Translating Learning Science into Learning Strategy

Cerego White Paper

May 1st 2016

Cerego

Iain M. Harlow, Paul T. Mumma & Andrew Smith Lewis

Introduction

What is Cerego?

Cerego is a learning platform that helps users learn material more efficiently and retain that knowledge for longer. Cerego does this by translating reliably demonstrated and effective results from learning and memory science into an adaptive, personalized learning tool, providing reviews to users at the moment they are most beneficial for learning. In this white paper, the science underlying Cerego's adaptive learning engine, how it is implemented, and how learning science influences other features of the platform, are explained.

First and foremost, the adaptive learning platform deployed by Cerego relies on a deep scientific understanding of human memory and learning. In **Learning Science Overview**, the core findings of more than one hundred years of research into how we learn are outlined. These include the advantages of retrieval practice rather than restudying, distributing learning across time rather than cramming, and understanding your own learning strategy and progress.

In **Cerego's Learning Engine** we describe the core functionality of Cerego's adaptive learning engine, and how it builds upon insights from learning and memory research to build long-lasting foundational knowledge as efficiently as possible.

Learning does not exist in a vacuum. Surrounding the core functionality of Cerego's adaptive learning engine are a number of important features. In particular, Cerego has a strong emphasis on **Measuring and Visualizing Knowledge**, giving users ownership of their learning progress and allowing instructors to identify and give targeted help to learners who need it.

Cerego's core aim is to support efficient learning of foundational knowledge, and this can be represented by many forms of **Structured Content**, from sentence completion and region identification to sequences and patterns. Learners build flexible and transferable knowledge by studying different types of material, in different formats, tested in different ways, and supported by notes, slides, videos or other media. Learners can also create or remix any of these item types themselves in the platform.

An important aim of Cerego is the **Integration** of adaptive learning into real life, and existing learning environments. Cerego provides scheduling tools that help guide learning over time; mobile apps that meet learners where they are and let them review conveniently; LTI integrations linking Cerego to Learning Management Systems that help instructors assign and track their users' progress; and deep integrations with content that embed Cerego directly into a primary information source such as an e-textbook.

As a scientifically driven company, developing and strenuously testing the efficacy of the platform is central to Cerego's mission. Cerego is also a mature product, with a history of use within different learning environments, an internal research and learning science team, as well as close and growing collaborations with academic researchers. In the final part of this white paper we highlight some example **Research and Case Studies**, including implementations from K-12 to large-scale online courses, as well as insights from internal research.

Learning Science Overview

Although human memory has been a subject of study and interest for thousands of years, in the last century the field of learning science has formalized research into the way we learn and retain information. This research has greatly enhanced our understanding of human memory in a wide range of ways, but among the most interesting results concern how learning and memory can be influenced by different strategies. These findings are especially important because they can be directly applied to improve learning outcomes in the real world, independently of a learner's inherent intellectual abilities.

In this section we summarize some of the most robust and important of these findings: The benefits of distributing learning across time, often functionally implemented using spaced repetition, and the benefits of retrieval practice, also termed the "testing effect". Both of these strategies have a substantial history of empirical research into their effectiveness, have been shown to enhance long-term retention, and have also shown promise in translating to effective learning outside the laboratory [*Dunlosky et al. 2013*]. In addition, we also highlight an important distinction between long-term retention and the current activation or availability of a memory, which lies at the root of why some techniques are more effective than others, and outline the important role of metacognition (awareness of our own cognitive processes) and learning strategies in general.

Distributed Learning

One of the most well-studied learning science phenomena is the benefit gained by spacing out learning across time [*Baddeley & Longman, 1978; Dempster, 1990; Sobel et al., 2007*]. The distributed learning or spaced repetition effect refers to the observation that spacing out learning episodes across time improves long-term retention relative to studying the same material in a single session, even when total study time is the same.

One of the key reasons that distributed learning has so much potential to impact learning in the real world is that it is not naturally intuitive to learners. Reviewing material shortly after learning it, while the information is still fresh and available, can confer a sense of fluency and ease that users mistake for learning. Similarly, reviewing material a short time after learning (instead of spacing out reviews) may increase performance a few seconds or minutes afterwards, even while significantly decreasing longer-term retention [*Ebbinghaus, 1885; Cepeda et al., 2006; Simon & Bjork, 2001; Kornell & Bjork, 2008*]. There is therefore considerable scope to improve long-term learning outcomes by shifting users from an intuitive non-spaced learning approach to a less intuitive, but much more effective, distributed learning strategy.

Defining the optimal spacing between reviews is a challenging empirical problem. Cepeda et al. [*2008*] carried out a series of long-term memory experiments to quantify the optimal review intervals for different scales of long-term learning, and demonstrated a relationship between the review interval and the length of time during which the memory was available. Shorter review intervals led to better retention on test administered a short time later, but memory over longer intervals was more robust when the review interval was also longer.

The optimal review interval, in practice, is likely to be a function of many factors, including the length of time the memory is required for, the difficulty and type of material being learned, and - perhaps most importantly - the characteristics of each individual learner. Cerego is designed to optimize this interval for every user, and for each item that user is learning.

Retrieval Practice

When trying to actively learn and review material, most people would intuitively re-study the information they are trying to learn, for example by rereading a crucial chapter before an exam. In fact, however, more than one hundred years of learning science has found strong and consistent evidence for what is variously termed the “testing effect”, “retrieval practice” or “test-enhanced learning”. Retrieval practice - actively attempting to recall previously studied material - is more effective for long-term retention than spending the same amount of time rereading or re-studying that same material [Carrier & Paschler, 1992; *Rawson & Dunlosky, 2011; Roediger & Butler, 2011*]. That is, rather than rereading the relevant chapter, a student would learn more spending their time attempting to answer practice questions on the subject.

Notably, the benefits of retrieval practice manifest primarily in long-term retention. Partly as a result of this, they are not necessarily obvious to learners, who often incorrectly judge their own learning to have been improved as much or more by study than retrieval practice, despite dramatic improvements in later retention after engaging in repeated retrieval [*Karpicke & Roediger, 2008*]. This mismatch between intuition and outcome is one reason why much recent work has focused on effectively translating the benefits of retrieval practice into learning environments, such as in higher education settings [*McDaniel et al., 2007; Dunlosky et al., 2013*]. Another reason retrieval practice is important is that it gives both students and instructors valuable direct feedback on their learning progress.

How and why does retrieval yield stronger memory than re-study? Many possible benefits and mechanisms have been proposed [*Roediger et al. 2011*]: Retrieval may strengthen and create a greater variety of critical connections in the related neural networks, enhancing the probability of later retrieval; it may help learners mentally organise content more effectively; it may improve learning on a later presentation of the material; it may enhance transfer to more general learning contexts; and it can help learners calibrate their metacognitive assessments of their memory. Although the mechanisms underlying the learning advantages of retrieval remain a rich focus of research, the practical benefits are certainly among the most robust and applicable of learning science effects [*Dunlosky et al, 2013*].

Activation and Retention

Both the retrieval practice and distributed learning effects strongly suggest that the more effort spent reactivating and retrieving some material at review, the more effective that review is at building long-term *retention* for the material being learned. That is, ideally a learner should schedule a retrieval practice for an item when its *activation* has decreased to the point where it is challenging, but still possible, to retrieve from memory.

How do we define and quantify *activation* and *retention*, and how are they related?

It is important to recognise that unlike memory in a computer, **human memories do not last indefinitely**. All memories fade and become less accurate over time. There is good reason for this: It would be impossible to store every single thing that we ever thought, saw, or encountered in our lives, and so the fact that old or unimportant memories fade and are replaced by new and more relevant ones keeps our knowledge about the world up to date.

We can refer to the current availability of a memory as its *activation*. Activation refers to the immediate availability of a memory: How accessible it is at a specific moment in time, and how easy it is to retrieve. The activation for some information will generally be high for things you have encountered recently: you might read a sentence in a book, close your eyes, and be able to repeat that sentence with high accuracy.

Not all memories fade at the same rate, however. Things you have encountered often are more likely to fade slowly from your memory, meaning that the things you remember are more likely to be things that are relevant and encountered often. *Retention* refers to this robustness of a memory over the longer term. It is unlikely you would be able to accurately recall the sentence from the book a week later, but you might easily recall the name of the book. That information has greater retention, for example because it has been encountered more frequently, and its activation decays more slowly as a result.

Building *retention* is key to learning and mastering knowledge.

Not all memory strategies are effective for building retention. For example, memory techniques such as the keyword mnemonic provide a boost to memory in the short term (i.e. they increase activation for a short time), but weaker and more fragile memories in the long term [*Wang et al., 1992*]. Underlining material as a mnemonic aid has also shown underwhelming results in studies, including a negative effect on building long-lasting, transferable, inferential knowledge [*Peterson, 1992*]. Similarly, studying or reviewing materials in bulk shortly before a test (“cramming”) is a common strategy for students - in part because it is a very easy strategy to schedule and employ - but also leads to poorer long-term retention than a distributed learning strategy [*Keppel 1964*]. In contrast, distributed learning and retrieval practice have both been demonstrated to enhance long-term learning [*Bahrick, 1979; Benjamin & Tullis, 2010; Rawson & Dunlosky, 2011; Roediger & Butler, 2011*], even in cases where they have smaller or negative effects on activation over the short term.

It is clear from these and other studies that the distinction between activation and long-term retention is critical to developing effective learning strategies and tools. Less obvious, but just as important, is the implication that the true effectiveness of study techniques may not be obvious to a learner while engaging in them, or shortly afterwards. This is why cramming for an exam feels like it helps: activation for the material is raised by recent exposure, and this activation makes the student feel like they've mastered the material. Cramming is less effective however for building retention, and the activation earned decays rapidly as a result, making cramming a poor choice of study technique for building lasting knowledge. Strategies that have low or even negative effects on long-term retention may therefor be (inaccurately) perceived by a learner as beneficial.

Activation can be achieved quickly, with a single study, and can also fade quickly. Retention, on the other hand - lasting benefit from learning - takes time, and an effective strategy, to build. Building retention pays off by keeping the activation and fluency of knowledge higher for longer; stronger memories with greater retention don't need to be refreshed as frequently to be available. For example, cramming for a mid-term may be less efficient than building retention, since the same information will have faded and require re-learning before a final exam. Retention is also critical for building flexible, inferential and transferable knowledge, since the slower rate of memory decay **allows more information to remain accessible and active at the same time**, and be combined or related in the ways that are critical to higher level mastery of a subject.

In terms of learning, the activation of a memory at review (i.e. how difficult it is to re-access) also plays a critical role in how effectively that review promotes long-term retention: When activation is high, such as shortly after seeing the material, reviewing provides less benefit to memory retention than when it is moderate, and more challenging to retrieve.

The Cerego learning engine uses a deep scientific understanding of human memory and learning, together with statistical analysis of each user's memory responses, to track and predict the activation and the retention of each item being learned. Each learner's personalized review schedule is guided by these predictions. The aim of the learning engine is to maximize retention as efficiently and effectively as possible by scheduling reviews of material for when it has decayed to a challenging level of activation. Cerego thereby supports and optimizes long-term learning by translating the most robustly demonstrated and effective memory strategies from the scientific literature (such as distributed learning and retrieval practice) into a personalized, adaptive learning engine.

Metacognition and Scheduling

The fields of learning and memory research are increasingly recognising the crucial role that metacognition - what we know about what we know - plays in effective learning, especially outside the controlled context of a laboratory [*Bjork & Yan, 2014*]. Having a limited amount of mental resources, such as available study time, means that being able to allocate these resources efficiently plays a large role in how much an individual is able to learn [*Metcalfe, 2009*]. When a learner engages with a source of content, a large number of strategic metacognitive decisions have to be made about the amount, timing and focus on different areas, all informed by the learners' knowledge of how well they know each part of the material, and all independent of the act of learning itself.

For example, a learner who recognises when they have learned some material effectively can move on and direct their attention and resources to learning new or refreshing more difficult material. A learner without such insight, or who mistakes the fluency of reviewing material they already know well for effective learning, may spend too little time on more difficult, but ultimately more important and effective reviews of material they know less well.

There is increasing evidence that individual differences in metacognition - the ability to accurately determine how well you know different material, and adopt a strategy that directs mental resources to where they are most needed - go a substantial way towards explaining individual differences in overall learning, independently of learners' IQ [*Veenman et al., 2004*]. For example, a student who correctly identifies their more weakly learned items and focuses greater study time on learning those will likely gain a stronger overall grasp of their coursework than an otherwise equally intelligent and diligent student who spends the bulk of their available time reviewing material they know well.

While a strategy of focusing on weaker items might seem quite obvious in the abstract, in practice it can be a challenging strategy for an individual to employ. Firstly, the requirement to accurately track strength of knowledge across different material being learned is not insubstantial. Secondly, there is a tension between strategies that feel intuitively helpful - often those which increase memory in the short term but not in the longer-term, such as mnemonic devices, cramming, or reviewing recently studied material - and those which support long-term retention, such as distributed learning and effortful retrieval practice. Studies have repeatedly shown that learners misjudge the effectiveness of memory strategies [e.g. Zechmeister & Shaughnessy, 1980; Karpicke & Roediger, 2008], mistakenly valuing positive feedback and sense of accomplishment for reviewing easier material ("I got that question right!") more highly than the more important, but less immediately tangible, long-term rewards of efficiently focusing learning where it is most needed. That is, easier retrieval tasks that benefit short-term learning can in fact have the opposite effect on long-term learning, and therefore lead individuals to feel as though they are studying effectively when in fact they are not. This tension, whereby tasks that seem difficult now may in fact result in the most significant improvements in long-term learning, has been termed the "Desirable Difficulty" effect [*Bjork, 1994*].

