuring three-dimensional learning requires a system of varied assessment occasions that locate learners on a continuum and yield next steps for instruction, woven through instruction rather than added at the end.
- *Science and Engineering for Grades 6–12: Investigation and Design at the Center* (NASEM, 2019) supplies the **instruction**: investigation and design should be the central approach, on evidence that they increase understanding and outperform transmission-oriented methods.

Joined to the ten principles, these four reports yield the full warrant on which SOLID rests: the principles tell us how people learn anything, and the reports tell us what learning in that way requires when the thing being learned is science. This is what I mean by designing for **coherence and endurance in learning** — coherence names how understanding is constructed and held together from the student's perspective, while endurance names whether that understanding lasts once instruction has ended.

5. A practical starting point: how standards define a unit

A model of unit design does not begin from a blank page; it begins from a group of standards a course is obligated to address, and the way those standards are grouped is the practical starting condition any model must accommodate. SOLID is deliberately built to work with the standards as a district actually receives them, whether they arrive as the integrated performance expectations of the NGSS or as the separate content, practice, and recurring-theme statements used in states such as Texas and Florida. In every case a group of standards specifies disciplinary content most directly, while the practices and crosscutting or organizing ideas are things students are meant to *do* and to *use*. This is precisely why investigation and design sit at the center of the model: the group of standards defines the destination, the investigation is the vehicle, and the practices and organizing ideas are enacted en route rather than delivered as a separate strand.

A practical consideration reinforces this. In many states, assessments pair more than one expectation within a single item and do not disclose in advance which standards will be tested together or when. In Texas, a shift from a single end-of-year test to three shorter assessments across the year (beginning in 2027–28) compounds the uncertainty, since a standard may now be assessed well before a

coherence-driven sequence would have reached it. These conditions favor units each built around a coherent disciplinary group of standards that can be taught in any order, so the standards a given assessment happens to address can be taught before that assessment arrives. SOLID recovers coherence from the students' perspective not by composing the unit from standards gathered across the disciplinary order, but at the level of the investigation *within* a unit — which is where the next section begins.

6. A SOLID unit at a glance

A SOLID unit is not a collection of lessons but a sequence in which each stage does a distinct job, and the shape of that sequence is shown in Figure 1. Read from left to right, it reveals a deliberate loop rather than a straight line: the unit opens and closes with the same assessment structure, and the work between is arranged so that ideas are constructed, consolidated, checked, and stabilized in turn. It is the loop, more than any single lesson, that allows the structure to carry both aspects of the evidence base at once.

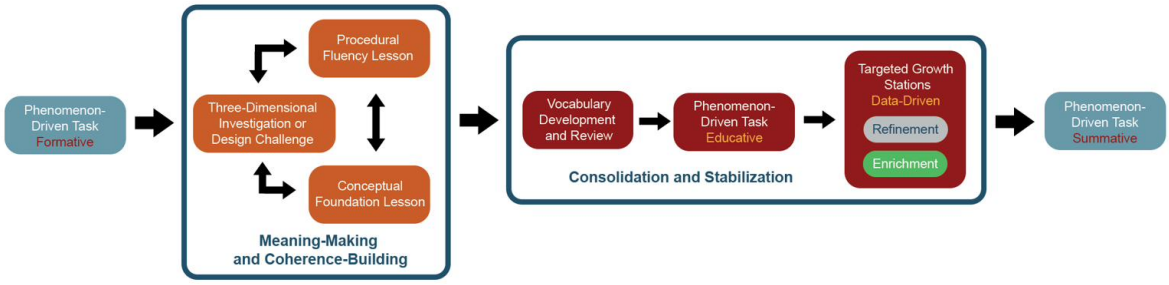


Figure 1. *The flow of a SOLID unit. A formative Phenomenon-Driven Task opens the unit and a summative one closes it. Between them sit two phases. In the Meaning-Making and Coherence-Building phase, a three-dimensional Investigation or Design Challenge serves as the integrative core, supported by a Conceptual Foundation lesson and a Procedural Fluency lesson. In the Consolidation and Stabilization phase, Vocabulary Development and Review, an educative Phenomenon-Driven Task, and data-driven Targeted Growth Stations (with Refinement and Enrichment pathways) help students consolidate and retain what they have learned.*

Two ideas make the rest of the structure legible. The first concerns the experience at the center of the unit. A three-dimensional investigation or design challenge functions as the **integrative core** — the experience that does a little of almost everything, because authentic inquiry and design naturally draw prior knowledge, motivation, collaborative reasoning, argument, and the beginnings of assessment into a single cycle. In this overview I draw the investigations and design challenges from the Argument-Driven Inquiry (ADI) instructional model, which meets the defining criterion of the core: the experience is three-dimensional, giving students the occasion to use science and engineering practices and crosscutting concepts together to make sense of disciplinary core ideas in order to figure something out.

