



Practicing Connections: A Framework to Guide Instructional Design for Developing Understanding in Complex Domains

Laura Fries¹  · Ji Y. Son²  · Karen B. Givvin¹  · James W. Stigler¹ 

Published online: 17 August 2020

© The Author(s) 2020

Abstract

Research suggests that expert understanding is characterized by coherent mental representations featuring a high level of connectedness. This paper advances the idea that educators can facilitate this level of understanding in students through the *practicing connections* framework: a practical framework to guide instructional design for developing deep understanding and transferable knowledge in complex academic domains. We start by reviewing what we know from learning sciences about the nature and development of transferable knowledge, arguing that *connectedness* is key to the coherent mental schemas that underlie deep understanding and transferable skills. We then propose features of instruction that might uniquely facilitate deep understanding and suggest that the connections between a domain's *core concepts*, *key representations*, and *contexts and practices of the world* must be made explicit and practiced, over time, in order for students to develop coherent understanding. We illustrate the practicing connections approach to instructional design in the context of a new online interactive introductory statistics textbook developed by the authors.

Keywords Learning in complex domains · Instruction · Learning theory · Statistics education · Instructional design · Transfer

Introduction

As we move rapidly through the twenty-first century, our goals for education become more and more ambitious. Although there may have been a time when rote learning of facts and procedures was sufficient as an outcome for education, that is certainly not the case today.

✉ James W. Stigler
Stigler@psych.ucla.edu

¹ Department of Psychology, University of California at Los Angeles, Box 951563, CA 90095-1563 Los Angeles, USA

² Department of Psychology, California State University, Los Angeles, CA, USA

Anyone with a phone can Google to find facts that they have forgotten. But gaps in thinking and understanding are not easily filled in by Internet searches. Increasingly, we value citizens who can think critically, coordinate different ideas together, solve novel problems, and apply their knowledge in all kinds of situations that do not look like ones they have previously encountered. In short, we want to produce students with deep understanding of the complex domains that constitute the modern knowledge landscape (National Academies of Sciences, Engineering, and Medicine 2018).

Although we have a long tradition of research in the learning sciences related to the problem of how to teach for understanding and transfer, there still is a wide gap between the research, on one hand, and the design and implementation of educational programs, on the other (Hiebert et al. 2002; Lagemann and Shulman 1999; Levin and O'Donnell 1999; Robinson 1998; Strauss 1998; Toth et al. 2000). There are likely a number of reasons for this gap, one of which must surely be the conditions under which much research is carried out. Understanding and deep domain knowledge develop slowly over long periods of time, especially for those domains that are hard to learn (Ericsson 2006). Research, on the other hand, is often carried out in laboratories (Richland et al. 2007), where undergraduate participants are generally available for an hour or so.

Because of this constraint, much of the research has focused on students' learning of independent facts and procedures—"bits" of knowledge—that can be mastered and tested over relatively short periods of time (e.g., Goldwater and Schalk 2016). We have learned a lot from this research; in particular, we know a lot about how to help students remember bits of knowledge. We know much less, however, about how to help students connect the bits together into a coherent and flexible representation of a complex domain (van Merriënboer 1997). And there is mounting evidence that just mastery of isolated pieces of information is not by itself enough to produce the kind of flexible domain understanding typical of experts (National Academies of Sciences, Engineering, and Medicine 2018; van Merriënboer 1997).

The over-emphasis on mastery of bits of information has been exacerbated by the advent of modern learning technologies, especially adaptive learning platforms. These platforms work by modeling domains as large numbers of skills and concepts to be mastered, then using sophisticated algorithms to provide learning resources to students based on their mastery of prior skills and concepts in an assumed learning progression (Tseng et al. 2008). Although such technologies can be an important part of the toolbox available to educators and instructional designers, we see students in our own classes who master all the bits (or micro-learning objectives) but fail to understand the domain in a deep way. Experts do not see their domains in terms of bits; they see the underlying structure of the domain that makes their knowledge flexible and transferable (Bransford and Stein 1984; Ericsson, Hoffman, & Kozbelt, 2018; Ginsburg 1977; Hiebert and Carpenter 1992).

Our goal in this paper is to present a framework—the *practicing connections* framework—to guide instructional design for understanding in complex domains. We base our framework on research in the learning sciences but find that we must go beyond the current research literature to get all the way to understanding. For example, although much research focuses on learning in short periods of time, we take as our challenge the design of learning experiences over longer periods of time such as those that make up the typical semester-long college course. And where research often focuses on specific variables, such as cognitive load, or spacing of practice, we are interested in starting with a detailed analysis of what it means to be an expert in a domain and then in charting the pathway(s) students might take to get there.

Other instructional design frameworks take this latter idea seriously, designing pathways that take students from novice all the way to expert participant in the authentic practices of a domain (e.g., Koedinger et al. 2012; van Merriënboer et al. 2002). One such framework is the *four component-instructional design* (4C/ID) system (van Merriënboer et al. 2002). The 4C/ID system starts with an analysis of the skills that comprise expert performance in a complex domain. It then guides the instructional designer through a series of design decisions, specifying the learning tasks, supportive information, just-in-time information, and part-task practice hypothesized to lead to expert performance. The 4C/ID framework is solidly grounded in research on how people learn (Sarfo and Elen 2007; Susilo et al. 2013).

Although frameworks such as 4C/ID are an excellent starting point, they tend to apply especially well to the development of complex *skills*. Many complex academic domains, however, require students not only to master complex skills but also to learn to reason with the abstract ideas that form the conceptual structure of the domain (van Merriënboer et al. 2002). Domains such as mathematics, science, and statistics are examples. Part of the domain of statistics, for example, is the complex set of skills that make up the practice of data analysis. But just mastery of these skills is not enough to ensure students understand the core concepts of the domain or that they can use statistical concepts in thinking across a variety of contexts (Richland et al. 2012; van Merriënboer et al. 2002).

The focus on skills over concepts has been noted as a trend in education that cuts across the normal ideological boundaries. On one hand, some educators define learning primarily as the mastery of a set of isolated skills, focusing primarily on students' repetitive practice of basic facts and step-by-step procedures (e.g., Stigler and Hiebert 2009). On the other hand, proponents of constructivist pedagogies focus more on embedding skills development within the authentic practices that define expert performance in a domain, assuming that the requisite conceptual understanding will evolve naturally from such participation (e.g., Savery and Duffy 1995; van Merriënboer et al. 2002). Neither end of the continuum, however, focuses on the teaching of a domain as a "body of knowledge" that is organized by principles, concepts, and theories (Handelsman et al. 2004; Hodson 1988; Kirschner et al. 2006).

In the *practicing connections* framework, we propose a complement to frameworks such as 4C/ID by setting core concepts and domain structure on an equal footing with routines designed to support the practice of increasingly complex and transferable skills. From our own experience as teachers, we have become convinced that there is a need, at least in complex domains in which knowledge develops slowly, to connect the development of skills to the deliberate teaching of core concepts and representations that underlie the domain. Gradual improvement in understanding of the core concepts of a domain makes knowledge more coherent, which in turn makes it more flexible and transferable to new situations (Hatano and Inagaki 1986; Richland et al. 2012).

Overview of the Paper

In the remainder of this paper, we outline our *practicing connections* framework. The ideas we present have been developed in the context of a project in which we have been designing, developing, and implementing an online interactive textbook for introductory statistics (Stigler et al. 2020). This paper is an attempt to make explicit the basis on which we are making design decisions and to justify those decisions based on theories and findings from the learning sciences. The examples we use to illustrate our points will be drawn from our online textbook.

We will start by discussing what we mean by understanding, and the role that understanding plays in the development of transferable knowledge. We argue, based on research, that transferable knowledge is primarily characterized by mental schemas with a high level of coherence and connections.

We then move on to the question of how we can help students develop connected and coherent knowledge in a domain. Our answer, in short, is that students must be provided opportunities to *practice making connections*.

In addressing the question of which connections in a domain are most important to practice, we focus on those that help support deep understanding and transferable knowledge. Specifically, we propose three types of connections that should be the focus of instructional design: *the world* (including practices and contexts); *core concepts*, which organize the domain; and *key representations* used for communication and thinking within a domain. Our hypothesis is that all of these need to be connected.

Then, having decided on which connections are most important, we lay out the kinds of experiences that will help students construct and strengthen these connections. We propose, based on our reading of the research, three principles to be considered when designing pedagogy: making connections explicit for students (as opposed to relying on self-discovery), engaging students in the hard work of incorporating the connections into their developing understanding of the domain, and providing opportunities for students to engage repeatedly in this process over time as they deepen and extend their domain knowledge.