Both of these challenges can be overcome within a controlled learning environment, by tracking memory strengths for different material, and by guiding the user along an appropriate strategy for their chosen learning goal (such as life-long retention, or an end-of-course exam). Even more valuably, an ideal learning platform should surface this strategy to learners, prompt assessment of their own memories, and give accurate feedback on how well material has been learned. By doing so, the platform can better support the development of strong metacognitive skills: helping users learn *how* to learn more effectively. In the following sections we describe how Cerego supports these crucial metacognitive tasks, to improve the effectiveness and efficiency of learning.

Translating Learning Science into Learning Strategy

The study of learning science has provided a solid understanding of the strategies and factors that improve long-term retention. Studies have consistently demonstrated more efficient and robust learning can be supported by:

- **Accurate metacognitive judgments** [*Metcalfe 2009*], that support the efficient distribution of study time and mental resources to material most in need of strengthening;

- **Distributed learning** [*Cepeda et al. 2008*], the presentation and re-learning of material at the optimal point in time; and

- **Retrieval practice** [*Bjork 1975*], that replaces passive re-study with effortful and varied retrieval.

In particular, a recent extensive review of the learning science literature [*Dunlosky et al., 2013*] examined the scientific evidence supporting ten different learning techniques, and assessed each one for potential effectiveness in the classroom. This review identified distributed learning and retrieval practice as the two most promising strategies for improving long-term learning.

Altogether, the adoption of a learning *strategy* that incorporates these elements has been demonstrated to be highly effective outside the laboratory [*Rawson et al., 2013*] - but an onerous and difficult task to set up manually for an individual learner, as well as a counterintuitive one to maintain, as discussed above. The crucial challenge then facing researchers, educators and users today is the effective translation of these strategies into real-world environments, such as a classroom, an online course, professional training or self-directed learning, and to do so for learners with different goals, backgrounds, and abilities, across varied types of learning content and subject areas. The Cerego platform, described in detail in the following sections, is designed to achieve this translation effectively and efficiently. For example:

- The Cerego learning engine implements adaptive **distributed learning** in the form of **retrieval practice**, scheduling each learner's type and time of review in order to optimize for **desirable review difficulty** and therefore greatest long-term **retention**.

- The Cerego memory bank visualizes each memory's current **activation** and **retention** for every learner, as well as their upcoming review schedule, promoting ownership and understanding of the learner's progress and the **learning strategy** being employed.

- Reviews in Cerego encourage learners to **effortfully recall** and **judge** their own memory before responding, and provide feedback to the learner, further helping to build and support the development of **metacognitive** skills and awareness of learning strategies.

- Cerego content is structured in sophisticated ways to support foundational knowledge that is **transferable**, **flexible**, and can represent useful information about the world.

- Cerego employs a range of different quiz types to add **depth** and **variation** to the retrieval and re-encoding process, helping to build knowledge that is **transferable**.

- Reviews of different types of material in Cerego can be **interleaved**, further promoting **desirable difficulties** and varying the context of retrieval.

- Learners in Cerego have access to content creation and remixing tools to **personalize** their learning experience and deepen **engagement** with the content.

- Learning in Cerego is integrated into daily life through mobile apps to allow for genuinely **distributed learning**, and can be embedded directly into primary source material or supplemented with supporting notes and media for **richer learning**.

Cerego's Learning Engine

Cerego's learning engine is built on a powerful combination of two of the most robust findings in learning science: **distributed learning** and **retrieval practice**. Following these two principles, the learning engine keeps track of each user's interactions with the material they are learning and adapts the review schedule and learning experience to that individual user. The goal of the learning engine is to combine scientific principles about memory dynamics, with user- and item-specific information from the system to predict the activation of each memory in the future, and schedule a review for when the memory has faded to the point of desirable difficulty. This overall strategy is referred to as *DARPA*: Distributed Adaptive Retrieval Practice Algorithm.

Distributed Learning in Cerego

At the core of Cerego's learning engine is a review scheduler, which tracks the activation and retention of each item in a user's library. The retention of an item refers to how robustly it has been learned. In contrast, the activation refers to how available the information is at the present moment: A recently-studied item will have high activation, and the activation will decay over time between reviews. Higher-retention items (those that have been learned to a greater degree) will tend to decay more slowly over time, and remain available for longer, making higher retention a desirable goal for learning.

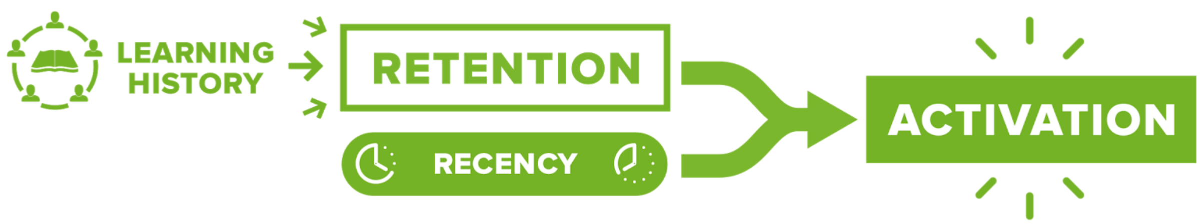

Figure 1: Retention and activation. The Cerego learning engine estimates the user's retention for some material by tracking the user's history of interaction with it. The activation of the material - how easy it is to access from memory right now - depends on how recently it was seen (more recently seen material is easier to bring to mind) and how much retention has been built (more retention means activation decays more slowly over time).

Every time a user studies or reviews some material in Cerego, the learning engine receives new information about how well the user is learning, how active their memories are currently, and how accurate the learning engine's predictions of activation actually were for that material. Using this new information, the learning engine adaptively updates the estimated retention and activation of that user's memories. Although the engine takes several different factors into account, generally speaking the retention is increased when the user demonstrates learning, and decreased when the user does not recall the material successfully, while the activation always increases after exposure since the material has been recently seen. This adaptive error-correcting process updates and recalculates the memory statistics, and as a consequence the review schedule, every time the user interacts with the material.

As a result of this, learners in Cerego will tend to notice their review schedules spreading out more for their items as they learn them, even when they initially started off quite similar to each other. When learners initially study some new items, the learning engine will generally schedule the intervals until each item's next review to be quite close to each other, since the main information available at that point is derived from the user's overall history with Cerego, and general scientific principles about memory activation decay. As the user returns and begins to review these items, however, the review intervals for each will start to diverge as more and more information becomes available about each memory a user is trying to build. In this way the learning engine iteratively improves its estimate of the retention and activation function for each individual memory.

The learning engine is constantly tracking the decaying activation for each memory, and schedules a review for when the activation is due to hit a predetermined threshold. This threshold is chosen so that the reviewed material is still able to be recalled with a relatively high probability (80-90%), but only with some effort - in other words, to achieve the ‘desirable difficulty’ that promotes deeper engagement with the material and better long-term learning.

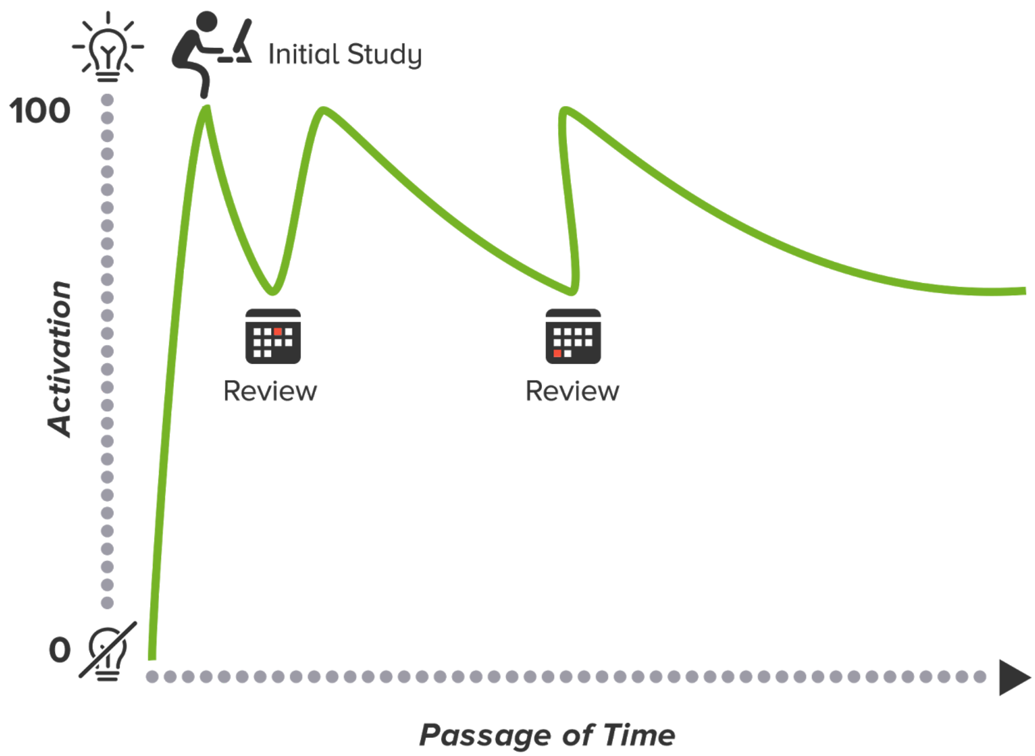

Figure 2: Change in activation of a memory over time. The activation of the memory is boosted by recent exposure, and decays between reviews. The decay rate depends on the memory’s retention, which generally increases over time as the user reviews the material.

After successfully reviewing the material, the user’s retention for that item will be higher, and so the activation will decay more slowly than before, increasing the amount of time before the next review is required. The review intervals therefore tend to expand over time, meaning Cerego can be understood to implement a form of distributed learning known as **expanded**

spaced repetition in the ideal case of a learner who always successfully retrieves the sought-after material.

Crucially, though, the review intervals are **adaptive** - that is, they do not only expand, but rather adjust over time according to the calculated (changing) retention and activation of the item being learned. The timing of the review is optimized not to an arbitrary interval, but to a specific **desirable review difficulty** designed to promote effortful retrieval and therefore more effective learning. Using the pattern of prior exposure, review accuracies and other data, identifying this optimal review time for each individual learner's memories is the primary function of the learning engine.

Retrieval Practice in Cerego

To promote effective learning, it is important that the user effortfully attempts to recall the information being learned [*Roediger & Karpicke, 2006*]. Accordingly, Cerego implements an effortful retrieval practice as the standard way of reviewing material. When an item that has been previously studied is presented for review, it takes the form of a recall probe (*Figure 3*). A cue element is presented and users are asked to recall an associated facet, with two choices available: "Don't Know It" (provides user the answer) and "Know It" (tests the user on their memory for the facet).

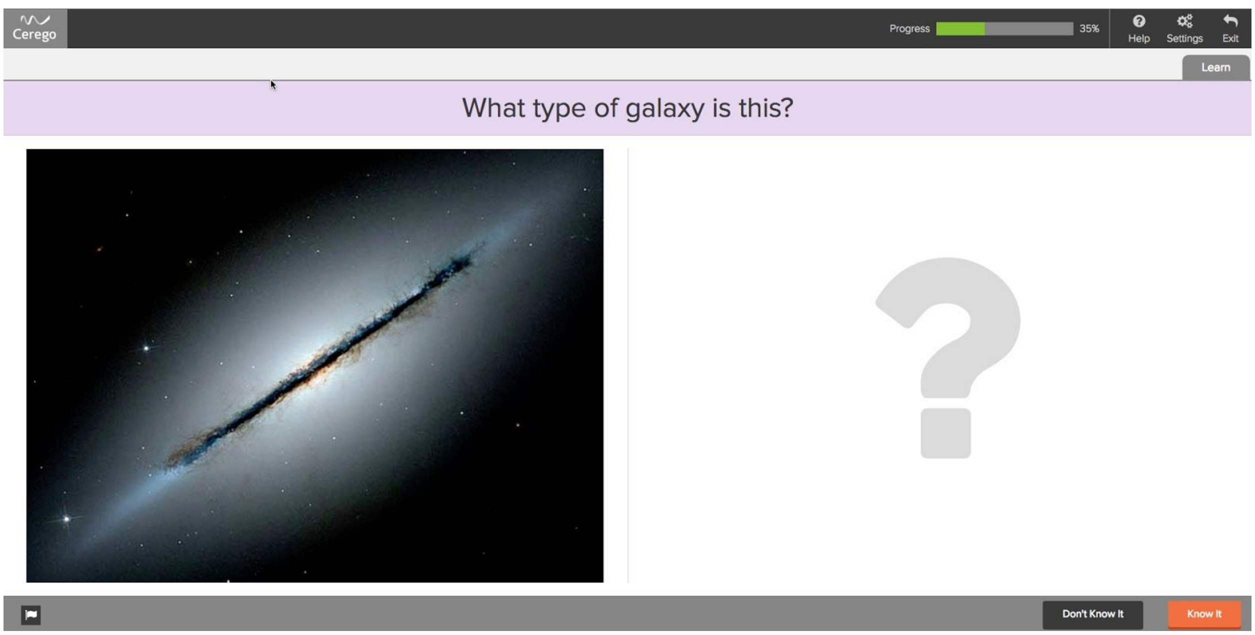

Figure 3: The recall screen in the Cerego Learn app. Users are given a review question and asked to try and recall the answer. Once they do so, they can proceed to the answer screen by clicking "Know It".

If users indicate that they don't know the associated facet, or indicate that they do but subsequently answer incorrectly, the learning engine updates its estimates of the retention and activation of the memory, and the user is presented with the original association for re-study along with any optional context or media linked with the item. The element is then rescheduled to appear later in the session, and the user progresses to the next item. In this way each item to be reviewed must be correctly retrieved at least once before the session is completed.

If users indicate that they do know the facet, they progress to a test screen (*Figure 4*), which varies depending on the type of content and type of test. Successful retrieval prompts an update (increase) of the retention and activation of the item, and reschedules it for some time in the future when the activation will have decayed again to an optimal level of difficulty.

