How Cognitive Psychology Can Inform Evidence-Based Education Reform

An Overview of *The Cambridge Handbook of Cognition and Education*

John Dunlosky and Katherine A. Rawson

Formal education has had a major and positive impact on society, but it is also true that not all students meet their learning aspirations. Many children and adults struggle to learn and many are left behind. The problems that undermine their efforts to succeed (and instructors' efforts to help them) arise from numerous sources; a short list includes poor nutrition, poor physical or mental health, a lack of motivation, boredom, social and interpersonal problems at school or at home, ineffective approaches to learning, learning disabilities, and poor access to educational resources. Successfully solving these problems will require many solutions and only a subset of them are targeted by cognitive psychologists. This subset of problems is nevertheless fundamental to education and, in general, includes the difficulties that many students have in effectively learning and understanding new ideas and concepts, correcting misconceptions, achieving proficiency in math and reading, and thinking critically. Even in the best of circumstances, many students will still struggle, and many of the efforts of cognitive and educational psychologists – and certainly those represented in this handbook – are aimed at helping students more effectively learn and teachers more effectively teach.

An Overview of *The Cambridge Handbook of Cognition and Education*

One implicit take-home message from the chapters of this handbook is that even with respect to the challenge of enhancing student learning, there are numerous problems and subsequently numerous solutions for them. One reason for such diversity pertains to the variety of content, concepts, and procedures that students are expected to learn. An effective tool for students who struggle with mathematics may not be as applicable for students trying to become proficient in a foreign language, and tools that are highly effective for learning a foreign language may not be as useful when students are learning to reason scientifically. For instance, effective reading may require techniques such as identifying main ideas, making predictive inferences, and summarizing (Stevens & Vaughn, Chapter 15, this volume). By contrast, learning how to solve math problems will involve a different set of tools, such as studying worked examples (van Gog, Rummel, & Renkl, Chapter 8, this volume). As another example, using appropriate gestures during instruction promises to improve math instruction (Wakefield & Goldin-Meadow, Chapter 9, this volume), but gestures will likely not be as relevant to a student who is struggling to integrate ideas across multiple sources about a historical event (Rouet, Britt, & Potocki, Chapter 14, this volume).

Although domain-specific or content-specific tools may be required to improve student achievement for some domains and tasks, a variety of tools are more domain-general and promise to benefit learning across many domains. For example, retrieval practice can be used to improve student learning in the classroom (McDaniel & Little, Chapter 19, this volume), and spaced schedules of practice can support durable learning across many domains (Wiseheart, Küpper-Tetzel, Weston, Kim, Kapler, & Foot-Seymour, Chapter 22, this volume). Both techniques are domain-general, and for anyone who has used, heard, and believed the expression "practice makes perfect," these two techniques are fundamental for much of the practice that results in mastery and expertise.

Improving potentially general skills (such as critical thinking, scientific reasoning, and metacognition) also promises to benefit students in many domains. For example, accurately monitoring progress toward a learning goal – a form of personalized formative evaluation – presumably can enhance the effectiveness of subsequent learning. Even so, it is clear that this general skill (i.e., accurately monitoring) is rather challenging, both for students when preparing for exams (Hacker & Bol, Chapter 25, this volume) and for teachers who are evaluating the progress of their individual students (Thiede, Oswalt, Brendefur, Carney, & Osguthorpe, Chapter 26, this volume). Even more general, Winne and Marzouk (Chapter 27, this volume) describe how ideal self-regulated learning involves engaged learners who must cope with a learning task by identifying relevant conditions within the environment (external and internal) that can influence performance. Students would then apply the best strategies (or operations) for the task at hand and subsequently compare progress against learning standards to decide if further effort is needed to obtain a learning goal.

Given the diversity in approaches to improving student achievement, important aims of this handbook were to showcase this diversity by inviting experts to discuss (1) approaches to promoting education within specific domains; (2) general strategies for improving student learning and thinking; and (3) the promise of improving domain-general skills (e.g., metacognition) toward helping students learn more effectively.