The second idea concerns the relationship between that core and the other experiences in the unit. Because the investigation or design challenge does a little of everything, it masters nothing on its own, and each of the other lesson types is best understood as focusing on a single goal the core can only touch, and carrying it toward mastery. Touching is not the same as mastering, and it is precisely the consolidation-and-stabilization aspect of learning, the making of ideas that endure, that most needs a dedicated experience rather than a single pass within a larger cycle.

One commitment runs through every one of these experiences: they are all **social**. This follows directly from the third learning principle, that people reason through interaction with one another, with ideas, and with the representations they build and refine together. The social character of learning operates above all through two mechanisms that recur across the unit — **modeling**, in which a more expert performance makes a way of thinking visible, and **feedback**, in which what a student produces is returned to them in a form they can use to revise. When a later lesson is built around worked examples or the collaborative construction of a knowledge map, these are not departures from the social character of learning but particular arrangements of modeling and feedback in the service of a specific goal.

7. Why each component exists

Because each component earns its place by doing work the others cannot, the most helpful way to understand a unit is to consider each in turn — what it is, why the science of learning recommends it, and what it looks like in practice. I take them in the order a teacher meets them, and I draw the examples from a single eighth-grade physical science unit on the properties and changes of matter, built with the ADI model, so that the division of labor among the components can be seen within one coherent unit.

Before the individual components, a word on why the three lessons that sit between the investigation and the assessment thread exist as a family. Proficiency of any kind requires practice, and an investigation, however rich, offers only a single pass through the ideas it draws upon. A single pass is enough to construct an idea but not to make it fluent or lasting. The Conceptual Foundation lesson, the Procedural Fluency lesson, and Vocabulary Development and Review are the unit's means of turning a constructed idea into a durable one, and they work through spaced practice, retrieval practice, interleaving, self-explanation, and metacognition. The investigation comes first for a reason rooted in the first principle: because understanding is built from what a learner already knows, learning is most productive when it begins from a concrete, shared experience that later, more abstract work can build upon.

7.1 The three-dimensional investigation or design challenge: Focus on the integrative core

What it is. The core experience asks students to confront a phenomenon or design problem, propose an explanation or solution, plan and carry out an investigation or design process, gather and analyze data, and construct a written argument from the evidence they generate. In this overview these investigations and design challenges come from the ADI model, which gives the experience its structure while meeting the three-dimensional criterion that defines the core.

Why it matters. Because the work is open and authentic, the core recruits most of the ten principles within a single cycle. It elicits the prior ideas students bring (principle 1); it holds student work to the standards of the scientific community through argumentation (principle 2); it makes peer talk and the refinement of shared representations do genuine cognitive work (principle 3); it gives students a real reason to engage and runs a full plan–monitor–reflect loop that supports self-regulation (principles 4 and 5); and its argument stage turns the appraisal of a claim into an occasion for feedback and learning (principle 6). Precisely because it does so many jobs at once, the core cannot carry any single one to mastery — and that limitation, far from being a weakness, is the reason the remaining components exist.

What it looks like. In this unit, the ADI investigation asks students to figure out what happens to the total mass of a closed system when three substances are mixed inside it. Students begin with a

phenomenon, recording what they notice as substances combine inside a sealed system; they commit to an initial model of what will happen to the mass, building from prior knowledge before any instruction. They plan their own measurements and secure teacher approval of the plan, which operationalizes self-regulation; they collect and analyze data; and they draft an argument consisting of a claim, evidence, and a justification, then defend it in an argumentation session that enacts the standards of the scientific community. The principles here are not features added to the investigation; they are the joints of its structure.

7.2 The Conceptual Foundation lesson: Focus on meaning and structure

What it is. Where the investigation touches many principles, the Conceptual Foundation lesson concentrates its attention, moving students from an individual model, to a small-group model, to a gallery of other groups' models, to a revised model, and finally to the account the scientific community would offer.