Understanding as Connected Knowledge

We have made the point that our interest is in the development of understanding and transferable knowledge in complex domains. We want students to develop robust and flexible knowledge that they can take with them into the world, coordinate with other knowledge, and apply to new problems and in new contexts. But what, exactly, is understanding? What makes knowledge broadly transferable? A useful place to start is with the large and growing body of research on the nature of expertise and expert performance (Ericsson et al. 2018; Ericsson et al. 2006; Ericsson and Pool 2016; Hatano and Inagaki 1986). What we want for our students is similar to the domain knowledge we see among experts, which, on the whole, can be characterized as coherent, connected, and relational.

Presumably, after some instruction, novices will achieve some level of expertise. But Hatano and Inagaki (1986), in a well-known paper, proposed a distinction between *adaptive* and *routine* expertise. *Routine experts* have a lot of knowledge, both declarative and procedural, which they have mastered well enough to perform with fluency up to some standard in familiar circumstances. In contrast, *adaptive experts* stand out for their ability to flexibly apply their knowledge in a wide range of contexts, both familiar and novel.

In some domains, routine expertise is sufficient, because conditions are stable and feedback is reliable (e.g., domains that Epstein 2019, refers to as “kind,” such as chess or music). But our focus here is on domains Epstein refers to as “wicked.” Wicked domains, which in our view include most complex academic domains, are marked by changing contexts and inconsistent or ambiguous feedback (Epstein 2019). Arguably, the aim of instruction is to create adaptive expertise which is presumably driven by what education psychologists call “transferable knowledge” (Bransford and Schwartz 1999; Greeno et al. 1993; Renkl et al. 1996). But well-documented failures of transfer in academic contexts (Bransford and Schwartz 1999;

Stigler et al. 2010) suggest that many common instructional practices leave students—at best—as routine experts, critically unable to adapt their knowledge to new circumstances.

In general, research on the nature of expert understanding paints a consistent picture: adaptive experts' knowledge is organized in a different way (Bilalic and Campitelli 2018; Carbonell et al. 2014; Chase and Simon 1973; Chi 2011; Chi and Koeske 1983; de Groot 1965; Ericsson and Charness 1994; Ericsson et al. 2018). Adaptive experts' knowledge organization is more coherent, interconnected, and reflective of the relational structure of the domain (Bilalic and Campitelli 2018; Carbonell et al. 2014; Kellman et al. 2010; McKeithen et al. 1981).

The building blocks of this transferable knowledge, sometimes referred to as *schemas*, emphasize the connections between abstract relations (such as hierarchies, embedded categories, and functional systems) rather than lists of discrete facts and procedures (Bedard and Chi 1992; Chi et al. 1981; Chi et al. 1982; Ericsson et al. 2018; Kellman et al. 2010; North et al. 2011). For example, expert physicists' schemas for a range of problem situations are organized by fundamental relationships (e.g., the work energy principle) rather than by superficial surface-level details specific to problem contexts (Chi et al. 1981). Experts also have fewer and more interconnected schemas that encompass more instances (Lachner et al. 2012).

Because their knowledge is highly organized and interconnected, adaptive experts perceive the world differently than routine experts or novices. They easily and quickly identify and attend to the relevant structural information critical to understanding the situation at hand, filtering out the irrelevant features of a problem to home in on a solution path (Campitelli and Gobet 2005; Endsley 2018). They anticipate how modifications to a system will influence outcomes and can explain why and how concepts from one scenario may apply to another (Carbonell et al. 2014; Hatano and Inagaki 1986; Holyoak 1991). Their knowledge structures prioritize connections among concepts, examples, and contexts (Bransford and Stein 1984; Ericsson et al. 2018; Ginsburg 1977; Hiebert and Carpenter 1992), such that adaptive experts are able to efficiently chunk information, leaving more working memory resources available (Ericsson 2018; Ericsson and Kintsch 1995; Glaser and Chi 1988). If information were simply stored as discrete entries in a mental list, lacking structural connections to lend coherence, it would be difficult to pick out relevant information and coordinate knowledge efficiently (e.g., Reed 1985). But because expert schemas encode a connected representation of the domain, the knowledge is rendered transferable (Bedard and Chi 1992; van Merriënboer 1997).

For our purposes, we equate deep understanding with the transferable knowledge of adaptive expertise. Thus, we offer the following characterization: understanding is characterized by the ability to perceive and make explicit the underlying structure of a domain, its connections and relations in the form of coherent mental schemas, which allows for transferable and flexible application of domain principles. (The transfer of principles is regarded as the highest form of transfer; see Barnett and Ceci (2002) for an in-depth discussion of the range of types of transfer.)

It is clear from this characterization of understanding that connections are important in order for knowledge to be transferable. Consider, as a thought experiment, the two individuals represented in Fig. 1, below. They both have the same pieces of knowledge, represented as A, B, C, D, E, F, G, and H, and they both are working to solve the same novel problem, the solution to which requires them to access and apply one specific piece of knowledge, B. Because the individual on the left has few connections among his bits of knowledge, he will not be able to transfer what he knows to this new situation. As he works to understand the problem, he keeps coming up with E (his strongest connection). He also tries to apply D. But

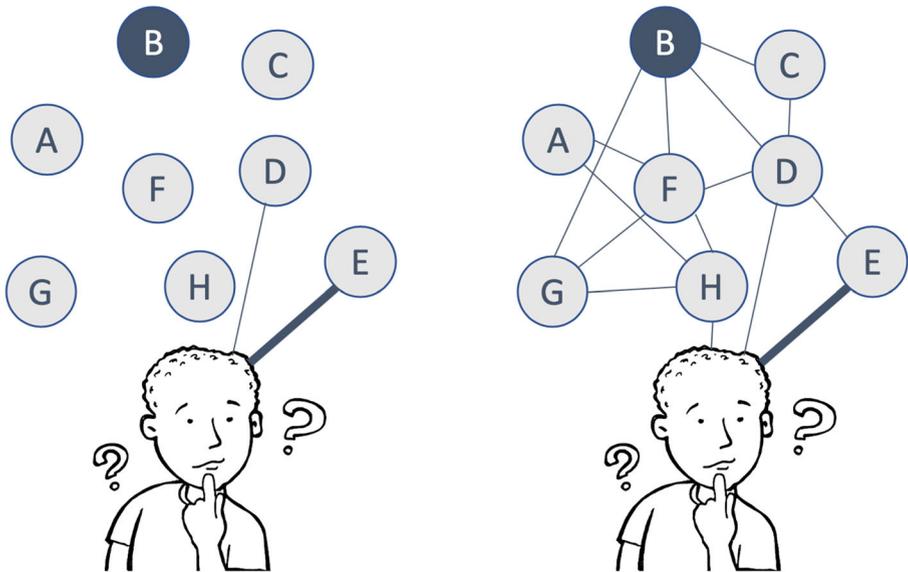


Fig. 1 The importance of connected knowledge to achieving transfer

he has no way to get to B, even though he “knows” it, because it is not connected with the other things he knows or with the problem he is currently facing.

The individual on the right (we can call him Mr. Right) approaches the same novel problem from a much stronger position. Like Mr. Left, his first thought is to retrieve and apply E to the situation at hand. He might then, like Mr. Left, think again and end up at D. But from D, he has many more options because of the interconnected nature of his knowledge. He could go from D right to B and then solve the problem. He could go from D to F and then B, from D to F to G to B, and so on. The point is that he has many ways to get to the knowledge he needs. Once he has arrived at B, he must, of course, coordinate B with other pieces of knowledge in order to craft a solution to the problem. He must also adjust his solution, once identified, to fit the unique context in which this novel problem presents itself. Because his knowledge is interconnected, we would say that he has understanding. Based on this understanding, he can retrieve, coordinate, and adapt what he knows to most any situation.

Much of education focuses on teaching students the “bits” of knowledge and skills—the A, B, C, and so on. One needs not go further than a standard math textbook to see lack of connection from one chapter to the next, much less connections made across content addressed at different grade levels. Although the bits are important to learn, and some of the bits require considerable time and effort to master, just learning the bits does not lead to understanding and transfer. Our practicing connections framework takes a different approach: instead of focusing our attention solely on the bits, we are trying to find ways to help students create the connections between the bits, with the goal of producing understanding and transferable knowledge.

We are not the first to propose this focus. Goldwater and Schalk (2016) proposed that relational categories, as distinct from the feature-based categories that are often the subject of cognitive psychology research, might be a bridge between learning sciences research and

education. Similarly, Schwartz and Goldstone (2015) suggest a “coordination” approach to learning, in which the goal of instruction is to strengthen skills in relation to other skills, not siloed in isolation. Skills practiced in isolation are destined to remain in isolation; even worse, they can become an obstacle to new learning (Woltz et al. 2000). Schwartz and Goldstone have characterized this kind of coordinated learning as “teaching the brain to dance.” Much of what we want students to understand in academic domains is relational.