Why this handbook and why now? As cognitive psychologists who are dedicated to helping students succeed, we (Dunlosky and Rawson) have been doing our best to keep informed about all the work relevant to our own areas of expertise and also trying to find out about advances being made by others who are exploring different issues. After all, we not only want to discover and evaluate strategies that may enhance student learning (one of our particular areas of interest) but also want to know about other discoveries and advances. Put simply, as educators, we want to know what works best and the strength of evidence for the effectiveness of different approaches. Given our assumption that many readers will share these interests, we encouraged authors to tell us what works best and, where possible, to offer prescriptions to students and instructors based on evidence demonstrating that the prescribed approach or technique is effective (e.g., by outperforming an often used but less effective approach in the area).

When making prescriptions, Robinson and Levin (Chapter 2, this volume) provocatively argue for caution, because even promising educational advances may "in fact [be] overpromises, in that the research from which the advances are derived is not based on replicable, scientifically 'credible' evidence" (p. 35). We agree, and apparently so do contributors to this volume who described the best evidence for an approach and evaluated its promise against well-established scientific criteria. In some cases, a great deal of evidence has experimentally established the efficacy of a particular approach. In other cases, however, evidence is promising but suggestive rather than definitive. For example, correlational research has revealed a meaningful relation between spatial skills, reasoning, and mathematics (Newcombe, Booth, & Gunderson, Chapter 5, this volume). These outcomes suggest that improving spatial skills may have broad benefits but they do not guarantee such benefits. In such cases, authors understandably did not offer strong prescriptions but instead highlighted viable future directions that could further establish the causal efficacy of an approach.

Reviewing the strength of evidence concerning the efficacy of any one approach can require considerable time. Despite some of our best attempts (e.g., Dunlosky et al., 2013), so much progress has been made recently that we have not been able to keep up. So, we selfishly discussed coediting a volume on cognition and education – why not have experts provide us with a bird's eye view on their areas? This selfishness is also reflected in our charge to the authors. We wanted each handbook chapter to satisfy both seasoned researchers who are interested in catching up on familiar areas of research and researchers who are looking for good entry points to begin new research within a less familiar area. For each chapter, we asked authors to start by defining their area of research and, if possible, by providing a brief history. We encouraged authors to then showcase some of their own favorite cutting-edge work, and the bulk of most chapters is aimed at helping readers develop a solid grasp of the current state of a particular area.

Finally, the origin of this particular handbook is significant. Keith Sawyer recently edited an outstanding volume, *The Cambridge handbook of the learning sciences*. The contribution of cognitive psychology is represented in that volume, but Dr. Sawyer also realized that much of the experimental work by cognitive psychologists was not included. As a consulting editor, he approached us to develop a companion handbook. We believe this handbook is a fitting companion, given that learning scientists and cognitive psychologists share the same primary goals of helping students succeed (for more on the learning sciences versus cognitive psychology, see Sawyer & Dunlosky, Chapter 1, this volume). We also hope that other companions to this handbook will be forthcoming; despite our desire to develop a handbook to survey all of the great work in cognitive psychology and education, much more excellent work is going on than could be covered in just a single volume. Even so, the current set of chapters does provide a representative and broad-ranging selection of topics from the intersection of cognition and education.

How the Handbook Is Organized

As suggested above, one way to categorize approaches to improving education is to separate them into those that focus on identifying strategies (or skills) that are effective for a specific domain and those that focus on identifying strategies that are more domain-general. An example of a relatively domain-specific strategy could involve developing an effective technique aimed at helping children to understand fractions, which is arguably an essential skill for developing math expertise (Sidney, Thompson, & Opfer, Chapter 7, this volume). An example of a domain-general approach is using multimedia presentations, which involve developing lessons that combine pictures and words. As Mayer's synthesis (Chapter 18, this volume) of multimedia research indicates, some approaches to using multimedia can undermine learning, whereas a properly designed multimedia presentation can have a meaningful impact on students' ability to gain knowledge and transfer it to solving problems in new contexts.