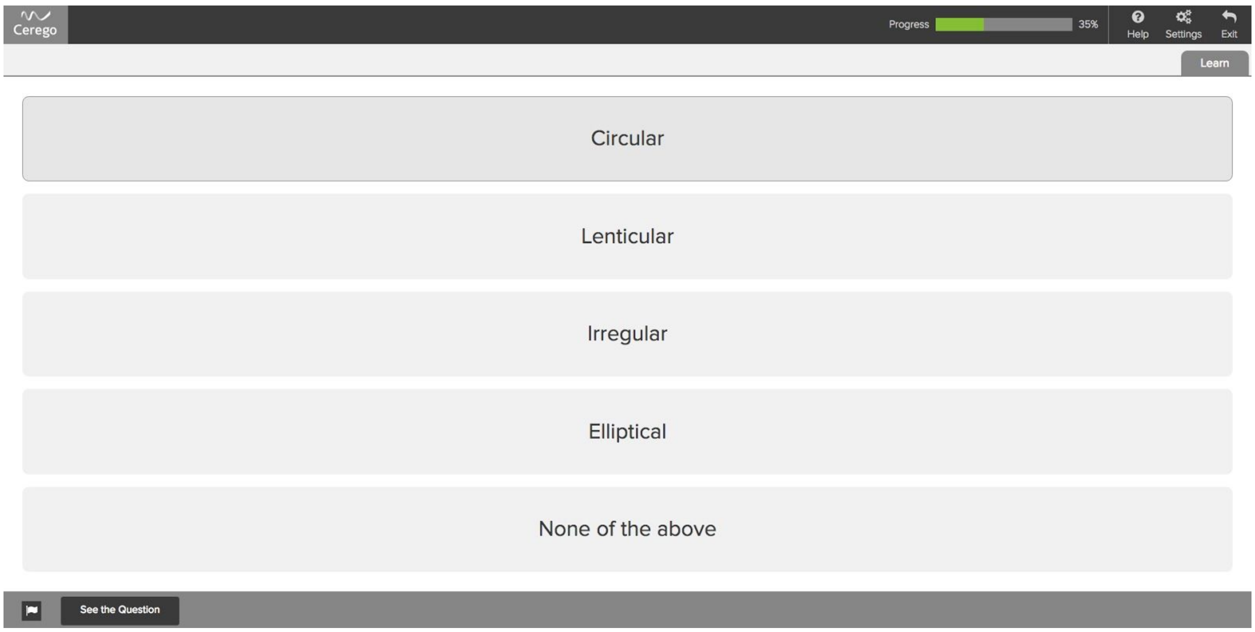

Figure 4: The test screen for the question presented in Figure 3. In this example, the review test takes the form of a multiple choice quiz. See Structured Content section for a discussion of the types of quizzes and content in Cerego.

In addition to optimizing the review interval for each user's memories, this review procedure replicates many of the key factors advocated by learning science research to improve long-term retention.

- Users are presented with a **recall** opportunity for every item being reviewed, promoting deeper and more effortful retrieval, as well as a simple **metacognitive assessment** of their own memory strength.

- The test question for each item is **variable**: the cue and target can be reversed, and the test type can involve receptive tests such as region selection and multiple choice, or productive tests such as fill-in-the-blank, sentence completion or sequence ordering. This further varies the retrieval context, promoting more robust and **transferable** learning.

- A correct response is always required for each item: initially missed trials are reviewed and retrieved later in the same session, ensuring users always **interact more deeply** with the correct answer by producing it themselves rather than only re-studying it.

- The correct answer is always presented after each trial to strengthen the association being learned and **reduce interference** with distractor items or false memories.

- Users are always incentivized to **effortfully retrieve** each memory, since accurate responses reduce both the amount of time spent in that session (because additional trials are required for every incorrect response) and in the future (since the following review interval for that item will expand when it is correctly retrieved, reflecting greater retention and slower memory decay).

Measuring and Visualizing Knowledge

Cerego's learning engine implements a distributed adaptive retrieval strategy to make learning more efficient and more durable for users. Mediating efficient learning, however, is just one purpose of the Cerego platform. Equally important is the ability to measure, visualize and demonstrate that learning, and Cerego provides a number of tools for this purpose.

The primary user-facing visualization is the Memory Bank, which shows the activation and retention of every item a user is studying in real-time. The memory bank also allows users to visually explore their review history, upcoming schedule, item difficulties and more.

Visualizing the learning progress of others is also a key feature of the Cerego platform. Learning is often undertaken in the context of a goal, qualification or training, and Cerego provides multiple ways - including a dedicated mobile app - for an instructor or course provider to keep track of the progress of their users.

Tracking Your Own Learning: The Memory Bank

A core feature of the Cerego platform is not simply to help people learn, but to be able to show them what they have learned. The Memory Bank is the primary way of visualizing this information. An important motivation for the Memory Bank is to give users ownership, control, and understanding of their learning progress and strategy. Cerego consciously replaces the "black-box" of a system that simply directs learners to study in an efficient way, with a transparent approach that also provides them with detailed, real-time information about their progress.

Strong metacognitive skills - such as knowing how strong and active one's memories are, and how much time and mental effort should be devoted to reviewing each one - are a key predictor of effective learning outside the laboratory [*Veenman et al., 2004; Metcalfe, 2009*]. While Cerego automatically implements an appropriate review strategy and schedule for each user's content, it is a considerably more interesting challenge to enhance a user's metacognitive skills, and understanding of their own learning, than it is to simply replace their intuitive strategy with a more efficient one. This is a key reason why Cerego places such a strong emphasis on user-facing visualizations: To surface this progress and strategy clearly to learners.

Another important outcome of this approach is to highlight the difference between activation and retention, a distinction that when poorly understood can lead to ineffective learning strategies: Keyword mnemonics, re-reading material, underlining and cramming can all confer a sense of achievement or fluency on a learner as they raise the activation of a memory, but may be ineffective for building retention in the form of robust, transferable knowledge [*for a review see Dunlosky et al, 2013*].

The Memory Bank (*Figure 5*) shows a learner both the activation and retention of each memory at once, using height for activation, and color and horizontal position for retention. Retention is tracked in terms of the rate of decay of the memory: higher retention items are active and available for longer. Distributed retrieval testing is a highly efficient way of learning in the longer term, since retention tends to increase exponentially with successful learning. This means that relatively few reviews, when spread out optimally across time, can lead to memories that last weeks, months or even years. Because exponential scales can be difficult to intuit, however, Cerego categorises memories into a linear scale of levels based on their retention. The levels correspond to ranges of retention time, as outlined in *Figure 6*.

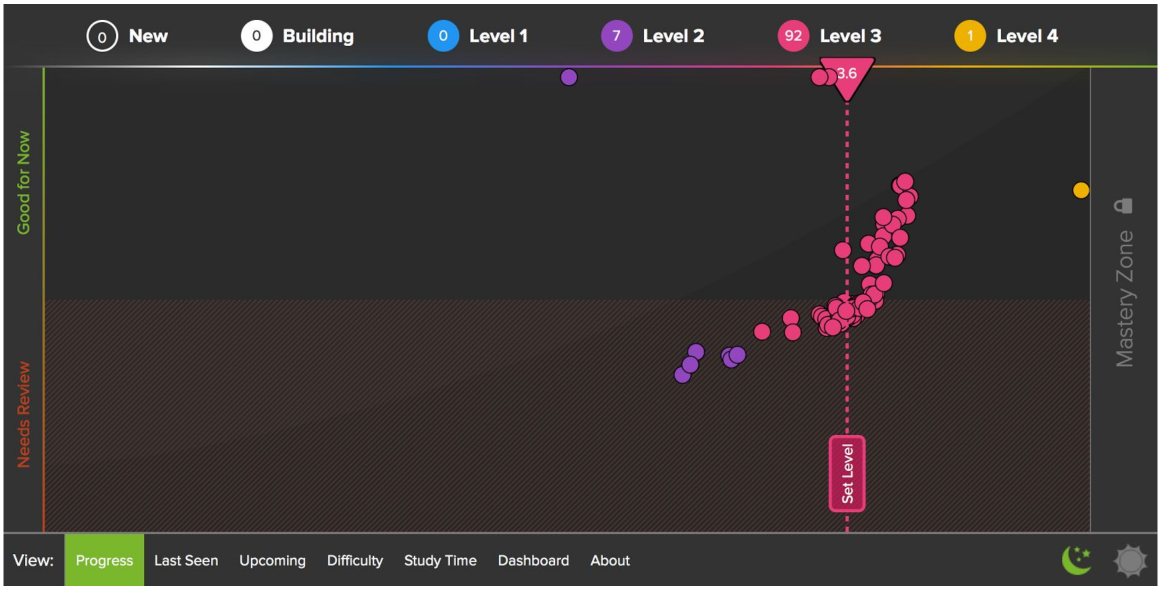

Figure 5: The Memory Bank Progress screen. Each memory is represented as an individual orb, where the color and horizontal position shows the long-term retention built for that memory, and the height shows how active it is. As memories fade over time, the orbs drop in height and the learning engine will schedule them for review. The Set Level shows the average level for all the items in the set, and gives an overview of how well the user has mastered that material.

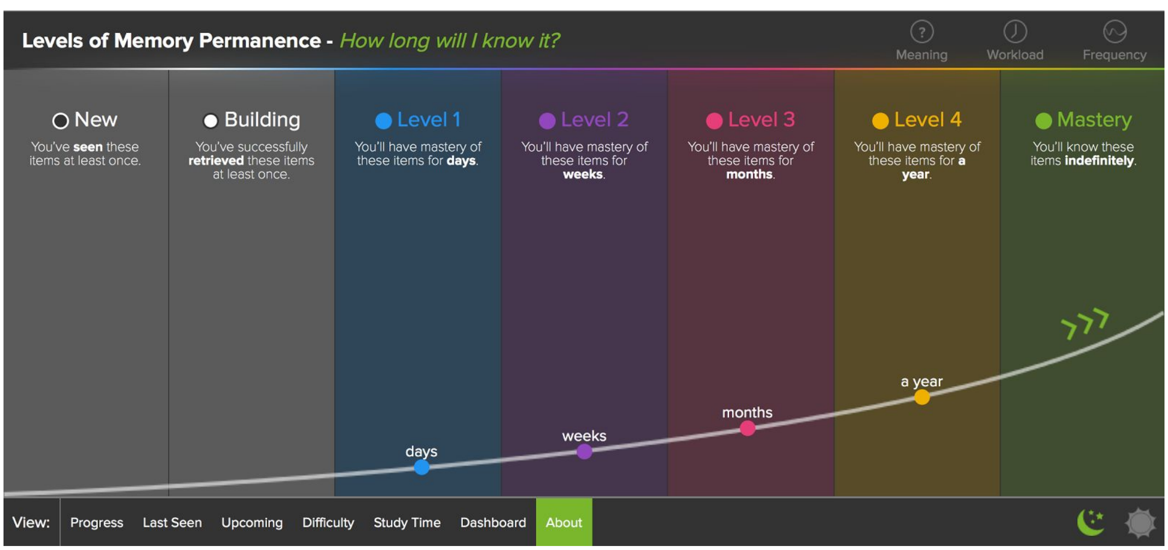

Figure 6: The levels of retention in Cerego correspond to different ranges of retention time. For example, Level 1 items are those which a user will generally remember a few days after last seeing them, while Level 3 items will often still be active several months since they were last encountered. New and Building memories have been recently studied for the first time, and will need to be reviewed soon to build enough retention for them to last for several days.

Learners can track which material is fading from their memory as they drop lower on the activation axis over time, falling from *Good For Now* into *Needs Review*. Simultaneously, the retention of each item highlights which of their memories they have built lasting strength for and which require more frequent reviews to stay active.

A *Set Level* is also shown on the memory bank, and provides an overall summary of how well the user has learned the material in a set. Set levels can be assigned as goals by content providers and instructors, requiring students to build a custom level of retention for the material.

The Memory Bank provides several additional views that give a learner information about, and control over, their learning. In the *Upcoming* screen (*Figure 7a*) learners can view their next scheduled review time for each individual piece of content they are learning, providing direct insight into their distributed learning strategy and helping them to plan their upcoming study time. While the learning engine will automatically schedule reviews for material that requires it, here users can directly see their learning schedule laid out, helping them to plan out study sessions or decide whether they have enough extra time to begin learning new material.

Difficulty (*Figure 7b*) displays each memory as a function of a learner's past history with that content, enabling learners to identify their most challenging material and guide their learning to focus on it. Additional screen show users their most recent past review for each item, as well as the total study time they have spent on it.

Learning often occurs in the context of a goal. In Cerego, the primary outcome is durable retention, and so goals are set in terms of retention. Learners can visualize their progress towards their goal directly on the memory bank: when a goal is assigned, the Goal, the learner's current aggregate retention (their Set Level), and their percentage progress are prominently displayed and updated directly on the memory bank before and after each study session (*Figure 8*).

Learning in the context of a goal generally exists in a more formal course environment, whether that is a class in a university of college, employee training, or a massive online open course. Cerego is designed to support and enhance, but not replace, the important role of the instructor in such a setting. The platform achieves this in part by promoting effective learning, frequent engagement and efficient building of foundational knowledge for users, but it also provides a number of tools specifically designed to support instructors, outlined in the following section.

a

b

Figure 7: Memory Bank screens. In each screen, as with the progress screen, individual orbs correspond to the memories being learned, and can be individually selected to bring up more details and allow users to explore their own learning history and strategy. a) Upcoming visualizes a user's review schedule for the set. b) Difficulty categorizes items by how challenging they are for the user.