We have argued for the value of understanding and transfer and that these things are the result of interconnected knowledge. What we have not yet addressed is how we expect students to gain that knowledge. Although we seek to move away from the learning of bits, the mechanism through which those bits are traditionally learned is one we value: practice. We will expand below on the type of practice we envision, but for now, our hypothesis, simply put, is this: students should *practice making connections*. If the goal of instruction is to support students in their practice making connections, two questions emerge, which we address in the rest of this paper: (1) What are the connections that need to be practiced? Obviously, this will be highly dependent on the domain, but we offer some principles to guide the instructional designer; and (2) how can we design instruction to give students more opportunities to practice the key connections that have been identified for the domain? In the following sections, we lay out our framework to answer these questions, illustrating our points with examples from the domain of introductory statistics.

Practicing Connections: What Connections to Practice?

We begin with connections. It is easy to make the case that students’ knowledge should be interconnected and coherent. But just saying this skirts an important question: Which connections, of all the possible connections students could practice making, should be the focus of instruction? Our work in statistics, and our reading of the research literature, leads us to propose that three types of connections are critical for learning in any complex domain. These are the following: (1) connections with the *contexts and practices in the world* to which the domain knowledge is intended to apply (Engle et al. 2012; van Merriënboer 1997), (2) connections with *core concepts* that serve to organize and lend coherence to the domain (e.g., National Council for Teachers of Mathematics 2000; National Governors Association 2010; Richland et al. 2012), and (3) connections with key *representations* used for thinking and communicating in the domain (e.g., Ainsworth 2008; Kozma 2003; Strauss 1998).

Connections #1: Contexts and Practices of the World

At its core, academic instruction should be motivated by the demands of the world beyond the classroom. The hope of instruction is that students will transfer what they are learning to authentic real-world situations, yet successful transfer has proven to be an elusive goal. Research suggests that continuously connecting learning in the classroom with the authentic practices of the domain may be one of the most effective ways of developing transferable knowledge (Barnett and Ceci 2002; Bransford and Schwartz 1999).

When developing our introductory statistics curriculum, we started by examining the practice of data analysis. There are many ways one might describe this practice. The important thing, in our view, is that some description of the world outside the classroom be made explicit—for the textbook authors, for students, and for instructors tasked with implementing

the curriculum. Our description of the practice of data analysis is illustrated in Fig. 2, which we often come back to as we teach the class. We tell students from the beginning that the goal of statistics is to explain variation in the world, which we break down into three core practices: (1) explore variation (in data), (2) model variation, and (3) evaluate and compare models.

We organized our textbook around these three practices, continually reminding students of where the specific skills and concepts they are studying connect with the work they might one day do if they pursue work as a data scientist. For instance, data scientists explore variation when they construct a graph to accompany a corporation's annual report. They model variation when they predict the effect of a particular variable on future sales. And they evaluate models when they compare competing explanations for customer satisfaction. By grounding instruction in the practice of data analysis, students can participate, from the beginning, in activities with clear connections to the goals and routines used by experts (Lave and Wenger 1991). Rather than teaching students isolated bits and hoping they will put them together to solve a larger problem later, students practice recognizing the need for their developing knowledge in the context of authentic tasks (National Academies of Sciences, Engineering, and Medicine 2018).

The world beyond cannot be fully described, though, just in relation to the practice of data analysis. It is also important to identify the range of situations and contexts to which students' developing skills and understanding should be applicable. Continually varying contexts (which in statistics could be as simple as varying data sets) help hone students' perception and understanding of which features of a situation are critical, and which superficial, for the application of domain knowledge (Gentner 1983; Kellman et al. 2010; Son et al. 2011). When students learn about an abstract idea in one context, it is as if a rubber band is tightly bound around that first learning context. As students experience more contexts, there is a gradual stretching of that rubber band to include an ever widening range of situations. By stretching to include more situations, the concept becomes more differentiated, coherent, and flexible and more likely to transfer to new situations.

One example from our own teaching of statistics is the concept of observational unit. Most of the data sets we use in psychology use people as the units of analysis; so, each row in the data table, or dot in the scatter plot, typically represents a person. But if students work only with data in which people are the units, they will have developed a limited concept of observational unit and will have trouble applying the concept in new situations. Following our rubber band analogy, we work from the beginning to give students practice with a variety

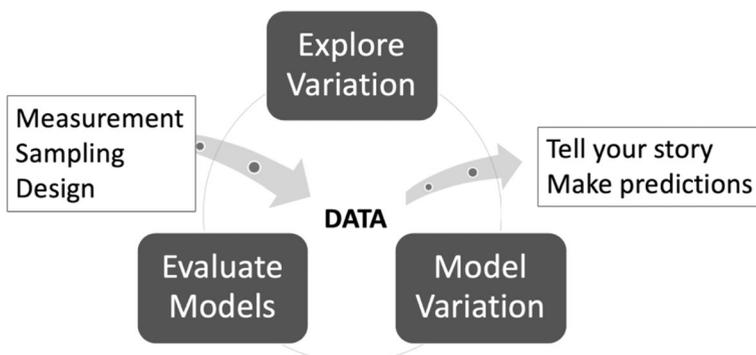


Fig. 2 The practice of data analysis

of data sets in which the rows are not people but instead represent families, states, countries, or companies. In this way, students' development of the concept of observational unit gets connected to a broader and more diverse set of contexts.

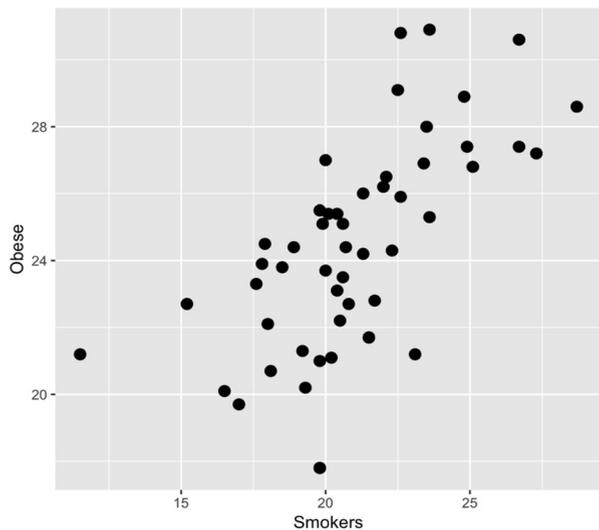
We want students to activate the concept of observational unit not only in the context of different data sets, but also within the different activities that comprise the practice of data analysis. For example, in the context of exploring variation in data, we present students with a scatter plot of data in which state (as in each of the USA) is the observational unit. With each dot representing a state, they can see a high correlation between the percentage of the state's population that is obese and the percentage that are smokers (see Fig. 3). Students at first mistakenly interpret this as evidence that smokers are more likely to be obese. To overcome this misconception, students must learn to connect the variation they see in the scatter plot with the units that are measured; they must learn to say, "States that are high in percentage of residents who smoke are also high in percentage who are obese."

Connections #2: Core Concepts That Organize the Domain

The second type of connection we want students to make is with core concepts that organize the domain. Studies show that expert knowledge in a domain is generally organized around a small set of core concepts (e.g., Lachner and Nückles 2015) that imbue coherence to even wicked domains. Because they are highly abstract and interconnected with other concepts, core concepts must be learned gradually, over extended periods of time and through extensive practice. As students practice connecting concepts with other concepts, contexts, and representations, these core concepts become more powerful and students' knowledge becomes more transferable (e.g., Baroody et al. 2007; National Council for Teachers of Mathematics 2000; Rittle-Johnson and Schneider 2015; Rittle-Johnson et al. 2001).

If only we could just ask an expert what the core concepts are, we could plan our statistics curriculum around those concepts. But the concepts most useful for novices in the early stages of learning might not be the same ones developed by experts over many years of experience (Kirschner et al. 2006). The design challenge is to select concepts that are accessible to

Fig. 3 Scatterplot showing percentage of a state's residents who are obese as a function of the percentage who smoke (from the US States data set)



novices, but recognizable to experts as having validity and utility in the domain. Critically, although the core concepts introduced to novices may be over-simplified in some ways—indeed, in complex domains, the concepts almost always need to be simplified for beginning students—it is important that the concepts not be simplified to the extent that misconceptions are introduced and that the simplification does not yield a collection of concepts that fail to form a coherent whole.