This domain-specific versus domain-general taxonomy is reflected in how we organized the chapters, grouping those about particular domains (*Science and Math; Reading and Writing*) and those about more domain-general strategies or skills (*General Strategies; Metacognition*). Of course, taxonomies can oversimplify, and we acknowledge the fuzzy boundaries between these groupings. For instance, although Luk and Kroll (Chapter 12, this volume) discuss some learning approaches that are specifically relevant to second-language acquisition (e.g., immersion, translanguaging), they also note the relevance of domain-general strategies such as interleaving. As another example, we placed Peverly and Wolf's overview (Chapter 13, this volume) on note-taking under *Reading and Writing* (given that taking notes can involve a great deal of both). Even so, their opening statement discloses the domain-general importance of good note-taking skills: "Note-taking is a pervasive and important activity that includes notes taken on lectures in classrooms but also notes taken in other contexts such as trials (jurors' notes), physician and clinicians' offices, and boardrooms, among others" (p. 320). Likewise, we included Griffin, Mielicki, and Wiley's contribution (Chapter 24) under *Metacognition*, which includes people's monitoring of their task progress and hence can be relevant to any ongoing activity. Despite the fact that such metacognitive skills are often viewed as domain-general (e.g., in the COPES model; Winne & Marzouk, Chapter 27, this volume), students likely will need to use different tools to accurately monitor their progress for different learning tasks and domains. In the present case, Griffin, Mielicki, and Wiley (Chapter 24, this volume) discuss current theory and evidence for helping students monitor their learning and understanding of *text materials* (and hence that chapter could also have been grouped within the *Reading and Writing* section).

Further blurring the boundaries, even domain-general approaches may need domain-specific adjustments to maximize their promise as a learning tool for particular domains. For example, Renkl and Eitel (Chapter 21, this volume) discuss research on self-explanation, which occurs when a student explains why a particular idea is true, why a particular solution is correct, why a particular procedure is appropriate, and so on. As these cases illustrate, self-explanation can be instantiated in many different ways and thus can be applied to many domains, such as when a student is trying to understand abstract principles of science writing or learning math using worked examples. Despite self-explanation being a rather general learning tool, Renkl and Eitel (Chapter 21, this volume) acknowledge that "It is highly probable that different types of self-explanations have different functions, lead to better learning via different mechanisms, and should not be regarded as a unitary construct when providing practice recommendations" (p. 531). Thus, applying this general learning technique may require task-specific adjustments.

Likewise, given that students may be limited in their ability to monitor their progress in some domains, Azevedo and colleagues have been systematically evaluating how computer-assisted learning systems can be developed to help students better regulate their learning by providing appropriate feedback about guiding their learning, using strategies, and monitoring progress (Azevedo, Mudrick, Taub, & Bradbury, Chapter 23, this volume). This general approach to developing learning systems that scaffold effective self-regulated learning has been successfully applied to a variety of specific domains, such as learning about the human circulatory system and learning to foster scientific reasoning. Certainly, depending on the targeted content or skills, we suspect that the most successful approach to fostering self-regulation skills will not be identical across domains. For instance, computer-assisted learning systems that help students regulate unwanted emotions during a math lesson may not be as critical for systems that help students learn about ecology.

In sum, although we organized the volume in an attempt to reflect domain-general versus domain-specific approaches, other arrangements would have been viable. We invite readers to keep an open mind, given that many chapters have relevance to both general and specific approaches.

Common Themes and Next Steps

A variety of themes emerged across chapters, including (1) understanding the psychological mechanisms that are relevant to a given activity; (2) exploring the promise of a particular learning technique through well-controlled investigations and for multiple learning outcomes; (3) establishing that an approach showing promise in the laboratory works in a real-world setting; and (4) mapping out the most important next steps for research in their field. We briefly consider these themes (and a few others) in the next sections.

Theory and application. A major goal of cognitive research is to reveal the basic cognitive mechanisms involved in skilled performance (e.g., reading and math). In some areas, a great deal of the theoretically motivated research is being conducted with the applied aim of directly improving student achievement. In some areas, however, a gap still exists between theory and application. For example, as aptly noted by Cook and O'Brien (Chapter 10, this volume) in the context of reading research:

> Since the early 1970s, the psychological study of reading has focused primarily on the basic cognitive processes and mechanisms involved in every stage of reading – from decoding to parsing to comprehension. Research on reading in education, however, has focused more on reading outcomes. Although research in reading is becoming increasingly interdisciplinary, the divide between psychological theory and educational practices remains wide . . . [T]he value of understanding basic comprehension processes is in the service of researchers whose aim is to develop improved methods of teaching reading and/or to develop effective interventions to assist struggling readers. (p. 237)

Cook and O'Brien argue for the importance of bridging the gap between theory and practice and offer examples of the positive impact of using theory to inform best practices in early literacy education.