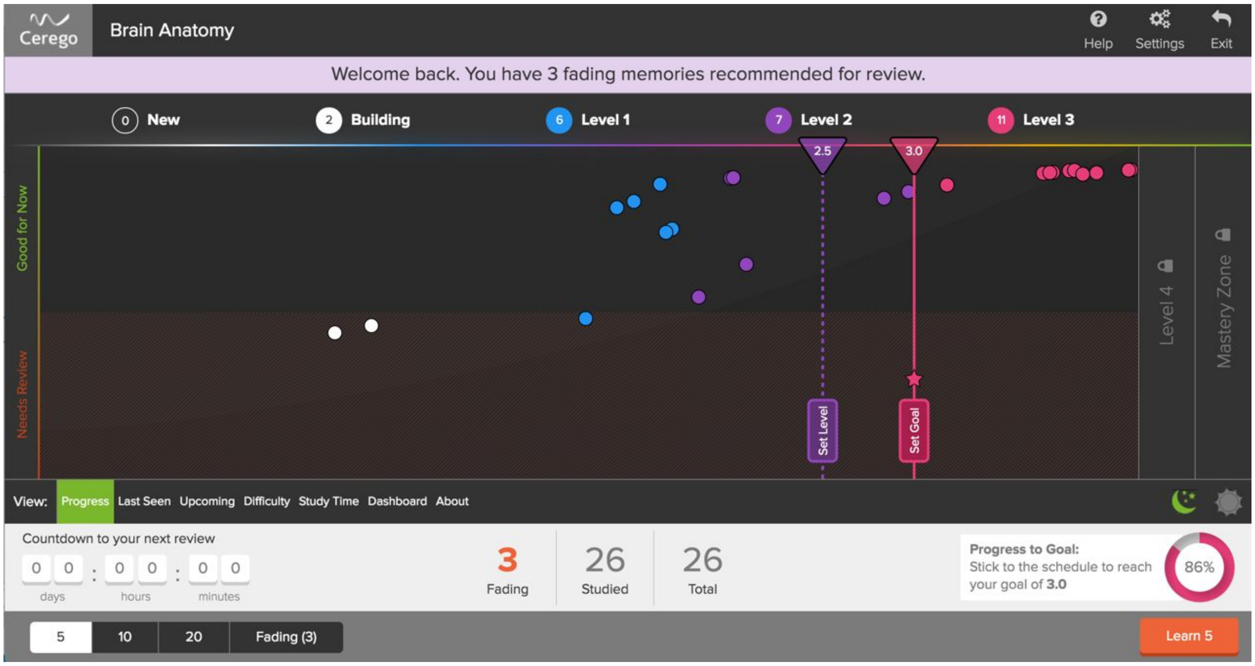

Figure 8: Memory Bank for a set with a goal assigned. Users reach the goal when their Set Level matches or exceed the goal target. Set levels are an average of item retention levels for the set, so to reach a goal of 3.0 a user would review the material until the items were just entering level 3 on average; corresponding to knowledge that lasts several weeks between reviews.

Tracking the Learning of Others: Instructor Analytics

While Cerego is widely used for independent, self-directed learning, a very common context for learning involves an instructor-learner relationship, which can vary all the way from one-on-one teaching to the co-ordination of a massive open online course. The Cerego platform is designed with this relationship in mind, and incorporates a number of tools designed to support instructors in different contexts.

Cerego supports the creation and administration of groups of learners with common content and learning goals, such as the students in a class, employees in a training program, or learners in an online course. An admin for a group in Cerego has access to a wide range of tools for viewing and understand their learners' progress, setting and updating learning goals, and tracking the quality and effectiveness of their content.

A key concept for learning in groups is that of progress towards a goal. Group admins are able to set a retention level for each set of material as a goal for their learners to reach, which is then made visible to the individual learners on their memory bank (*Figure 8*). Goals can be set - and adjusted - to any level the group administrator desires, but a converter is provided when setting the goal that gives an estimated length of time for completion (derived from empirical Cerego data on previous learning rates), and helps to guide appropriate goal selection (*Figure 9*). Progress towards this goal is reported from Cerego as a percentage, making it understandable and convenient for use in education, and allowing integration with a Learning Management System.

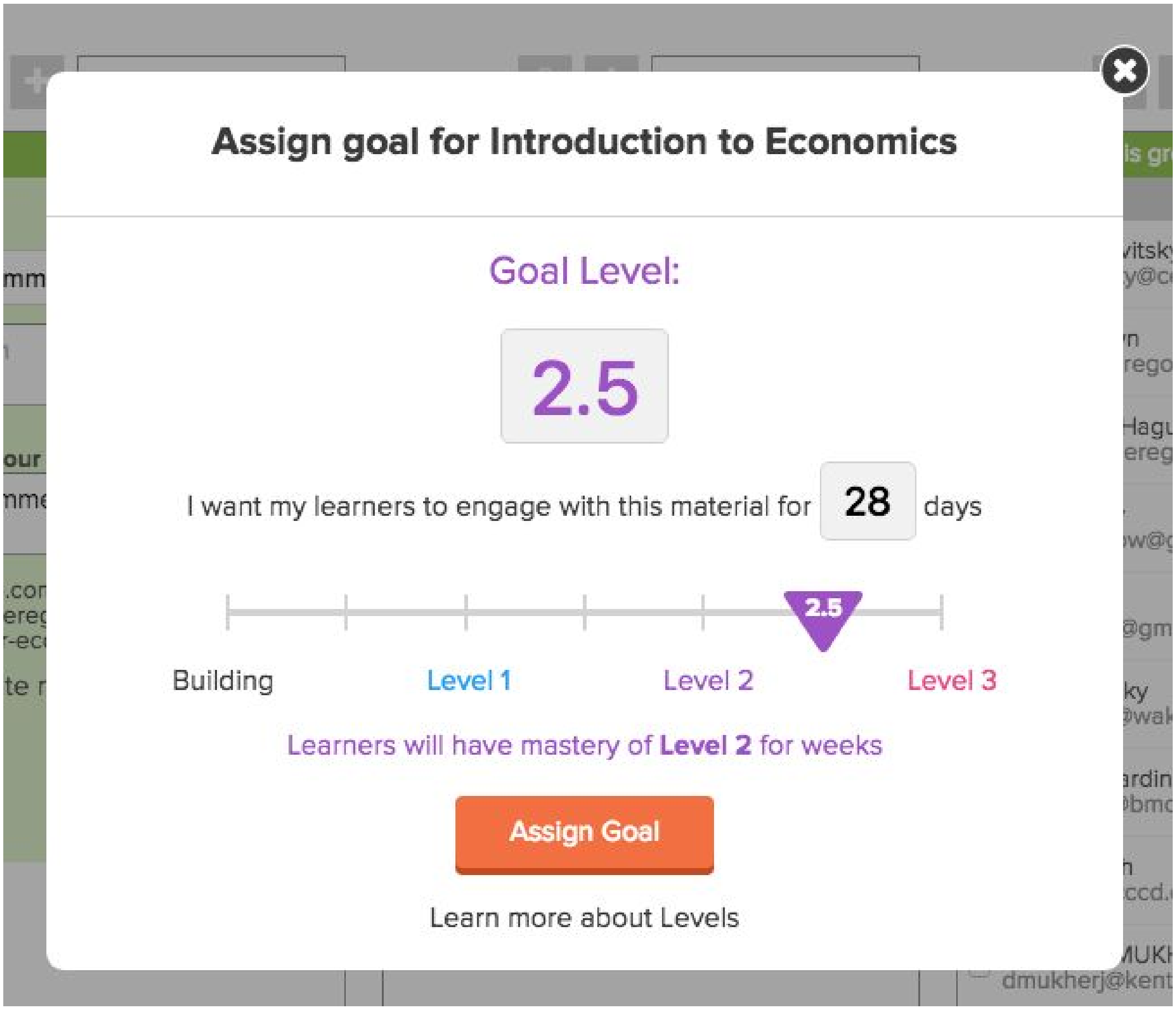

Figure 9: Goal setting tool. This tool allows instructors to define the period of time their students should be actively learning and engaging with a given set of material, such as the length of the relevant course module, and converts this length of time into a realistic retention level goal to set for students. Students will be able to continue learning the material more deeply after reaching the goal, if they wish.

The simplest possible group management involves setting a goal, and tracking the final progress score for each learner. Doing so can reduce the emphasis on single end-of-course assessments, promote consistent engagement with course materials, and free up instructor time for teaching by shifting both assessment and the learning of foundational concepts into an automatic and ongoing process.

Cerego, however, provides a powerful suite of analytics to help instructors and course admins dive deeper into their learners' progress and engagement, and to draw actionable conclusions about course material or individual learners that would benefit from additional instruction. The memory bank view is available to group admins for each set of material their learners are studying, allowing a quick visual overview of the progress of the group (*Figure 10*). From here group admins can also zoom in on individual users, giving them access to the same item-level view of the retention and activation as they would have for their own memories.

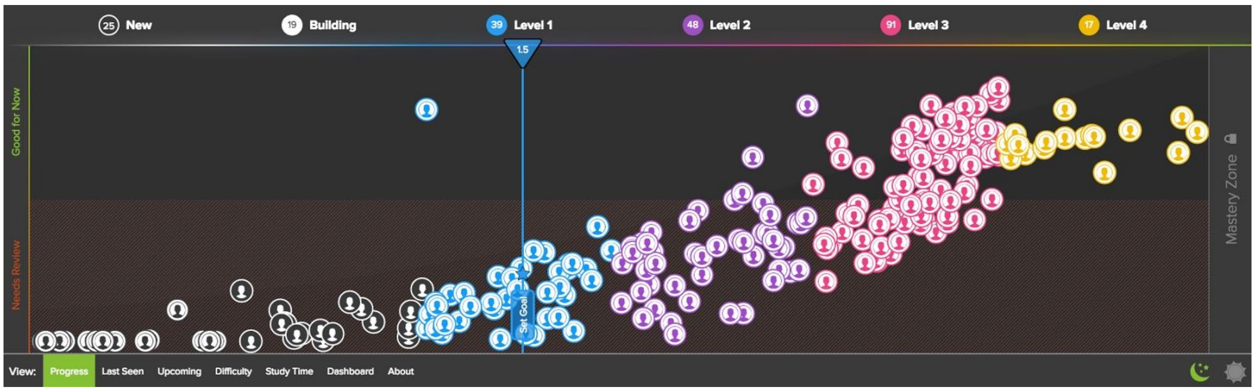

Figure 10: Group memory bank view. From here, instructors can visualize progress for each set of material they are teaching, and break out individual learners in item-level detail.

In order to track and address specific aspects of their course, instructors also have access to a rage of deeper quantitative analytics by running reports. These can be used to view metrics across learners, either within specific sets of material or over the course as a whole, and the data can be saved or exported. Metrics available include:

- **Distribution** of each learner's memories across retention levels
- **Proportion** of assigned items that have been studied by each learner
- Total **study time** within Cerego, plus a detailed breakdown by day, set and learner
- Goal assigned and **percentage progress** towards it
- **Changes** in progress across time

Instructors can also view content analytics, such as the number of reviews or the mean accuracy associated with each item of content in the course, as a way of tracking content quality and difficulty, and alerting them to areas of the course that require additional instruction.

Tracking Learning on the Go: Insights

As personalized adaptive learning tools make learning more efficient and less constrained to the classroom, the role of an expert or instructor can become more specialized and targeted. By reducing the amount of valuable instructor time spent teaching and reviewing fundamental concepts, more time is made available to focus on higher-level learning outcomes such as application and analysis, working through examples, or introducing richer and more demanding content. As so much of the learning experience moves out of the classroom, however, it becomes increasingly important that instructors are still able to rapidly identify the students, or content, that most require their support and attention.

A key tool aimed at solving this tension for instructors is the Cerego Insights app on iOS (Figure 11). Using Insights, an instructor or course admin can view the progress of each of their learners in detail, including their progress towards the assigned goal, breakdown of memories by retention level, total study time, and the interval since their last interaction with the each subset of content. Learners can also be quickly sorted by progress, name or total study time, and instructors can send a learner an email directly from the app.

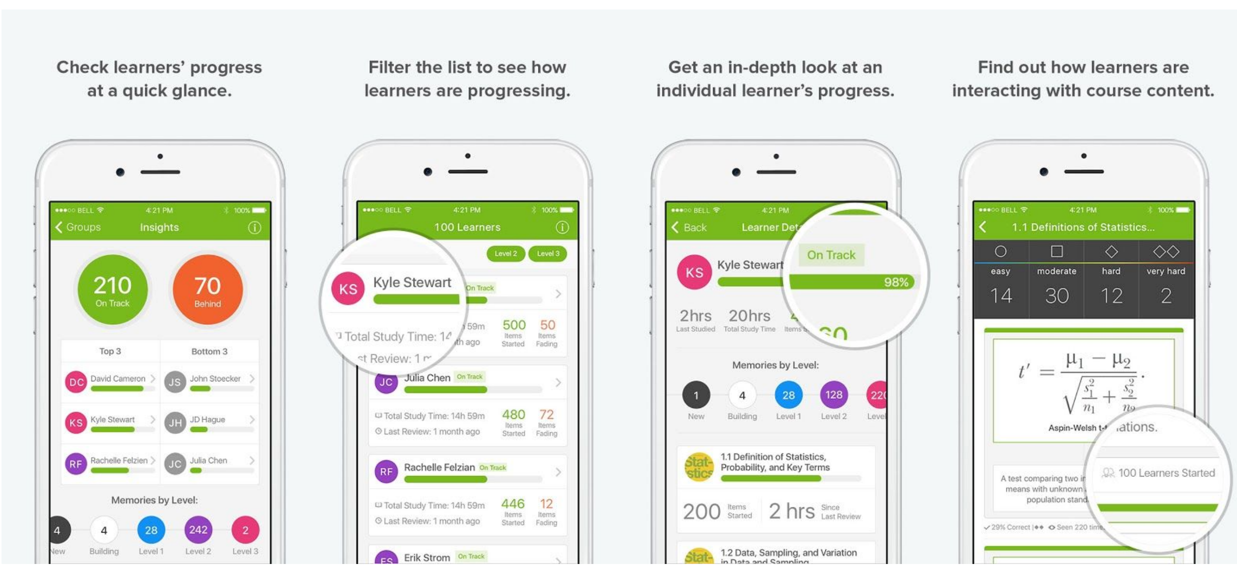

Figure 11: The Insights mobile app. Instructors have access to suite of tools and metrics focused on identifying users and content who need additional attention.