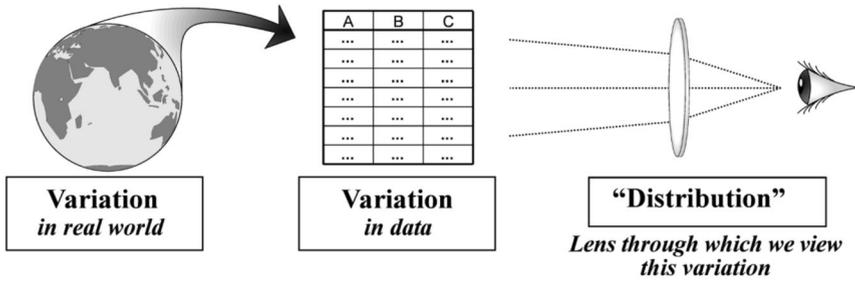
Through a lengthy process of study and discussion (which included testing by teaching), we chose to focus our textbook around three core concepts: *statistical model*, *distribution*, and *randomness*. We chose these concepts based on several criteria. They should be concepts (1) that we could continually connect to throughout the course, (2) that novices could engage with (i.e., within their zone of proximal development), (3) that would not have to be unlearned later as students progressed to higher levels of understanding, and (4) that, taken together, they would help students build a coherent understanding of the domain of statistics. (Although we will briefly describe the concepts we chose here, a more detailed development of our approach to core concepts in statistics is available in Son et al. [under review](#).)

Statistical Model Anyone who studies statistics at the advanced levels will be familiar with the concept of statistical model. Yet, the concept is almost never mentioned in the introductory course. Our view is that if we want students' knowledge of introductory statistics to transfer to advanced courses and to the wider practice of data analysis, we should connect what they are learning to the concept of modeling from the very beginning. Thus, we endeavor to do this in our introductory book. We start by conveying the overarching goal of data analysis as explaining variation in data. We then, from the beginning of the book to the end, connect everything students are learning to the statement $DATA = MODEL + ERROR$.

When we introduce the mean, for example, we conceptualize it as the simplest of all statistical models. If we use this simple model to predict each point in a distribution, most of our predictions will be wrong (though the mean will be unbiased and better than just a random guess). If the mean is a model, we can conceptualize error as the deviation from each predicted score to the actual score. Standard deviation is introduced, then, as a means of quantifying how much total error there is around the model predictions. As we add explanatory variables to the model, the accuracy of the predictions will go up, while the error goes down. Connecting the activities of data analysis to these fundamental abstract concepts yields knowledge that is more highly interconnected and therefore more usable.

Distribution Following Wild (2006), we adopted *distribution* as our second core concept. Wild defines distribution as “the pattern of variation in a variable” (cf. Garfield and Ben-Zvi 2005). The concept of distribution is the lens through which we view variation (see Fig. 4). In statistics, it is critical to situate reasoning relative to three distinct types of distributions: the sample data, the data generating process (DGP) that gave rise to that data (related to the concept of population), and sampling distributions, which are the imaginary distributions from which sample statistics are generated. Together, we refer to these three types of distributions as “the distribution triad.”

The concept of sampling distribution is exceedingly difficult for students to understand, yet critically important for the process of statistical inference. We spend time helping students practice the connections between different distributions and the questions they are best suited to answer. For example, asking about the probability that a group of people will have an average weight over some prescribed limit (e.g., 200 pounds) requires a sampling distribution.



Chris Wild

Fig. 4 Variation in the world is captured by measurement and turned into variation in data, which we then analyze through the lens of a distribution (figure from Wild 2006)

Asking about the probability of a single individual weighing more than 200 pounds, in contrast, will require us to use a distribution of sample data or a model of the data generating process constructed based on a sample of data.

Randomness The third core concept we emphasize is the concept of randomness as a data generating process. Students naturally think of causal explanations for variation. Seeing variation as caused by random processes, on the other hand, is not something students come to naturally (Batanero 2016; Batanero et al. 1998; Kaplan et al. 2014). Randomness, of course, plays an important role in statistical thinking, mainly because we know how to model random processes. Students start out associating randomness with “unpredictability” (e.g., a single randomly generated number between 1 and 10 is hard to predict). Through computational techniques such as simulation, bootstrapping, and randomization, we give students the tools and experiences they can use for thinking about randomness as a process that yields predictable patterns of variation over the long-run even though a single sample of data remains unpredictable.

These three core concepts—*DATA = MODEL + ERROR*, the *distribution triad*, and *randomness as a process*—provide an organizing framework to lend connections and coherence to statistical routines and their purposes. For example, although univariate ANOVA and simple regression are typically taught as two separate concepts in most introductory statistics courses, we emphasize that both are examples of linear models. Both kinds of models generate predictions, and both measure error in the same way (i.e., the difference between predicted and actual scores). In both, we compare a more complex model to a simple one in which the resulting distribution is the result of a random data generating process (traditionally called the null hypothesis). We hypothesize that when these analytic techniques are taught as examples of the same thing, students’ knowledge of statistics will be more coherent and more likely to transfer to new situations.

Connections #3: Key Representations for Thinking and Communicating

In addition to building a connected understanding of the core concepts that comprise a domain, we also want students to learn to use key representations that embody those concepts and that can represent explicitly the relational structure of a domain (see Ainsworth 2008; Gentner and Rattermann 1991; Star and Rittle-Johnson 2009; Uttal

et al. 1997). Evidence suggests that teaching with multiple representations produces deeper, more flexible learning (Ainsworth 2008; Ainsworth et al. 2002; Brenner et al. 1997; Pape and Tchoshanov 2001). Also, using and understanding symbolic representations are necessary for communicating and developing higher order skills (e.g., Gilbert and Treagust 2009). Because learning to use a representational system requires an investment of time and effort (Ainsworth 2008; Star and Rittle-Johnson 2009), however, we must decide which representations to focus on for novices. In making this selection, we want to find representations that are accessible to novices (Vygotsky 1980), important for the field (Tabachneck-Schijf et al. 1997; Tsui and Treagust 2013), and most productive for making connections (Kaput et al. 2017).

In our introductory statistics course, we decided to focus on five key representations: verbal descriptions, visualizations, word equations, GLM (general linear model) notation, and R code. Accordingly, throughout our course, we repeatedly ask students to translate and connect using verbal descriptions, visualizations, word equations, GLM notation, and R code. Note that we did not choose algebra, in the form of formulas and equations, as one of our key representations, even though it is emphasized in most textbooks. Our reason for this is that our students do not typically find algebra to be readily accessible or useful, a consequence, we surmise, of K-12 mathematics education in the USA.

Graphs and visualizations have always been important to understanding in statistics and are even more important in the age of data science, in which professionals are often called on to create publication-ready figures and graphs. Coding (e.g., in R) is also a representation that has become increasingly important in statistics as the field itself has become more computational in nature. R also gives students a means of participating in the emerging routines of data analysis that place increasing value on producing and sharing reproducible analyses.

In developing students' understanding of statistical models, we start by having students describe simple models in words (e.g., "knowing someone's height will help us make a better but not perfect prediction of their thumb length"). We then teach them to write word equations (e.g., Thumb Length = Height + other stuff), which helps them begin to take their initial insight (that statistical predictions are not perfect) and connect it with the concept of error (represented in the word equation as "other stuff").

Later, students are asked to map word equations to a more general idea represented in a similar way (i.e., DATA = MODEL + ERROR) and then to use the same structure in writing R code to generate a graph (e.g., scatterplot(Thumb ~ Height, data = data set) (see Pruijm et al. 2017, and their resulting visualizations). This body of connections then supports learning the mathematical notation of the GLM (e.g., $\text{Thumb}_i = b_0 + b_1\text{Height}_i + e_i$ and more generally $Y_i = b_0 + b_1X_i + e_i$).

We want to expand here on the choice to use the notation of the GLM. It is a difficult representational system for students to learn but exemplifies the features that support connections: generalizability (it can represent many situations, even those that are not present in the course) and alignability (it aligns with the other representations in the curriculum). Furthermore, because it is used by many professionals and frequently published in papers, it is highly ecologically valid. For instance, the notation of the GLM explicitly represents the structural similarity between group models such as ANOVA (e.g., $\text{Outcome}_i = b_0 + b_1\text{Group}_i + e_i$) and regression models (e.g., Outcome_i

$= b_0 + b_1(\text{Quantity}_i + e_i)$, supporting students as they work to recognize the structural similarities across the two types of models.