The volume is packed with chapters that discuss current theories of how people perform various educationally relevant tasks, which have implications for improving student learning and task performance. For instance, the value of understanding basic processes is exemplified by Carvalho and Goldstone (Chapter 16, this volume), who consider the potential benefit of interleaving practice over blocked practice. Although interleaving has been shown to boost performance in learning concepts, their Sequential Attention Theory indicates that interleaving will not always be better than blocking practice and, more important, it describes those conditions in which each technique will benefit learning the most. Another example comes from Klahr, Zimmerman, and Matlen (Chapter 4, this volume). They describe the Scientific Discovery as Dual Search model that both defines and explains scientific thinking and has direct implications for improving people's scientific reasoning. They even highlight studies focusing on children's domain-specific and domain-general knowledge, which affirms our observation above that (despite how we organized chapters in this handbook) many chapters have relevance to more than one domain. Importantly, the theories offered in these and other chapters can be used to guide future theoretical and applied work, and we hope readers will be inspired by these theoretical advances and seek to test their implications in the laboratory and in authentic educational settings.

For a final example, Nokes-Malach, Zepeda, Richey, and Gadgil (Chapter 20, this volume) review the benefits and costs of collaborative learning. Importantly, they describe the mechanisms that result in these benefits and costs, and then subsequently appeal to these mechanisms for offering prescriptions to educators on how to maximize benefits and minimize costs when students collaborate while learning. Certainly, these prescriptions will require further evaluation in classroom settings, but as Nokes-Malach and colleagues emphasize, "There is also extensive evidence from the educational psychology approach for the benefits of learning in a group versus learning individually, especially outside of the laboratory in classroom contexts" (p. 508). The key point here is that revealing the underlying mechanisms can sharpen prescriptions and lead to systematic research aimed at (1) evaluating and improving an approach to learning (in this case, collaborative learning) and (2) potentially revising and updating the theory.

What techniques work and what counts as working? Although many of the approaches to learning or instruction offered in these chapters are partly (or entirely) informed by well-vetted theories, at least some approaches have been shown to work before researchers have fully understood why they work. Notable examples here include the use of retrieval practice (in which theory development has been relatively recent) and spaced practice (in which theoretical debates continue). Thus, even though theories often have implications for improving the efficacy of particular techniques or interventions, theory is not essential for establishing that a particular technique works.

Of course, this claim may lead one to ask, what counts as "working"? Part of the answer to this question is that a technique is working when students perform better on the targeted educational outcome. But this immediately begs a subsequent question: Better than what? Effectiveness studies often involve comparing a new technique or approach to business as usual. However, as argued by Hattie (2009), almost any intervention in the classroom is bound to produce some effect, partly because students and teachers get appropriately excited by any changes that occur and by new and enthusiastic faces (i.e., education researchers) visiting the classroom. Thus, appropriate comparison conditions or groups are needed to establish that the particular approach being evaluated (and not just social engagement or interest in trying something new) is responsible for any improvements.

The search for such evidence is analogous to the concern raised by Samuel Hopkins Adams (1905) in *Collier's Weekly* about the patent medicine business and subsequent calls for experimental evidence to establish that new medical treatments actually help people. In contrast to the Food and Drug Administration that oversees medical recommendations in the United States, there is no similar agency to ensure that educational interventions marketed to and adopted by schools and educators actually work.[1] Firmly establishing that an intervention works in authentic educational settings raises many challenges, and this handbook includes innovative examples of how this is done (for some specific recommendations on how to meet these challenges, see Dunlosky et al., in press). Moreover, in an address to the American Education Research Association, Dr. Whitehurst (2003) (Director of the Institute of Education Sciences [IES] at the time) emphasized that the position of IES is that "randomized trials are the only sure method for determining the effectiveness of education programs and practices." Importantly, this position does not mean that other approaches are invalid or uninformative, because randomized trials do not address all relevant educational issues and other approaches (e.g., case studies and

[1] However, see the What Works Clearing House, https://ies.ed.gov/ncee/wwc/

surveys) can mutually inform why a particular approach works and/or how to improve it. Readers will find a multiplicity of methodological approaches in the current handbook, all aimed at revealing how people learn and how to improve their learning.

