# Organizing Instruction and Study to Improve Student Learning

IES Practice Guide

NCER 2007-2004

U.S. DEPARTMENT OF EDUCATION

NATIONAL CENTER FOR EDUCATION RESEARCH

Institute of Education Sciences

# Organizing Instruction and Study to Improve Student Learning

IES Practice Guide

**SEPTEMBER 2007**


**Harold Pashler (Chair)**
University of California, San Diego

**Patrice M. Bain**
Columbia Middle School, Illinois

**Brian A. Bottge**
University of Wisconsin–Madison

**Arthur Graesser**
University of Memphis

**Kenneth Koedinger**
Carnegie Mellon University

**Mark McDaniel**
Washington University in St. Louis

**Janet Metcalfe**
Columbia University



NCER 2007-2004

**U.S. DEPARTMENT OF EDUCATION**

NATIONAL CENTER FOR EDUCATION RESEARCH
Institute of Education Sciences

This report was prepared for the National Center for Education Research, Institute of Education Sciences, under contract no. ED-05-CO-0026 to Optimal Solutions Group, LLC.

**Disclaimer**

The opinions and positions expressed in this practice guide are the authors' and do not necessarily represent the opinions and positions of the Institute of Education Sciences or the U.S. Department of Education. This practice guide should be reviewed and applied according to the specific needs of the educators and education agencies using it and with full realization that it represents only one approach that might be taken, based on the research that was available at the time of publication. This practice guide should be used as a tool to assist in decision-making rather than as a "cookbook." Any references within the document to specific education products are illustrative and do not imply endorsement of these products to the exclusion of other products that are not referenced.

**U.S. Department of Education**
Margaret Spellings
*Secretary*

**Institute of Education Sciences**
Grover J. Whitehurst
*Director*

**National Center for Education Research**
Lynn Okagaki
*Commissioner*

**September 2007**



Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., and Metcalfe, J. (2007) *Organizing Instruction and Study to Improve Student Learning* (NCER 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ncer.ed.gov.

This report is available for download on the IES website at http://ncer.ed.gov.

**Alternative Formats**
On request, this publication can be made available in alternative formats, such as Braille, large print, audiotape, or computer diskette. For more information, call the Alternative Format Center at (202) 205-8113.

# Organizing Instruction and Study to Improve Student Learning

## Contents

## List of Tables

# Preamble from the Institute of Education Sciences

## What is a practice guide?

The health care professions have embraced a mechanism for assembling and communicating evidence-based advice to practitioners about care for specific clinical conditions. Variously called practice guidelines, treatment protocols, critical pathways, best practice guides, or simply practice guides, these documents are systematically developed recommendations about the course of care for frequently encountered problems, ranging from physical conditions such as foot ulcers to psychosocial conditions such as adolescent development.[1]

Practice guides are similar to the products of typical expert consensus panels in reflecting the views of those serving on the panel and the social decisions that come into play as the positions of individual panel members are forged into statements that all are willing to endorse. However, practice guides are generated under three constraints that do not typically apply to consensus panels. The first is that a practice guide consists of a list of discrete recommendations that are intended to be *actionable*. The second is that those recommendations taken together are intended to be a *coherent* approach to a multifaceted problem. The third, which is most important, is that each recommendation is explicitly connected to *the level of evidence* supporting it, with the level represented by a grade, e.g., strong, moderate, and low. The levels of evidence, or grades, are usually constructed around the value of particular types of studies for drawing causal conclusions about what works. Thus one typically finds that the top level of evidence is drawn from a body of randomized controlled trials, the middle level from well-designed studies that do not involve randomization, and the bottom level from the opinions of respected authorities (see table 1). Levels of evidence can also be constructed around the value of particular types of studies for other goals, such as the reliability and validity of assessments.

Practice guides can also be distinguished from systematic reviews or meta-analyses, which employ statistical methods to summarize the results of studies obtained from a rule-based search of the literature. Authors of practice guides seldom conduct the types of systematic literature searches that are the backbone of a meta-analysis, though they take advantage of such work when it is already published. Instead, they use their expertise to identify the most important research with respect to their recommendations, augmented by a search of recent publications to assure that the research citations are up-to-date. Further, the characterization of the quality and direction of the evidence underlying a recommendation in a practice guide relies less on a tight set of rules and statistical algorithms and more on the judgment of the authors than would be the case in a high quality meta-analysis. Another distinction is that a practice guide, because it aims for a comprehensive and coherent approach, operates with more numerous and more contextualized statements of what works than does a typical meta-analysis.

Thus, practice guides sit somewhere between consensus reports and meta-analyses in the degree to which systematic processes are used for locating relevant research and characterizing its meaning. Practice guides are more like consensus panel reports than meta-analyses in the breadth and complexity of the topic that is addressed. Practice guides are different from both consensus reports and meta-analyses in providing advice at the level of specific action steps along a pathway that represents a more or less coherent and comprehensive approach to a multifaceted problem.

## Practice guides in education at the Institute of Education Sciences

The Institute of Education Sciences (IES) publishes practice guides in education to bring the best available evidence and expertise to bear on the types of systemic challenges that cannot currently be addressed by single interventions or programs. Although IES has taken advantage of the history of practice guides in healthcare to provide models of how to proceed in education, education is different from health-care in ways that may require that practice guides in education have somewhat different designs. Even within health care, where practice guides now number in the

[1] Field and Lohr (1990).

**Table 1. Institute of Education Sciences Levels of Evidence**

| Level | Description |
|---|---|
| **Strong** | In general, characterization of the evidence for a recommendation as strong requires both studies with high internal validity (i.e., studies whose designs can support causal conclusions), as well as studies with high external validity (i.e., studies that in total include enough of the range of participants and settings on which the recommendation is focused to support the conclusion that the results can be generalized to those participants and settings). Strong evidence for this practice guide is operationalized as:<br>• A systematic review of research that generally meets the standards of the What Works Clearinghouse (see http://ies.ed.gov/ncee/wwc/) and supports the effectiveness of a program, practice, or approach with no contradictory evidence of similar quality; OR<br>• Several well-designed, randomized, controlled trials or well-designed quasi-experiments that generally meet the standards of the What Works Clearinghouse and support the effectiveness of a program, practice, or approach, with no contradictory evidence of similar quality; OR<br>• One large, well-designed, randomized, controlled, multisite trial that meets the standards of "the What Works Clearinghouse" and supports the effectiveness of a program, practice, or approach, with no contradictory evidence of similar quality; OR<br>• For assessments, evidence of reliability and validity that meets The Standards for Educational and Psychological Testing.[2] |
| **Moderate** | In general, characterization of the evidence for a recommendation as moderate requires studies with high internal validity but moderate external validity, or studies with high external validity but moderate internal validity. In other words, moderate evidence is derived from studies that support strong causal conclusions but where generalization is uncertain, or studies that support the generality of a relationship but where the causality is uncertain. Moderate evidence for this practice guide is operationalized as:<br>• Experiments or quasi-experiments generally meeting the standards of the What Works Clearinghouse and supporting the effectiveness of a program, practice, or approach with small sample sizes and/or other conditions of implementation or analysis that limit generalizability, and no contrary evidence; OR<br>• Comparison group studies that do not demonstrate equivalence of groups at pretest and therefore do not meet the standards of the What Works Clearinghouse but that (a) consistently show enhanced outcomes for participants experiencing a particular program, practice, or approach and (b) have no major flaws related to internal validity other than lack of demonstrated equivalence at pretest (e.g., only one teacher or one class per condition, unequal amounts of instructional time, highly biased outcome measures); OR<br>• Correlational research with strong statistical controls for selection bias and for discerning influence of endogenous factors and no contrary evidence; OR<br>• For assessments, evidence of reliability that meets The Standards for Educational and Psychological Testing but with evidence of validity from samples not adequately representative of the population on which the recommendation is focused. |
| **Low** | In general, characterization of the evidence for a recommendation as low means that the recommendation is based on expert opinion derived from strong findings or theories in related areas and/or expert opinion buttressed by direct evidence that does not rise to the moderate or strong levels. Low evidence is operationalized as evidence not meeting the standards for the moderate or high levels. |

[2] American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999)

thousands, there is no single template in use. Rather, one finds descriptions of general design features that permit substantial variation in the realization of practice guides across subspecialties and panels of experts.[3] Accordingly, the templates for IES practice guides may vary across practice guides and change over time and with experience.

The steps involved in producing an IES-sponsored practice guide are first to select a topic, which is informed by formal surveys of practitioners and requests. Next, a panel chair is recruited who has a national reputation and up-to-date expertise in the topic. Third, the chair, working in collaboration with IES, selects a small number of panelists to co-author the practice guide. These are people the chair believes can work well together and have the requisite expertise to be a convincing source of recommendations. IES recommends that at least one of the panelists be a practitioner with experience relevant to the topic being addressed. The chair and the panelists are provided a general template for a practice guide along the lines of the information provided in this preamble. They are also provided with examples of practice guides. The practice guide panel works under a short deadline of 6-9 months to produce a draft document. The expert panel interacts with and receives feedback from staff at IES during the development of the practice guide, but they understand that they are the authors and thus responsible for the final product.

One unique feature of IES-sponsored practice guides is that they are subjected to rigorous external peer review through the same office that is responsible for independent review of other IES publications. A critical task of the peer reviewers of a practice guide is to determine whether the evidence cited in support of particular recommendations is up-to-date and that studies of similar or better quality that point in a different direction have not been ignored. Peer reviewers are also asked to evaluate whether the evidence grade assigned to particular recommendations by the practice guide authors is appropriate. A practice guide is revised as necessary to meet the concerns of external peer reviews and gain the approval of the standards and review staff at IES. The process of external peer review is carried out independent of the office and staff within IES that instigated the practice guide.

Because practice guides depend on the expertise of their authors and their group decision-making, the content of a practice guide is not and should not, be viewed as, a set of recommendations that in every case depends on and flows inevitably from scientific research. It is not only possible, but also likely, that two teams of recognized experts working independently to produce a practice guide on the same topic would generate products that differ in important respects. Thus, consumers of practice guides need to understand that they are, in effect, getting the advice of consultants. These consultants should, on average, provide substantially better advice than an individual school district might obtain on its own because the authors are national authorities who have to achieve consensus among themselves, justify their recommendations in terms of supporting evidence, and undergo rigorous independent peer review of their product.

[3] E.g., American Psychological Association (2002).

# About the authors

Dr. Harold Pashler (Chair) is a professor in the Department of Psychology at the University of California, San Diego (Ph.D. from University of Pennsylvania). His research interests are in learning, memory and attention. He is the author of *The Psychology of Attention* (MIT Press, 1998) and was the editor-in-chief of the *Stevens Handbook of Experimental Psychology* (Wiley, 2001).

Mrs. Patrice Murphy Bain received a B.S. from the University of Iowa and M.Ed. and Ed.S. from Southern Illinois University Edwardsville. Mrs. Bain has taught in the public schools for 14 years; she was a finalist for Illinois Teacher of the Year and a Fulbright Scholar in Russia. She is currently teaching middle school social studies at Columbia Middle School in Columbia, IL.

Dr. Brian A. Bottge is a Professor of Special Education in the Department of Rehabilitation Psychology and Special Education at the University of Wisconsin-Madison (Ed.D. from Vanderbilt University). He has combined his extensive classroom experience with learning theory to develop and test technology-based curricula for improving the mathematics learning of low-achieving students in middle and high schools.

Dr. Arthur Graesser is a professor in the Department of Psychology, an adjunct professor in Computer Science, and co-Director of the Institute for Intelligent Systems at the University of Memphis (Ph.D. from the University of California, San Diego). His primary research interests are in learning sciences, cognitive science, and discourse processing, with specific interests in text comprehension, tutoring, conversation, question asking and answering, and the design of advanced learning environments with computer tutors (including AutoTutor). Dr. Graesser is editor of the *Journal of Educational Psychology*, former editor of the journal *Discourse Processes*, and was senior editor of the *Handbook of Discourse Processes*.

Dr. Ken Koedinger is a professor in the Human-Computer Interaction Institute, School of Computer Science and Psychology Department, at Carnegie Mellon University (Ph.D. from Carnegie Mellon University). He is also the Carnegie Mellon University director of the Pittsburgh Science of Learning Center (PSLC). His research has contributed new principles and techniques for the design of educational software and has produced basic cognitive science research results on the nature of mathematical thinking and learning. Dr. Koedinger serves on the editorial board of *Cognition and Instruction*.

Dr. Mark McDaniel is a Professor of Psychology at Washington University in St. Louis (Ph.D. from University of Colorado). His research is in the general area of human learning and memory, with an emphasis on encoding and retrieval processes in memory and applications to educational contexts. He has published over 160 journal articles, book chapters, and edited books on human learning and memory. Dr. McDaniel serves on the editorial board of *Cognitive Psychology* and *Journal of Experimental Psychology*: *Learning, Memory, and Cognition*.

Dr. Janet Metcalfe is a Professor of Psychology and of Neurobiology and Behavior at Columbia University (Ph. D. from University of Toronto). She has conducted studies applying Principles of Cognitive Science to Enhance Learning with at-risk inner city children for over 10 years. Her current research centers on how people – both children and adults – know what they know, that is, their metacognitive skills and abilities, and whether they use these abilities efficaciously – for effective self-control. Dr. Metcalfe serves on the editorial board of *Psychological Review* and *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

# Disclosures of potential conflicts of interest

Practice guide panels are composed of individuals who are nationally recognized experts on the topics about which they are rendering recommendations. IES expects that such experts will be involved professionally in a variety of matters that relate to their work as a panel. Panel members are asked to disclose their professional involvements and to institute deliberative processes that encourage critical examination the views of panel members as they relate to the content of the practice guide. The potential influence of panel members' professional engagements is further muted by the requirement that they ground their recommendations in evidence that is documented in the practice guide. In addition, the practice guide is subjected to independent external peer review prior to publication, with particular focus on whether the evidence related to the recommendations in the practice guide has been has been appropriately presented.

The professional engagements reported by each panel members that appear most closely associated with the panel recommendations are noted below.

Dr. Koedinger has researched practices discussed in this guide such as self-explanation and worked examples. He is a shareholder and receives royalties from Carnegie Learning, Inc. Cognitive Tutor, a product of Carnegie Learning, Inc., makes use of some of the practices described in this Guide. Cognitive Tutor is not referenced in this Guide.

Dr. Bottge has conducted several studies using Enhanced Anchored Instruction with middle school and high school students and has reported the findings in journal articles and book chapters. Dr. Bottge has provided professional development to teachers on implementing Enhanced Anchored Instruction.

# Organizing instruction and study to improve student learning

## Overview

Much of teaching is about helping students master new knowledge and skills and then helping students not to forget what they have learned. The recommendations in this practice guide are intended to provide teachers with specific strategies for organizing both instruction and students' studying of material to facilitate learning and remembering information, and to enable students to use what they have learned in new situations.

One distinguishing characteristic of our recommendations is a relatively high degree of concreteness. Concrete questions about how to promote learning were the main focus of the earliest work in educational psychology during the first half of the 20th Century.[4] However, concrete choices about procedures and timing received much less attention in the later part of the 20th Century. In the past 5 years or so, partly due to support from the Institute of Education Sciences, there has been a flurry of new interest in these topics, and the empirical research base has grown rapidly.

The seven recommendations in this practice guide reflect our panel's consensus on some of the most important concrete and applicable principles to emerge from research on learning and memory (see table 2). The first recommendation about the spacing of key course content is an overarching principle that teachers should attend to as they plan out sequences of instruction. This recommendation provides advice that is intended to help students remember information longer. Our second, third, and fourth recommendations relate to how different forms of instruction should be combined: worked example solutions and new problems posed to the student (in Recommendation 2), graphical and verbal descriptions of concepts and mechanisms (Recommendation 3), and abstract and concrete representations of a concept (Recommendation 4). Recommendation 5 reflects our ongoing concern with memory. In these days of high-stakes tests, teachers are often reminded of how often students appear to have mastered information and concepts in December or February, only to have forgotten them by June. As well as using spacing to mitigate forgetting, a substantial body of work recommends that teachers use quizzing, both formal and informal, as a tool to help students remember. Although forgetting is a reality of life, its effects can be somewhat mitigated through appropriate use of what we call "spaced" learning and through strategic use of quizzing.

Recommendation 6 relates to students' ability to judge how well they have learned new knowledge or skills—psychologists refer to this ability as "metacognition." We recognize that this recommendation may strike the reader as a bit exotic. It is our belief, however, that students' ability to manage their own studying is one of the more important skills that students need to learn, with consequences that will be felt throughout their lives. Psychological research has documented the fact that accurately assessing one's own degree of learning is not something that comes naturally to our species, and fostering this ability is a useful, albeit neglected, component of education.