Perhaps most importantly, the Insights app automatically calculates a single categorical "On-track/Behind" metric for every learner, designed specifically to identify those learners most at risk of failing to reach their learning objective. The metric is based on the learner's interactions with the content including progress towards the assigned goal, recency and frequency of studying, and review accuracy. This provides instructors with a simple, automatic and convenient way of identifying and, if they choose, directly contacting students who may be struggling or falling behind their peers.

Insights can also be used to track the content in a course. The number of views, average accuracy and difficulty of every item can be checked separately, providing a quick way of identifying gaps in student knowledge and adapting or improving teaching materials in a targeted way.

Structured Content

Cerego provides a platform to efficiently build lasting foundational knowledge. Importantly, knowledge can be represented by many structures. Expanding the range of types of content that can be learned adaptively is a key step towards translating learning science from the laboratory to real life. Cerego supports the creation and customization of a wide variety of content types, modalities and quizzing methods, to help users build flexible and transferable knowledge.

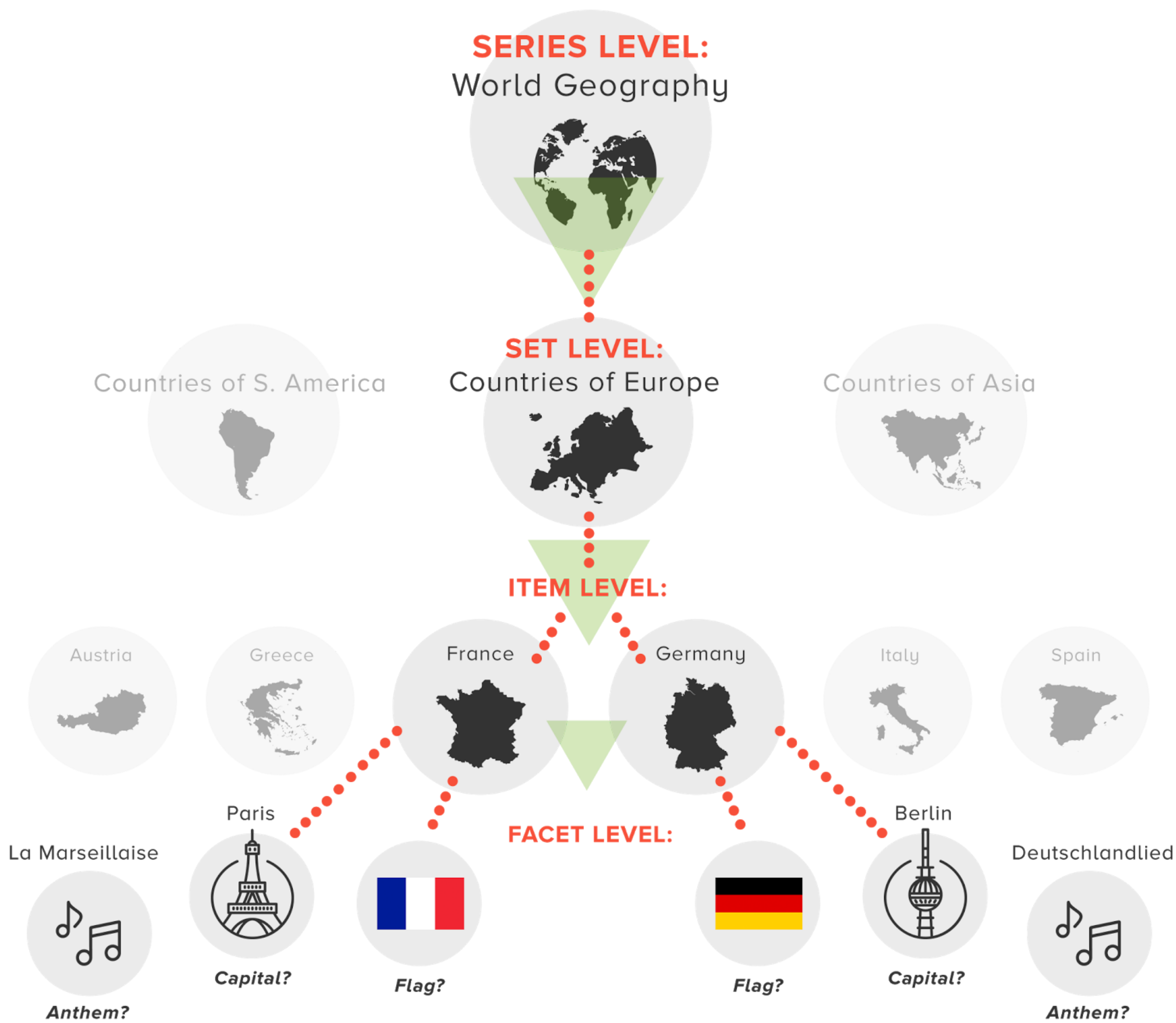

Figure 12: An example of hierarchical content in Cerego. Sets are groups of related items, each of which is built from one or more facets. Facets are the associations, patterns and concepts directly learned and tracked in Cerego, and can be quizzed in multiple ways.

Knowledge can also be thought of as hierarchical: core related concepts leading to higher level understanding. Cerego's approach to structuring content is similarly hierarchical (*Figure 12*).

The smallest piece of information - learned, tracked and reviewed as a single memory - is known as a *facet*. An example of this might be that "**Paris** is the capital of **France**". In this structure the relationship between Paris and France - the fact they are associated and the nature of this association, one being the capital city of the other - is one memory to be learned.

Facets are the individual components of a memory *item*. An item may have one single facet, or it may have many. For example, an item might be "France", and the facets might include the capital city (Paris), the population (66 million), the national anthem (La Marseillaise as an audio clip and lyrics) and the national flag (an image of the Tricolore). Each of these facets may be learned, forgotten or reviewed independently, but they are linked together into an item that represents basic knowledge about the country of France. Similarly, a function in a coding language might be built as an item, with the description, function arguments, or usage examples each built as a facet. Each item can also be enriched with *notes*: additional information on a separate tab that supports learning or provides context, and is always available to the learner. Notes are very flexible and might include videos, broader context, further reading, references or lecture slides.

Items are in turn grouped into *sets*, where each set represents a higher level grouping of related knowledge, that might for example map to a book chapter or a toolbox for a coding language. In our example above, we could extend the set by adding similar items for other countries, such as Germany, Spain and Belgium, to build a set themed around European countries.

The highest-level grouping in Cerego is the *series*. A series consists of a number of sets, generally arranged in a specific learning order, and often maps to a larger body of knowledge such as a course or a book. By structuring content in this way, foundational concepts can be built into - and subsequently tracked and measured in terms of - higher level knowledge.

Content Types

A key objective of learning is knowledge that is **transferable**, that is to say, that can be extended from one context to another [*Byrnes, 1996*]. In some sense this property can be thought of as a key difference between **education**, the preparation of students for a complex and changing outside world, and **training** in a specific and more limited set of skills [*Broudy, 1977*]. Transfer and abstraction of knowledge is also critical for building foundational concepts into higher levels of understanding, analysis and interpretation [*Bransford et al., 2000*].

One important way in which Cerego is designed to support development of these higher level functions is by building long-term retention for a broad base of foundational knowledge, meaning that students have simultaneous access to a wide range of concepts rather than a narrower set of recently-learned material (gained for example through cramming, and forgotten shortly afterwards). Access to this wide and consistently active base of knowledge, rather than a narrow subset, allows for the inferential, combinatorial and relational reasoning essential to building abstracted and deeper understanding. Strongly-learned foundational knowledge is in itself necessary for transferable or flexible knowledge to be built.

A second, very direct way in which Cerego supports the learning of transferable information is through variation and complexity in both the structure of content, and the ways in which it is reviewed and retrieved. By varying the retrieval experience and context, and by representing learning concepts in multiple different structures, the resulting learning should be less strongly tied to one specific context, leading to better abstraction and transfer. The concept of transfer may be one factor in why a **desirable difficulty** exists for learning reviews: retrieving and applying information in a new context may be more difficult in the moment, even while the knowledge being reviewed is being generalized and abstracted.

To build varied and abstractable foundational knowledge, Cerego allows for a range of content types and structures, including an explicitly abstracted item type, *patterns*, designed to promote transferable and flexible learning directly. Here we outline some examples; for greater detail on the different content types supported by Cerego and how to create them see the content creation guidelines at cerego.com/content-creation-guide.

Associations

Associations are the simplest item type available in Cerego. Each item has an anchor - a core concept - and any number of linked facets, or associations (*Figure 13*). Each of the associated facets relates some piece of information to the anchor; such as a definition, a fact, an image, or a sound clip. Each of these facets is tracked as a separate memory, can be tested bi-directionally and in multiple ways, and can be labeled with its own custom question.

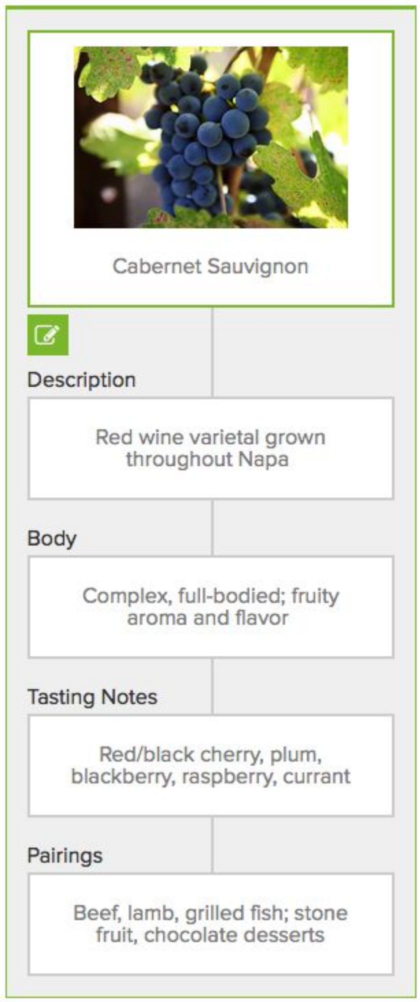

Figure 13: An example of an association item in Cerego. The core concept (Cabernet Sauvignon) can be paired with any number of labeled facets, which each form a memory that is learned, tracked and reviewed separately. Both the anchor and associated elements can include images, audio clips, rich text (including formulae) or any combination.

Languages

Cerego's first adaptive learning product was originally developed for language learning, and language or vocabulary items include a number of features specifically helpful for learning languages. Each anchor is a word, and is associated with a definition, making it in some sense analogous to a single-facet association item. In addition, however, language items can be tagged with the part of speech (e.g. noun, verb), associated with a pronunciation audio file, given a transliteration for non-Latin languages, and tested in context using example sentences.

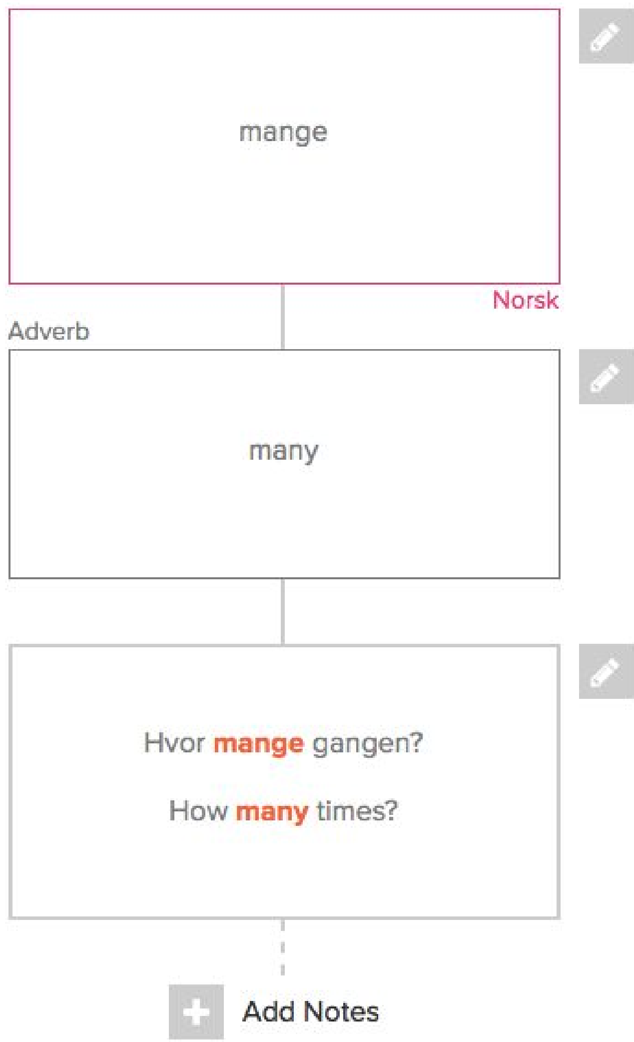

Figure 14: A Norwegian-English vocabulary item being created. Included are the Norwegian word, the English meaning, the part of speech and a sample sentence with translation. The item could also include audio pronunciation files, images, additional sentences, notes, and in the case of non-Latin languages a transliteration into Latin characters.

Passages

Passages of text are well-suited for building conceptual knowledge, and can also serve as a variation of vocabulary assessment. The learnable facets in a passage are important words or phrases that employ context to drive relational learning.