Given the range of educational content areas represented in this volume, it is perhaps no surprise that the targeted educational outcomes vary greatly. A common theme pertains to the degree to which a particular intervention produces transfer to new contexts, tasks, and so forth (for a general framework of transfer, see Barnett & Ceci, 2002). Halpern and Butler's chapter (Chapter 3, this volume) captures the key question in their own domain: "Can students enhance their critical thinking skills in ways that endure over time and transfer across domains?" Spoiler alert: their emphatic answer to this question is "Yes", and they provide concrete recommendations on how to teach critical thinking to promote its widespread transfer. In this (and most other) cases, transfer is a good thing – it's often what we strive for when developing educational interventions. Thus, it may be surprising that some forms of transfer can be detrimental, causing problems instead of solving them. In the domain of math, Sidney, Thompson, and Opfer (Chapter 7, this volume) convincingly argue for the importance of children's fraction knowledge for subsequent achievement in mathematical cognition. Unfortunately, however, children's knowledge of whole numbers can negatively bias their conceptions of fractions! Thus, an important goal for all education research will be to understand how to maximize positive transfer and minimize negative transfer.

In our own work exploring interventions to boost learning (e.g., Rawson & Dunlosky, 2011), we have argued that researchers and practitioners should be concerned not only with the level and durability of learning achieved but also with the efficiency of using a given educational tool. Students are required to learn a vast amount of information and to master many tasks and procedures, but the amount of time and effort they have to expend is necessarily limited. Being a successful student can be overwhelming, so recommending an approach that takes a great deal of time to implement may not be feasible and, even if it is, students may simply not want to use that much time. Certainly, learning is difficult and takes time – thinking scientifically, understanding fractions, integrating information from multiple texts, and so forth is likely going to take a great deal of time and effort to do well. Nevertheless, all else equal, more efficient approaches to obtaining a learning goal are obviously preferable. In discussing best practices for correcting student misconceptions, Marsh and Eliseev (Chapter 17, this volume) note that providing feedback after a student responds correctly is rather inert and thus simply a waste of time. Likewise, as van Gog, Rummel, and Renkl (Chapter 8, this volume) emphasize, studying worked examples rather than attempting to solve problems may be the best way to begin for novice learners – not only do worked examples often yield higher levels of learning than problem-solving but they also typically have a significant advantage with respect to efficiency. In this case, novice learners can obtain the same or better level of performance and do so using less time.

Another important factor to consider when evaluating approaches to improving education is usability. Even a well-vetted technique that has been shown to boost student learning above business as usual can be inert if students or instructors cannot or do not use it with fidelity. Many educational tools discussed in the present handbook show promise partly because they are inexpensive and presumably easy-to-use, which means that the use of the tool can be widespread and inclusive. Along these lines, one exciting discovery is that using simple hand gestures can improve instruction and learning. Wakefield and Goldin-Meadow (Chapter 9, this volume) note that "A good teaching tool is one that can be implemented broadly," and they provide evidence that supplementing math instruction with the right kind of gestures can improve students' understanding of difficult math lessons (e.g., equivalence). Other approaches may be broad and cost-free, yet students may need instruction and practice using the tool to reap its benefits. For example, in the context of writing to learn, Klein and Van Dijk (Chapter 11, this volume) argue that elementary students will "need explicit instruction and examples of how to use strategies, as well as prompting to actually apply them . . . [The importance of strategy instruction] has also been independently demonstrated with respect to genres such as summary, discourse synthesis, and argumentation" (p. 284).