Designing Learning Experiences to Support Practicing Connections

Having laid out the types of connections students need to work on making, we turn now to the question of how, as educators, we can design instruction that will facilitate this process. Based on our reading of the research, we believe there are three key principles that should guide instructional design. First, the connections identified above (core concepts, key representations, and the world) must be made explicit at some point during the instruction. Second, students must be engaged in productive struggle. Connections, as it were, must be earned through hard work, as much as we would like to be able just to give them to students. And finally, opportunities to engage in the work of forging connections must be offered repeatedly to students over sustained periods of time as they deepen and extend their domain knowledge. We will discuss each of these principles next.

Principle #1: Make Connections Explicit

Recognizing structural relatedness is critical to understanding in complex domains (Bassok 2001; Chen and Klahr 1999; Novick 1988; Novick and Holyoak 1991; Reed 1985; Ross 1987). However, research suggests that connections between structurally similar problems often go unrecognized by learners, unless the connections are pointed out explicitly (Gick and Holyoak 1983). Learners are often unable to summon relevant prior knowledge when needed (Reeves and Weisberg 1994). And novice learners, especially, often get stuck on the surface features of a problem and need help to connect the surface representation to deeper underlying concepts (Ross 1987).

These results are consistent with findings from the voluminous research literature on discovery learning. Although it is likely that many high-level experts have acquired their domain knowledge on their own, based mainly on their own persistence and lengthy experience, the research suggests that most students do not benefit from simply being given the opportunity to discover structure on their own (Alfieri et al. 2011). As we will see in the next section, effort is most certainly required on the part of the learner. But effort alone is usually not enough (Mayer 2004). At every age level, *guided discovery* is more effective than simple discovery learning (e.g., Alfieri et al. 2011; Klahr and Nigam 2004; Weisberg et al. 2015).

A number of research paradigms from cognitive psychology demonstrate techniques for making connections explicit and the benefits of such techniques. Research on learning from analogies (superficially dissimilar but structurally parallel instances) provides one example. When learners draw parallels between two cases and practice aligning similar elements across two systems (Gentner et al. 2003; Son et al. 2011), they are better able to transfer their knowledge to superficially dissimilar novel problems (Alfieri et al. 2013). Learning tasks that make relations explicit, while stripping away distracting details, results in more portable and generalizable knowledge (Gentner and Markman 1997; Kaminski et al. 2008; Son et al. 2008; Uttal et al. 2009).

In our online textbook, we use two specific techniques for making connections explicit. The first is to structure the book so as to afford students' seeing of connections and then to explicitly point out these connections in the text itself. The constant connections to the concept

of statistical modeling in our book exemplify the first strategy. Having decided on modeling as a core concept, we structured the book around the enterprise of modeling—and then testing and comparing models of—variation in data. As statistical concepts and procedures are introduced, each is related explicitly to the statement $DATA = MODEL + ERROR$.

When the arithmetic mean is introduced, for example, we point out explicitly that the mean is an example of a function that could fill the MODEL slot in the abstract statement $DATA = MODEL + ERROR$. Residuals, sum of squares (SS), variance, and standard deviation, similarly, are explicitly connected to the concept of ERROR. When sum of squares is introduced, we point out that this particular measure of error is minimized at the mean and relate this fact to the overall goal of minimizing error as a means of increasing the power of a model. These connections increase coherence across the domain by recasting what were previously thought of as separate topics (e.g., “measures of central tendency” and “measures of variation”) as important concepts that relate to the overall work of statistical modeling.

The second strategy for making connections explicit is to repeatedly ask questions of students, the answers to which are the explicit connections we are trying to strengthen. We constantly ask students to explain, for example, how core concepts and representations relate to each other (e.g., Chi 2000; Lombrozo 2006). Across various situations and representations, we repeatedly ask students to explicitly identify DATA, MODEL, and ERROR. We ask students to connect a data point in a scatter plot to a particular value in a table of raw data. After asking them to fit a regression model, we ask them to save the model predictions back into the data frame (using R) and then plot the model predictions as a function of the explanatory variable. We ask them to explain why the model predictions seem to coincide with a linear regression line, while the actual data points do not. We ask students to continually contextualize what they see in data in terms of the overall modeling enterprise.

The same principle of making connections explicit can be seen in the research on worked examples, in which students learn by examining an expert’s problem solution (Atkinson et al. 2000; Paas et al. 2003). When students are asked to solve a problem, much of their effort goes toward generating and executing a solution method, with few attentional resources left over for reflecting on the concepts that underlie their solution. When given the solution and asked to study it, however, learners can reflect on *why* the solution works, which has the potential to make clear how the solution relates to the core underlying principles of the domain.

The literature on worked examples is closely related to cognitive load theory (CLT; Paas and van Gog 2006; Paas et al. 2010). CLT distinguishes three types of load: *intrinsic*, *germane*, and *extraneous*. Intrinsic load is determined by the task at hand; if you want to do the task, you will need to contend with the intrinsic load. For learning, the goal is to decrease extraneous load (the optional extra stuff that surrounds the task) and increase germane load (Sweller 1988; van Merriënboer and Sweller 2005). In our conceptualization, germane load is the part of attention and effort focused on critical connections in the domain. A desire to increase germane load means accommodating a need to leave more cognitive resources available for conscious reflection on principles and concepts that underlie the deep structure of the domain.

One of the reasons we interleave R throughout our curriculum is that R takes care of the calculations quickly, thus reducing extraneous load and leaving students with more mental resources available for thinking about what the results of the calculations mean. For example, using a simple R function to calculate residuals leaves students with more resources they can direct toward seeing that residuals, as ERROR, can be expressed as $DATA - MODEL$, an equivalent form of our core conceptual framework.

Instead of asking students to engage in repeated subtraction, students run simple R code (e.g., `resid(model)`) and then focus on explaining the meaning of what they find, such as why some residuals are negative and some are positive, how the residuals relate to the mean and the original data, and how residuals relate to aggregate error (e.g., SS). Making these connections is difficult but part of the germane load of fostering understanding.

Principle #2: Engage Students in Productive Struggle

As important as it is to make explicit the connections that lead to understanding in a domain, it is clear that simply making connections explicit is not sufficient to produce deep understanding in students. In our culture, there is a widespread belief in the myth that understanding comes suddenly, like a lightning strike, in an “a-ha!” moment. If only it were that simple. In reality, understanding is something that must be earned, gradually over time, through the hard work of each individual learner. If a student experiences learning as easy and effortless, this generally means that what was learned will not be retained (Koriat and Bjork 2005). We have all had the experience of watching a teacher solve a problem, thinking, “That looks easy. I can do that.” But in the end, when we are on our own, we often find that we cannot “do that.”

This “no pain, no gain” perspective on learning is consistent with a broad body of research in the learning sciences. Many studies demonstrate the superiority of active learning over passive learning (Bean 2011; Michael 2006; Prince 2004). Neuroscientist Stanislas Dehaene (2020), in a wide-ranging review of the literature, calls *active engagement* one of the four pillars of learning. Dehaene writes: “A passive organism learns almost nothing, because learning requires an active generation of hypotheses, with motivation and curiosity (p. xxvii).” Bjork and Bjork (2011) have coined the term “desirable difficulties” to refer to the struggle that necessarily and productively accompanies lasting learning. Research on the *testing effect*, for example, shows that students learn more from trying to answer test questions than they do from being re-presented with the same information (McDaniel et al. 2007; Rowland 2014).

Productive struggle not only is important in its own right but also plays a role in students’ learning from explicit connections. Across a variety of research paradigms, it has been shown that giving students opportunities to struggle in solving a challenging task *before* presenting them with solutions is more effective than the typical sequence in which direct instruction precedes practice (Hiebert et al. 1996; Vygotsky 1980). When they grapple with a challenging task, learners inevitably work to connect features of the problem with their prior knowledge. Not only does this help them to identify gaps in their knowledge but also prepares them for subsequent instruction that explicitly connects their prior knowledge to the core concepts and representations of a domain (Capon and Kuhn 2004; Lawson et al. 2019a; Schwartz and Bransford 1998).

In our online book, we design each page with a “struggle first” pedagogy, asking students to answer questions—often in an open-response format—before we have presented them with the information and explicit connections they might need in order to give a well-formulated answer. For example, before we discuss the properties of the mean and median, we present students with a distribution of 5 data points and ask: “In what sense might the median be a better model for this distribution? In what sense might the mean be a better model?” Putting the question before the answer in this way is counter to what students expect, and sometimes they will point out to us, in an attempt to be helpful, that we have mistakenly “gotten the order wrong.”