Finally, we have included a seventh recommendation that targets ways to shape instruction as students gain expertise in a particular domain. After students have acquired some basic skill and conceptual knowledge of a topic, we recommend that teachers selectively ask students to try to answer "deep" questions that focus on underlying causal and explanatory principles. A sizable body of research shows that this activity can facilitate learners' mastery of a domain.

In sum, we recommend a set of actions that teachers can take that reflect the process of teaching and learning, and that recognizes the ways in which instruction must respond to the state of the learner. It also reflects our central organizing principle that learning depends upon memory, and that memory of skills and concepts can be strengthened by relatively concrete—and in some cases quite nonobvious strategies—and we hope that the users of this guide will find these recommendations to be of some value in their vital work.

[4] E.g., Mace (1932); Starch (1927).

| Table 2. Recommendations and corresponding Level of Evidence to support each | |
|---|---|
| **Recommendation** | **Level of Evidence** |
| 1. Space learning over time. *Arrange to review key elements of course content after a delay of several weeks to several months after initial presentation.* | **Moderate** |
| 2. Interleave worked example solutions with problem solving exercises. *Have students alternate between reading already worked solutions and trying to solve problems on their own.* | **Moderate** |
| 3. Combine graphics with verbal descriptions. *Combine graphical presentations (e.g., graphs, figures) that illustrate key processes and procedures with verbal descriptions.* | **Moderate** |
| 4. Connect and integrate abstract and concrete representations of concepts. *Connect and integrate abstract representations of a concept with concrete representations of the same concept.* | **Moderate** |
| 5. Use quizzing to promote learning. *Use quizzing with active retrieval of information at all phases of the learning process to exploit the ability of retrieval directly to facilitate long-lasting memory traces.*<br>5a. *Use pre-questions to introduce a new topic*<br>5b. *Use quizzes to re-expose students to key content* | **5a. Low**<br>**5b. Strong** |
| 6. Help students allocate study time efficiently. *Assist students in identifying what material they know well, and what needs further study, by teaching children how to judge what they have learned.*<br>6a. *Teach students how to use delayed judgments of learning to identify content that needs further study*<br>6b. *Use Tests and Quizzes to Identify Content that Needs to be Learned* | **6a. Low**<br>**6b. Low** |
| 7. Ask deep explanatory questions. *Use instructional prompts that encourage students to pose and answer "deep-level" questions on course material. These questions enable students to respond with explanations and supports deep understanding of taught material.* | **7. Strong** |

# Scope of the practice guide

The purpose of the present practice guide is to provide evidence-based recommendations on the organization of study and instruction. These recommendations are intended to suggest ways that teachers can organize their instructional time and help students structure their use of study time to promote faster learning and better retention of knowledge across a broad range of subject matters.

The primary intended audience for this practice guide consists of teachers and guidance counselors in elementary, middle, and high schools. However, some of the issues and recommendations discussed here are also relevant to the decisions made by publishers of textbooks and designers of educational technologies, because these kinds of products exert an important influence on how study and instructional time are organized. Although the findings described here are probably as pertinent to college instruction as to lower grades, our most direct concern in producing this guide has been education from 3rd through 12th grade.

Although our recommendations are directed to professional educators, we believe that some of the information presented in this practice guide includes valuable information that students themselves should be aware of. Thus, it is our hope that the present recommendations may help educators not only when they set about to decide questions such as "How shall I use my class time?" and "What should I include in my homework assignments?," but also when they consider "What advice should I give to students who ask me how best to study for my class?" We have also included a checklist for teachers to assist them in carrying out the recommendations (see page 4).

The recommendations described here reflect research carried out in the fields of cognitive science, experimental psychology, education, and educational technology. The backgrounds of the panelists encompassed all of these fields. Our primary goal here has been to identify relatively concrete actions relating to the use of instructional and study time that are generally applicable to subjects that demand a great deal of content learning. Social studies and science instruction are obvious examples, but the recommendations are by no means limited to those areas.

As pointed out in a preceding section, a distinctive feature of IES Practice Guides is that they provide an explicit assessment of the degree of empirical support enjoyed by each of the recommendations offered. When we stated that a recommendation is backed up by "strong" evidence this generally meant that it received considerable support from randomized experimental studies, both in well-controlled laboratory contexts and within the context of schools. Strength levels of "moderate" and "low" imply a correspondingly weaker and narrower evidence base. When the evidence level fell short of "strong" this was usually because although the evidence was experimental in character, it was limited to laboratory studies, thus making the applicability of the results to other situations (e.g., classroom instruction) less certain.

In classifying levels of empirical support for the effectiveness of our recommendations, we have been mindful not only to the issue of whether a study meets the "gold-standard" of a randomized trial, but also to the question "Effective as compared to what?" Virtually any educational manipulation that involves exposing students to subject content, regardless of how this exposure is provided, is likely to provide some benefit when compared against no exposure at all. To recommend it, however, the question becomes "Is it more effective than the alternative it would likely replace?" In laboratory studies, the nature of instruction in the control group is usually quite well defined, but in classroom studies, it is often much less clear. In assessing classroom studies, we have placed most value on studies that involve a baseline that seems reasonably likely to approximate what might be the "ordinary practice default".

## Checklist for carrying out the recommendations

**Recommendation 1:**
**Space learning over time.**

- ☐ Identify key concepts, terms, and skills to be taught and learned.
- ☐ Arrange for students to be exposed to each main elements of material on at least two occasions, separated by a period of at least several weeks—and preferably several months.
- ☐ Arrange homework, quizzes, and exams in a way that promotes *delayed* reviewing of important course content.

**Recommendation 2:**
**Interleave worked example solutions with problem-solving exercises.**

- ☐ Have students alternate between reading already worked solutions and trying to solve problems on their own.
- ☐ As students develop greater expertise, reduce the number of worked examples provided and increase the number of problems that students solve independently.

**Recommendation 3:**
**Combine graphics with verbal descriptions.**

- ☐ Use graphical presentations (e.g., graphs, figures) that illustrate key processes and procedures. This integration leads to better learning than simply presenting text alone.
- ☐ When possible, present the verbal description in an audio format rather than as written text. Students can then use visual and auditory processing capacities of the brain separately rather than potentially overloading the visual processing capacity by viewing both the visualization and the written text.

**Recommendation 4:**
**Connect and integrate abstract and concrete representations of concepts.**

- ☐ Connect and integrate abstract and concrete representations of concepts, making sure to highlight the relevant features across all forms of the representation.

**Recommendation 5:**
**Use quizzing to promote learning.**

- ☐ Prepare pre-questions, and require students to answer the questions, before introducing a new topic.
- ☐ Use quizzes for retrieval practice and spaced exposure, thereby reducing forgetting.
- ☐ Use game-like quizzes as a fun way to provide additional exposure to material.

**Recommendation 6:**
**Help students allocate study time efficiently.**

- ☐ Conduct regular study sessions where students are taught how to judge whether or not they have learned key concepts in order to promote effective study habits.
- ☐ Teach students that the best time to figure out if they have learned something is *not* immediately after they have finished studying, but rather after a delay. Only after some time away from the material will they be able to determine if the key concepts are well learned or require further study.
- ☐ Remind students to complete judgments of learning without the answers in front of them.
- ☐ Teach students how to use these delayed judgments of learning techniques after completing assigned reading materials, as well as when they are studying for tests.
- ☐ Use quizzes to alert learners to which items are not well learned.
- ☐ Provide corrective feedback to students, or show students where to find the answers to questions, when they are not able to generate correct answers independently.

**Recommendation 7:**
**Ask deep explanatory questions.**

- ☐ Encourage students to "think aloud" in speaking or writing their explanations as they study; feedback is beneficial.
- ☐ Ask deep questions when teaching, and provide students with opportunities to answer deep questions, such as: *What caused Y? How did X occur? What if? How does X compare to Y?*
- ☐ Challenge students with problems that stimulate thought, encourage explanations, and support the consideration of deep questions.

## Recommendation 1: Space learning over time.

To help students remember key facts, concepts, and knowledge, we recommend that teachers arrange for students to be exposed to key course concepts on at least two occasions—separated by a period of several weeks to several months. Research has shown that delayed re-exposure to course material often markedly increases the amount of information that students remember. The delayed re-exposure to the material can be promoted through homework assignments, in-class reviews, quizzes (see Recommendation 3), or other instructional exercises. In certain classes, important content is automatically reviewed as the learner progresses through the standard curriculum (e.g., students use single-digit addition nearly every day in second grade math), and this recommendation may be unnecessary in courses where this is the case. This recommendation applies to those (very common) course situations in which important knowledge and skills are not automatically reviewed.

### Level of evidence: Moderate

The panel judges the level of evidence supporting this recommendation to be *moderate* based on three experimental classroom studies examining the effects of this practice for improving school-aged students' performance on academic content (e.g., mathematics, spelling),[5] two experimental classroom studies that examined the effect of this strategy for improving college students' academic performance,[6] and the hundreds of laboratory experiments which have been completed examining the effects of massed versus distributed practice on memory.[7]

### Brief summary of evidence to support the recommendation

Hundreds of laboratory experiments have been carried out which present materials to learners on two separate occasions. Then, following a delay, the learners are given some sort of recall test on the material. Although a few inconsistencies have been found, by far the most common finding is that when the time between study sessions is very brief relative to the amount of time to the final test, students do not do as well on the final test.[8] Students typically remember much more when they have been exposed to information on two occasions, rather than one, and when the interval between these two occasions is not less than about 5 percent of the interval during which the information has to be retained. In the studies that have tested this principle of delayed review, researchers have kept constant the amount of time that students have to learn the information;[9] thus, the observed improvement in learning is not a result of learners having more time to study the material. Delaying of reviews produces an actual increase in the efficiency of learning. Having too long a temporal spacing separating learning sessions has been found to produce a small decrease in final memory performance as compared to an optimal spacing, but the cost of "overshooting" the right spacing is consistently found to be much smaller than the cost of having very short spacing. Thus, the practical implication is that it makes sense to be sure to have enough spacing, but it rarely makes sense to worry about having too much.

---

5 Rea and Modigliani (1985); Bloom and Shuell (1981); Carpenter, Pashler, Cepeda, et al. (2007).

6 Rohrer and Taylor (2006); Bahrick, Bahrick, Bahrick, et al. (1993).

7 See Cepeda, Pashler, Vul, et al. (2006) for a review.

8 Examples of what is meant by a brief interval relative between study sessions would be a 10-second interval when the test occurs a half hour later, or a one-day delay when the test occurs months later.

9 For example, one group of students might spend 20 minutes learning the definitions of a list of words and then have a test on those words ten days later. These students would be compared to a group of students who spend 10 minutes on one day learning the definitions and then 10 minutes on another day reviewing the definitions.

Research on the delayed review of materials has examined learning of (a) mathematical skills,[10] (b) foreign language vocabulary,[11] and (c) historical and other facts.[12] Although the research literature primarily involves well-controlled laboratory studies, there are a number of classroom-based studies that have shown similar results. One recent study examined memory for historical facts by eighth-graders enrolled in a U.S. history class.[13] The study compared the effect of a review given 1 week after initial presentation, versus 16 weeks after. On a final test given 9 months after the review session, the 16-week delay group showed significantly greater performance (almost 100 percent increase) as compared to the 1-week delayed group.

One limitation of the literature is that few studies have examined acquisition of complex bodies of structured information.[14] For measurement reasons, researchers have mostly focused on acquisition of isolated bits of information (e.g., facts or definitions of vocabulary words). The acquisition of facts and definitions of terms is certainly an essential component of mastering any complex content domain, and may have broad cultural utility,[15] but the panel recognizes that acquiring facts and key definitions is merely one goal of schooling.[16] There does not appear to be any evidence to suggest that spacing benefits are confined to isolated elements of course content.

## How to carry out the recommendation

The key action recommended here is for teachers to make sure that important curriculum content is reviewed at least several weeks, and ideally several months, after the time that it was first encountered by the students. Research shows that a delayed review typically has a large positive impact on the amount of information that is remembered much later. The benefit of a delayed review seems to be much greater than the same amount of time spent reviewing shortly after initial learning. This review can occur in a variety of ways, including those described below.

***1. Use class time to review important curriculum content.***

For example, every other week a high school social studies teacher spends half a class period reviewing facts that were covered several weeks earlier in the class.

***2. Use homework assignments as opportunities for students to have spaced practice of key skills and content.***

For example, in every homework assignment, a junior high school math teacher intentionally includes a few problems covering the material presented in class 1 or 2 months earlier.

***3. Give cumulative midterm and final examinations.***

When teachers give their students cumulative midterm and final examinations, students are provided with a strong incentive to study all course material at widely separated points in time.

## Possible roadblocks and solutions

***Roadblock 1.1.*** Most textbooks contain reviews and problem sets that deal only with the most recently taught material.

***Solution.*** Teachers can supplement problem sets provided in the textbook with at least a "sprinkling" of problems relating to material covered much earlier in the course. One may hope that in the future, textbook publishers will respond to the growing body of research on spacing of learning and develop textbooks that directly promote spaced review of key concepts and procedures.

***Roadblock 1.2.*** Teachers may frequently become discouraged during a review session to discover that many students appear to have forgotten what they appeared to have mastered several weeks earlier.

***Solution.*** By implementing our recommended practice of spacing over time, teachers will find that students

---

[10] E.g., Rohrer and Taylor (2006, in press).
[11] E.g., Dempster (1987); Bahrick, Bahrick, Bahrick, et al. (1993).
[12] E.g., Carpenter, Pashler, Cepeda, et al. (2007); Pashler, Rohrer, Cepeda, et al. (2007).
[13] Carpenter, Pashler, Cepeda, et al. (2007).
[14] Ausubel and Youssef (1965) showed benefits of delayed review on memory for a coherent passage on endocrinology, but the comparison was with a procedure that lacked the delayed review (rather than one that included a review at a short lag).
[15] See Hirsch (1987) for a discussion.
[16] See Bloom (1956) for a well-known discussion.

remember more. At the beginning of this process, teachers should expect to see substantial forgetting. Although this initial forgetting may be discouraging, the panel reminds our readers that research shows that even when students cannot recall previously learned material, reawakening of the knowledge through reviewing is more easily accomplished than was the original learning (psychologists refer to this as "savings"), and the final result of the delayed review is a marked reduction in the rate of subsequent forgetting.[17] Thus, by implementing spaced review, the teacher can not only repair the forgetting that will have happened since initial learning, but also, to some degree, inoculate against subsequent forgetting.

[17] Berger, Hall, and Bahrick (1999).

## Recommendation 2: Interleave worked example solutions and problem-solving exercises.

When teaching mathematical or science problem solving, we recommend that teachers interleave worked example solutions and problem-solving exercises—literally alternating between worked examples demonstrating one possible solution path and problems that the student is asked to solve for himself or herself—because research has shown that this interleaving markedly enhances student learning.

### Level of evidence: Moderate

The panel judges the level of evidence supporting this recommendation to be *moderate*. Numerous laboratory experiments provide support for the benefits of interleaving worked example solutions and problem-solving exercises.[18] Some classroom experiments provide further evidence that the recommendation can be practically and effectively implemented in real courses at the K-12 and college levels.[19] These experiments have explored these techniques in a variety of content domains, particularly in mathematics, science, and technology.

### Brief summary of evidence to support the recommendation

A large number of laboratory experiments and a smaller number of classroom studies have demonstrated that students learn more by alternating between studying examples of worked-out problem solutions and solving similar problems on their own than they do when just given problems to solve on their own. For example, in a series of laboratory experiments in the domain of algebra,[20] researchers had 8th and 9th grade students in the treatment condition alternate between four pairs of solution examples and problems. Students in the control condition were simply asked to solve eight problems, as one might typically ask students in a homework assignment. Students in the interleaved example/problem treatment condition not only took less time to complete the eight problems, but also performed better on the post-test. Another study, in the domain of computer programming,[21] found that if students are given all six examples before all six problems, they learn significantly less than if the examples and problems are interleaved, with the same six examples and six problems alternating in order. An early classroom study[22] compared conventional mathematics instruction with a similar instruction in which some class activities, particularly lectures, were replaced with worked example study. The results showed a dramatic acceleration in learning such that students finished a 3-year course sequence in 2 years with as good or better final test performance. The benefits of interleaving examples and problems for improving learning efficiency and learning outcomes

[18] E.g., Catrambone (1996; 1998); Cooper and Sweller (1987); Kalyuga, Chandler, and Sweller (2001); Kalyuga, Chandler, Tuovinen, et al. (2001); Paas and van Merriënboer (1994); Renkl (1997; 2002); Renkl, Atkinson, and Große (2004); Renkl, Atkinson, Maier, et al, (2002); Renkl, Stark, Gruber, et al. (1998); Schwonke, Wittmer, Aleven, et al. (2007); Schworm and Renkl (2002); Sweller (1999); Sweller and Cooper (1985); Trafton and Reiser (1993); Ward and Sweller (1990).