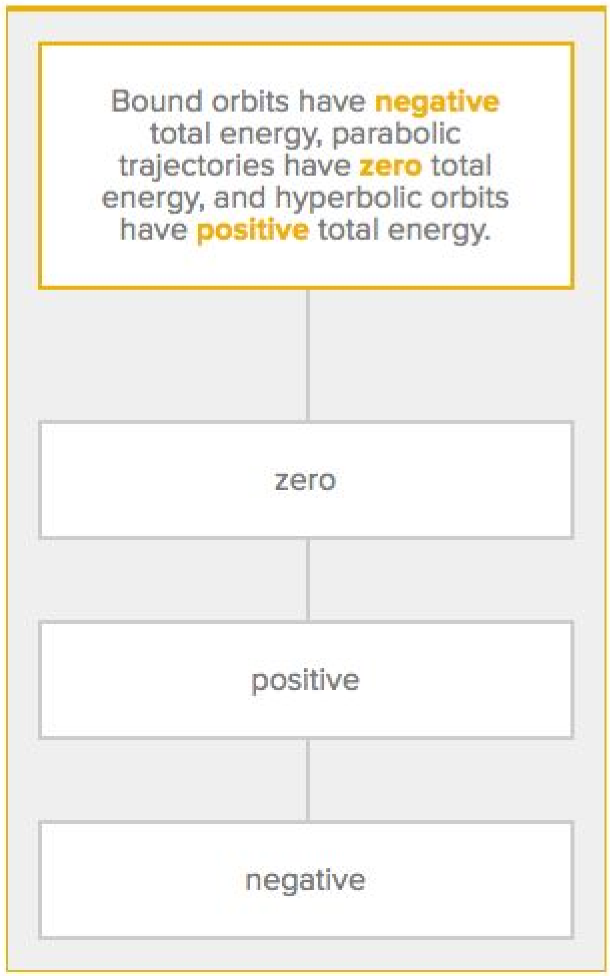

Figure 15: A passage item with three facets. Each facet is learned within the larger context of the passage structure, and in relation to the other facets.

Regions

Region items are a very flexible visual way of learning a complex system. Maps lend themselves well to region items, but anything that can be represented using a spatial layout or as elements in a larger image can be represented in this way: for example charts and graphs, anatomical diagrams, or flowcharts and systems. Within this larger structure, each identifiable element is a facet; region items are analogous to an association item with a spatial structure. Regions are also an excellent way of learning visual differences between similar facets, such as mathematical functions, painting techniques or bird species, by presenting each as a separate labeled region on the same larger image.

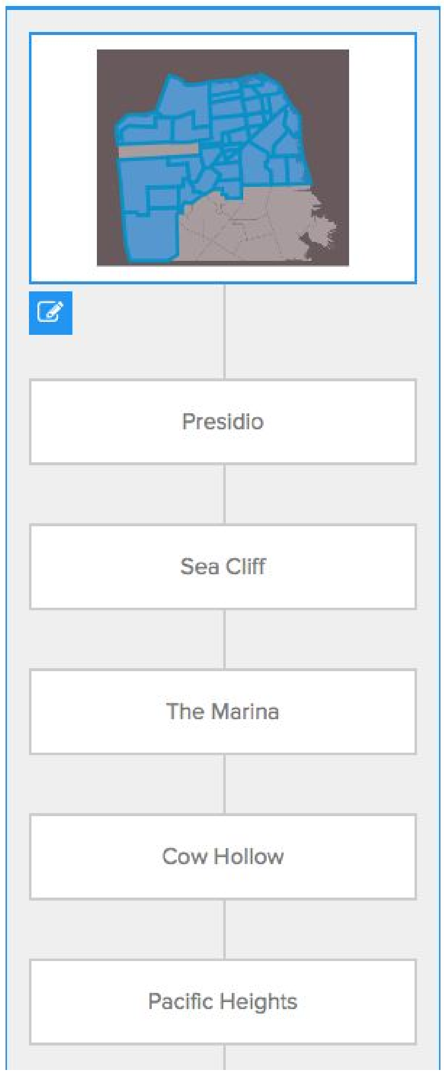

Figure 16: Part of a region item used for learning the different areas of the San Francisco peninsula. As with most items in Cerego, notes can be attached to add context to the information being learned.

Sequences

Region items allow spatially structured knowledge to be learned within Cerego; similarly, sequences enable knowledge to be embedded in a temporal structure. Sequences can be used to represent historical events, recipes, methods, techniques and procedures. Each facet has a position in an overall sequence order.

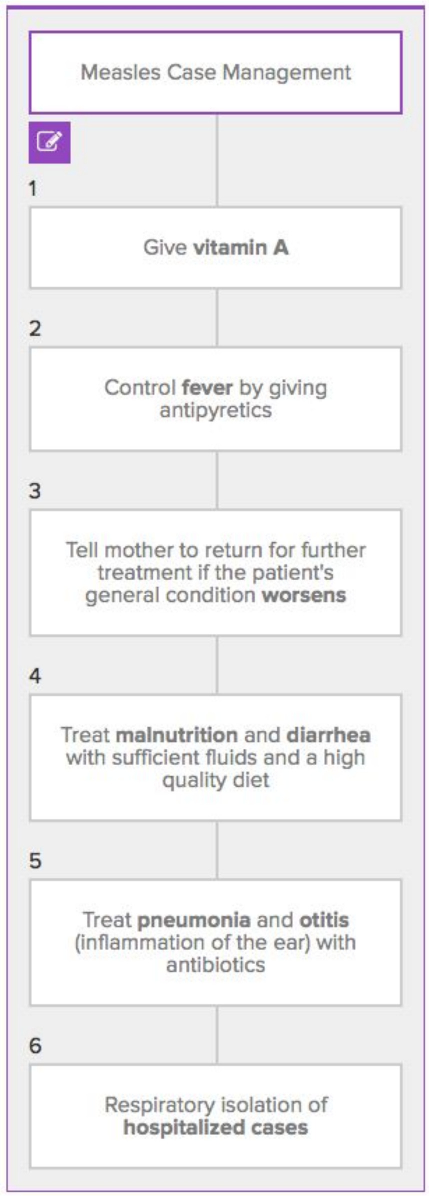

Figure 17: An example of a sequence item, in this case procedural instructions for treating measles cases. Learners review the material by dragging the particular step into its correct position in the overall sequence.

Patterns

Patterns are a powerful item type that allow for learning of an association that shows variability. Examples might include identification of a bird or plant species, classification of clouds or architectural styles, or diagnosis from a list of symptoms or medical scan. In each of these cases, the learning objective is not to associate a single discrete pair of items, but to be able to classify, identify or diagnose a pattern - a crucial skill in the real world. In a pattern item, each core concept or anchor is associated with many possible examples, but unlike associations all of these anchor-example relations are bound to a single memory, linking the core concept with the shared features of those examples. Patterns are ideally suited for training recognition and classification and building flexible, transferable knowledge.

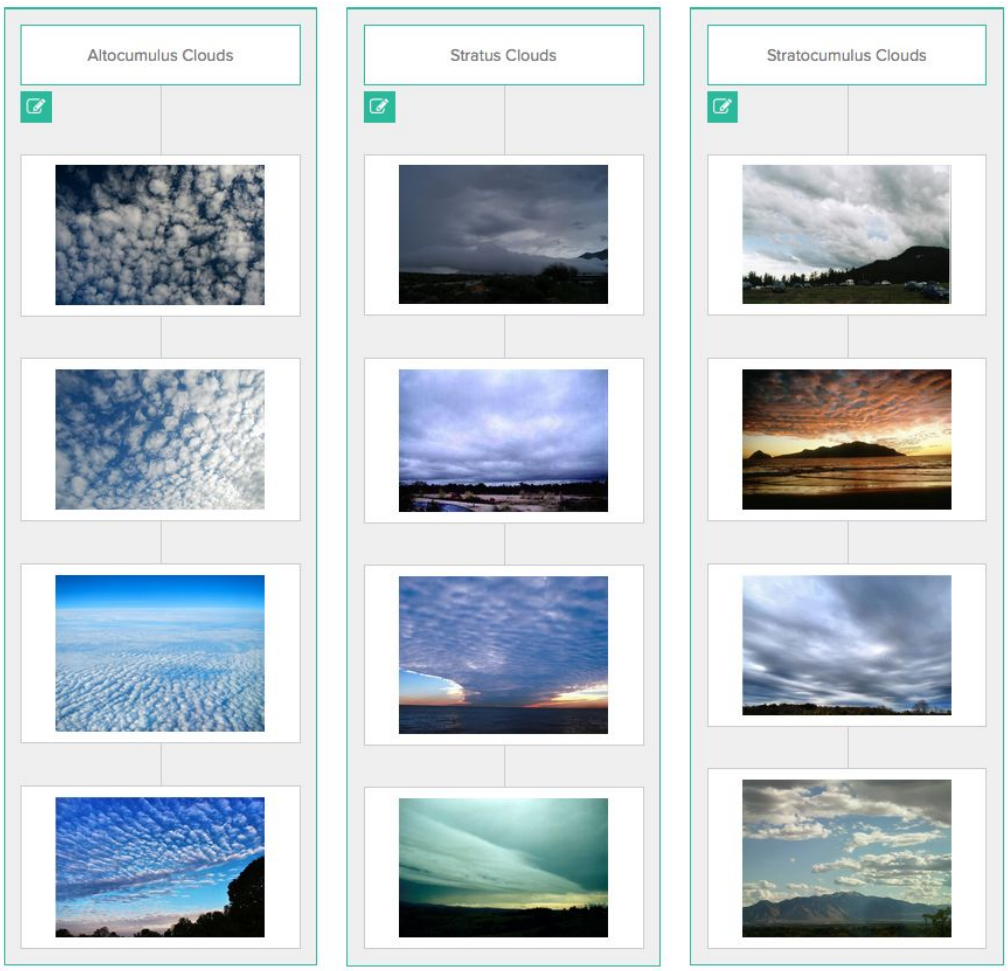

Figure 18: Three pattern items used to learn cloud identification. Users learn to identify and classify the images into each type, and are reviewed on random examples (which they may have never seen before).

Quiz Types

Matching the structure of content to the structure of the knowledge being learned helps bridge the gap between the foundational learning materials and their application in the real world. Another key aspect of content presentation, critical to building deeper, more flexible and transferable knowledge, is variety in the type of review experience [*Bransford et al., 2000; Gick & Holyoak, 1983*]. Cerego supports more than a dozen different test types across the range of content, ranging from simple multiple choice responses to sentence completion, region identification, sequence re-ordering, or fill-in-the-blank. The same individual memory can be tested in multiple ways on different reviews, varying the experience and helping to promote transferable and flexible learning.

Facets can also, in most cases, be tested bi-directionally: That is, an anchor can provide a cue to retrieve an association, or the reverse. While the system will automatically generate quizzes based on the information structured into item templates, content creators can also select the types of test to be provided for each facet at the creation stage, allowing them significant control over the learning experience.

For multiple choice testing, each facet being tested requires a minimum number of appropriate distractor facets to provide a suitably challenging review. Content creators can define these distractors directly, giving them even closer control over the review experience for learners.

Alternatively, however, the content creation tool also provides intelligent distractor generation. With this option selected, distractors will be selected for each facet from answers among all the other facets in the set. These distractors are not selected at random, but on the basis of their similarity to the correct facet, as measured by lexical distance, media type, creator-assigned label and other aspects. For example, when generating distractors for a multiple choice testing knowledge of the French flag, the algorithm will know to select the German flag as a distractor rather than the German national anthem. This is important, since highly dissimilar distractors could lead to excessively easy reviews and miss out on the desirable difficulties most helpful to long-term retention.

Further, the actual distractors used in a given question are drawn randomly from a larger pool of appropriate distractors, meaning that even within the multiple choice question format the context at retrieval will vary.

The learning engine takes the difficulty of a particular quiz type into account when updating the memory estimates for each facet. If a user successfully recalls and types out the correct answer in a fill-in-the-blank quiz, the learning engine will generally interpret that as stronger evidence of memory than if the same answer was provided from a multiple choice test, and will increment the memory retention for that user and facet by a greater amount.

Ultimately, content in Cerego generally benefits from a greater variety in quiz types. Variety in review context makes learning more transferable and deepens interaction by requiring different approaches to the learned information at reviews.

Content Creation & Remixing

In Cerego, any user can create their own entirely personalized content for learning, or adapt and customize content created by others. Besides allowing Cerego's learn engine to help users learn any content they wish, content creation may help learning in and of itself. Specifically, the creation or customization of content generally involves a deeper and more thoughtful interaction with the to-be-learned material than studying does, especially if the creation involves generating custom questions and distractors, or abstracting concepts from source material into the item types in Cerego.

The ability to build, edit and learn bespoke sets of material using the Cerego system is also integral to the aim of personalizing learning, and making efficient, adaptive learning strategies widely available. Whether a suite of programming commands, the names and faces of colleagues at a new job, or carefully chosen phrases ahead of a foreign trip, defining a custom learning objective may enhance intrinsic motivation [*Bransford et. al, 2000*].

Cerego also allows for the remixing and personalization of existing content. A user can fork a set they have access to, creating copies of the items which they can edit, adjust and personalize to better fit their own learning objectives. Instructors and learners can use remixing to create private versions of public sets; remove, add or or update content from a standard set to better fit their own preference; derive subsets of material from a larger 'parent' set, for example to give more flexibility to individual instructors; or simply speed up the creation of their own content by starting from an existing framework.

When new items are cloned from old ones using remixing, memory estimates are copied across to the new item, so learners don't lose any memory progress and begin the new item at the appropriate level of retention. Any set of created material can also be maintained for private access only, or made available publicly. A full guide to content creation can be found at cerego.com/content-creation-guide.

Integration

Adaptive learning tools have the potential to shift evidence-backed learning strategies out from the research lab and into mainstream use. Discussed extensively in this white paper are the principles Cerego draw from the learning science literature, and the approach taken to implement them in practice for learners. Just as important as making these principles work, however, is making them available and practical for users in the real world.

In this section, consequently, the important issue of integration is discussed. Integration in this context means, firstly, integration into daily life and routines, through scheduling and planning tools that gives users control and ownership over their learning time investment, to the availability of mobile apps to make the promise of distributed learning available and realistic in practice. Secondly, Cerego exists as part of a broader ecosystem of increasingly specialized educational tools, and is therefore constructed to allow integration with content providers, learning management systems, or indeed to be linked and embedded in any wider system. An example of this principle in action, in the form of the Bookshelf GPS collaboration with VitalSource Technologies, is outlined.