Another strategy we use to increase productive struggle (and germane cognitive load) is to remove opportunities for calculation from the book. We want students to expend their energy grappling with key concepts and the connections between them, but because US college students tend to equate doing math with doing calculations, they often will, if given the chance, start calculating before they have had a chance to think about a problem, how it relates to core concepts, and even which calculations might be most appropriate given the situation. One way to stave off premature calculations and reserve attention for conceptual connections is by giving students problems that do not have any numbers in them, making calculations impossible (Givvin et al. 2019; Lawson et al. 2019b). In our book, we hold off on presenting any formulas or calculations until chapter 5, asking students instead to work on developing their intuitive ideas about models and model comparison by examining and discussing graphical representations of data (e.g., histograms or box plots).

For example, in chapter 4, students create faceted histograms to compare distributions of restaurant tips between two randomly assigned conditions in an experiment, one in which the server puts a smiley face on the check and another where they do not. Instead of calculating or displaying the means of the two groups, we ask students whether they think variation in tips is explained by the experimental manipulation. This leads to a rich discussion of what it means to “explain” variation. Working with an intuitive definition of “explain,” we reformulate the question as: “Does knowing which condition a table was in help us make a better prediction of the tip?” Thinking about this question helps students to see that both the central tendency and the variation are important in answering the question. Much of this discussion would have been cut off if students had rushed to calculate the mean difference between the two groups.

Principle #3: Provide Opportunities for Deliberate Practice

Our first two principles work together to form the in-the-moment experience of a specific learning opportunity. Struggling productively to make connections between problems in a domain and the core conceptual structure of the domain is the raw material from which understanding is forged. But developing transferable knowledge in a complex domain is a long-term prospect (Ericsson 2018; Ericsson et al. 2007) and happens in fits and starts (Felder and Silverman 1988). For students to develop a deep understanding of a domain, we will need to find ways to provide repeated opportunities for struggling with important connections over long periods of time. In other words, students must practice the connections. It is generally accepted that skill learning requires practice; we argue that understanding requires practice, too.

This leads us to the third principle in our framework: *deliberate practice*. Deliberate practice is a term that comes from the expertise literature (Ericsson 2017; Ericsson et al. 1993). Although it is normally discussed in relation to activities like music, chess, and motor behavior, much of what is posited about what makes practice effective should apply equally well when the task is to develop proficiency with concepts and making connections between them, as it does when applied to skill acquisition.

It is helpful to begin with the clarification that deliberate practice is distinct from repetitive practice. Repetitive practice leads to fluency, creating in learners the feeling that a concept or skill is getting easier over time (e.g., Bjork et al. 2013). Automaticity, which is the goal of repetitive practice, can also be its downfall. The feeling of fluency is a sign that learners have plateaued, not that they are deepening their understanding (Ericsson 2008).

The key challenge for aspiring expert performers is to avoid the arrested development associated with automaticity and to acquire cognitive skills to support their continued learning and improvement. By actively seeking out demanding tasks – often provided by their teachers and coaches – that force the performers to engage in problem solving and to stretch their performance, the expert performers overcome the detrimental effects of automaticity and actively acquire and refine cognitive mechanisms to support continued learning and improvement” (Ericsson 2006, p. 696)

Although we acknowledge the need for certain skills to be readily available in a student’s repertoire, our ultimate aim is that students be capable of calling upon them in a way that is not routinized, so that they can make use of them in novel situations.

Deliberate practice requires that instructional materials maintain a constant (and high) level of challenge as abilities develop. This is accomplished by raising the difficulty of the connections being practiced to meet an individual’s growing capability, all in accordance with a desired learning trajectory. As content designers (and instructors), we must help define and design the tasks that are practiced, “to correct some specific weakness while preserving other successful aspects of function” (Ericsson 2006, p. 700). We must keep students practicing connections in increasingly challenging situations, spending time where weaknesses are greatest. Returning to our rubber band analogy from earlier in the paper, our book must continually expand the scope of the rubber band in order to keep ahead of learners’ developing understanding in the domain while continuing to encompass what was initially contained within it.

One of the most effective ways of increasing challenge (and expanding the rubber band) is to continually vary the context, forcing learners to continually adapt their conceptual understanding to new situations. One way to do so in a statistics curriculum is by constantly introducing new data sets to which students must apply their developing knowledge. In our online book, we use a very limited number of data sets. But when we teach the class, we introduce new data at almost every class session. We leave this to the teacher because statistics is taught in multiple academic departments and the students in those departments have different interests. Data is the most direct way of connecting students’ practice of data analysis with the things they care about most. We want students to come to see statistics as a tool they can apply generally, across a variety of contexts, to help make sense of the world.

If the principles of deliberate practice, alone, were to be applied to statistics learning, one might find a student gradually learning a series of data analysis skills with ever-increasing complexity. Even within a single statistical test—ANOVA for instance—they might begin with a simple 2×2 factorial and gradually learn how to manage experimental designs with covariates, repeated measures, and nested levels. All of that might be accomplished, however, without attention paid to conceptual connections across these special cases and how those concepts might apply to other statistical tests. So while we add deliberate practice to our triad of principles that inform the design of learning experiences, applying it in combination with the first two is critical.

It is not just skill acquisition we are after; our goal is to support students’ ability to see connections across contexts and practices, core concepts of the domain, and key representations. It is these connections that will make it possible for students to coordinate their skills and adapt them to the needs of new contexts and situations. And we believe that the only way to increase the strength and robustness of these connections is through the experience of

productive struggle. Simply put, if we want students to have knowledge that is flexible and transferable to new and even unforeseen situations, we want to create opportunities for deliberate practice of transfer itself.

Putting It All Together

In summary, it is useful to situate the practicing connections framework in a broader pedagogical landscape. A simple two-by-two table helps us to do this (see Fig. 5). On one dimension, we have productive struggle, on the other, explicit connections. Where neither feature is present, as in the upper left corner, we have rote memorization of disconnected facts and procedures. Students just listen to the teacher and then repeat back what the teacher has said or the steps that have been demonstrated.

If we have struggle but no attempt to make explicit connections with the core concepts or representations of the domain, we end up in the lower left corner. This is the province of discovery learning. If we have explicit connections but no struggle, we have a well-formed lecture. A lecture gives the feeling of fluency because a good lecturer lays out the structure explicitly and clearly. Yet, unless the student is engaged in the active work of internalizing connections into her own mental model, understanding will not result. The sweet spot, as we call it, is in the lower right quadrant. Here is where we have both productive struggle and explicit connections.

Just finding the sweet spot once or even occasionally, however, is not sufficient. Deliberate practice adds in the dimension of time. Understanding unfolds slowly. In complex academic domains, such as statistics, understanding is achieved over weeks, months, and years. The instructional designer, therefore, needs to think carefully about how to sequence instructional activities so as to support the gradual deepening of conceptual understanding. Much has been written about the development of complex skills (van Merriënboer 1997). Our focus is specifically on the understanding part. Understanding is at the root of far transfer and flexible knowledge (Barnett and Ceci 2002; Bedard and Chi 1992). We are just beginning to think explicitly about designing instruction that specifically takes the growth of conceptual understanding as its target.

Summary

In this paper, we have proposed a practical framework for instructional design in complex academic domains. Our specific focus is on the development of understanding. Following the research on expertise, we are proposing what we think it would take to produce students who “see the structure” of the domain and are able to transfer their knowledge to new and

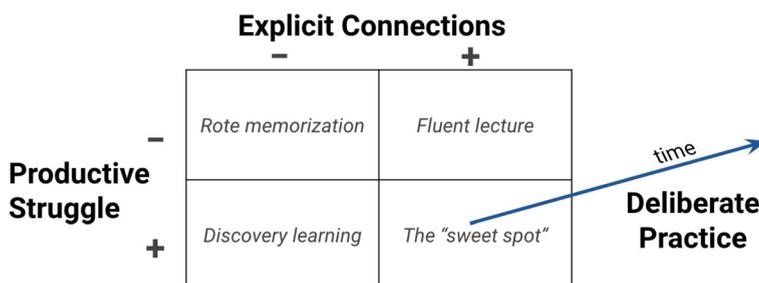


Fig. 5 The pedagogical landscape for practicing connections

unforeseen situations. We call our framework the *practicing connections* framework. Many students learn the bits of knowledge—facts and skill—that make up a domain but fail to see how they all fit together. Our hypothesis is that it is the interconnectedness of knowledge that makes it transferable. As instructional designers, we need to give students opportunities, over extended periods of time, to practice making connections.

There are two parts to our proposed framework, each part answering an important question. The first question we address is this: If connections are important for understanding, what specific types of connections, of all the possible ones, do we want students to make? We proposed three types of connections: with core concepts, with key representations, and with the world to which the domain knowledge is expected to apply. We reviewed the rationale for each type of connection and then gave some examples from our own experience trying to apply the framework to the design and implementation of an introductory statistics textbook.