[19] E.g., McLaren, Lim, Gagnon, et al. (2006); Zhu and Simon (1987).

[20] Sweller and Cooper (1985).

[21] Trafton and Reiser (1993).

[22] Zhu and Simon (1987).

have been demonstrated in many other laboratory studies[23] and some other classroom studies.[24]

The amount of guidance and annotation that should be included in worked examples presented to students probably varies depending on the situation and the student. But at least in some studies, worked examples that did not include instructional explanations of the steps were found to be most effective.[25] Other studies have found that labeling groups of steps within a problem solution according to what goal they seek to achieve can be helpful.[26]

As students develop greater expertise, decreased example use and correspondingly increased problem solving appears to improve learning.[27] Gradually "fading" examples into problems, by giving early steps in a problem and requiring students to provide more and more of the later steps as they acquire more expertise with the problem type, also seems to benefit student learning.[28]

Finally, using worked examples and problems that involve greater variability from one example or problem to the next (e.g., changing both the values included in the problem and the problem formats), after students receive instruction on a mathematical concept, puts greater demands on students during study but pays off in better learning and post-test performance.[29]

### How to carry out the recommendation

Instead of giving students a list of problems to solve as a homework assignment, teachers should provide a worked out solution for every other problem on the list.[30] Here is an example of what we mean.

Consider a typical homework or seatwork assignment involving eight math problems. Following the interleaving principle, the teacher might take the same eight math problems and provide students with the worked out solution for every other problem.

Let's say that the even-numbered items would be usual problems, like the following algebra problem.

Solve $5 + 3x = 20$ for $x$

The odd numbered problems, come with solutions, like this:

**Below is an example solution to the problem:**

"Solve $12 + 2x = 15$ for $x$"

**Study each step in this solution, so that you can better solve the next problem on your own:**

$$12 + 2x = 15$$
$$2x = 15\text{-}12$$
$$2x = 3$$
$$x = 3/2$$
$$x = 1.5$$

Which approach, asking for solutions to all eight problems or interleaving four examples with four problems, will be lead to better student learning? Intuitively, one might think that because solving eight problems gives students more practice, or because students might ignore the examples, that assigning eight problems would lead to more learning. But, as discussed in the previous section, much research has shown that students typically learn more deeply and more easily from the second approach, when examples are interleaved between problems.

In whole classroom situations, a teacher might implement this recommendation by beginning with a class or small group discussion around an example solution followed by small groups or individuals solving a problem (just one!) on their own. The

23 E.g., Cooper and Sweller (1987); Kirshner, Sweller, and Clark (2006); Renkl (1997).
24 For example, see Ward and Sweller (1990).
25 Hausmann and VanLehn (in press); Schworm and Renkl (2002).
26 Catrambone (1996; 1998).
27 Kalyuga, Chandler, and Sweller (2001); Kalyuga, Chandler, Touvinen, et al. (2001).
28 Renkl, Atkinson, and Große (2004); Renkl, Atkinson, Maier, et al. (2002); Schwonke, Wittwer, Aleven, et al. (2007).
29 Paas and van Merriënboer (1994); Renkl, Stark, Gruber, et al. (1998).
30 The example provided below is based on Sweller and Cooper (1985).

teacher then directs the class back to studying an example, for instance, by having students present their solutions and having others attempt to explain the steps (see Recommendation 7). After studying this worked example, the students are given a second problem to solve. Again, this follows the principle of interleaving worked examples with problems to solve.

### Potential roadblocks and solutions

***Roadblock 2.1.*** Curricular materials do not often provide teachers with large numbers of worked example solutions.

***Solution.*** Teachers can work together on teams to prepare homework sets that interleave worked examples with problems for students to solve. Teachers can take worked examples included in the instructional section of the textbook and interleave them into the assigned homework problem sets.

***Roadblock 2.2.*** Teachers may be concerned that by providing large numbers of worked-out examples to students, they will memorize the solution sequences and not attain mastery of the underlying concepts being taught and reinforced through this interleaving technique.

***Solution.*** By having problems to solve in between the worked examples, students are motivated to pay more attention to the worked example because it helps them prepare for the next problem and/or resolve a question from the past problem. Having problems to solve helps students recognize what they do not understand. Students are notoriously poor at identifying what they do not understand (see Recommendation 6 for a discussion of learners' "illusion of knowing"). By interleaving worked examples with problems to solve, students are less inclined to skim the example because they believe that the answer is obvious or they already know how to solve this type of problem.

## Recommendation 3: Combine graphics with verbal descriptions.

We recommend that teachers combine graphical presentations (e.g., graphs, figures) that illustrate key processes and concepts with verbal descriptions of those processes and concepts in order to facilitate student learning.

### Level of evidence: Moderate

The panel judges the level of evidence supporting this recommendation to be *moderate*. Many laboratory experiments provide support for the benefits of combining graphical presentations and verbal descriptions of key processes and concepts.[31] Some classroom experiments and quasi-experiments provide further evidence that the recommendation can be practically and effectively implemented in real courses at the K-12 and college levels.[32] Again, it is important to note that these experiments have explored these techniques in a variety of content domains particularly in mathematics, science, and technology.

### Brief summary of evidence to support the recommendation

Many studies have demonstrated that adding relevant graphical presentations to text descriptions can lead to better learning than text alone.[33] Most of these studies have focused on scientific processes, for example, how things work (e.g., lightning, disk brakes, bike pumps, volcanic eruptions). These studies emphasize that it is important that text descriptions appear near the relevant elements in visual representations to best enhance learning.[34] In addition, students learn more when the verbal description is presented in audio form rather than in written text,[35] probably because a learner cannot read text and scrutinize an accompanying graphic at the same time. It should be noted that current evidence suggests that a well chosen sequence of still pictures with accompanying prose can be just as effective in enhancing learning as narrated animations.[36]

The benefits of interleaving graphics and verbal description have also been demonstrated for certain kinds of mathematics instruction. Researchers have found that adding a number-line visualization to mathematics instruction significantly improved learning.[37] Students required to use a number line while performing addition and subtraction of signed numbers showed better learning than students who solved equations without the number line. Classroom studies of this approach have demonstrated large student learning improvements in mathematics at the elementary, middle, and high school levels.[38]

---

31 E.g., Clark and Mayer (2003); Mayer (2001); Mayer and Anderson (1991; 1992); Mayer and Moreno (1998); Moreno and Mayer (1999a); Mousavi, Low, and Sweller (1995).

32 E.g., Griffin, Case, and Siegler (1994); Kalchman, Moss, and Case (2001); Kalchman and Koedinger (2005); Moss (2005).

33 See Mayer (2001) and Sweller (1999) for reviews.

34 For example, see Moreno and Mayer (1999a).

35 Clark and Mayer (2003); Mayer (2001); Mayer and Anderson (1991; 1992); Mayer and Moreno (1998); Moreno and Mayer (1999a); Mousavi, Low, and Sweller (1995).

36 Mayer, Hegarty, Mayer, et al. (2005); Pane, Corbett, and John (1996).

37 Moreno and Mayer (1999b).

38 For example, see Griffin, Case, and Siegler, 1994; Kalchman, Moss, and Case, 2001; Kalchman and Koedinger, 2005; Moss, 2005.

### How to carry out the recommendation

***When teaching students about processes and procedures that can be well represented through pictures, figures, charts, video clips, or other graphic formats, teachers should combine verbal description of the key steps in a process with graphical representations that illustrate these steps.***

Here is an example of what we mean. Consider the task of teaching a science topic, such as what causes the seasons or how lightning works. Providing visual representations that illustrate how such processes unfold can enhance learning. Such visual representations should be integrated with verbal descriptions that help students focus on where to look and on what is being illustrated. When visual representations are used in text materials or written handouts, they should include brief text that labels unfamiliar objects and describes steps in the process being illustrated. These descriptions should be positioned as close as possible to the parts of the visualization being described and help students identify what specifically they should be looking at. When visual representations are used in lecture or multimedia, it is useful to describe the objects and processes in speech while simultaneously indicating the relevant parts of the visual representations.

To enhance learning, teachers should choose pictures, graphs, or other visual representations carefully. The visual representations need to be relevant to the processes or concepts that are being taught. For instance, a picture of a high school football player whose football helmet has been scarred by lightning is interesting, but it may well detract from learning about how lightning works.

Graphics do not have to be completely realistic to be useful. Sometimes a more abstract or schematic picture will best illustrate a key idea, whereas a more photorealistic graphic may actually distract the learner with details that are irrelevant to the main point. For example, students may learn better about the two loops of the human circulatory system (heart to body and heart to lungs) from a more abstract visualization of the heart chambers than from a realistic illustration of the heart. Animations may sometimes add interest value, but a well-chosen sequence of still pictures is often as, or more, effective in enhancing learning.

Graphics in mathematics can help students make connections between mathematical symbols and procedures and the quantities and relations they represent. Such connections are the basis for conceptual understanding. For example, the use of number lines can help students master a wide range of mathematics topics including basic counting, place value, rational numbers, and integers. It is important to make regular integrative connections between steps in the symbolic procedures and how they are represented in visual representations.

Finally, graphics can be used to help students understand abstract ideas. For example, using multiple representations (e.g., symbols, graphs, pictures, or real objects) of the same abstract concept allows students to see that the concept can be depicted in many different ways. Authentic situations can be portrayed through stories, real world problem scenarios, or movie clips and used to convey abstract concepts. When using multiple visual representations of an abstract concept, teachers should draw students' attention to the components of the visualization that are relevant to the abstract concept so that students understand that the same core idea is being expressed in multiple ways.

### Potential roadblocks and solutions

***Roadblock 3.1.*** Instructional materials may present verbal descriptions of a graphic or figure on a different page of the text, or alternatively not include a verbal description that aligns with the graphic or figure.

***Solution.*** Teachers should preview the instructional materials that their students will be learning from and make sure to draw the students' attention to the verbal description that maps onto the graph or figure. In addition, when preparing instructional materials or homework assignments, teachers should attend to the physical alignment of the graphs or figures and their matching verbal description.

## Recommendation 4: Connect and integrate abstract and concrete representations of concepts.

We recommend that teachers connect and integrate abstract representations of a concept with concrete representations of the same concept. Connecting different forms of representations helps students master the concept being taught and improves the likelihood that students will use it appropriately across a range of different contexts.

### Level of evidence: Moderate

The panel judges the level of evidence supporting this recommendation to be *moderate*. A substantial number of laboratory experiments provide support for the benefits of connecting and interleaving both abstract and concrete representations of problems.[39] A growing number of classroom experiments and quasi-experiments provide further evidence that the recommendation can be practically and effectively implemented in courses at the K-12 and college levels, and with students of different abilities.[40] These research efforts have explored these techniques in a variety of content domains particularly in mathematics, science, and technology.

### Brief summary of evidence to support the recommendation

Many experimental laboratory studies and a growing number of classroom based quasi-experiments have found that teaching students about key principles or concepts using only abstract or only concrete representations of those concepts leads to less flexible knowledge acquisition and use than teaching students to recognize and use those key principles across a range of different situations. Although some classroom research suggested that young learners, or learners being taught a new concept, benefited from using concrete objects—such as blocks to solve problems—other research finds that learning with concrete objects supports initial understanding of the instructed concept, but does not support the transfer of that knowledge to novel but relevant contexts.[41] Experimental research with both college students and K-12 learners finds that although students have an easier time acquiring an initial understanding of a concept presented in a concrete form, those same students are unable to use that knowledge in a different context (e.g., to solve a problem with the same underlying structure).[42] On the other hand, when students are initially introduced to a concept using a more abstract representation, those students struggle slightly more to master the concept initially, but are then able to use their new understanding successfully in a different context. It seems that the greater initial difficulty in comprehending abstract instruction is compensated for by a greater ability to apply the concept to very different situations. Thus, teachers need to be aware of both the limits and benefits of providing initial instruction using concrete representations.

An emerging set of research is examining the best ways for teachers to incorporate the use of both concrete

[39] E.g., Goldstone and Sakamoto (2003); Goldstone and Son (2005); Kaminski, Sloutsky, and Heckler (2006a; 2006b); Richland, Zur, and Holyoak (2007); Sloutsky, Kaminiski, & Heckler (2005).

[40] E.g., Bottge (1999); Bottge, Heinrichs, Chan, et al. (2001); Bottge, Heinrichs, Mehta, et al. (2002); Bottge, Rueda, Serlin, et al. (2007); Bottge, Rueda, and Skivington (2006); Bottge, Rueda, LaRoque, et al. (2007).

[41] Resnick and Omanson (1987); Amaya, Uttal, and DeLoache (submitted).

[42] Kaminiski, Sloutsky, and Heckler (2006); Sloutsky, Kaminski, and Heckler (2005).

instantiations and abstract representations during instruction. In an attempt to capitalize on the benefits of initial learning using a concrete representation and to support students' understanding of an abstract representation of the same concept or principle, one line of research suggests using a technique they call "concreteness fading."[43] Concreteness fading is a process by which initial learning with a concrete representation occurs, and then over time, key components of the concrete representation are replaced by more idealized and abstract representations. For example, in one line of experimental work examining the use of computer simulations in science instruction,[44] initial learning about the ecological concept of *competitive specialization* used images of ants seeking food. In the concreteness fading condition, as students gain experience with the concept, the ants were replaced with black dots, and the food sources with green patches. Students in the concreteness fading condition outperformed students in the concrete-only or abstract-only condition, both on measures of initial learning and on their ability to use the principle of competitive specialization to understand a different problem. In classroom instruction, this technique of using concrete representation to introduce a concept or principle, and then systematically replacing relevant components of the concrete representation with abstract representations, can also be used, and holds promise for helping learners with a range of abilities and prior knowledge to master abstract representation of the concept.

A second line of research indicates that explicit marking of the relationships between different types of representations supports learning.[45] As in the research described above, a critical aspect of using both concrete and abstract representations of a concept seems to be the role of the instructor in drawing students' attention to the relevant and shared components of the concrete and abstract representation. When students are not provided guidance, it is difficult for learners to identify which components of the problem can be transferred to new problems.[46] On the other hand, when teachers explicitly identify the critical components of a representation and draw students' attention to those critical components, their performance improves.[47] Finally, there is evidence that lower-achieving students demonstrate improved learning when they are asked to solve hands-on or authentic problems that require the use of these underlying key concepts or principles.[48] Again, the role of the teacher or peers in guiding the student to note the critical components of the concept or problem across the different representational forms is important.

Taken together, these research findings support our recommendation that teachers use both abstract and concrete representations of key concepts and highlight the critical aspects of the concept to be learned (e.g., pointing out to the student which variables in the mathematical function being taught are related to which aspects of the word problem). This process of interleaving and connecting both concrete and abstract representations has been shown to support better mastery of the taught principle, as well as transfer to other tasks that require students to use the same principle or concept.

## How to carry out the recommendation

When teaching students about an abstract principle or skill, such as a mathematical function, teachers should connect those abstract ideas to relevant concrete representations and situations, making sure to highlight the relevant features across all forms of the representation of the function. An abstract idea, like a mathematical function, can be expressed in many different ways: Concisely in mathematical symbols like "$y = 2x$"; visually in a line graph that starts at 0 and goes by 2 units for every 1 unit over, discretely in a table showing that 0 goes to 0, 1 goes to 2, 2 goes to 4, and so on; practically in a real world scenario like making $2 for every mile you walk in a walkathon; and, physically by walking at two miles per hour. By showing students the same idea in different forms, teachers can demonstrate that although the "surface" form may vary, it is the "deep" structure—what does not change—that is the essence of the idea. Teachers

[43] Goldstone and Son (2005).
[44] Goldstone and Sakamoto (2003).
[45] E.g., Ainsworth, Bibby, and Wood (2002); Bottge, Rueda, LaRoque, et al. (2007); Richland, Zur, and Holyoak (2007).
[46] Ainsworth, Bibby, and Wood (2002).
[47] Richland, Zur, and Holyoak (2007).
[48] Bottge, Rueda, LaRoque, et al. (2007).

can also get at the deep structure by showing how variations in an idea, like receiving $3 per mile in a walkathon, lead to particular variations in each representation: The graph is now steeper; the table has $y$ values that go up by 3 rather than 2; and, the equation is now $y = 3x$.

