Toward the end of the 19th century, the German scientist Hermann Ebbinghaus concocted an experiment that countless children have unwittingly replicated ever since, over a morning bowl of Alpha-Bits. Ebbinghaus took consonants from the alphabet, slapped a random vowel between them, and, voila! some 2,300 nonsense syllables were born. For years, Ebbinghaus practiced these syllables at random, learning and re-learning until he had mastered the material. In 1885, he recorded his observations in *Memory: A Contribution to Experimental Psychology* — a seminal work that countless psychologists have wittingly read ever since, over morning bowls of cereal or otherwise.

Near the end of his monograph, Ebbinghaus mentioned a “noteworthy” detail from his learning trials. He found that a particular 12-syllable series could be conquered in two ways: by cramming 68 repetitions into a single day before testing, or by spacing 38 repetitions across several days. The difference, he wrote, was significant. “It makes the assumption probable that with any considerable number of repetitions a suitable distribution of them over a space of time is decidedly more advantageous than the massing of them at a single time.”

Ebbinghaus’s tests had a sample size of one: himself. But time and again, using far more rigorous empirical settings, psychologists confirmed the potency of this “spacing effect.” The method would seem to lend itself to immediate real-world application; what teacher or student would not want to enhance learning while limiting study-time? A century after Ebbinghaus, however, the spacing effect had still failed to grip the general culture. In a 1988 article in *American Psychologist*, Frank Dempster wondered why educators had not embraced “one of the most remarkable phenomena to emerge from laboratory research on learning.” He examined nine potential reasons not to adopt the practice; finding none of them compelling enough to offset potential benefits, Dempster recommended immediate classroom application.

Yet 20 years after Dempster, and more than 100 after Ebbinghaus, the spacing effect remains as invisible inside schools as it is unimpeachable in the minds of psychologists. But a recent spate of behavioral research that challenges educational conventions has pushed techniques like the spacing effect to the forefront of learning conversations. The act of testing, not studying, plays a superlative role in learning. The use of generic examples, not concrete ones, improves the understanding of math concepts. And spacing, instead of cramming, continues to prove its superiority in ways that surprise even long-time educational psychologists.
Meanwhile the past decade has also seen improved efforts at extending these practices into the public. The Department of Education’s Institute of Education Sciences has taken some steps to infuse such methods into a U.S. education system recently ranked 18th out of 24 nations. Last September, the institute published a comprehensive guide for teachers and textbook publishers that outlined the latest scientific findings.

Together the new research and the heightened application efforts have created hope that better approaches to learning will soon become classroom mainstays.

“The public at large has a strong feeling that schooling could be more effective,” says APS Past President Robert Bjork of the University of California, Los Angeles. “I think few things are more urgent.”

The Ideal Spacing Interval

In August of 2007, Doug Rohrer and APS Fellow Harold Pashler published a survey of recent spacing-effect literature in *Current Directions in Psychological Science*. As heartily as it endorses spacing, it shows the inefficiency of cramming. Rohrer and Pashler also listed a few “simple, concrete principles” that teachers could incorporate almost immediately. Grade-school teachers could save a pinch of time during weekly vocabulary sessions to review a few words from weeks before. College professors could give cumulative exams that ensured students would return to early material. In short, the superiority of spacing and the practicability of applying it are abundantly clear.

“Until last fall I never got a single [media] call about my research in 20 years,” says Rohrer, University of South Florida, who witnessed the strategy’s success during five years as a high school math teacher. “I've had maybe 20 calls since.”

Part of the spacing effect’s allure is its simplicity. Instead of cramming for an entire week before an exam, a student might learn the same material with two shorter, properly spaced study sessions. But the strategy has an obvious limit. With only so many hours in a day, students and teachers cannot review all the material they learned earlier. There must be some critical juncture between final study session and test date that optimizes time and effort.

Recently Rohrer, Pashler, and several colleagues completed a massive investigation to pinpoint this ideal distance. To decipher this Rosetta Stone of the spacing effect, the researchers taught some 1,350 subjects 32 semi-obscure facts (Q: Who invented snow golf? A: Rudyard Kipling). The subjects then reviewed what they learned after a break of anywhere from three minutes to three-and-a-half months. After this review study session the subjects endured an additional gap, up to a year long, before taking a final exam.