The second question addressed by the framework is: How, given that we have specified what the important connections are for a given domain, do we create opportunities for students to practice these connections? Although the literature on teaching and learning is vast, we tried to present a simple organizing framework. Students need to experience productive struggle with explicit connections, and they need to do this continually over extended periods of time (deliberate practice). There is a sweet spot for the design of learning experiences, and we believe this spot is well-supported by research. Yes, we have glossed over many details. But in a real sense, these details need to be filled in by educators and designers working in the trenches to produce deep learning in all students across a wide variety of complex domains.

Challenges of Implementation

Implementing this framework will not be easy in practice, and we want to acknowledge that. First, identifying the core concepts that organize a domain is hard. It requires deep content knowledge as well as pedagogical content knowledge. Second, identifying which connections to make explicit as the course unfolds is not a trivial task. This critical element of our pedagogical approach is one of the things that sets it apart from more traditional teaching, and there is no ready source for a list of connections in the same way that there is for a list of “bits.” Third, progress through a curriculum that results from applying the three design principles we have outlined and offering students sufficient opportunity to practice connections is a slow process. We are convinced, though, that the “slow and sticky” (Hess and Azuma 1991) pedagogy that results contributes to students’ ability to retain their learning over time.

We know the difficulty of implementation from firsthand experience as instructors who tried to break down a wicked domain (statistics) into the relevant connections and then get our students to practice making those connections many times over a school term, all while dealing with broken projectors, exams, office hours, and everything else that goes into teaching a course. We acknowledge also that, although grounded in the experience of having gone from core concept identification all the way through to the implementation of a full curriculum, our practicing connections framework has been applied in only one domain. We invite other educators to take on the challenge of doing the same in other content areas. In doing so, they will no doubt contribute to refining and improving the framework. Our experiences have deepened our belief that this framework can help to guide us toward more successful instructional design.

Funding Information This project has been made possible in part by a grant from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation (DRL-1229004), and a grant from the California Governor's Office of Planning and Research (OPR18115).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: theory and practice in science education. Models and modeling in science education* (Vol. 3, pp. 191–203). Dordrecht: Springer.
- Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *The Journal of the Learning Sciences*, *11*(1), 25–61.
- Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18.
- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: a meta-analytic review. *Educational Psychologist*, *48*(2), 87–113.
- Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181–214.
- 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.
- Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education*, 115–131.
- Bassok, M. (2001). Semantic alignments in mathematical word problems. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: perspectives from cognitive science* (pp. 401–433). Cambridge, MA, US: The MIT Press.
- Batanero, C. (2016). Understanding randomness: challenges for research and teaching. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (pp. 34–49). Prague: European Society for Research in Mathematics Education.
- Batanero, C., Green, D. R., & Serrano, L. R. (1998). Randomness, its meanings and educational implications. *International Journal of Mathematical Education in Science and Technology*, *29*(1), 113–123.
- Bean, J. C. (2011). *Engaging ideas: the professor's guide to integrating writing, critical thinking, and active learning in the classroom*. John Wiley & Sons.
- Bedard, J., & Chi, M. T. (1992). Expertise. *Current Directions in Psychological Science*, *1*(4), 135–139.
- Bilalic, M., & Campitelli, G. (2018). Studies of the activation and structural changes of the brain associated with expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 233–249). Cambridge University Press.
- Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning. In M. A. Gernsbacher & J. Pomerantz (Eds.), *Psychology and the real world: essays illustrating fundamental contributions to society* (2nd ed., pp. 55–64). New York, NY: Worth.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*(1), 417–444.
- Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: rethinking transfer: a simple proposal with multiple implications. *Review of Research in Education*, *24*(1), 61–100.

- Bransford, J. D., & Stein, B. (1984). *The IDEAL Problem Solver: a guide for improving thinking, learning, and creativity*. New York: W.H. Freeman.
- Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Durán, R., Reed, B. S., & Webb, D. (1997). Learning by understanding: the role of multiple representations in learning algebra. *American Educational Research Journal*, 34(4), 663–689.
- Campitelli, G., & Gobet, F. (2005). The mind's eye in blindfold chess. *European Journal of Cognitive Psychology*, 17(1), 23–45.
- Capon, N., & Kuhn, D. (2004). What's so good about problem-based learning? *Cognition and Instruction*, 22(1), 61–79.
- Carbonell, K. B., Stalmeijer, R. E., Kónings, K. D., Segers, M., & van Merriënboer, J. J. (2014). How experts deal with novel situations: a review of adaptive expertise. *Educational Research Review*, 12, 14–29.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4(1), 55–81.
- Chen, Z., & Klahr, D. (1999). All other things being equal: acquisition and transfer of the control of variables strategy. *Child Development*, 70(5), 1098–1120.
- Chi, M. T. (2000). Self-explaining expository texts: the dual processes of generating inferences and repairing mental models. *Advances in Instructional Psychology*, 5, 161–238.
- Chi, M. T. (2011). Theoretical perspectives, methodological approaches, and trends in the study of expertise. In Y. Li & G. Kaiser (Eds.), *Expertise in mathematics instruction* (pp. 17–39). Springer.
- Chi, M. T., & Koeske, R. D. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology*, 19(1), 29–39.
- Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121–152.
- Chi, M. T., Glaser, R., & Rees, E. (1982). Expertise in problem solving: advances in the psychology of human intelligence. *Erlbaum*, 1–75.
- de Groot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton (Original work published 1946).
- Dehaene, S. (2020). *How we learn: why brains learn better than any machine... for now*. Penguin.
- Endsley, M. R. (2018). Expertise and situation awareness. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance*. Cambridge University Press. <https://doi.org/10.1017/9781316480748.036>.
- Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist*, 47(3), 215–231.
- Epstein, D. J. (2019). *Range: why generalists triumph in a specialized world*. New York, NY: Riverhead Books.
- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, and R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (p. 685–706). Cambridge University Press.
- Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: a general overview. *Academic Emergency Medicine*, 15(11), 988–994.
- Ericsson, K. A. (2017). Expertise and individual differences: the search for the structure and acquisition of experts' superior performance. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1–2), e1382.
- Ericsson, K. A. (2018). Superior working memory in experts. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (p. 696–713). Cambridge University Press. <https://doi.org/10.1017/9781316480748.036>.
- Ericsson, K. A., & Charness, N. (1994). Expert performance: its structure and acquisition. *American Psychologist*, 49(8), 725–747.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211–245.
- Ericsson, K. A., & Pool, R. (2016). *Peak: secrets from the new science of expertise*. New York, NY: Houghton Mifflin Harcourt.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406. <https://doi.org/10.1037/0033-295X.100.3.363>.
- Ericsson, K. A., Hoffman, R. R., Kozbelt, A., & Williams, A. M. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. New York, NY: Cambridge University Press.
- Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, 85(7/8), 11hoff4.
- Ericsson, K. A., Hoffman, R. R., & Kozbelt, A. (Eds.). (2018). *The Cambridge handbook of expertise and expert performance*. Cambridge University Press.
- Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering Education*, 78(7), 674–681.
- Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, 4(1), 92–99 <http://www.stat.auckland.ac.nz/ser>.

- Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cognitive Science*, 7(2), 155–170.
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45–56.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. *Perspectives on Language and Thought: Interrelations in Development*, 225.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: a general role for analogical encoding. *Journal of Educational Psychology*, 95(2), 393–408.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1–38.
- Gilbert, J. K., & Treagust, D. F. (2009). Introduction: macro, submicro and symbolic representations and the relationship between them: key models in chemical education. In *Multiple representations in chemical education* (pp. 1–8). Dordrecht: Springer.
- Ginsburg, H. (1977). *Children's arithmetic: the learning process*. D. van Nostrand.
- Givvin, K. B., Moroz, V., Loftus, W., & Stigler, J. W. (2019). Removing opportunities to calculate improves students' performance on subsequent word problems. *Cognitive Research: Principles and Implications*, 4(1), 24.
- Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. R. Chi, Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 15–27). Hillsdale, NJ: Erlbaum.
- Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin*, 142(7), 742–757.
- Greeno, J. G., Moore, J. L., & Smith, D. R. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: intelligence, cognition, and instruction* (pp. 99–127). Norwood, NJ: Ablex.
- Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S. M., & Wood, W. B. (2004). Scientific teaching. *Science*, 304(6056), 521–522.
- Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York, NY: Freeman.
- Hess, R. D., & Azuma, H. (1991). Cultural support for schooling: contrasts between Japan and the United States. *Educational Researcher*, 20(9), 2–9.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 65–97).
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: the case of mathematics. *Educational Researcher*, 25(4), 12–21.
- Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: what would it look like and how can we get one? *Educational Researcher*, 31(5), 3–15.
- Hodson, D. (1988). Toward a philosophically more valid science curriculum. *Science Education*, 72(1), 19–40.
- Holyoak, K. J. (1991). 12 symbolic connectionism: toward third-generation theories of expertise. *Toward a general theory of expertise: Prospects and limits*, 301.
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, 320(5875), 454–455.
- Kaplan, J. J., Rogness, N. T., & Fisher, D. G. (2014). Exploiting lexical ambiguity to help students understand the meaning of random. *Statistics Education Research Journal*, 13(1).
- Kaput, J. J., Blanton, M. L., & Moreno, L. (2017). Algebra from a symbolization point of view. In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 19–56). Routledge.
- Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, 2(2), 285–305.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86.
- Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning. *Psychological Science*, 15(10), 661–667.
- Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction framework: bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, 36(5), 757–798.
- Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2), 187.
- Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, 13(2), 205–226.
- Lachner, A., & Nückles, M. (2015). Bothered by abstractness or engaged by cohesion? Experts' explanations enhance novices' deep-learning. *Journal of Experimental Psychology: Applied*, 21(1), 101–115.
- Lachner, A., Gurlitt, J., & Nückles, M. (2012). A graph-oriented approach to measuring expertise-detecting structural differences between experts and intermediates. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 34(34), 653–658.