Scheduling

The Cerego method helps users to devote all their available cognitive workload to learning by optimally scheduling the presentation of items due for review. The principle underlying the scheduling algorithm is expanded spaced repetition, which adapts to each user as they interact with the material they are learning.

Scheduling involves more than simply tracking memory decay, however. Users can elect to receive emails and mobile notifications when they have a number of fading memories ready for review, and these are provided at a moderate frequency designed to integrate learning with daily life. Cerego also provides a “Review All” option for learners to review the most necessary fading memories from across multiple sets, making it easier to keep on top of all of their learning objectives while also interleaving reviews from different contexts to enhance the **desirable difficulty** of the retrieval.

While Cerego handles the scheduling of reviews for each learner’s items, it also provides a tool for planning a longer-term learning schedule: the Workload Calculator. The adaptive distributed learning approach employed by Cerego build lasting memories efficiently, but the learning effort required for each memory is frontloaded since the interval between reviews expands as memory retention is added. This means that an effective high-level schedule for learning should generally distribute the studying of new material across time, adding new items as the previously studied items move into higher retention and require less and less maintenance.

Using the workload calculator, a learner can plan out their longer-term learning schedule in this way (*Figure 19*). Learners define their learning objective (e.g. 180 items) and a preferred amount of available study time per day (e.g. 30 mins). The workload calculator will use these parameters to generate a personalized study plan for the learner that introduces new content at an appropriate pace, keeping the daily study time below the user’s chosen level.

An instructor can also use the workload calculator to set appropriate goals for their students, or plan out the introduction of new material in a course. A dedicated tool for choosing appropriate retention goals for a given course length is also provided for instructors and course administrators.

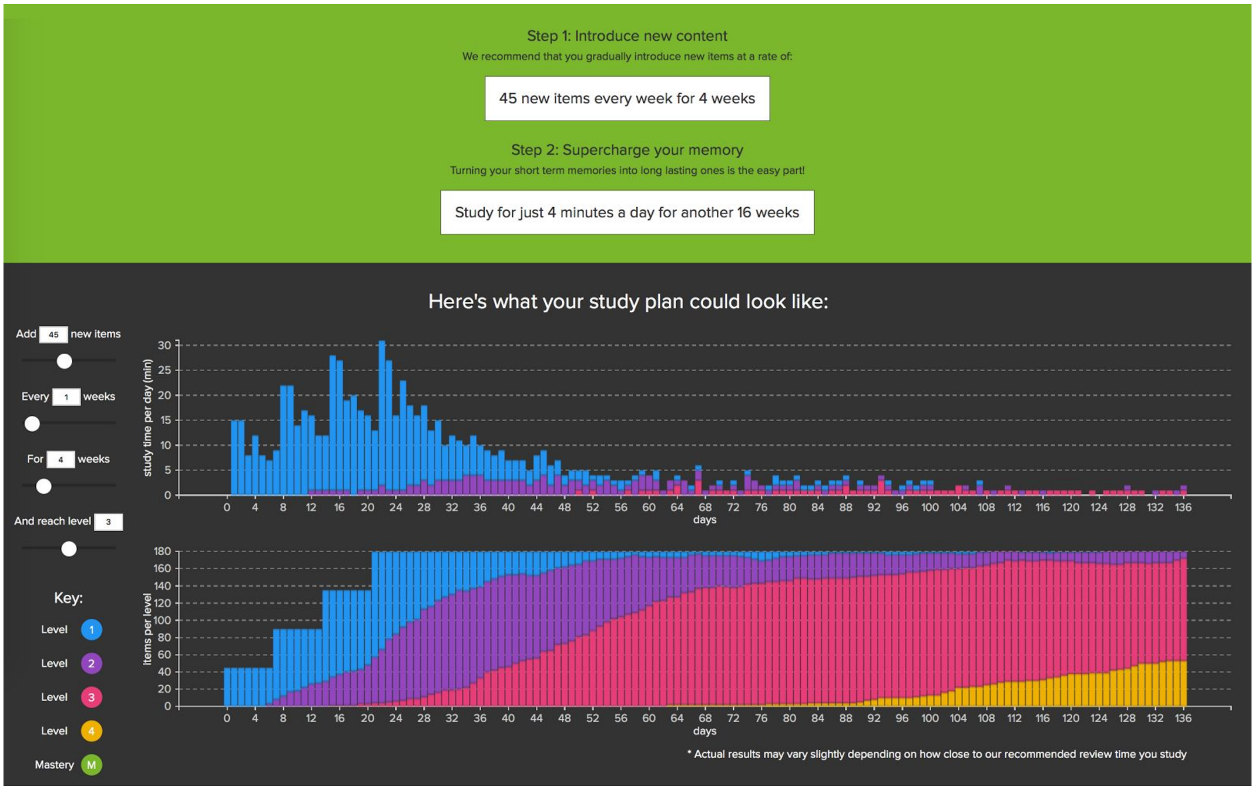

Figure 19: The workload calculator can be used to develop a personalized study plan that fits within the daily study time available to a learner. The output shows the expected daily time required to learn and review all of the chosen material, and the expected progress from unstarted items to high-retention, robust long term memories. Note that after the initial effort learning, efficiently building retention for the material means very little effort is then required to maintain and deepen the memories.

Platforms & Availability

Cerego aims to make learning easier and more efficient by distributing short reviews optimally across time. As learning moves out of the classroom and the library, and can be incorporated into daily life, the wide availability of mobile devices makes this mode of learning increasingly practical. Cerego provides mobile apps on iOS and Android to support convenient learning and review, including offline study when an internet connection is unavailable. The Insights app offers further mobile availability to instructors and content creators.

Cerego can also be linked closely to learning materials and other platforms via APIs to the learning engine, or full LTI integration. The platform is therefore flexibly integrated with primary learning materials and other forms of assessment: An entire interactive course can be built directly in the Cerego learn app using notes to embed videos, slides, images, audio or text; alternatively the entire source material and grading output can take place outside of Cerego with only the core adaptive learning experience added to this existing course. In the next section we describe one example of the Cerego platform being directly integrated with a set of online learning materials: Bookshelf GPS.

Integrated Learning: Cerego and Bookshelf GPS

An important challenge for learning applications is establishing a meaningful link between primary content and the learning experience. An example of how Cerego seeks to bridge that gap can be found in Bookshelf GPS, a partnership between Cerego and VitalSource Technologies, providers of the world's largest e-textbook platform with over one million publications from 750 publishers and 12 million learners worldwide. With Bookshelf GPS the Cerego learning experience is embedded in an online textbook, and directly supports the creation, learning and measurement of foundational knowledge from this primary source material.

Significantly, customized Cerego content can be created directly and easily from the textbook by an instructor (*Figure 20*) and then presented and learned in the same environment by a student. Instead of simply reading the material and answering questions later, the learning objectives for each section are built into a Cerego set linked directly to that section of the book.

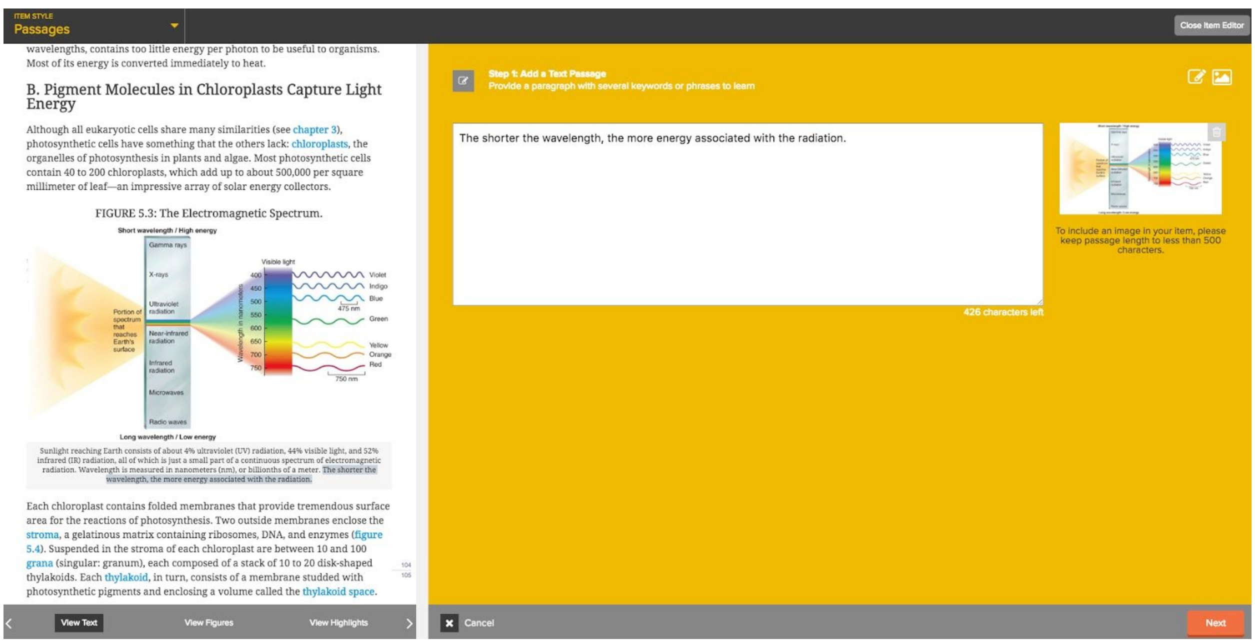

Figure 20: A Cerego adaptive learning item being generated from the associated section in the primary material. Content creators can quickly build items from the text, figures, and automatically generated highlights of the textbook, and create sets that correspond to - and appear in - the relevant chapter. Students can learn or review individual sets, or the entire book, and track their learning for each section.

Cerego items are created by copying key words, ideas and facts into item templates. Bookshelf GPS automatically generates assessments to help students learn and review the information, in the same environment that it was learned and created.

Learners access both the primary content and the Cerego content the same way, through Bookshelf. By entering GPS, learners will have access to all of the Cerego sets created, aligned with the table of contents for the book. By embedding a Cerego learning experience into the book, students can learn the foundational concepts more effectively than by only reading the material, and also keep track of their own learning progress.

Instructors in Bookshelf GPS have access to the detailed reports, metrics and visualizations described earlier in this paper, allowing them to view each student's learning experience, export detailed content-specific breakdowns of student progress for grading, or simply to better understand how the material and the Cerego items they created are being used.

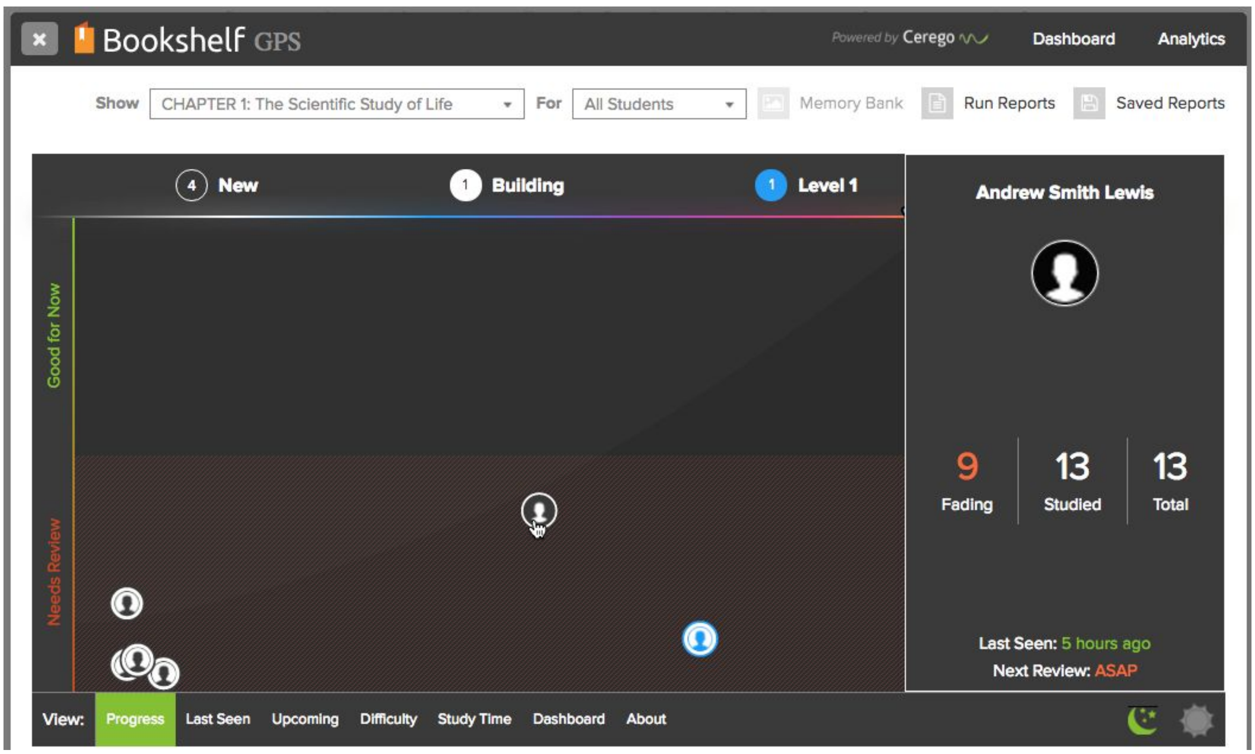

Figure 21: From Bookshelf GPS, an instructor can view detailed information about their learners and content. Here the instructor has hovered over a learner icon, and a summary card for "Andrew Smith Lewis" appears on the right. Clicking on the learner icon opens the learner's memory bank (Figure 8).

Internal Research

At its most fundamental, Cerego implements core principles from the learning science literature to help people learn effectively in the real world. Durable memories are built efficiently using a process of **distributed learning**, strengthened by actively engaging in **retrieval practice**, and adaptively scheduled by tracking and **predicting the activation** of each memory into the future.