Why it matters. It is a strong carrier of building from prior knowledge (principle 1), because it opens by asking students to articulate and explain an idea in their own terms before instruction, and of interaction and the refinement of representations (principle 3), because the revision genuinely depends on what students take from other groups. Its closing move, which sets the group's account beside the community's, carries the community-standards principle (principle 2), and its central work of relating ideas into a clean mental structure serves organized, connected knowledge at the scale of a single concept (principle 10). Because that opening articulation is itself an act of retrieval, the lesson makes an early contribution to reconstruction over time (principle 8).

What it looks like. In this unit, the Conceptual Foundation lesson asks students to model what happens to atoms during a chemical reaction. Students examine a chemical reaction equation and record what they notice; they build an individual model, using spheres to represent atoms; they consolidate individual models into a shared group model; they study other groups' models with the support of productive-talk stems and then revise their own; and only then do they compare their account with the one the scientific community offers. The revision step is the one on which the lesson turns, because it makes another group's reasoning a necessary input to one's own.

7.3 The Procedural Fluency lesson: Focus on durability of learning

What it is. If the Conceptual Foundation lesson builds the structure, the Procedural Fluency lesson makes a procedure fluent and lasting, and it is built almost entirely from the consolidation-and-stabilization aspect of the evidence base. It opens with a brief act of reconstruction; it pairs a worked example with the target procedure; it fades support from a demonstrated solution, to a solution a student talks a partner through, to independent practice; it asks students to diagnose worked solutions that contain errors; and it closes with self-paced retrieval practice on a graduated ramp of difficulty.

Why it matters. It is the primary carrier of reconstruction over time (principle 8), because its retrieval practice returns students to the procedure across occasions, and it serves context-dependent, cued reconstruction (principle 9) as squarely, because the graduated ramp varies the cues under which the procedure must be rebuilt. Its worked examples and error diagnosis are arrangements of modeling and feedback, and because the ramp moves from familiar to unfamiliar problems, it engineers near and mid-range transfer (principle 7). It is here, more than anywhere else in the unit, that the endurance of learning is not merely present but deliberately engineered.

What it looks like. In this unit, students learn to identify the atoms on each side of a chemical reaction equation and determine whether mass is conserved. They first retrieve and explain, in their own words, the idea that mass is conserved; they study a fully worked example; they solve a similar

problem with a partner; they work through problems on a shared placemat, solving each individually before reconciling as a group; they analyze two erroneous worked solutions, deciding for each whether an error was made and what should have been done instead; and they close with self-paced practice drawn from problems labeled mild, medium, and spicy. Endurance here is not a matter of repetition but of a sequenced fading of support across a widening range of cues.

7.4 Vocabulary Development and Review: Focus on knowledge structure

What it is. Late in the unit, after students have met the full vocabulary of the topic in context, Vocabulary Development and Review asks them to map every key term into a single representation of terms, pictures, and labeled linking words — first individually, then as a group, then individually again on the strength of the group's input. It is concept mapping in all but name.

Why it matters. It is a strong carrier of organized, connected knowledge (principle 10), but at the scale of the whole unit rather than a single concept, asking not what one idea means but how all of the unit's ideas connect; the requirement that each link be *named* forces the relationship between two ideas into the open, which is what distinguishes a map from a glossary. Because it arrives near the end of instruction and requires every term to be pulled from memory, it does a second job as a cumulative reconstruction-and-integration event, quietly serving reconstruction over time (principle 8) — which is why the model places it within the Consolidation and Stabilization phase.

What it looks like. In this unit, the review lesson gathers terms ranging from element, compound, and mixture to conservation of mass, chemical reaction equation, coefficient, and subscript into a single connected representation. Each student first builds a representation showing what the terms mean and how they relate, using pictures, definitions, and linking words; the group consolidates these into a shared representation; and each student then produces a final one on the strength of the group's work. The labeled linking words are the heart of the design, because a link a student must name forces a relationship into the open.

7.5 The Phenomenon-Driven Task thread: Focus on spacing and assessment as learning

What it is. The same task structure appears three times across the unit, varying only its phenomenon: a formative administration before instruction, an educative administration at the midpoint, and a summative administration at the end. What looks like repetition is in fact the unit's principal mechanism for endurance. The first and third administrations sit outside the learning process, making student thinking visible to the teacher — once to set a baseline and inform placement, once to measure how far learning has carried. It is the second, educative administration that does the pedagogical work.

Why it matters. Because the thread's target ideas are retrieved on more than one occasion separated in time, it creates the spaced reconstruction under which memory consolidates, making it a primary carrier of reconstruction over time (principle 8); because each task asks students to explain their reasoning rather than select an answer, the task itself teaches and returns feedback (principle 6); and because each administration presents a fresh phenomenon, the same ideas must be reconstructed under a new cue, which serves context-dependent reconstruction (principle 9) and presses gently on transfer (principle 7). The thread is the only true spacing mechanism in the unit, which is why the three tasks are deliberately distributed across it.