- Lagemann, E. C., & Shulman, L. S. (1999). *Issues in education research: problems and possibilities*. San Francisco, CA: Jossey-Bass Inc.
- Lave, J., & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. New York, NY: Cambridge University press.
- Lawson, A. P., Davis, C., & Son, J. Y. (2019a). Not all flipped classes are the same: using learning science to design flipped classrooms. *Journal of Scholarship in Teaching and Learning*, 19(5), 77–104.
- Lawson, A. P., Mirinjian, A., & Son, J. Y. (2019b). Can preventing calculations help students learn math? *Journal of Cognitive Education and Psychology*, 17(2), 178–197.
- Levin, J. R., & O'Donnell, A. M. (1999). What to do about educational research's credibility gaps? *Issues in Education*, 5(2), 177–229.
- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, 10(10), 464–470.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14–19.
- McDaniel, M. A., Roediger, H. L., & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, 14(2), 200–206.
- McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13(3), 307–325.
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30(4), 159–167.
- National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: learners, contexts, and cultures*. National Academies Press.
- National Council for Teachers of Mathematics. (2000). *Principles and standards for school mathematics* (3rd ed.). National Council of Teachers of Mathematics.
- National Governors Association. (2010). *Common core state standards*. DC: Washington.
- North, J. S., Ward, P., Ericsson, A., & Williams, A. M. (2011). Mechanisms underlying skilled anticipation and recognition in a dynamic and temporally constrained domain. *Memory*, 19(2), 155–168.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 510.
- Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(3), 398.
- Paas, F., & van Gog, T. (2006). Optimising worked example instruction: different ways to increase germane cognitive load. *Learning and Instruction*, 16(2), 87–91.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educational Psychologist*, 38(1), 1–4.
- Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, 22(2), 115–121.
- Pape, S. J., & Tchoshanov, M. A. (2001). The role of representation(s) in developing mathematical understanding. *Theory Into Practice*, 40(2), 118–127.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231.
- Pruim, R., Kaplan, D. T., & Horton, N. J. (2017). The mosaic package: helping students to “think with data” using R. *The R Journal*, 9(1), 77–102.
- Reed, S. K. (1985). Effect of computer graphics on improving estimates to algebra word problems. *Journal of Educational Psychology*, 77(3), 285–298.
- Reeves, L., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, 115(3), 381–400.
- Renkl, A., Mandl, H., & Gruber, H. (1996). Inert knowledge: analyses and remedies. *Educational Psychologist*, 31(2), 115–121.
- Richland, L. E., Linn, M. C., & Bjork, R. A. (2007). Cognition and instruction: bridging laboratory and classroom settings. In *Handbook of Applied Cognition* (pp. 555–584).
- Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, 47(3), 189–203.
- Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. Cohen Kadosh & A. Dowker (Eds.), *Oxford handbook of numerical cognition*. Oxford University Press.
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: an iterative process. *Journal of Educational Psychology*, 93(2), 346–362. <https://doi.org/10.1037/0022-0663.93.2.346>.
- Robinson, V. M. (1998). Methodology and the research-practice gap. *Educational Researcher*, 27(1), 17–26.
- Ross, B. H. (1987). This is like that: the use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(4), 629–639. <https://doi.org/10.1037/0278-7393.13.4.629>.

- Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. <https://doi.org/10.1037/a0037559>.
- Sarfo, F. K., & Elen, J. (2007). Developing technical expertise in secondary technical schools: the effect of 4C/ID learning environments. *Learning Environments Research*, *10*(3), 207–221.
- Savery, J. R., & Duffy, T. M. (1995). Problem based learning: an instructional model and its constructivist framework. *Educational Technology*, *35*(5), 31–38.
- Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–5223.
- Schwartz, D. L., & Goldstone, R. (2015). Learning as coordination. In *Handbook of Educational Psychology* (pp. 61–75).
- Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: short-cutting abstraction in children's object categorizations. *Cognition*, *108*(3), 626–638.
- Son, J. Y., Smith, L. B., & Goldstone, R. L. (2011). Connecting instances to promote children's relational reasoning. *Journal of Experimental Child Psychology*, *108*(2), 260–277.
- Son, J. Y., Blake, A., Fries, L., & Stigler, J. W. (under review). Modeling first: applying learning science to the teaching of introductory statistics.
- Star, J. R., & Rittle-Johnson, B. (2009). Making algebra work: instructional strategies that deepen student understanding, within and between representations. *ERS Spectrum*, *27*(2), 11–18.
- Stigler, J. W., & Hiebert, J. (2009). *The teaching gap: best ideas from the world's teachers for improving education in the classroom*. Simon and Schuster.
- Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator*, *1*(3), 4–16.
- Stigler, J. W., Son, J. Y., Givvin, K. B., Blake, A., Fries, L., Shaw, S. T., & Tucker, M. C. (2020). The *Better Book* approach for education research and development. *Teachers College Record*.
- Strauss, S. (1998). Cognitive development and science education: toward a middle level model. In W. Damon (Series Ed.) and I. Sigel, K. A. Renninger (Vol. Eds.), *Handbook of child psychology: Vol. 4. Child psychology in practice* (5th edn., pp. 357–400). New York: Wiley.
- Susilo, A. P., van Merriënboer, J., van Dalen, J., Claramita, M., & Scherpbier, A. (2013). From lecture to learning tasks: use of the 4C/ID model in a communication skills course in a continuing professional education context. *The Journal of Continuing Education in Nursing*, *44*(6), 278–284.
- Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, *12*(2), 257–285.
- Tabachneck-Schijf, H. J., Leonardo, A. M., & Simon, H. A. (1997). CaMeRa: a computational model of multiple representations. *Cognitive Science*, *21*(3), 305–350.
- Toth, E. E., Klahr, D., & Chen, Z. (2000). Bridging research and practice: a cognitively based classroom intervention for teaching experimentation skills to elementary school children. *Cognition and Instruction*, *18*(4), 423–459.
- Tseng, S. S., Su, J. M., Hwang, G. J., Hwang, G. H., Tsai, C. C., & Tsai, C. J. (2008). An object-oriented course framework for developing adaptive learning systems. *Journal of Educational Technology & Society*, *11*(2), 171–191.
- Tsui, C. Y., & Treagust, D. F. (2013). Introduction to multiple representations: their importance in biology and biological education. In D. Treagust & C. Y. Tsui (Eds.), *Multiple representations in biological education. Models and modeling in Science Education* (Vol. 7). Dordrecht: Springer.
- Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: a new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology*, *18*(1), 37–54.
- Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives*, *3*(3), 156–159.
- van Merriënboer, J. J. (1997). *Training complex cognitive skills: a four-component instructional design model for technical training*. Educational Technology Publications.
- van Merriënboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177.
- van Merriënboer, J. J., Clark, R. E., & de Croock, M. B. (2002). Blueprints for complex learning: the 4C/ID-model. *Educational Technology Research and Development*, *50*(2), 39–61.
- Vygotsky, L. S. (1980). *Mind in society: the development of higher psychological processes*. Harvard University Press.
- Weisberg, D. S., Kittredge, A. K., Hirsh-Pasek, K., Golinkoff, R. M., & Klahr, D. (2015). Making play work for education. *Phi Delta Kappan*, *96*(8), 8–13.
- Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal*, *5*(2), 10–26.
- Woltz, D. J., Gardner, M. K., & Bell, B. G. (2000). Negative transfer errors in sequential cognitive skills: strong-but-wrong sequence application. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 601–625.