When the gap between initial learning and test date was a week, the optimum review took place a day after initial learning, the group reports in a paper published this month in *Psychological Science*. With a month gap, the ideal review occurred after about a week; with a year, the prime review came three weeks after learning.
The results don’t reveal some secret, cure-all review date. But when the researchers compared subjects whose review fell on the “ideal” gaps to those whose review immediately succeeded initial learning — the cramming scenario — they found “enormous” differences. When the exam fell a year after the review, ideal studiers beat crammers by 77 percent; when the exam fell seventy days after review, this improvement soared to 111 percent. For people who know how long they would like to retain knowledge — a semester for college students, a week for businessmen planning an overseas visit — the results point clearly to an ideal moment for reviewing what you want to know.

“Not only does spacing help, but the amount of spacing makes a difference,” says Rohrer. “The longer the test delay — the time between the final study and the test — the longer the ideal spacing gap.”

The spacing effect is so overwhelming that Bjork, in what he called an effort to “be fair,” recently designed a study in which he figured cramming would aid students. Although spacing clearly reigns supreme when learning facts, Bjork and UCLA colleague Nate Kornell thought cramming might improve “inductive” learning — that is, learning a concept or style through a slew of examples.

Bjork and Kornell showed subjects paintings by a dozen relatively unknown artists. Some learned the artistic styles by spacing, others by cramming. “In marked contrast to our expectations,” the authors concluded in the June 2008 issue of *Psychological Science*, spacing significantly beat cramming in every condition.

**When Testing Becomes Studying**

Regardless of how long or how often a person chooses to study, conventional wisdom says that the act of studying itself prompts learning to occur. We read information, we store some of what we read in our brains, we retrieve some of what we’ve stored during an exam, and we move on.

But a few recent studies, as powerful as they are contrarian, show that this widely-held belief is far from certain. Some learning indeed occurs during the study phase — but not nearly as much as occurs during testing.

“Traditionally people have focused at looking at processes when you study, and retrieval is just accessing” the information, says Jeffrey Karpicke of Purdue University. “We’re showing that something does happen when you retrieve some piece of information.”

To challenge these long-standing assumptions, Karpicke and APS Past President Henry L. Roediger III, Washington University in St. Louis, recently taught four groups of subjects 40 Swahili-English word-pairs. In the first phase of the experiment, each of the four groups went through a different study-test routine until each word was correctly learned once. The routines were designed to distinguish precisely which aspect, studying or testing, was more responsible for learning the word pair.
In the first group, subjects studied all 40 pairs of words, then were tested on all 40 pairs — repeating this routine until each word pair had been mastered once. The second group stopped studying a pair once it was recalled correctly, but continued being tested on all 40 pairs. Group 3 studied all 40 pairs each time but stopped being tested on pairs once they were correctly identified. Subjects in the last group, which represents typical schooling habits, stopped studying and being tested on each pair once it had been learned.

The second phase of the experiment occurred a week later, when the students took yet another exam of the word-pairs. “The results show that testing (and not studying) is the critical factor for promoting long-term recall,” the researchers conclude in the February 15, 2008, issue of Science. Indeed, the students who were tested on all the word pairs during their routines — Groups 1 and 2 — recalled nearly 80 percent on the final exam. Students who stopped studying words once they learned them a first time recalled only about a third of the pairs on the exam — even Group 3, which continued to study all 40 words throughout the duration of the learning phase.

“It’s actually very functional that our memories work this way,” Karpicke says. “You’d want to build a memory system, if you could, that when retrieves some piece of information, it kind of registers somehow ‘This is important, I need to get this again in the future.’”

The results fly in the face of the standard principle of learning through repetition. They also offer a clear guide to enhancing study habits. Instead of reviewing or transcribing lecture notes, students might be better off inventing potential test questions, recalling key points from memory, or, at the very least, making flashcards.

“Through college, kids think studying is re-reading and sometimes highlighting,” says Pashler, University of California, San Diego, who was not part of the Science study. “I think their work shows that if you force active retrieval it’s much more effective.”

Concrete and Blocks

Ultimately, students decide their own study approaches. But other recent research suggests that teachers might improve learning inside the classroom by tweaking their habits just slightly. After completing a unit on, say, long division, a teacher typically assigns a block of long division problems. If a few problems were swapped with questions from an old multiplication unit — a strategy known as “interleaving” practice items, as opposed to “blocking” them en masse — students might benefit in the long-run. Professors could accomplish a similar goal by redesigning syllabi to include short reviews of previous lessons at the end of each class.