When students first encounter a new idea they may pick up on the wrong features of the examples we give them. They might think that averages are about sports if we give them mostly sports examples or, more subtly, that an average is a ratio between two numbers (e.g., hits to at-bats) rather than a ratio of a sum of measures to the number of those measures (e.g., $(2 + 6 + 4)/3$). Or, they might think the seasons are caused by how the tilt of the earth brings parts of the earth closer to the sun rather than by how the tilt causes sunlight to be spread out in some places and more concentrated in others. A variety of representations and explicit discussion of the connections between them can help students avoid such misconceptions.

Instruction may often move too quickly to the use of new terms or symbols before students have had a chance to understand the meanings of those new terms or symbols through drawing connections to multiple familiar objects or situations that the terms represent. Students may be able to memorize new terms and their definitions, or learn how to manipulate new symbolic forms (like mathematical equations), without ever drawing such connections. However, that knowledge may end up being "inert" in the sense that a student cannot easily apply it beyond the specific examples or situations used in instruction. Again, asking students to apply this knowledge across multiple examples that vary in their relative concreteness or abstractness ensures that students acquire a more flexible understanding of the key concept.

Another technique involves connecting or "anchoring" new ideas in stories or problem scenarios that are interesting and familiar to students. Thus, students not only have more motivation to learn, but have a strong base on which to build the new idea and on which to return later if they forget. Further, by using a variety of successively more abstract representations of the new idea, students can develop conceptions that get beyond the surface features of those early examples and get to the deep features and core concepts that are the essence of the idea.

## Potential roadblocks and solutions

***Roadblock 4.1.*** Explicit connections between abstract concepts and their concrete representations are not always made in textbooks, nor in instructional materials prepared to support teachers.

***Solution.*** When preparing examples and instructional materials, textbook publishers and teachers should clearly identify which aspects of an abstract representation and its concrete instantiation are connected. We believe that having these relationships clearly identified ahead of time can support the use of this recommended technique during instruction.

## Recommendation 5: Use quizzing to promote learning.

The process of taking a quiz or test can directly promote learning in the context of classroom instruction, and reduce the rate at which information is forgotten. In Recommendation 5, we recommend two ways of using quizzing to help students learn: (a) using "pre-questions" to activate prior knowledge and focus students' attention on the material that will be presented in class; and (b) using quizzes to re-expose students to key course content. Recommendation 6 includes a third way to use quizzing to help students make decisions about allocating study time.

### Recommendation 5a: Use pre-questions to introduce a new topic.

We recommend that teachers use pre-questions as a way to introduce a new topic. Pre-questions (or pre-tests) help students identify what material they do not yet know, and hence need to study. In addition, responding to pre-questions automatically activates any relevant prior knowledge in the student's mind. These processes contribute to improved student learning.

### Level of evidence: Low

The panel judges the level of evidence supporting this recommendation to be *low* based on a series of laboratory experiments primarily carried out on learning from reading written text.[49] Most of this research has been completed with college students. It has not yet been tested as a component of regular classroom instruction.

### Brief summary of evidence to support the recommendation

A body of experimental studies on learning from written text has established that when students are given pre-questions to answer prior to reading both expository and narrative text, they learn more from the text than when they do not respond to such pre-questions. Some of these studies have used actual classroom material (e.g., textbook chapters). However, there is little or no published experimental evidence regarding whether pre-questions will promote the learning of orally presented classroom content as well. Accordingly, even though the evidence is reasonably consistent, it has not been demonstrated using methods of delivery common in classroom practice, and thus we cannot characterize it as "strong" or "moderate."

There is one important caveat to this recommendation. In some experiments in which students were not explicitly discouraged from reading the text selectively based on the pre-questions, pre-questions tended to reduce learning for non-questioned material. However, researchers have demonstrated that pre-questions do not hinder learning of non-questioned material when learners are explicitly required to read all of the material. Moreover, in such cases the advantages shown for learning of pre-questioned material remains. These results suggest that when pre-questions are used to preview the content for assigned readings that students are encouraged (or required) to read in full, there likely will be gains in learning for pre-questioned material and no penalty for non-questioned material.

---

[49] E.g., Beck, McKeown, Hamilton, and Kucan (1997); Craig, Sullins, Witherspoon and Gholson (2006); Driscoll, Craig, Gholson, Ventura, and Graesser (2003); Gholson and Craig (2006); King (1994, 1996, 2006); Rosenshine, Meister, and Chapman (1996); Wisher and Graesser (2007).

### How to carry out the recommendation

To carry out this recommendation, teachers should use pre-questions as a way to introduce new topics. The pre-questions should address a few of the important concepts that are covered in the new material. Because one purpose of pre-questions is to direct students' attention to key facts and concepts, teachers should avoid creating pre-questions that highlight extraneous information.

***Although existing research has focused on learning from reading, we believe that a reasonable extrapolation from this research is the use of pre-questions when teachers begin oral instruction on topics.***

For example, a middle school social studies teacher might begin a section on World War II by asking the following pre-question (among others): "Why were people imprisoned in concentration camps?" In a college neuroscience course, a teacher might pose the pre-question, "Normally, the left hemisphere of the brain processes what kind of information?"

One simple way to use pre-questions, particularly with middle and high school students, is to prepare several pre-questions that can be copied onto a sheet of paper, and placed on the students' desks for them to answer immediately upon taking their seats. These questions should be relatively quick to answer, but should require students to describe or explain their responses to the questions. For example, in a sixth-grade social studies class, in preparation for a lesson on "Bodies of Water" in a unit devoted to the geography of the United States, students could be asked to list the major bodies of water in the United States, to define the term "drainage basin", and to answer this question: "What is the main cause of ocean currents?" These questions serve to preview the classroom instruction for the day.

### Potential roadblocks and solutions

***Roadblock 5a.1.*** A teacher might wonder how they will get their students to attend to the non-questioned as well as the pre-questioned material, particularly when quizzes and spaced study also focus on key concepts.

***Solution.*** To avoid students' attending to only pre-questioned material, teachers can emphasize (emphatically) that it is important for the students to attend to all of the daily lesson and all of the reading. This idea could be reinforced by noting that the pre-questions could not cover all of the important concepts that students would be expected to learn.

***Roadblock 5a.2.*** Some teachers might object that this is 'giving students the answer' before they have even covered the new material—that no mental work is left for the student, and that this is simply feeding into the frenzy of, 'Just tell me what I need to know so I can do well on the test,' with little regard left for sparking a student's intrinsic motivation to learn.

***Solution.*** To foster students' involvement in learning, teachers could focus class discussion on explaining correct and incorrect alternatives to the pre-questions. For example, a pre-question used in a middle school social studies class on ancient Egypt was "What were the ancient doctors NOT able to do?", with the alternatives "give shots", "cure illnesses", "measure heart beats", and "fix broken bones." The teacher could use the alternatives to stimulate discussion on why some medical practices in ancient Egypt were possible and others not.

Further, to encourage learning of a complex fact, rather than learning of a particular answer when given a particular question, teachers could change the wording of test items from those used for pre-questions. A concept from a college-level course covered by the quiz question, "All preganglionic axons, whether sympathetic or parasympathetic, release what neurotransmitter?" could be tested with the question "What axons, whether sympathetic or parasympathetic, release acetylcholine as a neurotransmitter?"

## Recommendation 5b: Use quizzes to re-expose students to information.

We recommend that teachers use "closed-book" quizzes or tests as one method for re-exposing students to key course content. As indicated in Recommendation 1, a delayed re-exposure to course content helps students remember key information longer. In addition, quizzes or tests that require students to actively recall specific information (e.g., questions that use fill-in-the-blank or short-answer formats, as opposed to multiple-choice items) directly promote learning and help students remember information longer. To use quizzes or tests to promote learning and retention of information, correct-answer feedback should be provided.

### Level of evidence: Strong

The panel judges the level of evidence supporting this recommendation to be *strong* based on nine experimental studies examining the effects of this practice for improving K-12 students' performance on academic content or classroom performance, over 30 experimental studies that examined the effect of this strategy for improving college students' academic performance, and the large number of carefully controlled laboratory experiments that have examined the testing effect.[50]

### Brief summary of evidence to support the recommendation

Laboratory experiments have repeatedly demonstrated that taking a test on studied material promotes remembering that material on a final test, a phenomenon called the "testing effect." Experimental memory research has established that the testing effect is very robust.[51] The testing effect generalizes across a wide range of materials, including word lists, pictorial information, and prose material. Testing effects are observed across a range of ages from elementary school children to college students. Testing effects surface when the intervening tests are different from the final tests; for example, intervening tests with fill-in-the-blank items improve subsequent performance on tests that use multiple-choice and true/false items and vice versa.

Perhaps most importantly, researchers have found that having students take a test is almost always a more potent learning device than having students spend additional time studying the target material.[52] This is especially true when the test requires students to actively recall information (e.g., providing answers to fill-in-the-blank or short-answer/essay type items). That is, the act of recalling information from memory helps to cement the information to memory and thereby reduces forgetting. By answering the questions on the quiz, the student is practicing the act of recalling specific information from memory. For example, a recent study examined the effect of quizzing on the performance of college students enrolled in a web-based Brain and Behavior course.[53] After completing the week's reading, students either (a) took multiple-choice or short-answer quizzes, (b) re-read the key facts, or (c) did not revisit the key facts presented during that week's reading. After completing the quizzes, students received feedback that included a restatement of the quiz question and the correct answer. This process was followed throughout the semester, and students took both unit tests and a cumulative final test. Facts that students had been re-exposed to through the quizzes

---

[50] See Roediger and Karpicke (2006a) for a recent review and synthesis of both laboratory and classroom research that empirically examines the testing effect.

[51] For examples of recent research on the "testing effect" see Butler and Roediger (2007); McDaniel, Roediger, and McDermott (2007); Bjork (1988).

[52] E.g., Gates (1917); McDaniel and Fisher (1991); Carrier and Pashler (1992); Roediger and Karpicke (2006b).

were more likely to be remembered correctly on the unit tests as compared to facts that students had simply re-read (or restudied). In addition, the benefit of completing short-answer questions on the weekly quizzes extended to performance on the final test.

Moreover, several recent studies have shown that testing not only enhances learning—it also reduces the rate at which information is forgotten.[54] One recent high school-based study showed that a quiz format review of historical facts reduced forgetting over the subsequent 16 weeks, when compared to a review that presented the same content to students without requiring them to actually retrieve the facts.[55]

### How to carry out the recommendation

To carry out this recommendation, teachers should give students closed-book quizzes between the initial exposure to the material and the final assessment at the end of the semester or end of the year. Note that the quizzes can be both formal quizzes and informal testing situations, such as playing a Jeopardy-like game. The principle is that requiring students to actively recall information from memory gives them opportunities to practice recalling or retrieving that information from memory, and this practice helps to solidify that knowledge in the student's memory. Is it harmful for a learner to produce an answer that has a high likelihood of being an error? If so, should efforts be taken to discourage production of incorrect responses? Not surprisingly, in the absence of corrective feedback, any errors produced on one test will remain present, and will reappear on subsequent tests.[56] However, guessing when unsure has not so far been shown to have detrimental effects, at least with memory for facts and vocabulary.[57] In sum, then, our recommendation is to take every opportunity to prompt students to retrieve information, and whenever a substantial number of errors are expected, to be sure to make corrective feedback available.

### Potential roadblocks and solutions

***Roadblock 5b.1.*** Teachers may feel that they do not have the out-of-class time to prepare and grade additional short-answer quizzes.

***Solution.*** With the advent of technology, there are websites available to teachers that allow them to create quizzes quickly using content specified by the teacher. For example, on the website www.quia.com, teachers can create quizzes or puzzles that provide students with the opportunity to test themselves on their mastery of key facts and concepts. Such sites provide immediate feedback and the opportunity for students to actively recall the material. Teachers should also explore websites that accompany assigned textbooks and, as appropriate, require students to use them during study. Most K-12 academic publishing sites include automatically graded self-check quizzes, flashcards, and other types of self-testing opportunities that students can use in an online format.

---

53 McDaniel, Anderson, Derbish, et al. (2007).
54 Roediger and Karpicke, 2006b; Carpenter, Pashler, Wixted, et al. (in press).
55 Carpenter, Pashler, Cepeda, et al. (2007).
56 Pashler, Zarow, and Triplett (2003); Butterfield and Metcalfe (2001).
57 Pashler, Rohrer, Cepeda, et al. (2007).

## Recommendation 6: Help students allocate study time efficiently.

To promote efficient and effective study habits, we recommend that teachers help students more accurately assess what they know and do not know, and to use this information to more efficiently allocate their study time. Teachers can help students break the "illusion of knowing" that often impedes accurate assessment of knowledge in two ways.

### Recommendation 6a: Teach students how to use delayed judgment of learning techniques to identify concepts that need further study.

First, teachers can teach students how to create accurate "judgments of learning" during study. Second, students can use their performance on closed book quizzes to identify what material they need to re-study in order to master all critical course content (see recommendation 6b).

### Level of evidence: Low

The panel judges the level of evidence supporting this recommendation to be *low* because the body of evidence supporting this recommendation is primarily experimental research completed in the laboratory using academic content rather than in classroom experiments. The research provides direct evidence supporting causal links between delayed judgments of learning and accurate assessments of knowledge, delayed keyword generation and accurate assessments of knowledge, and links between accurate assessments of knowledge, study behavior, and improved performance on tests.[58] Research has been completed both with college students and school-aged children.

### Brief summary of evidence to support the recommendation

Much research has been conducted on the ability to judge how well one has learned new knowledge or skills—what psychologists call "metacognition."[59] This research finds that, without training, most learners cannot accurately judge what they do and don't know, and typically overestimate how well they have mastered material when they are finished studying. This "illusion of knowing" is reflected in the assertion that many students make after they receive a poor grade on a test: "But I studied so hard. I thought I really knew the material cold. How could I have failed?" Fortunately, a growing body of work has identified how to improve learners' ability to judge what they do and do not know after a study session. Recent research provides techniques teachers can use to teach students in overcoming this illusion of knowing so that they may spend their time studying material they have not yet mastered.

A key technique for breaking the illusion of knowing and being able to more accurately assess whether or not one knows the information is known in psychological literature as the cue-only delayed judgment of learning procedure. This technique has three critical features: one, students should test their mastery of a set of concepts, not right after they have finished studying the material, but after a meaningful delay (an hour, a day, or a week); two, when testing whether they know the concepts or not, students should only have access to the cue and not the correct answer; and three, students should judge how likely they are to get the correct answer on a quiz, rather than only attempting to generate the answer. Studies completed with learners of different ages find that when students are required to make their judgments after there has been a delay since studying the material, and without the target or the answer being present (e.g., looking at a question

[58] Hertzog, Kidder, Moman-Powell, et al., (2002); Metcalfe and Finn (in press); Thiede, Anderson, and Therriault (2003).
[59] See Metcalfe and Dunlosky (in press) for a recent review of research on metacognition.

and deciding if one knows the answer without having multiple-choice answers on the same page to choose from), they are highly accurate in their ability to judge whether or not they know the correct answer.[60] The materials used in these studies include general information questions, foreign language translations, and vocabulary terms. People, then, at all age levels, can determine accurately what they know and what they don't know, when they make the judgments in this way.

Students can also use a highly similar technique to help them judge how well they have understood a text that they have recently read. Students often read a section from a textbook, get to the end of the chapter, close the book, and believe that they have understood the material. When they try to participate in class the next day, they are often chagrined to realize that they cannot remember what they read the night before. As in studying key terms and questions, psychologists have found techniques that readers can use to help them make sure that their time spent reading course material is productive. The key components of this technique include trying to remember what was read after a delay, and generating keywords or sentences that summarize the main points of the chapter they just read. For example, when people were asked to type into the computer the definitions for four central terms from a text after a delay, and were then asked to evaluate how well they understood what they just read, they were more able to accurately identify their level of understanding.[61] Furthermore, in a similar study,[62] when students used this technique to identify their level of understanding a text, their study of the text and eventual test performance also improved. Thus, when learners attempt to second-guess the test, in order to isolate the key components, they break the illusion of knowing and are better able to spend their study time focusing on the items they don't know. This, in turn, leads to improved performance on the final test.