What it looks like. In this unit, the Phenomenon-Driven Task presents students with observable changes in matter — the behavior of water in a graduated cylinder, or the outcome of combining substances — and asks them to classify what they see and explain their reasoning using models and the properties of the substances involved. The task asks for explanation rather than selection, and the same structure is administered three times, each built around a different set of changes in matter. The

structural fact that matters is the spacing: the three administrations are deliberately separated, so that each is a retrieval occasion as much as a measurement, and the second is positioned to feed the reteaching that follows it.

7.6 The Targeted Growth Stations: Focus on responding to assessment

What it is. Between the educative and summative tasks sit a set of self-checking stations that reteach the specific ideas the educative task has brought into view, organized into two pathways — **Refinement** for students who need a further pass, and **Enrichment** for students ready to extend. The stations are selected by the evidence the educative task provides rather than assigned uniformly, which is what makes the component a response to assessment rather than another round of undifferentiated practice.

Why it matters. This is the component that closes the loop, turning the task thread from an assessment *of* learning into an assessment *for* learning. It is a primary carrier of assessment as feedback (principle 6); because each station offers immediate, idea-specific feedback, it adds a further spaced occasion to reconstruct an idea under a new cue (principles 8 and 9); and because it meets students wherever the educative task finds them, it supports motivation and self-regulation by offering work matched to where a student actually is (principles 4 and 5). It is also the component at which the model's commitment to equity becomes most concrete.

What it looks like. Following the educative task in the changes-of-matter unit, the educative task sorts ideas into those a student holds securely, such as the conservation of mass, and those still forming, such as identifying the atoms on each side of an equation. Each Refinement station addresses one still-forming idea with immediate, idea-specific feedback, while an Enrichment station offers students who are secure a problem that reaches toward a new context. The design signature is that the stations are selected by the data rather than assigned by a label — which distinguishes this component both from undifferentiated practice and from differentiation as it is conventionally understood.

8. Coverage at a glance

The account developed across Section 7 can be summarized in a single display. Table 2 arrays the ten principles against the six components so that the division of labor can be seen at a glance. Reading down a column shows what one component contributes; reading across a row shows which components carry a given principle. *Strong* marks a primary and deliberate carrier of a principle, *Partial* a meaningful secondary contribution, *Near* (on the transfer row) an engineered near- or mid-range transfer, and a dash a principle a component is simply not built to serve. These are analytic judgments about a designed structure, offered to make the model's internal logic inspectable rather than as measurements of what a lesson produces in a classroom.

Two patterns carry the argument. The first is that every principle has at least one Strong carrier, with the single exception of transfer, which the table marks as the unit's open frontier rather than a job left undone. The second becomes visible when the table is read by aspect rather than by component: the five consolidation-and-stabilization principles are carried most strongly by exactly the components with no close equivalent in current meaning-making practice — the Procedural Fluency lesson, Vocabulary Development and Review, the Phenomenon-Driven Task thread, and the Targeted Growth Stations. That alignment is the heart of the matter. A SOLID unit keeps the meaning-making aspect the field already designs well and adds the consolidation-and-stabilization aspect that prevailing approaches, and the instruments used to evaluate them, have had little occasion to require.

Table 2. *Coverage of the ten principles across the six components of a SOLID unit.*

Learning principle	Investigation / Design Challenge	Conceptual Foundation	Procedural Fluency	Vocabulary Dev. & Review	Phenomenon-Driven Task thread	Targeted Growth Stations
1. Build from prior knowledge	**Strong**	**Strong**	Partial	Partial	**Strong**	Partial
2. Culture and community standards	**Strong**	**Strong**	Partial	Partial	—	Partial
3. Interaction and refinement of representations	**Strong**	**Strong**	**Strong**	**Strong**	—	**Strong**
4. Motivation	**Strong**	Partial	Partial	—	Partial	**Strong**
5. Self-regulation	**Strong**	Partial	**Strong**	Partial	Partial	**Strong**
6. Assessment as feedback for learning	**Strong**	Partial	Partial	—	**Strong**	**Strong**
7. Transfer (usefulness)	Near	—	Near	—	Near	Near
8. Reconstruction over time (spacing)	Partial	**Strong**	**Strong**	Partial	**Strong**	**Strong**
9. Context-dependent, cued reconstruction	Partial	Partial	**Strong**	Partial	**Strong**	Partial
10. Organized, connected knowledge	Partial	**Strong**	Partial	**Strong**	Partial	Partial

Note: On the Phenomenon-Driven Task thread, the ratings reflect the educative (midpoint) administration, the one positioned inside the learning process; the formative and summative administrations sit outside it and make thinking visible to the teacher. Taken together, the three administrations form the spaced thread on which the unit's reconstruction-over-time depends.