Even the beloved approach of using concrete, real-world examples to explain abstract concepts has been called into question. In a study published in the April 25, 2008, issue of Science, cognitive scientists from Ohio State taught subjects the concept of modular arithmetic — the type used to tell time on a clock. If you add two hours to a clock at 11,
for example, it becomes 1, not 13. (The irony of needing a concrete example to explain this study was not lost on this reporter.)

Some subjects learned the modular concept through generic examples, such as shapes, whereas others learned it using concrete examples, such as pizza slices or tennis balls. After learning the concept, the subjects had to apply it to a children’s game whose rules functioned on the basic modular properties just learned. They were then asked questions about the game’s rules; the idea was that subjects who had truly mastered the concept would be able to apply their understanding to this new setting.

Sure enough, students who had learned the modular principle through generic examples significantly outperformed the tennis ball and pizza crowd. Responses by the latter group, in fact, did not exceed chance. In another experiment, students who learned the concept through generic examples bested those who learned it using both types of examples. Concrete examples, it seems, may even hurt a student’s ability to learn.

In the concrete scenarios, “we think attention is getting shifted to the superficial, and not focused on the mathematical structure,” says Jennifer Kaminski, the paper’s lead author. Though students might grasp concrete examples more quickly, in the long-run they will be limiting their ability to apply their knowledge to new settings. “What if you’ve got to wrestle with something that looks very different [from the concrete example], but the structure is the same?” says Kaminski. “That’s the challenge.”

A Tough Battle

About seven or eight years ago, Harold Pashler sat down to write a paper with a clear purpose: He would advise educators of the latest experimental findings that might directly aid classroom learning. Try as he might, Pashler recalled recently, “there wasn’t enough to say.”

It was about that time that the federal Institute of Education Sciences was established. From the start, the institute emphasized tough empirical standards and aspired to bridge the gap between laboratory and classroom. “In education there’s a bottleneck between research and practice,” says Rohrer. “IES has made this a particular aim.”

In September 2007, this aim hit its target: an IES Practice Guide, “Organizing Instruction and Study to Improve Student Learning,” similar in spirit to Pashler’s abandoned paper from years before. In fact, Pashler was chair of the seven researchers who oversaw the guide.

“I have always thought it’s a little bit strange and slightly embarrassing how little impact the field has had on even nuts-and-bolts educational kinds of content,” he says. “But it wasn’t until IES got going that I — and many other psychologists — could get stable funding” to do this type of research.

The 63-page practice guide outlines seven major learning strategies — from spacing study sessions to interleaving practice problems to embracing generic examples — and
offers teachers tips on applying the approaches in their own classrooms. The guide also ranks the science behind each strategy from “strong” to “low,” perhaps in an effort to help convince wary school administrators.

“There was enough tangible content to start urging practitioners to pay notice of it,” Pashler says. “We’re hoping teachers will make use of the guide and the Web site. I also hope publishers will pay some attention.” Rohrer and others envision software programs that track student performance and tailor some of these strategies accordingly.

Though the prognosis for classroom application looks far rosier than it did in Dempster’s day, the obstacles remain considerable. Adopting many of these changes would overhaul the entire classroom experience, perhaps leading to frustrated students, parents, and principals, says Bjork, who recently completed a three-year, IES-funded project called IDDEAS, or Introducing Desirable Difficulties for Educational Applications in Science. The term “desirable difficulties,” coined by Bjork, rightly suggests that though strategies like spacing or self-testing will pay off in the long-run, they require great discipline and commitment. In a culture bent on immediate success, such longitudinal thinking doesn’t always fly.

An additional barrier comes from our own brains. Research on metacognition by APS Fellow Janet Metcalfe, Columbia University, and others has shown that students are not very good at judging the effectiveness of their own study habits. In Bjork’s recent study on artistic styles, for example, the subjects overwhelmingly learned better by spacing techniques — yet when asked which approach had benefited them more, the subjects overwhelmingly said cramming.

“Our intuitions are not a good guide for how to optimize learning instruction, nor are standard practices from years past,” says Bjork. “It’s the research that should be the guide, but it’s a tough battle to introduce and make some of those changes.”

Related Links


[http://www.psychologicalscience.org/observer/getArticle.cfm?id=2356](http://www.psychologicalscience.org/observer/getArticle.cfm?id=2356)

[http://www.psychologicalscience.org/observer/getArticle.cfm?id=1953](http://www.psychologicalscience.org/observer/getArticle.cfm?id=1953)

1 Italics are the author’s. Ebbinghaus, H. (1913) **Memory: A contribution to experimental psychology.** (p.89) New York: Columbia University.