There have been only a few studies investigating the development of study strategies based on children's knowledge of what they do and do not know.[63] But the results are consistent in indicating that while younger children may know what they know, they often fail to put that knowledge to use in their choice of what to study. In one study, 7-year-old, 9-year-old, and college-aged students studied pictures, recalled, and then chose half of the pictures for restudy.[64] While the 9-year-olds and college students chose items not recalled correctly on the first trial, the younger children ignored their first trial performance in their selection. Similar age-related differences in the allocation of study time have also been seen in other studies with children.[65] Finally, when children in grades 3 and 5 were asked to make judgments of learning first and to then choose what they wanted to study—even though both groups of children were able to identify what they did and did not know—the younger children were more likely to choose items for restudy randomly, while the older children chose those items they knew they did not know.[66] These researchers included a second part to the experiment—either honoring or dishonoring the children's own choices. Honoring the choices only helped their eventual test performance when the children had chosen appropriately; namely, when they chose to study the items they knew they did not know. Thus, accurate identification of items that are known and unknown can be used to improve learning over time.

### How to carry out the recommendation

The development of effective study skills depends crucially on the learner being able to assess what they know and do not know. Such accurate knowledge, then, needs to be used to allow the individual to study appropriately. Unfortunately, people at all grade levels, from the youngest children through advanced college students, experience failures of metacognition and illusions of knowing in which they think they know something but do not. The illusion of knowing—thinking you know something when you don't—leads to ineffective and dysfunctional study. How can you

[60] Dunlosky and Nelson (1992); Schneider, Visé, Lockl, et al. (2000); Metcalfe and Finn (in press).
[61] Dunlosky, Rawson, and Middleton (2005).
[62] Thiede, Anderson, and Therriault (2003); Thiede, Dunlosky, Griffin, et al. (2005).
[63] See Metcalfe and Kornell (2005) for a recent discussion of those studies.
[64] Masur, McIntyre, and Flavell (1973).
[65] Dufresne and Kobasigawa (1989); See Lockl and Schneider (2002) for another example of age-related differences in metacomprehension and allocation of study time.
[66] Metcalfe and Finn (in press).

use the techniques described in the section above to help students study efficiently and effectively?

***1. Use the cue-only judgment of learning procedure in a study session during class time.***

Suppose you have a grade 3, 4 or 5 social studies class and there are central points that the children are expected to learn and understand. Schedule a study session for some days after you have taught a new section of the to-be-learned material. In order for students to make accurate judgments of learning, it is important that judgments be made sometime after initial instruction. Prepare a selection of about 10 key questions that capture the central meaning of the material.

Then give the students the questions, one at a time, *without the answers*, and ask them to judge whether or not they know the answer for each question. When completing judgments of learning, students rate how likely they think it is that they will be able to get the answer to this question right, on a scale from 0 to 100, when given a test the next day. Tell the students to try to answer the question and then make the judgment. Teachers should remind the students that people forget, and that their judgments should reflect the chance that they might forget the answer by the time of the test.

After making these judgments, students should be told to review the material and find out the answers for every question that they did not give a score of 100. If they do not know the answer, they can ask the teacher or look for the answer in their texts. They should write these answers down on a separate piece of paper and think about them, and, if necessary, discuss them with the teacher for clarification. This "cue-only delayed judgment of learning" procedure should be used repeatedly, on the same materials, over the semester. The child can use the initial set of questions, using their judgments of what they do and do not know to determine what to study further during each of the study sessions.

***2. After several opportunities to use this study technique during class, then teach students how to use the cue-only judgment of learning procedure when they are studying independently at home or in the library.***

Students can also use this technique when they are studying independently. Imagine that a student is taking a Spanish I class and needs to remember a list of 25 vocabulary words for a test next week. Using the "judgment of learning" technique described above, a student should prepare a stack of flashcards. After studying the words, the student should then go through the stack one word at a time, and should try to generate the translation. If the student is immediately able to generate the translation, say from "*leche*" to "milk," the student should rate the likelihood of answering the quiz question correctly as 100 percent. If, when presented with the word "*pan*," the student hesitates and isn't immediately sure as to what the translation is, the learner might rate their likelihood of getting that quiz question correct at 60 percent. Finally, if the student looks at the word "*huevo*" and cannot generate a response at all, the student might rate the likelihood of answering that question at 0 percent. By stopping, and assessing their likelihood of a correct response, students are better able to allocate their study time to items that need more study.

***3. Teach students how to use the delayed keyword technique to identify whether they have understood their assigned reading.***

Finally, students should also be encouraged to use the delayed keyword technique after they complete the assigned reading every evening. When completing the evening's social studies homework, for example, students should be told to read the assigned text, then complete an assignment for a different class (say complete their math problems), then they should take out a blank piece of paper and try to write down the four key terms and definitions from their social studies reading without looking back at the text. If the students find that they cannot generate four key terms, the students should re-read the text, and then follow the delayed keyword technique again. This process should ensure that the students remember more of what they read during study.

## Recommendation 6b: Use tests and quizzes to identify content that needs to be learned.

We recommend that teachers use tests or quizzes after the presentation of new material to help students identify content that requires further study. As a companion technique to Recommendation 4a, we recommend that teachers and students use tests and quizzes with feedback identifying incorrect responses, providing correct answers to those incorrect responses as tools for helping students identify content that they have not yet mastered during study.

### Level of evidence: Low

The panel judges the level of evidence supporting this recommendation to be *low* based on three experimental studies that examined the effect of this strategy for improving college students' performance on academic content (text material, vocabulary),[67] and a handful of laboratory experiments which have been completed examining the impact of testing on learners' subsequent study activities.[68] To date, no experimental studies have been completed examining this question with K-12 learners or in the context of classroom instruction.

### Brief summary of evidence to support the recommendation

Researchers have found that learners typically cannot accurately judge how well they will remember information they have previously studied (e.g., content read from texts).[69] Not surprisingly, being unable to accurately predict what has been learned negatively affects students' ability to implement effective study strategies. As was described above, in laboratory studies, researchers have found that learners tend to allocate study time to items they think they have not learned well and to discontinue study for items they think they have learned. When their judgments of learning are inaccurate, learners cannot make wise choices about how to allocate their study time.

Quizzes may help students identify which items are not well learned, and thus enable more effective allocation of study time. Consistent with this implication, quizzes improve the learning of foreign vocabulary words better than extra study or review time.[70] Similarly, re-reading material produces more gains in learning when a test is interposed between the reading sessions.[71] Well-controlled research on this issue is at present limited to laboratory experiments. As noted, however, some of this work has used educationally relevant materials. We believe there is little reason to doubt that in the classroom setting, quizzes improve learners' awareness of what they know and do not know, thereby helping to increase the effectiveness of study activities.

### How to carry out the recommendation

***1. Teachers should provide "closed-book" quizzes after presentation of material to provide students with the opportunity to check their learning.***

Such quizzes can be either oral or written; in addition, they can be formal or informal (e.g., games in class).

---

[67] Dunlosky, Rawson, and McDonald (2002); Karpicke (2007).
[68] E.g., Thompson, Wegner, and Bartling (1978).
[69] E.g., Dunlosky and Nelson (1994); Koriat (1997); Jang and Nelson (2005); Meeter and Nelson (2003).
[70] Karpicke (2007).
[71] Amlund, Kardash, and Kulhavy (1986).

***2. Teachers can do "spot checks" in which they begin a class session with a very short quiz on the previous night's reading assignment, or on material covered during prior class sessions.***

These questions can be included in the seatwork example described in the 'pre-question' section. Immediately following the quiz, the teacher elicits the correct responses from the class and explicitly tells students that if they did not know the correct answer, they should study specific pages in their text where the answers can be found. Spot checks can also include a question that was previously covered several weeks ago (see Recommendation 1). Even though the spot check is a closed-book quiz, teachers can identify the pages in the text that pertain to the question so that students can easily find the material that needs to be studied.

The basic principle is to use quizzes to help students more accurately identify the material and concepts that they need to study further.

### Potential roadblocks and solutions

***Roadblock 6b.1.*** A teacher might object that there is not enough class time available for quizzing and testing—that he or she will end up having to sacrifice content in order to make time for this.

***Solution.*** Some of this quizzing can be completed as homework. Students can be encouraged to make use of online self-checking quizzes that are frequently available on websites tied to textbooks. These websites grade the quizzes immediately, and often identify pages in the text where the concept, principle, or skill was taught, or where the students can locate a worked example.

***Roadblock 6b.2.*** A teacher might object that there is already an overemphasis on testing in the school system.

***Solution.*** Typically, testing in schools involves using tests to assign grades or using standardized tests to assess students' achievement. That is, tests are used to assess what a student knows. The testing recommended here is to help students better identify what they have *not* learned, so that students can more effectively guide their study for material not yet mastered. It is important for teachers to help students re-envision quizzes as tools to help them learn, rather than as tools used to evaluate their performance.

## Recommendation 7: Help students build explanations by asking and answering deep questions.

When students have acquired a basic set of knowledge about a particular topic of study and are ready to build a more complex understanding of a topic, we recommend that teachers find opportunities to ask questions and model answers to these questions, in order to help students build deep explanations of key concepts. By *deep* explanations we mean explanations that appeal to causal mechanisms, planning, well-reasoned arguments, and logic. Examples of deep explanations include those that inquire about causes and consequences of historical events, motivations of people involved in historical events, scientific evidence for particular theories, and logical justifications for the steps of a mathematical proof. Examples of the types of questions that prompt deep explanations are *why, why-not, how, what-if, how does X compare to Y, and what is the evidence for X*? These questions and explanations can occur both during classroom instruction, class discussion, and during independent study.

### Level of evidence: Strong

The panel judges the level of evidence supporting this recommendation to be *strong* based on over a dozen experimental studies examining the effects of this practice for improving K-12 students' academic performance, over a dozen experimental studies that examined the effect of this strategy for improving college students' academic performance, and the large number of laboratory experiments which have been completed examining the use of deep questions to build explanations and deep understanding.[72]

### Brief summary of evidence to support the recommendation

Many experiments have been conducted in which students are randomly assigned either to conditions that encourage deep explanations or to comparison conditions that expose the students to the similar content, but without the process of building explanations. The students in these studies have typically ranged from fourth grade to college and have been carried out in both laboratory and classroom contexts. The materials have included science, history, different types of informational text, stories, argumentative text, mathematics, and statistics. Student performance has been measured using essays, open-ended questions, multiple-choice questions, and other assessments that tap both shallow and deep knowledge. Shallow knowledge taps basic factual or skill knowledge, whereas deep knowledge is expressed when learners are able to answer "why" questions and describe causal relationships between facts or concepts.

A meta-analysis[73] of dozens of studies support the claim that comprehension and learning improves from interventions that explicitly train students how to ask deep-level questions while reading text, listening to lectures, or studying material.[74] The research has involved classroom discussion, workbooks that provided didactic training with definitions of question types and examples, and pedagogical agents on computers (i.e., talking heads) that modeled question asking and answering. These manipulations have been found to increase the rate of student questions, the depth of their questions, and/or their comprehension of the material.

---

[72] E.g., Beck, McKeown, Hamilton, et al. (1997); Craig, Sullins, Witherspoon, et al. (2006); Driscoll, Craig, Gholson, et al. (2003); Gholson and Craig (2006); King (1992; 1994); Rosenshine, Meister, and Chapman (1996); Wisher and Graesser (2007).

[73] A meta-analysis is a statistical technique used to combine data gathered across multiple research studies and is used to ascertain whether impacts of an intervention are consistent across different studies.

[74] Rosenshine, Meister, and Chapman (1996).

## How to carry out the recommendation

Teachers should identify deep-level questions that they can use to prompt students to reason about underlying explanatory principles relating to the course content. Posing these questions during instruction is usually most useful after students have mastered the more basic factual content relating to a topic. Instructors should keep in mind that students typically need considerably more time to answer deep questions than they would to answer more superficial questions. In addition, instructors may often need to model for students the process of trying to answer deep questions. Research offers a number of pointers for how deep questioning can be used most effectively:

***1. Periodically encourage students to "think aloud" in speaking or writing their explanations as they study the material.[75] After presenting their explanations, it is beneficial for them to get feedback by observing good explanations of peers, tutors, teachers, and computer environments.***

When students think aloud, they often include subjective explanations that go beyond the explicit material and that link the material to personal knowledge and experiences. Better learners produce deep explanations rather than simply repeating the explicit material or expressing loose associations. Explanations can be elicited in a classroom environment when the teacher presents a challenging story and invites a student to think aloud while the student reads the story. Students can respond to other students' explanations and give explanations of their own. The quality of self-explanations improves when students are exposed to high-quality explanations provided by teachers and peers. These examples give students feedback and guidance on appropriate content. Instead of there being only one explanation, some subject matter may allow multiple valid explanations, allowing students to be exposed to different points of view and paths of reasoning. Other subject matter has only one good explanation, which should emerge from discussion. It is important to give a student enough time to think and prepare responses rather than quickly providing the correct answer. After modeling the "think aloud" process during classroom instruction and discussion, students should be encouraged to use this technique as they read class material independently.

***2. Ask questions that elicit explanations, such as those with the following question stems: why, what caused X, how did X occur, what if, what-if-not, how does X compare to Y, what is the evidence for X, and why is X important?[76]***

Questions that stimulate deep explanations are needed when students have trouble expressing explanations on their own. For example, students in a biology class might be asked "How do bees pollinate flowers?" or "Why will the destruction of bees threaten life on planet earth?" The answers and classroom discussion should prompt students to consider underlying biological mechanisms and principles. The process of asking and answering deeper questions often needs to be modeled by teachers, tutors, knowledgeable peers, or computer environments because students may have had relatively little exposure to it. Typically, they will have had far more experience with instructors asking classes many relatively easy and shallow questions that elicit quick short-answers, such as "What organism pollinates flowers?" or "How many types of bees are there?" Again, after learning this question-asking and answering technique by observing the modeling of the process, students can be encouraged to ask deep questions of the material that they are studying outside of class.

***3. Ask questions that challenge students' prior beliefs and assumptions, thereby promoting more intensive and deeper reasoning.[77]***

Problems that challenge students' assumptions or prior beliefs, or which bring to light some puzzling or paradoxical state of affairs, seem to be particularly effective in stimulating students to construct deep questions. An example of a potentially thought provoking question in a middle school science classroom would be "Why is it good for a forest to periodically have forest fires?" Most students naturally assume that because forest fires are dangerous and destructive, they can only be wholly undesirable. Raising the possibility that a forest fire can sometimes produce benefits for a

---

75 Chi, deLeew, Chiu, et al. (1994); McNamara (2004); Pressley and Afflerbach (1995); Trabasso and Magliano (1996).

76 Rosenshine, Meister, and Chapman (1996).

77 Dillon (1988); Festinger (1957); Graesser and McMahen (1993); Graesser and Olde (2003); Otero and Graesser (2001).

forest challenges the student to consider deeper explanatory mechanisms and principles of ecology.

### Potential roadblocks and solutions

***Roadblock 7.1.*** Some students do not have sufficient subject knowledge to construct an explanation, ask a deep question, or answer a question. Consequently, the learning is disappointing and/or the student loses motivation to learn.

***Solution.*** Teachers will need to determine when their students have acquired sufficient subject knowledge in order to benefit from participating in the deep question-asking and answering process. Teachers can use the quizzing techniques described in Recommendations 5 and 6 to help with ascertaining how well foundational subject knowledge has been acquired. Teachers should also provide ample opportunities for students to observe modeling of the question-asking and answering techniques described in this recommendation prior to asking students to use these techniques independently.

***Roadblock 7.2.*** In response to teacher prompts, some students may generate explanations, questions, or answers that are shallow or tangential to the problem at hand.

***Solution.*** When the student generates shallow content, the teacher or learning environment can give feedback or present examples that model the desired explanations, questions, or answers.

***Roadblock 7.3.*** Some students are not motivated to invest the cognitive effort to generate deep explanations, questions, or answers.

***Solution.*** Teachers can present problems that challenge students' beliefs. Another approach is to present problems that are anchored in the real world for which there is some obvious utility in solving the problem. For example, students in a high school chemistry course may be challenged to figure out how to reduce the calcium, chlorine, or pollutants in a water system. They would need to know why such substances are a hazard, how to measure the concentration of the chemicals, and methods for lowering the concentration. The potential hazards and solutions would motivate some students because it solves a problem in the community and/or may challenge the government, a corporation, or some authority.

# Conclusion

This practice guide has attempted to distill some of the more well supported and actionable educational recommendations to emerge from recent (and sometimes not-so-recent) research in the fields of cognitive science and cognitive psychology. These recommendations are meant to shed light on how educators can facilitate not only initial learning and understanding, but—equally importantly—the long-term retention of information and skills taught in schools.