9. The one honest limit: far transfer

It would be easy, and it would be a disservice, to claim that the model delivers all ten principles with equal completeness, and naming the one place where it does not is part of describing it honestly. Of the ten principles, transfer is the one a unit serves least fully. The distinction that matters is between **near transfer**, the application of an idea to a new problem that resembles the one in which it was learned, and **far transfer**, the application of an idea in a genuinely different domain, at a considerable remove in time, and without prompting. The model builds deliberately for near and mid-range transfer — most visibly in the Procedural Fluency lesson's graduated ramp and in the Phenomenon-Driven Task thread's successive administrations. Far transfer is a different matter.

The limit is not that far transfer is mysterious or that the means of engineering it are unknown. It is a straightforward consequence of scope. A unit, by definition, works within a single disciplinary domain, and far transfer requires two domains in order to occur at all: a setting in which an idea is learned and a genuinely different setting in which it is later applied. A single unit does not contain the second setting, and so it cannot, on its own, occasion the far transfer of the very ideas it teaches. This is a structural feature of the grain size at which a unit operates, not a shortfall in the design of any lesson within it.

The opportunity for far transfer therefore lives at the scale of the **course** rather than the unit, and the architecture for it is already present in the vision the field has adopted. Because the same crosscutting concepts and science and engineering practices recur from one unit to the next, every later unit is a potential second domain into which the disciplinary core ideas of an earlier unit can be deliberately drawn. A model of conservation of mass built to explain a chemical reaction in one unit can be summoned again, weeks later, to reason about an unfamiliar system in another. The move that turns

this latent structure into genuine far transfer is a deliberate one — the pulling-forward of earlier core ideas into the reasoning demands of later units — and it is a decision about the sequencing of a whole course. I name this frontier not as a concession but as an invitation: locating the limit precisely, at the boundary of the unit rather than in the craft of the lesson, is what lets a curriculum be designed across units for the far transfer that no unit can engineer alone.

10. Designing for endurance is designing for equity

Equity is woven through the model rather than added to it, because the *Framework* commits to three-dimensional proficiency for every student, and a model that takes the *Framework* as its vision inherits that commitment as a design requirement. Several of the model's features are equity moves in their own right.

The first and most general is the attention to whether learning endures. As noted earlier, a student able to reconstruct a fragile idea at home can absorb a curriculum that does not design for endurance; a student without those resources cannot. A model that builds consolidation and stabilization into the unit, rather than leaving it to be supplied afterward, distributes the benefit of instruction most widely to the students least able to supplement it elsewhere.

The second concerns how access is provided within a learning experience. Access to a rigorous task can be distributed unevenly, as accommodations triggered by a student's label, or built into the task itself as features available to every student. There is good reason to prefer the second: the labels by which access is conventionally routed often encode the assumption that instruction should be matched to a preferred learning style, an assumption repeatedly found to lack empirical support. When the investigation provides a planning organizer, the Conceptual Foundation lesson supplies productive-talk stems, and the Procedural Fluency lesson offers worked examples and staged practice, these supports are features of the task available to every student. Every student works on the same rigorous, three-dimensional science; what varies is not the task but the resources each student draws on to engage with it.

It is against this stance that the Targeted Growth Stations must be examined, because they are the one component in which students are routed, on the basis of evidence, to different work — and a careful reader will rightly ask how that differs from the differentiation the stance sets itself against. The answer turns on three distinctions. First, the stations are driven by **evidence rather than labels**: a student is directed to a Refinement station not because of a classification attached to them, but because a specific idea, identified by that student's own response to the educative task, is still forming. Second, the routing is **temporary and diagnostic** rather than fixed, addressing a particular idea at a particular moment rather than assigning a student to a track. Third, and most importantly, the stations do not offer easier or lesser versions of the science: a Refinement station reteaches a still-forming idea to the same standard the unit holds for everyone, and the Enrichment path extends rather than replaces that standard. What the stations distribute unevenly is not the rigor of the work but the specific feedback each student needs, which is exactly what an assessment designed to inform instruction is for.