Another key theme that is prevalent in this handbook is that some strategies are effective for some students but not others. That is, individual differences on a variety of dimensions may moderate the impact of a given educational intervention (i.e., *Treatment X Aptitude* interactions). A non-exhaustive list of factors that can moderate the effectiveness of a given technique includes the grade level of students (Stevens & Vaughn, Chapter 15, this volume), the diversity of materials being used within an intervention (Halpern & Butler, Chapter 3, this volume), differences in the structure of to-be-learned concepts (Carvalho & Goldstone, Chapter 16, this volume), and the level of prior knowledge that students have about a targeted domain (van Gog, Rummel, & Renkl, Chapter 8, this volume). Prior knowledge is a particularly likely suspect, given its long history of moderating effects in many different literatures. For example, prior knowledge has long been known to moderate how well students learn from individual texts (e.g., McNamara et al., 1996), and Rouet, Britt, and Potocki (Chapter 14, this volume) conclude that prior knowledge also is relevant to integrating and comprehending content across multiple texts:

> Our brief review of the literature demonstrates that some specific task instructions are effective at enhancing students' multiple text comprehension. However, it is worth noting that several of these studies also showed that not all readers respond in the same way . . . Indeed, [prior knowledge] of the readers themselves [is] likely to influence multiple text processing and to mediate the effects of task instructions. (p. 370)

We invite readers to consider this and other factors that may moderate the benefits of each educational intervention or strategy, because discovering these moderators is an important goal for education researchers.

In summary, our major thesis in this section is simply that a particular technique or approach can work (or fail to do so) for many reasons. Of course, if it simply fails to promote better learning than business as usual, then adopting it likely makes no sense. Nevertheless, even approaches that have been shown to promote learning may not work because (1) they are prohibitively inefficient; (2) they are too difficult to use; or (3) students and/or teachers hold misconceptions about the effectiveness of the strategy or technique (for details, see Bjork, Dunlosky, & Kornell, 2013). In such cases, students and teachers may be unlikely to adopt an effective strategy. Thus, another avenue for research will be to understand how to sidestep these barriers, so as to better inform students and teachers about what works best and to modify effective strategies to make them more useable.

Exploring the impact of educational interventions in the wild. Many cognitive and education researchers conduct a great deal of research in their laboratories and some of us are a bit afraid to venture beyond them. Nevertheless, regardless of how well a particular intervention performs in the laboratory, it may not have the same impact when implemented in a classroom or in environments in which students are regulating their own learning. In reviewing the chapters in this handbook, we were gratified to find that the efficacy of many approaches is being evaluated outside of the laboratory. Conducting classroom-based research (or any research outside of the laboratory) poses challenges, especially if one aims to firmly establish efficacy using randomized trials; it can often take a great deal of time and resources to complete. Such investigations and demonstrations are critical, and we briefly consider two (perhaps obvious) reasons next.

First, realizing that experimental outcomes will not always transfer outside of the laboratory, researchers should use caution when making prescriptions to students and teachers in general until some questions are empirically addressed. Does the intervention work for the intended material and students when delivered in a classroom setting? Does the intervention produce educationally meaningful gains on the targeted knowledge or skills, whether that involves high-stakes exams, the quality of student projects, how well students write, and so forth? Does the intervention work only when the research team delivers it to a particular class or will it benefit students even when implemented by teachers or by students on their own? These are just a few of the important questions that can be addressed when one wishes to establish the efficacy of an approach outside the laboratory. Note, however, we are not arguing that all of these questions must be answered before making prescriptions to educators and students. Instead, we encourage researchers to address them when possible, and when they cannot, to consider qualifying their prescriptions by explaining the breadth and strength of the evidence supporting them.

Second, experimental evidence that a given intervention or strategy works outside the laboratory (e.g., boosts performance above an appropriate comparison group or control in an authentic educational context) may be helpful in convincing students and teachers to adopt the technique. We suspect that enthusiastic and charismatic champions of a particular approach can make a great deal of headway in promoting buy-in by administrators, teachers, and students. Like patent medicine in the early 1900s, however, stakeholders in education reform can be persuaded to adopt practices that can be costly (at minimum, adopting a new approach will require extra time) but do not work well. One high-profile example is how an emphasis on learning