We have sought to make the recommendations here as concrete as possible, and we can recap the essence of our recommendations in very plain and concrete terms: We think it is useful (but not presently very common) to provide reviews of important information after a substantial delay (weeks or even months) from the time when this content was first taught. We think that teachers can increase their effectiveness by viewing testing as not merely something that one does on a few special occasions in order to assess and to motivate, but rather as something done frequently in order to directly promote learning. We have tried to urge instructors to take advantage of the fact that active retrieval of information promotes better and more durable learning than does the more passive sorts of studying that often occupy the majority of a student's study time in and out of the classroom.

A number of our recommendations try to provide some guidance about the most effective way to present subjects that have considerable intellectual depth and complexity, such as science or history. Research shows that asking students to try to answer deep questions about these kinds of material is very helpful in prompting them to identify causal connections and encode underlying principles into their memory—all of which facilitates retention as well as understanding. In many cases, an educator will face the choice between a number of potentially complementary types of instruction to promote understanding of a topic—e.g., practice with concrete procedures for solving some kinds of math problems, and discussion of the abstract significance of these procedures. We have described the considerable body of research that shows that it is often a good idea not to choose one or the other of these complementary forms of instruction, but rather to provide all of them, and to interleave them rather than having students experience uninterrupted blocks of one form or another.

Finally, we have described how research shows that student study practices can be misdirected by the natural human tendency to overestimate our own degree of learning. Again, introducing quizzing and testing throughout the learning process can help students develop a better appreciation of what they have and have not learned, and thus better manage their own study time.

While the authors of this report hope and believe that these research-based recommendations are more concrete than a great deal of advice that has been offered to educators over the years, we are aware there will be many occasions on which it will be far from obvious how to apply them. Or worse, occasions in which they trade off against each other, so that adopting one recommendation requires dismissing another. Like medicine, teaching remains an art even as it seeks to ground its practices more heavily on scientifically collected evidence. Moreover, the evidence that emerges from research is never etched in stone, but continues to evolve and to have its boundaries tested and clarified. Thus, we would not be surprised at all if further research provides important qualifications of the recommendations described here. (This expectation is reflected, of course, in our ratings of evidence strength, which often fell below maximum strength.) As with professionals in other fields, such as medicine, that also seek to rely on a base of evidence and yet must deal with important practical problems on a daily basis, educators must make the best use they can of the current knowledge as it is, even while being mindful of its imperfection.

Despite the inherently tentative nature of our conclusions, it is the hope of the authors of this report that the recommendations provided here will be useful in many classrooms. We also hope that as educators seek to make use of some of these ideas, their experiences will help to bring new attention to important variables that need further investigation, in order that the body of evidence underlying the recommendations described here can continue to grow in depth and comprehensiveness.

# Appendix: Technical information on the studies

## Recommendation 1:

### Space learning over time.

### Level of evidence: Moderate

The panel rated the level of evidence as *moderate*. Hundreds of laboratory experiments have been reported that (a) randomly assigned subjects to conditions in which they studied some material in either a "massed" presentation (two study episodes on each bit of material, separated by a very short time interval, e.g., with two back-to-back study sessions) or in a more temporally distributed fashion (two study episodes on each bit of material, separated by a more considerable intervening time period), while (b) equating the total amount of study time across the different spacing conditions. Almost all of these studies have reported improvements in final memory recall measures for greater inter-study spacing as compared to the more massed presentations.[78] The stimuli and learning tasks used in these studies are quite variable,[79] and include simple laboratory type stimuli (word lists, paired associates), facts,[80] foreign language and English vocabulary,[81] pictures,[82] and prose material.[83] In addition to recall measures, similar benefits have been found for acquisition and retention of a variety of different skills, including surgical skills,[84] complex video games,[85] and typing skills.[86]

There is a significant amount of research involving memory tasks indicating that the optimal amount of spacing tends to increase as the retention interval (time from the second study to the final test) is lengthened, and thus that if a person needs to retain information for a long time, he or she is well advised to increase the degree of spacing. Increases in spacing beyond this optimum have generally been found to produce some decline in final-test memory, but the degree of decline is much smaller than the gain achieved by increasing spacing from the massed condition to the optimum spacing value. This pattern of results has been found over time periods ranging from minutes,[87] to days,[88] to the 6-month / 1 year range,[89] and, in one study, periods of several years.[90] However, the optimum value of interstudy interval appears not to be a fixed constant of the retention interval, but rather a proportion that slowly declines as retention interval is increased. The literature would suggest that optimum performance is achieved by an interstudy interval equal to roughly 100 percent of the retention interval when that retention interval is just a few minutes, declining to 20 percent of the retention interval when the interval is 1 week, and shrinking still further to 5-10 percent of a 1-year retention interval. Based on this pattern, it would appear that whenever it is desired that the learner retain information for many years, it is advisable to utilize spacing of at least several months—and spacing even greater than that would seem more likely to improve retention over the long term than to reduce it.

Spacing effects appear to be large in magnitude. Some meta-analyses have used measures based on within-group variability and pronounced that the typical spacing effect is only moderate in size.[91] However, only a handful of the studies covered in that review involved retention intervals of even 1 week or more, and recent studies show that the increase in recall probability becomes dramatically larger at longer retention intervals. When the retention interval is 6

[78] See Cepeda, Pashler, Vul, et al. (2006) for a recent meta-analysis that examined over 400 studies involving free or cued recall of verbally presented materials.
[79] See Dempster (1996).
[80] E.g., Pashler, Rohrer, Cepeda, et al. (2007).
[81] E.g., Dempster (1987). Bahrick; Bahrick, Bahrick, et al. (1993).
[82] E.g., Paivio (1974).
[83] E.g., Krug, Davis, and Glover (1990).
[84] Moulton, Dubrowski, MacRae, et al. (2006).
[85] Goettl, Yadrick, Connolly-Gomez, et al. (1996).
[86] Baddeley and Longman (1978).
[87] Peterson, Wampler, Kirkpatrick, et al. (1963); Glenberg (1976).
[88] Glenberg and Lehmann (1980).
[89] Pashler, Rohrer, Cepeda, et al. (2007); Pashler, Cepeda, Rohrer, et al. (2004).
[90] Bahrick, Bahrick, Bahrick et al. (1993).
[91] Donovan and Radosevich (1999).

months, spacing has sometimes been found to produce more than a 100 percent increase in recall.[92] Very large effects of spacing were also recently found in retention of a mathematical procedure over a 1-month retention interval[93] and in some studies on very long-term retention of foreign language vocabulary.[94] To sum up: In general, large effects seem to be the rule with studies involving meaningful retention intervals. Importantly, the conditions that produce best performance on the final test are often those that produce worst performance on earlier learning sessions.[95]

***Examples of classroom studies examining the spacing of learning over time.***

These kinds of results have mostly been obtained with young adults in laboratory studies. One limitation of this literature is the paucity of studies conducted in a classroom context using children, but there have been a few such studies. In one classroom study, 44 third-grade students were assigned either to a group that learned multiplication facts and spelling words all at once, or to a group that learned the same facts spaced over time. This experiment found superior memory of multiplication facts and spelling lists for students who were taught with spacing (this finding was made despite the use of an immediate test).[96] In another study, high school students were taught 20 French vocabulary words. Half of the students were taught the words in one 30-minute session; the other students were taught the words in three 10-minute sessions on successive days.[97] There was a 35 percent advantage for spacing over massing, as assessed 4 days later. A longer-term study conducted with 8th graders enrolled in a U.S. History class, examined the effects of delayed review.[98] Students experienced a written review of historical facts contained in the textbook and classroom presentations they had been exposed to. This review occurred either 1 week or 16 weeks after the initial learning. After an additional 9 months, a test was given, requiring them to answer factual questions. The 16-week-delayed review group performed almost twice as well as the 1-week-delayed review group (although, not surprisingly, both groups showed a tremendous amount of forgetting).

# Recommendation 2:

## Interleave worked problem solutions with problem-solving exercises.

### Level of evidence: Moderate

The panel judges the level of evidence supporting this recommendation to be *moderate*. Numerous laboratory experiments provide support for this recommendation.[99] Some classroom experiments provide further evidence that the recommendation can be practically and effectively implemented in real courses at the K-12 and college levels.[100] These experiments have explored these techniques in a variety of content domains particularly in mathematics, science, and technology.

***Example of an experimental study examining the impact of interleaving worked examples and problem-solving practice.***

A large number of laboratory experiments have demonstrated that students learn more by alternating between studying examples of worked-out problem solutions and solving similar problems on their own than they do when just given problems to solve on their own. A classic example of this type of research can be seen in a series of laboratory experiments in

[92] Pashler, Rohrer, Cepeda, et al. (2007); Pashler, Cepeda, Rohrer, et al. (2004).
[93] Rohrer and Taylor (2006).
[94] Bahrick, Bahrick, Bahrick, et al. (1993).
[95] See Schmidt and Bjork (1992); and Bjork and Bjork (2006) for illuminating discussions of this general theme.
[96] Rea and Modigliani (1985).
[97] Bloom and Shuell (1981).
[98] Carpenter, Pashler, Cepeda, et al. (2007).
[99] E.g., Catrambone (1996; 1998); Cooper and Sweller (1987); Kalyuga, Chandler, and Sweller (2001); Kalyuga, Chandler, Tuovinenand, et al. (2001); Mousavi, Low, and Sweller (1995); Paas and van Merriënboer (1994); Renkl (1997; 2002); Renkl, Atkinson, and Große (2004); Renkl, Atkinson, Maier, et al. (2002); Renkl, Stark, Gruber, et al. (1998); Schwonke, Wittmer, Aleven, et al. (2007); Schworm and Renkl (2002); Sweller (1999); Sweller and Cooper (1985); Trafton and Reiser (1993); Ward and Sweller (1990).
[100] E.g., McLaren, Lim, Gagnon, et al. (2006); Zhu and Simon (1987).

the domain of algebra.[101] In this series of experiments, eighth- and ninth-grade students in the treatment condition were asked to alternate between four pairs of solution examples and problems. Students in the control condition were simply asked to solve eight problems, as one might typically ask students in a homework assignment. Note that in both cases, students were first given general instruction on the relevant algebra principles and initial examples of each. Students in the interleaved example/problem treatment condition not only took less time during instruction, but also performed better on the post-test following instruction. For instance, in one study the students in the interleaved example/problem condition required less than half of the study time and made half as many errors on the post-test as the conventional problem-solving condition.

## Recommendation 3:

### Combine graphics with verbal descriptions.

#### Level of evidence: Moderate

The panel judges the level of evidence supporting this recommendation to be *moderate*. Many laboratory experiments have provided data consistent with this recommendation.[102] Some classroom experiments and quasi-experiments provide further evidence that the recommendation can be practically and effectively implemented in real courses at the K-12 and college levels.[103]

***Example of experimental studies examining how to integrate text and visualization in support of learning.***

Many studies have demonstrated how adding relevant visual representations to text descriptions leads to better learning than text alone.[104] Most of these studies have focused on scientific processes—for example, how things work, like lightning, disk breaks, bike pumps, and volcanic eruptions. A series of studies carried out by Richard Mayer and his colleagues examining these questions informed our recommendation to integrate visual representations and verbal descriptions.[105]

In these studies, college students were randomly assigned to conditions where they watched an animation of a bicycle tire pump and were asked to subsequently answer a series of problem-solving questions about ways to improve the pump's effectiveness. Students in one condition watched the visual presentation of the operation of the bicycle tire pump with simultaneous verbal narration of the actions; students in the other condition heard the verbal narration prior to viewing the visualization. In all three of the experiments, students who heard the verbal narration at the same time as they watched the visualization of the bicycle tire pump outperformed students on the problem-solving task completed after they watched the visualization.

## Recommendation 4:

### Interleave abstract and concrete material.

#### Level of evidence: Moderate

The panel judges the level of evidence supporting this recommendation to be *moderate*. Both laboratory experiments[106] and classroom quasi-experiments[107] provide evidence to support this recommendation.

***Example of a laboratory study examining effects of concrete and abstract representations on learning.***

As discussed in the main recommendation, there is an emerging body of experimental research that has examined the conditions under which concrete and abstract representations of concepts support either initial learning of a concept and its subsequent use

---

[101] Sweller and Cooper (1985).
[102] E.g., Clark and Mayer (2003); Griffin, Case, and Siegler (1994); Kalchman and Koedinger (2005); Kalchman, Moss, and Case (2001); Mayer (2001); Mayer and Anderson (1991; 1992); Mayer and Moreno (1998); Moreno and Mayer (1999a; 1999b); Moss (2005).
[103] E.g., Griffin, Case, and Siegler (1994); Kalchman, Moss, and Case (2001); Kalchman and Koedinger (2005); Moss (2005).
[104] See Mayer (2001) and Sweller (1999) for reviews.
[105] Mayer and Anderson (1991).
[106] Goldstone and Sakamoto (2003); Goldstone and Son (2005); Kaminiski, Sloutsky, and Heckler (2006a; 2006b); Sloutsky, Kaminski, and Heckler (2005).
[107] E.g., Ainsworth, Bibby, and Wood (2002); Bottge (1999); Bottge, Heinrichs, Chan, and Serlin (2001); Bottge, Heinrichs, Mehta, and Hung (2002); Bottge, Rueda, and Skivington (2006); Bottge, Rueda, LaRoque, et al. (2007); Bottge, Rueda, Serlin, et al. (2007).

to solve other conceptually related problems. The majority of this research has been completed in the context of mathematics learning. In one recent study, 19 sixth-grade students were taught a novel mathematical concept using either a generic or concrete representation of the problems.[108] Students were randomly assigned to one of two conditions. Students in the generic situation were told that the concept was a symbolic language and that there are specific rules that constrain the combination of the three symbols. Students who learned the concept using a concrete representation were taught using measuring cups that could be filled with different levels of liquid. The rules for combining the three possible levels of the measuring cup were the same as the rules for combining the symbols. After being trained and tested in the same domain in which they were trained, students were then asked to solve a parallel set of problems in a third domain. Students were not explicitly taught the rules of the new situation.

Two interesting findings emerged from this study. First, students who learned using the relevant concrete situation showed marginally better initial learning of the concept than their colleagues who learned the generic symbolic domain. However, students taught with the more abstract symbolic system also learned the concept. Second, when students were tested in the third domain and asked to transfer the skills they learned either in the relevant concrete situation or the generic symbolic situation, students who learned in the generic situation were substantially more likely to solve the transfer problems in the third domain successfully. Indeed, students who learned in the relevant concreteness condition performed at chance levels on the transfer task.

This study is complemented by other experimental studies with adults that demonstrate how learning a concept in a concrete context can hinder transfer to novel situations.[109] The take-home message from this line of work is that while learning in a concrete context may support initial learning, the concrete context alone does not support learning. Thus, classroom instruction designed to promote the use of knowledge across different contexts should include instruction in the abstract or generic representations of the concept being taught, and teachers should not expect students to be able to infer the underlying symbolic or abstract representation of a problem from learning how to solve a problem using a single concrete instantiation of the problem.

***Example of classroom studies examining effects of using anchored instruction to support learning of abstract concepts.***

Some recent classroom-based quasi-experimental research has been examining whether connecting abstract to concrete representations and authentic situations supports learning in low-achieving students. Instruction involves the use of video-based problems that directly immerse students in the problems, in contrast to traditional problem formats (e.g., word problems).[110] Each video anchor presents a realistic scenario consisting of several subproblems and it typically takes students 1 to 2 weeks to solve the entire problem. As in authentic tasks (e.g. "real-life" tasks that students might need to use mathematical skills to solve), students must first understand the problems, "unpack" the relevant pieces of information for solving them, and then "repack" them into a solution that makes sense. It is this interleaving between the understanding of the concrete nature of the problem, identifying the underlying and abstract principles relevant to solving the concrete problem, and then integrating those abstract principles into the solution of the concrete problem that is the focus of our recommendation.

Current research extends this type of anchoring by affording students additional opportunities to practice their skills and deepen their understanding. Students using Enhanced Anchored Instruction (EAI) are expected to solve new but analogous problems in applied contexts (e.g., designing, building, and riding hovercrafts). These projects help students create more vivid mental models of the problem situations presented in the video-based anchors. Adding the interleaved enhancements provides students several contexts in which to apply their concepts and skills, acknowledging that highly contextualized learning may

[108] Kaminski, Sloutsky, and Heckler (2006).

[109] Goldstone and Sakamoto (2003); Kaminiski, Sloutsky, and Heckler (2005).