Finally, a word on the relationship between this model and the body of work that has placed justice at the center of science teaching. That work argues that equity requires more than access to rigorous content — it requires sustaining students' cultural and linguistic identities and drawing on community knowledge as a resource for science. SOLID is not a substitute for that work. Its contribution is structural: a unit design that carries the full range of how people learn without sorting students by label. That is itself a condition of equitable instruction, even as it is not the whole of it. A unit built on

this model still requires a designer and a teacher to choose phenomena that matter to the students in the room and to treat what students bring as an asset rather than a deficit.

11. What this means for you

The value of the model is finally practical, and it looks a little different depending on where you sit.

For teachers

A SOLID unit gives you a structure in which each lesson type has a clear and distinct job, so that the effort you already invest in helping students make sense of phenomena is joined by lessons designed to make that understanding last. You can expect the investigation to do the heavy lifting of sensemaking, the Conceptual Foundation and Procedural Fluency lessons to turn what students construct into something clean and durable, and the Phenomenon-Driven Task thread and Targeted Growth Stations to tell you — before the summative measure — exactly which ideas still need attention and for whom. Students will struggle at first with tasks that ask them to reason and explain rather than recall, as students always do with unfamiliar work; that struggle gives way to fluency with practice, feedback, and opportunity.

For curriculum leaders and adopters

The coverage analysis in Table 2 is meant to double as a checklist. When you evaluate materials, it is worth asking not only whether a unit anchors in a phenomenon and builds coherence from the students' perspective — which strong current materials already do — but also whether it spaces the reconstruction of its core ideas, varies the cues under which students rebuild them, diagnoses and reteaches before its summative measure, and builds procedural fluency deliberately. A unit can earn top marks against a familiar rubric and still leave learning more fragile than intended, precisely because those questions belong to the consolidation aspect the rubric was not built to ask about.

> *Units organized around coherence from the students' perspective are one faithful route to the vision of the Framework. A unit built to draw on both bodies of the learning-sciences evidence is another — distinguished by building the endurance of learning in from the start rather than leaving it to be supplied afterward.*

For instructional-materials designers

The model offers a way to realize the vision of the *Framework* that builds endurance into a unit from the start rather than leaving it to be supplied by teachers afterward. The diagnose–reteach–remeasure logic of the Phenomenon-Driven Task thread and the Targeted Growth Stations shows how the system of classroom-embedded assessment the consensus reports call for can be built into the architecture of a unit rather than added around its edges, with the midpoint task positioned so that what it reveals actually redirects instruction before the summative measure.

A note on the status of the claim is owed to every reader. This is a design argument: it establishes that a unit built on this model would draw on both bodies of the learning-sciences evidence, not that doing so demonstrably improves outcomes. The validation the argument most needs is empirical — a study comparing retention, measured weeks after instruction, between units designed for meaning-making alone and units designed, on this model, for coherence and endurance together. Until that study is

done, the case is a case for the plausibility and the warrant of the design, and it should be read as an invitation to that test.

To design this way is to design for coherence and for endurance at once, and it is, for the same reasons, to design for equity, because a unit that makes learning last without sorting students by label extends the benefit of instruction most fully to the students least able to secure it elsewhere. If the field is to realize the vision it has set for itself, for every student, then the design of its units will need to draw on the whole of what is known about how people learn. The work that remains is to build such units, to study whether they deliver what the design intends, and to refine the model against what that study reveals.

Selected references

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. Psychological Bulletin, 128(4), 612–637.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. Psychological Bulletin, 132(3), 354–380.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques. Psychological Science in the Public Interest, 14(1), 4–58.

Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. Science Education, 92(1), 1–32.

National Academies of Sciences, Engineering, and Medicine. (2018). How people learn II: Learners, contexts, and cultures. The National Academies Press.

National Academies of Sciences, Engineering, and Medicine. (2019). Science and engineering for grades 6–12: Investigation and design at the center. The National Academies Press.

National Research Council. (2012). A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. The National Academies Press.

National Research Council. (2014). Developing assessments for the Next Generation Science Standards. The National Academies Press.

NGSS Lead States. (2013). Next Generation Science Standards: For states, by states. The National Academies Press.

Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them. Florida Institute for Human and Machine Cognition.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. Psychological Science in the Public Interest, 9(3), 105–119.

Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. Journal of Science Teacher Education, 32(7), 805–829.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science, 17(3), 249–255.

Windschitl, M., Thompson, J., & Braaten, M. (2018). Ambitious science teaching. Harvard Education Press.