[110] Cawley, Parmar, Foley, et al. (2001).

actually hamper learning transfer. Quasi-experimental tests comparing student outcomes after being taught the same course concepts through the use of EAI versus traditional instruction finds that Enhanced Anchored Instruction engages hard-to-teach adolescents in middle schools, high schools, and alternative settings and improves their problem solving skills.[111]

# Recommendation 5:

### Use quizzing to promote learning.

## Recommendation 5a:

### Use pre-questions to introduce new topics.

### Level of evidence: Low

The panel rated the level of evidence supporting this recommendation as *low*. We were able to locate two high-quality studies completed in the laboratory examining the use of pre-questions prior to reading text.[112] In addition, research completed on the use of advanced organizers during reading lends additional support to this recommendation.[113]

***Example of a study on pre-questions.***

For possible application to classroom practice, two especially important findings have been reported in a well-controlled experiment conducted in the laboratory (random assignment of participants to conditions and demonstrating equivalence of groups at pre-test were both used).[114] First, only learners who were required to answer pre-questions prior to reading a text showed gains in acquisition of content; learners who read but did not answer the pre-questions did not show significant gains relative to learners not given pre-questions. Second, the attempt to answer pre-questions was beneficial regardless of whether the learners provided a correct answer. That is, answering pre-questions incorrectly did not eliminate the pre-question advantage—the advantage simply required an attempt to answer the pre-questions.

## Recommendation 5b:

### Use quizzing to promote learning.

### Level of evidence: Strong

The panel judged the level of evidence supporting this recommendation to be *strong* based on nine experimental studies examining the effects of this practice for improving K-12 students' performance on academic content or classroom performance, over 30 experimental studies that examined the effect of this strategy for improving college students' academic performance, and the large number of carefully controlled laboratory experiments that have examined the testing effect.[115]

***Two experimental studies using quizzes with classroom materials to reduce forgetting.***

In one study,[116] a laboratory experiment using college level art-history lectures found that multiple-choice and short-answer quizzes administered immediately after viewing the lectures substantially improved performance on a test 30 days later relative to when no quizzes were given. In addition, short-answer but not multiple-choice quizzes improved performance on the test 30 days later relative to a condition in which the target facts were given for additional study immediately after viewing the lecture.

In a well-controlled experiment conducted in a college class on brain and behavior[117] content that was quizzed with multiple-choice quizzes and short-answer quizzes—both with corrective feedback for students' answers—was remembered significantly better on exams than non-quizzed content. Also, short-answer quizzing produced significantly better performance on exams than did providing the target facts for extra study.

---

[111] E.g., Bottge (1999); Bottge, Heinrichs, Chan, et al. (2001); Bottge, Heinrichs, Mehta, et al. (2002); Bottge, Rueda, and Skivington (2006); Bottge, Rueda, LaRoque, et al. (2007); Bottge, Rueda, Serlin, et al. (2007).

[112] Pressley, Tannebaum, McDaniel, et al. (1990); Rickards (1976).

[113] E.g., Rickards (1975-1976).

[114] Pressley, Tannebaum, McDaniel, et al. (1990).

[115] See Roediger and Karpicke (2006a) for a recent review and synthesis of both laboratory and classroom research that empirically examines the testing effect (see also Dempster and Perkins, 1993). See McDaniel, Roediger, and McDermott (2007) for a discussion of how the laboratory research generalizes to classroom use of the testing effect.

[116] Butler and Roediger (2007).

[117] McDaniel, Anderson, Derbish, et al. (2007).

# Recommendation 6:

**Help students allocate study time efficiently.**

## Recommendation 6a:

**Teach students how to use delayed judgment of learning techniques to identify concepts that need further study.**

### Level of evidence: Low

The panel judged the level of evidence supporting this recommendation to be *low* because the body of evidence supporting this recommendation is primarily experimental research completed in the laboratory using academic content. The research provides direct evidence supporting causal links between delayed judgments of learning and accurate assessments of knowledge, delayed keyword generation and accurate assessments of knowledge, and links between accurate assessments of knowledge, study behavior, and improved performance on tests.[118] Research has been completed both with college students and school-aged children.

***Example of an experiment using delayed judgments of learning to improve study.***[119]

Thirty third- and fifth-grade students attending a public elementary school in New York City recently participated in an experiment that examined how well the strategy of using a delayed judgment of learning task, to guide study and restudy, worked to improve learning. The students studied 54 definition-word pairs drawn from school textbooks and online vocabulary resources over a 4-week time span. The definition-word pair items were studied in clusters of six, and a total of 18 pairs were studied each week. After studying each set of six items, the students were asked to make judgments of how well they had learned each of the definition-word pairs. Then, the students saw all six words in a circular arrangement and were asked to choose three of those items to restudy. This process was repeated until all 18 items had been studied.

The critical experimental manipulation occurred at this point in the study. Students were either asked to restudy the words they had identified as most in need of restudy (honor choice), to restudy the words that had identified as NOT in need of restudy (dishonor choice), or to restudy the words that they had rated with the highest judgments of learning (e.g., the ones they thought they knew the best). During the fourth week, the students were tested on all definition-word pair items.

The researchers found that when the fifth graders' choices were honored—and they were allowed to restudy the items that they had identified with a low judgment of learning and needed to restudy—their final test performance was substantially improved. Dishonoring their choices or asking them to restudy words with their highest judgments of learning did not lead to improved test performance. However, honoring the third graders' choices did not lead to improved test performance because the third graders did not choose to restudy items to which they had given low judgments of learning. Their identification of which words needed to be restudied appeared to be random.

This study illustrates that fifth-grade children can use the delayed judgment of learning task to accurately identify items they need to spend additional time learning, and that spending time restudying those items leads to improved final test performance. On the other hand, while third graders were found to be able to accurately identify items they did not know well, they did not use that knowledge to choose items to restudy. Together, these findings suggest that the delayed judgment of learning task has promise as a tool that students can be taught to use to guide their study.

***Example of an experiment using delayed keywords to improve learning from reading.***

In a recent series of experiments completed by Thiede and his colleagues,[120] college students were asked to read seven expository texts adopted from encyclopedia articles on different topics and to generate keywords. All participants were asked to rate how well they understood the text, and then to answer test questions about what they read. In the experiments, the students were presented with the title of the article and asked to write down five keywords that captured the essence of the text identified in the title. Across the different

[118] Dunlosky, Hertzog, Kennedy, et al. (2005); Hertzog, Kidder, Moman-Powell, et al. (2002); Thiede, Anderson, and Therriault (2003); Thiede, Dunlosky, Griffin, et al. (2005).

[119] Metcalfe and Finn (in press).

[120] Thiede, Dunlosky, Griffin, et al. (2005).

experiments, the delay between reading the text, generating the keywords, and making judgments of learning varied systematically across participants, as did the order in which students were asked to complete these tasks. The results of these experiments indicate that both generating one's own keywords and including a delay between reading a text and generating the keywords are the critical components to include when using this technique to improve students' ability to identify how well they have understood a text.

### Recommendation 6b:
### Use quizzes to help students identify content requiring further study.

#### Level of evidence: Low

The panel judged the level of evidence supporting this recommendation to be *low* based on three experimental studies that examined the effect of this strategy for improving college students' performance on academic content (text material, vocabulary),[121] and a handful of laboratory experiments which have been completed examining the impact of testing on learners' subsequent study activities.[122] To date, no experimental studies have been completed examining this question with K-12 learners or in the context of classroom instruction.

***Example of an experiment on using quizzes (tests) to help guide further study.***

In a laboratory experiment,[123] college students were given a list of unknown foreign vocabulary to learn (presented as foreign vocabulary—English translation pairs). One group studied the list five times before being given the final test (recall English translation given the vocabulary item), and another group was given two quizzes (tests) interleaved between three study trials before the final test. In both groups, learners allocated more study times for individual items that the learners judged were not well learned. However, judgments of learning made during each study trial became substantially more accurate across study trials in the quiz group, not the study-only group. Thus, learners in the quiz group were allocating their study time more effectively and more in line with their intentions than learners in the study-only group. This pattern was reflected in significantly higher performance on the final test for the quiz group than the study group (which may have also reflected the direct benefits of quizzing noted in the previous section).

## Recommendation 7:
### Help students ask deep question in order to build explanations.

#### Level of evidence: Strong

The panel judged the level of evidence supporting this recommendation to be *strong* based on over a dozen experimental studies examining the effects of this practice for improving K-12 students' academic performance, over a dozen experimental studies that examined the effect of this strategy for improving college students' academic performance, and the large number of laboratory experiments which have been completed examining the use of deep questions to build explanations and deep understanding.[124]

Dozens of studies in the cognitive and learning sciences have conducted experiments that manipulate the process of students constructing explanations by themselves or by interacting with peers, tutors, teachers, or computers. In these experiments, students are randomly assigned either to conditions that encourage deep explanations or to comparison conditions that expose the students to the similar content, but without the process of building explanations. The students in these studies have typically ranged from fourth grade to college in both laboratory and classroom contexts. Learning gains have been documented in these studies that manipulate the construction of explanations by the students themselves,[125] or by construction of

---

[121] Dunlosky, Rawson, and McDonald (2002); Karpicke (2007).
[122] E.g., Thompson, Wegner, and Bartling (1978).
[123] Karpicke (2007).
[124] Beck, McKeown, Hamilton, et al. (1997); Craig, Sullins, Witherspoon, et al. (2006); Driscoll, Craig, Gholson, et al. (2003); Gholson and Craig (2006); Rosenshine, Meister, and Chapman (1996); Wisher and Graesser (2007). [1]
[25] Laboratory: Chi, de Leeuw, Chiu, et al. (1994); Pressley and Afflerbach (1995); Classroom: King (1992, 1994); McNamara (2004); Pressley and Afflerbach (1995); Pressley, Wood, Woloshyn, et al. (1992).

explanations with human or computer tutors.[126] The materials have included science, history, informational text, stories, argumentative text, mathematics, and statistics. The dependent measures have included essays, open-ended questions, multiple-choice questions, and other assessments that tap both shallow and deep knowledge. In correlational studies that assess individual differences, students who construct more explanations also tend to have better memory, problem solving, and reasoning for the material.[127]

***Summaries of the findings from two recent literature reviews examining specific techniques to support asking deep questions and constructing deep explanations.***

A recent literature review reported seven experiments on learning Newtonian physics that compared interactive construction of explanations with a human or computer tutor versus reading explanations or merely reading subject matter in textbooks.[128] The tutoring sessions involved students answering challenging physics questions that are known to elicit persistent misconceptions, such as: "When a car without headrests on the seat is struck from behind, the passengers often experience neck injuries. Why do passengers experience neck injuries in this situation?" The dependent measures were well-constructed multiple-choice tests and essays that answer difficult conceptual physics problems. The reading of subject matter from textbooks had substantially lower learning gains for college students than the reading of explanations and the interactive construction of explanations. The interactive construction of explanations showed advantages over merely reading explanations, but primarily for those students in which the problems were at the zone of proximal development (i.e., not too easy or too hard).

Dozens of studies support the claim that comprehension and learning improves from interventions that improve question-asking and answering skills. However, it is an uphill battle to get students to generate deep questions because this is not a natural proclivity for most students.[129] A typical student in a classroom asks only one question per 7 hours and most of the questions are shallow (e.g., *who, what, when, where.*[130]). One approach to training students to ask good questions is through modeling. An alternative approach to stimulate inquiry is to give them a problem that challenges their beliefs, that stimulates thought at their zone of proximal development, or that creates cognitive dissonance through obstacles, impasses, contradictions, anomalies, or uncertainty.[131] Indeed, explicit training on students' asking deep-level questions has been shown to improve comprehension and learning from texts and classroom lectures in student populations of fourth grade through college.[132] Rosenshine, Meister, and Chapman, (1996) provided the most comprehensive analysis of the impact of question-generation training (QGT) on learning in their meta-analysis of 26 empirical studies that compared QGT to learning conditions with appropriate controls. The outcome measures included standardized tests, short-answer or multiple-choice questions prepared by experimenters, and summaries of the texts. The review revealed that effects were greatest for experimenter-generated tests and summary tests, and substantial, although smaller, for standardized tests. One informative result of this meta-analysis was that the question format was important when training the learners how to ask questions. The analysis compared training with signal words (*who, what, when, where, why, and how*), training with generic question stems (*How is X like Y?, Why is X important?*, *What conclusions can you draw about X?*), and training with main idea prompts (*What is the main idea of the paragraph*). The generic-question stems were most effective, perhaps because they give the learner more direction, are more concrete, and are easier to teach and apply.

---

[126] Laboratory: Chi, Siler, Jeong, et al. (2001); Cohen, Kulik, and Kulik (1982); Graesser, Lu, Jackson, et al. (2004); VanLehn, Graesser, Jackson, et al. (2007); Classroom: Aleven and Koedinger (2002); Cohen, Kulik, and Kulik (1982); Graesser, Lu, Jackson, et al. (2004); Hunt and Minstrell (1996); McNamara, O'Reilly, Best, et al. (2006); VanLehn, Graesser, Jackson et al. (2007).

[127] Chi (2000); Chi, Bassok, Lewis, et al. (1989); Graesser and Person (1994); Trabasso and Magliano (1996).

[128] VanLehn, Graesser, Jackson, et al. (2007).

[129] Wisher and Graesser (2007).

[130] Dillon (1988); Graesser and Person (1994).

[131] Dillon (1988); Festinger (1957); Graesser and McMahen (1993); Graesser and Olde (2003); Otero and Graesser (2001).

[132] King (1992, 1994); Palincsar and Brown (1984); Rosenshine, Meister, and Chapman (1996).

# References


Ainsworth, S., Bibby, P., and Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *The Journal of the Learning Sciences*, *11*, 25–61.

Aleven, V., and Koedinger, K.R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26*, 147-179.

Amaya, M.M., Uttal, D.H., and DeLoache, J.S. (2007). *Procedural knowledge in two-digit subtraction: Comparing concrete and abstract.* Manuscript submitted for publication.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: AERA Publications.

American Psychological Association. (2002). Criteria for practice guideline development and evaluation. *American Psychologist, 57*, 1048-1051.

Amlund, J.T., Kardash, C.A.M., and Kulhavy, R.W. (1986). Repetitive reading and recall of expository text. *Reading Research Quarterly, 21,* 49-58.

Ausubel, D.P., and Youssef, M. (1965). The effect of spaced repetition on meaningful retention. *The Journal of General Psychology*, *73*, 147-150.

Baddeley, A.D., and Longman, D.J.A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, *21*, 627-635.

Bahrick, H.P., Bahrick, L.E., Bahrick, A.S., and Bahrick, P.E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, *4*, 316-321.

Beck, I.L., McKeown, M.G., Hamilton, R.L., and Kucan, L. (1997). *Questioning the Author: An approach for enhancing student engagement with text*. Delaware: International Reading Association.

Berger, S.A., Hall, L.K., and Bahrick, H.P. (1999). Stabilizing access to marginal and submarginal knowledge. *Journal of Experimental Psychology: Applied, 5*, 438-447.

Bjork, R.A. (1988). Retrieval practice and the maintenance of knowledge. In M.M. Gruneberg, P.E. Morris, and R.N. Sykes (Eds.), *Practical aspects of memory II* (pp. 396-401). New York: Wiley.

Bjork, R.A., and Bjork, E.L. (2006). Optimizing treatment and instruction: Implications of a new theory of disuse. In L.-G. Nilsson and N. Ohta (Eds.), *Memory and society: Psychological perspectives* (pp. 116-140). New York: Psychology Press.

Bloom, B.S. (1956). *Taxonomy of educational objectives: Handbook I. The cognitive domain.* New York: David McKay Publications.

Bloom, K.C., and Shuell, T.J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*, *74*, 245-248.

Bottge, B.A. (1999). Effects of contextualized math instruction on problem solving of average and below-average achieving students. *Journal of Special Education, 33,* 81-92.

Bottge, B.A., Heinrichs, M., Chan, S., and Serlin, R. (2001). Anchoring adoelscents' understanding of math concepts in rich problem solving environments. *Remedial and Special Education, 22,* 299-314.

Bottge, B.A., Heinrichs, M., Mehta, Z.D., and Hung, Y.-H. (2002). Weighting the benefits of anchored math instruction for students with disabilities in general education classes. *Journal of Special Education, 35,* 186-200.

Bottge, B.A., Rueda, E., LaRoque, P.T., Serlin, R.C., and Kwon, J. (2007). Integrating reform-oriented math instruction in special education settings. *Learning Disabilities Research & Practice, 22*, 96-109.

Bottge, B.A., Rueda, E., Serlin, R., Hung, Y.-H., and Kwon, J. (2007). Shrinking achievement differences with anchored math problems: Challenges and possibilities. *Journal of Special Education, 41*, 31-49.

Bottge, B.A., Rueda, E., and Skivington. M. (2006). Situating math instruction in rich problem-solving contexts: Effects on adolescents with challenging behaviors. *Behavioral Disorders, 31,* 394-407.

Butler, A.C., and Roediger, H.L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19,* 514-527.

Butterfield, B., and Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1491-1494.

Carpenter, S.K., Pashler, H., Cepeda, N.J., and Alvarez, D. (2007). Applying the principles of testing and spacing to classroom learning. In D.S. McNamara and J.G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society*, (p. 19). Nashville, TN: Cognitive Science Society.

Carpenter, S.K., Pashler, H., Wixted, J.T., and Vul, E. (in press). The effects of tests on learning and forgetting. *Memory & Cognition.*

Carrier, M., and Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20,* 632-642.

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1020-1031.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127*, 355-376.

Cawley, J., Parmar, R., Foley, T.E., Salmon, S., and Roy, S. (2001). Arithmetic performance of students: Implications for standards and programming. *Exceptional Children, 67*, 311-328.

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

Chi, M.T.H. (2000). Self-explaining: The dual processes of generating and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology* (pp. 161-238). Mahwah, N.J.: Erlbaum.

Chi, M.T.H., Bassok, M., Lewis, M., Reimann, P., and Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chi, M.T.H., de Leeuw, N., Chiu, M., and LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

Chi, M.T.H., Siler, S., Jeong, H., Yamauchi, T., and Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science, 25*, 471-533.

Clark, R.C., and Mayer, R.E. (2003). *e-Learning and the Science of Instruction : Proven Guidelines for Consumers and Designers of Multimedia Learning*. San Francisco: Jossey-Bass.

Cohen, P.A., Kulik, J.A., and Kulik, C.C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*, 237-248.

Cooper, G., and Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79,* 347–362.

Craig, S.D., Sullins, J., Witherspoon, A., and Gholson, B. (2006). The deep-level-reasoning-question effect: The role of dialogue and deep-level-reasoning questions during vicarious learning. *Cognition and Instruction, 24*, 565-591.

Dempster, F.N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology, 79*, 162-170.

Dempster, F.N. (1996). Distributing and managing the conditions of encoding and practice. In E.L. Bjork and R.A. Bjork (Eds.), *Handbook of perception and cognition* (*Vol. 10: Memory*, pp. 317-344). San Diego, CA: Academic Press.

Dempster, F.N. and Perkins, P.G. (1993). Revitalizing classroom assessment: Using tests to promote learning. *Journal of Instructional Psychology*, *20*, 197–203.

Dillon, T.J. (1988). *Questioning and teaching: A manual of practice.* New York: Teachers College Press.

Donovan, J.J., and Radosevich, D.J. (1999). A meta-analytic review of the distribution of practice effect. *Journal of Applied Psychology*, *84*, 795-805.

Driscoll, D., Craig, S.D., Gholson, B., Ventura, M., and Graesser, A. (2003). Vicarious learning: Effects of overhearing dialog and monologue-like discourse in a virtual tutoring session. *Journal of Educational Computing Research, 29*, 431-450.

Dufresne, A., and Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology, 47,* 274-296.

Dunlosky, J., Hertzog, C., Kennedy, M., and Thiede, K. (2005). The self-monitoring approach for effective learning. *Cognitive Technology, 10,* 4-11.

Dunlosky, J., and Nelson, T.O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20,* 374-380.

Dunlosky, J., and Nelson, T.O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33,* 545-565.

Dunlosky, J., Rawson, K.A., and McDonald, S.L. (2002). Influence of practice tests on the accuracy of predicting memory performance for paired associates, sentences, and text material. In T.J. Perfect and B.L. Schwartz (Eds.), *Applied Metacognition* (pp. 68-92). Cambridge, UK: Cambridge University Press.

Dunlosky, J., Rawson, K.A., and Middleton, E.L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language. Special Issue: Metamemory, 52*, 551-565.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Field, M.J., and Lohr, K.N. (Eds.). (1990). *Clinical Practice Guidelines: Directions for a New Program, at 8*. Institute of Medicine.

Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, *6*(40), 104.

Gholson, B., and Craig, S.D. (2006). Promoting constructive activities that support vicarious learning during computer-based instruction. *Educational Psychology Review, 18*, 119-139.

Glenberg, A.M., and Lehmann, T.S. (1980). Spacing repetitions over 1 week. *Memory & Cognition*, *8*, 528-538.

Goettl, B.P., Yadrick, R.M., Connolly-Gomez, C., Regian, W., and Shebilske, W.L. (1996). Alternating task modules in isochronal distributed training of complex tasks. *Human Factors*, 38, 330-346.

Goldstone, R.L., and Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology, 46*, 414-466.

Goldstone, R.L, and Son, J.Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*, 69-110.

Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H., Ventura, M., Olney, A., and Louwerse, M.M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers, 36*, 180-193.

Graesser, A.C., and McMahen, C.L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology*, *85*, 136-151.

Graesser, A.C., and Olde, B.A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology, 95*, 524-536.

Graesser, A.C., and Person, N.K. (1994). Question asking during tutoring. *American Educational Research Journal, 31*, 104-137.

Griffin, S., Case, R., and Siegler, R.S. (1994). Rightstart: Providing the central conceptual prerequisities for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and the classroom* (pp. 25-50). Cambridge, MA: MIT Press.

Hausmann, R.G.M., and VanLehn, K. (in press). Explaining self-explaining: A contrast between content and generation. *13th International Conference on Artificial Intelligence in Education*, Marina del Rey, CA.

Hertzog, C., Kidder, D., Moman-Powell, A., and Dunlosky, J. (2002). Monitoring associative learning: What determines the accuracy of metacognitive judgments. *Psychology and Aging, 17*, 209-225.

Hirsch, E.D., Jr. (1987). *Cultural literacy: what every American needs to know.* Boston: Houghton Mifflin.

Hunt, E., and Minstrell, J. (1996). A collaborative classroom for teaching conceptual physics. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and the classroom* (pp. 51-74). Cambridge, MA: MIT Press.

Jang, Y., and Nelson, T.O. (2005). How many dimensions underlie judgments of learning and recall? Evidence from state-trace methodology. *Journal of Experimental Psychology: General, 134,* 308-326.

Kalchman, M., and Koedinger, K.R. (2005). Teaching and learning functions. In Donovan, S. and Bransford, J. (Eds.) *How students learn: History, mathematics and science in the classroom* (pp. 351-396). Washington, D.C.: National Academy Press.

Kalchman, M., Moss, J., and Case, R. (2001). Psychological models for development of mathematical understanding: Rational numbers and functions. In S. Carver and D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 1-38). Mahwah, NJ: Erlbaum.

Kalyuga, S., Chandler, P., and Sweller, J. (2001). Learner experience and efficiency of instructional guidance. *Educational Psychology, 21*, 5–23.

Kalyuga, S., Chandler, P., Tuovinen, J., and Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*, 579–588.

Kaminiski, J.A., Sloutsky, V.M., and Heckler, A.F. (2006a). Do children need concrete instantiations to learn an abstract concept? In R. Sun and N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 411-416). Mahwah, NJ: Erlbaum.

Kaminiski, J.A., Sloutsky, V.M., and Heckler, A.F. (2006b). Effects of concreteness on representation: An explanation for differential transfer. In R. Sun and N. Miyake (Eds.), *Proceedings of the XXVIII Annual Conference of the Cognitive Science Society* (pp. 1581-1586). Mahwah, NJ: Erlbaum.

Karpicke, J.D. (2007). *Students' use of self-testing as a strategy to enhance learning*. Unpublished doctoral dissertation, Washington University, St. Louis, MO.

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Journal, 29,* 303-323.

King A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal*, *31*, 338-368.

Kirschner, P.A., Sweller, J., and 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*, 75–86.

Koriat, A. (1997). Monitoring one's knowledge during study: A cue-utilization framework to judgments of learning. *Journal of Experimental Psychology: General, 126,* 349-370.

Krug, D., Davis, T.B., and Glover, J. (1990). Massed versus distributed repeated reading: A case of forgetting helping recall? *Journal of Educational Psychology, 82*, 366-371.

Lockl, K., and Schneider, W. (2002). Developmental trends in children's feeling-of-knowing judgements. *International Journal of Behavioral Development, 26,* 327-333.

*Mace, C.A. (1932). The psychology of study. London: Methuen.*

Masur, E.F., McIntyre, C.W., and Flavell, J.H. (1973). Developmental changes in apportionment of study time among items in a multitrial free recall task. *Journal of Experimental Child Psychology, 15,* 237-246.

Mayer, R.E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mayer, R.E., and Anderson, R. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83,* 484–490.

Mayer, R.E., and Anderson, R. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84,* 444–452.

Mayer, R.E., Hegarty, M., Mayer, S., and Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruct. *Journal of Experimental Psychology: Applied,* 11, 256-265.

Mayer, R.E., and Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual-processing systems in working memory. *Journal of Educational Psychology, 90,* 312–320.

McDaniel, M.A., Anderson, J.L., Derbish, M.H., and Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19,* 494-513.

McDaniel, M.A., and Fisher, R.P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology, 16,* 192-201.

McDaniel, M.A., Roediger, H.L., and McDermott, K.B. (2007). Generalizing test enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review, 14,* 200-206.

McLaren, B.M., Lim, S., Gagnon, F., Yaron, D., and Koedinger, K.R. (2006). Studying the effects of personalized language and worked examples in the context of a web-based intelligent tutor. In *The Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (pp. 318-328). New York: Springer.

McNamara, D.S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*, 1-30.

McNamara, D.S., O'Reilly, T., Best, R., and Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research, 34*, 147-171.

Meeter, M., and Nelson, T.O. (2003). Multiple study trials and judgments of learning. *Acta Psychologica, 113,* 123-132.

Metcalfe, J., and Dunlosky, J. (in press). Metamemory. In J.H. Byrne (Ed.), *Learning and Memory: A Comprehensive Reference*. Oxford: Elsevier.

Metcalfe, J., and Finn, B. (in press). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review.*

Metcalfe, J., and Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language, 52*, 463-477.

Moreno, R., and Mayer, R.C. (1999a). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*, 358-368.

Moreno, R., and Mayer, R.C. (1999b). Multimedia-supported metaphors for meaning making in mathematics. *Cognition and Instruction, 17*, 215-248.

Moss, J. (2005). Pipes, tubs, and beakers: New approaches to teaching the rational-number system. In M.S. Donovan and J. D.Bransford (Eds.), *How students learn: History, math, and science in the classroom* (pp. 309-349). Washington DC: National Academies Press.

Moulton, C., Dubrowski, A., MacRae, H., Graham, B., Grober, E., and Reznick, R. (2006). Teaching surgical skills: What kind of practice makes perfect? *Annals of Surgery*, *244*, 400-409.

Mousavi, S.Y., Low, R., and Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87,* 319-334.

Otero, J., and Graesser, A.C. (2001). PREG: Elements of a model of question asking. *Cognition and Instruction, 19*, 143-175.

Paas, F., and van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86,* 122–133.

Paivio, A. (1974). Spacing of repetitions in the incidental and intentional free recall of pictures and words. *Journal of Verbal Learning and Verbal Behavior*, *13*, 497-511.

Palincsar, A.S., and Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*, 117-175.

Pane, J.F., Corbett, A.T., and John, B.E. (1996). Assessing dynamics in computer-based instruction. In M. Tauber (Ed.), *Proceedings of ACM CHI'96 Conference on Human Factors in Computing Systems* (pp. 197-204). Addison-Wesley Publishers.

Pashler, H., Cepeda, N., Rohrer, D., and Wixted, J.T. (2004). *The spacing effect: Useful or just interesting?* Paper presented at the 45th Annual Meeting of the Psychonomic Society, Minneapolis, MN.

Pashler, H., Rohrer, D., Cepeda, N.J., and Carpenter, S.K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 19,* 187-193.

Pashler, H., Zarow, G., and Triplett, B. (2003). Is temporal spacing of tests helpful even when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29,* 1051-1057.

Peterson, L.R., Wampler, R., Kirkpatrick, M., and Saltzman, D. (1963). Effect of spacing presentations on retention of a paired associate over short intervals. *Journal of Experimental Psychology*, *66*, 206-209.

Pressley, M., and Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.

Pressley, M., Tannebaum, R., McDaniel, M. A., and Wood, E. (1990). What happens when university students try to answer prequestions that accompany textbook materials? *Contemporary Educational Psychology*, *15*, 27-35.

Pressley, M., Wood, E., Woloshyn, V.E., Martin, V., King, A., and Menk, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27,* 91-110.

Rea, C.P., and Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning*, *4*, 11-18.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21,* 1–29.

Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction, 12*, 529-556.

Renkl, A., Atkinson, R.K., and Große, C.S. (2004) How fading worked solution steps works – a cognitive load perspective. *Instructional Science, 32*, 59-82.

Renkl, A., Atkinson, R., Maier, U., and Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education, 70,* 293-315.

Renkl, A., Stark, R., Gruber, H., and Mandl, H. (1998). Learning from worked-out examples: the effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23,* 90-108.

Resnick, L.B., and Omanson, S.F. (1987). Learning to understand arithmetic. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 3, pp. 41-95). Hillsdale, NJ: Erlbaum.

Richland, L.E., Zur, O., and Holyoak, K.J. (2007). Cognitive supports for analogy in the mathematics classroom. *Science, 316*, 1128-1129.

Rickards, J.P. (1975-1976). Processing effects of advance organizers interspersed in text. *Reading Research Quarterly,* 11, 599-622.

Rickards, J.P. (1976). Interaction of position and conceptual level of adjunct questions on immediate and delayed retention of text. *Journal of Educational Psychology, 68,* 210-217.

Roediger, H.L., and Karpicke, J.D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181-210.

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

Rohrer, D., and Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology, 20*, 1209-1224.

Rohrer, D., and Taylor, K. (in press). The shuffling of mathematics practice problems boosts learning. *Instructional Science.* (Available online at http://www.springerlink.com).

Rosenshine, B., Meister, C., and Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66,* 181-221.

Schmidt, R.A., and Bjork, R.A. (1992). New conceptualization of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207-217.

Schneider, W., Vise, M., Lockl, K., and Nelson, T.O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning (JOL) task. *Cognitive Development*, *15*, 115-134.

Schwonke, R., Wittwer, J., Aleven, V., Salden, R.J.C.M., Krieg, C., and Renkl, A. (2007). *Can tutored problem solving benefit from faded worked-out examples?* Paper presented at The European Cognitive Science Conference, Delphi, Greece.

Schworm, S., and Renkl, A. (2002). Learning by solved example problems: Instructional explanations reduce self-explanation activity. In W.D. Gray and C.D. Schunn (Eds.), *Proceeding of the 24th Annual Conference of the Cognitive Science Society* (pp. 816-821). Mahwah, NJ: Erlbaum.

Sloutsky, V.M., Kaminski, J.A., and Heckler, A.F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin & Review, 12*, 508-513.

*Starch, D. (1927). Educational psychology.* New York: *MacMillan.*

Sweller, J. (1999). *Instructional design in technical areas.* Victoria, Australia: Australian Council for Education Press.

Sweller, J., and Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2,* 59–89.

Thiede, K.W., Anderson, M.C.M., and Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95,* 66-73.

Thiede, K.W., Dunlosky, J., Griffin, T.D., and Wiley, J. (2005). Understanding the delayed-keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1267-1280.

Thompson, C.P., Wegner, S.K., and Bartling, C.A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 210-221.

Trabasso, T., and Magliano, J.P. (1996). Conscious understanding during comprehension. *Discourse Processes, 21*, 255-287.

Trafton, J.G., and Reiser, B.J. (1993). The contributions of studying examples and solving problems to skill acquisition. In M. Polson (Ed.), *Proceedings of the 15th Annual Conference of the Cognitive Science Society* (pp. 1017-1022). Hillsdale, NJ: Erlbaum.

VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., and Rose, C.P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science, 31*, 3-62.

Ward, M., and Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7*, 1-39.

Wisher, R.A., and Graesser, A.C. (2007). Question asking in advanced distributed learning environments. In S.M. Fiore and E. Salas (Eds.), *Toward a science of distributed learning and training* (pp. 209-234). Washington, DC: American Psychological Association.

Zhu, X., and Simon, H.A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4*, 137-166.