How do these principles support learning in practice? Cerego is an evolving and empirically-driven platform. In this section three brief snapshots are presented from ongoing internal research at Cerego, demonstrating how each principle works in the real world. All of the anonymized data used in the following three sections are drawn from newly signed up Cerego users, who took part in courses between Fall 2015 and Spring 2016. In particular, the studies provide direct evidence that:

- The adaptive learning engine **predictions** of memory activation accurately reflect the actual **review difficulty**, which is crucial for review schedules to be optimized for desirable difficulty.

- Reviewing each memory close to the time suggested by Cerego leads to **better learning**, supporting both the principle of distributed learning as well as the effectiveness of Cerego's specific **schedule**.

- Longer time spent trying to **recall** answers before making a response leads to **higher accuracy on later trials**, supporting the long-term retention advantage of deeper metacognitive engagement and effortful retrieval when learning.

Tracking Memory Accurately

In order to effectively implement distributed learning, and adapt an individual learner's schedule to review items at the optimal difficulty, **it is critical to accurately predict the difficulty of a memory in a future test**. The Cerego learning engine does so by combining a learner's previous performance for each item with theoretical memory decay models, to predict the activation of each memory at a given moment in the future.

To test the accuracy of this approach, we examined around 1.7 million reviews in Cerego, and plotted the relationship between Cerego's measure of activation at the moment of review, and the user's actual accuracy on that trial. The two are strongly linearly related; that is **the activation calculated by Cerego is an accurate predictor of the difficulty on a review**.

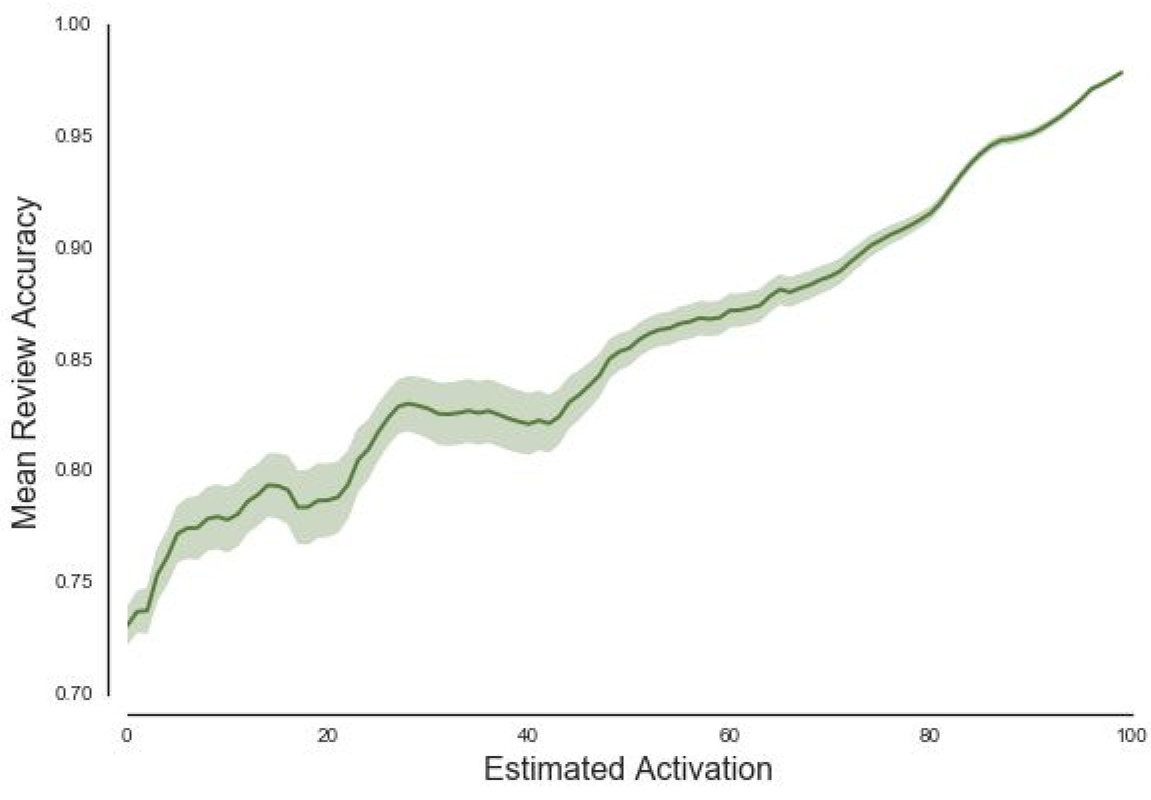

Figure 22: Cerego schedules items based on their estimated activation, with the goal of predicting when reviews should be 'desirably difficult' and therefore lead to better long term retention. It is therefore crucial that the activation value estimated for each memory is, as shown here, an accurate prediction of the relative difficulty of a review. Range shows 95% confidence interval.

Reviewing at Desirable Difficulties

Being able to accurately predict the difficulty of each learner's reviews allows Cerego to adjust review schedules to tune this difficulty to an optimal value for learning: a **desirable difficulty** that is challenging enough to engage effortful retrieval, but not so difficult that the memory has faded and the learner is unable to retrieve it altogether.

Cerego estimates this optimal difficulty from past data, and schedules reviews for when activation fades to this level. Does this lead to more durable memories, and better long-term learning? In other words, **do learners perform better on *future* reviews if they stick to the suggested schedule from Cerego and review when prompted?** We tested this directly by checking whether learners who reviewed close to Cerego's suggested time (had a high **schedule compliance**) showed stronger learning. This was indeed the case:

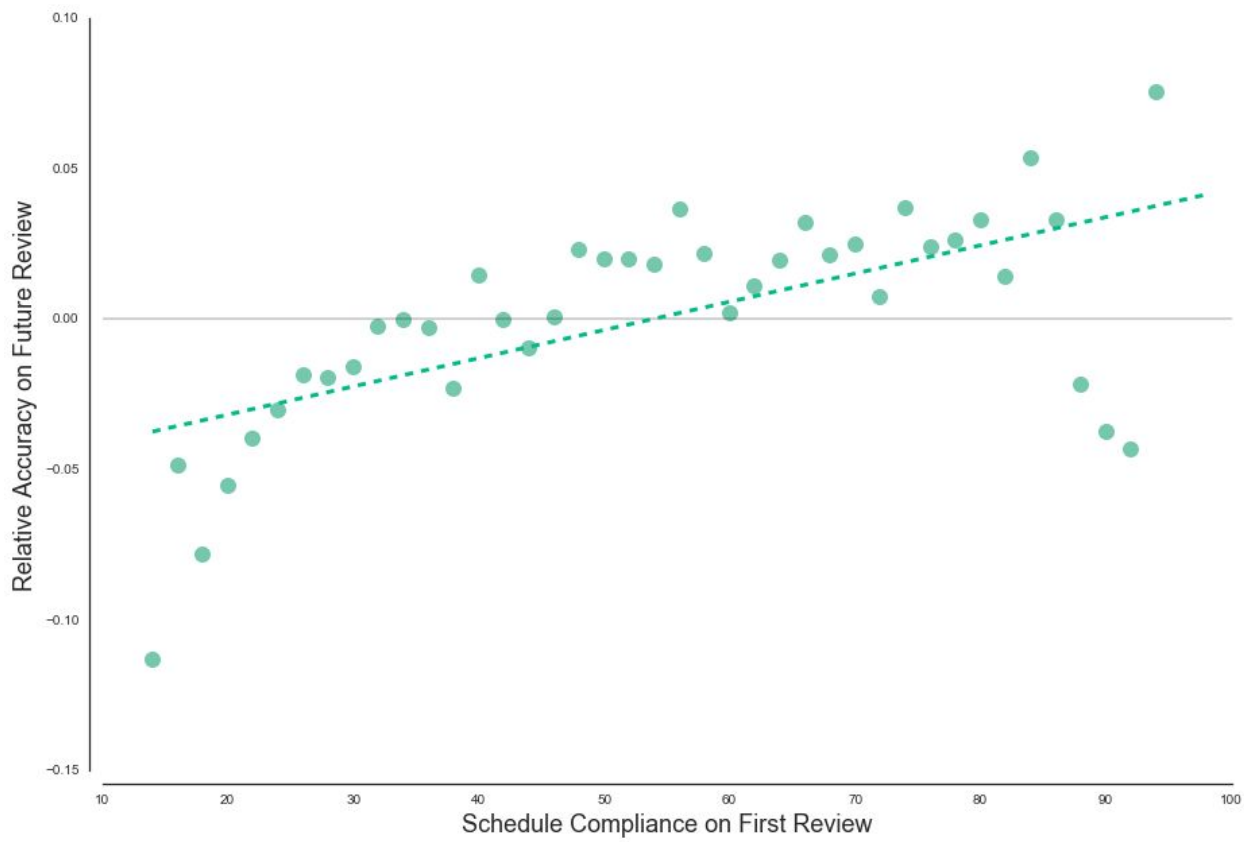

Figure 23: Closer compliance with Cerego's suggested schedule leads to better accuracy on later trials. Relative accuracy is corrected for activation at the moment of review; it reflects the % accuracy on trials relative to the average (so higher numbers mean the material has been better learned). Higher compliance values mean the user reviewed closer to the suggested time for their items. Users are binned according to their compliance on the first review.

This is not simply because users who tended to have high compliance just tended to be the users who performed better in Cerego. We also factored out each user's overall ability in Cerego and ran a multivariate regularized regression model, across users. The predictors were:

- The **schedule compliance**: a metric denoting how close to Cerego's suggested time each user actually came back and reviewed each item. We looked only at the first review for each item.

- The average **accuracy** on the **first review** for each item.

The two predictors (accuracy and compliance on the first review) were then regressed against each user's **accuracy** on **future** reviews for the same items. In this way we can address the question: Does following Cerego's suggested distributed learning schedule lead to better long-term *learning* outcomes, as measured by improvements in accuracy *beyond that expected from the user's initial accuracy*?

In fact they do. Users who stuck more closely to the suggested review schedule for their first reviews showed significantly higher later accuracy for those items, $t(1603) = 6.69; p<0.001$, even after accounting for their activation at later reviews, and their accuracy on the first review [*overall model* $F(2, 1603) = 460.20; p<0.001$]. Including additional predictors in the model such as total number of items studied or median session length did not substantially affect the estimate of the compliance parameter.

Overall this analysis provides strong evidence that **closely following the specific distributed learning schedule in Cerego benefits long-term learning**, compared to not following the schedule or following it less closely.

Retrieval Practice and Engagement

Every review in Cerego centers around an effortful retrieval. The advantage of testing over restudying in the laboratory has been well-established [*Bjork, 1994*] and **deeper and more effortful engagement at retrieval leads to better longer-term learning**. Further, probing and assessing one's own memory (i.e. metacognitive judgments) may also lead to greater long-term retention [*Metcalfe, 2009*].

Consequently, reviews in Cerego always begin with a **recall screen** (*Figure 24*), in which learners are prompted with the review question but have the opportunity to attempt to recall the information before advancing to the quiz screen (where they can make their response). Although all reviews in Cerego involve a retrieval test, we can indirectly assess engagement, memory strength introspection, and recall effort by examining **recall time** - the length of time learners spend on this recall screen for each question before indicating whether they know the answer.

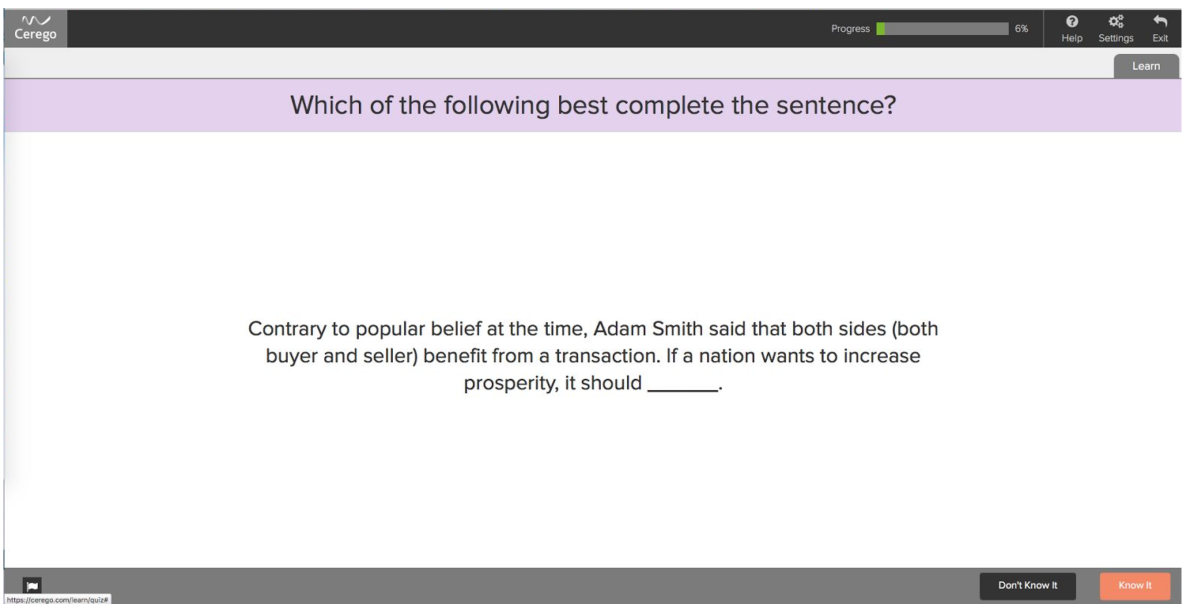

Figure 24: Recall screen. Learners can progress from this screen by pressing "Know It", at which point they are presented with the opportunity to enter a response. The length of time spent on this screen is termed recall time and can be interpreted as a measure of the effort made to retrieve the memory being probed.

In particular, if greater effort and engagement are indeed beneficial for long-term learning, then longer recall time should lead to better future performance on reviews for that item. *Figure 25* shows that this is indeed the case: The improvement in memory performance between the first review and all later reviews as a function of recall time on the first review is present whether or not the first review was accurate.

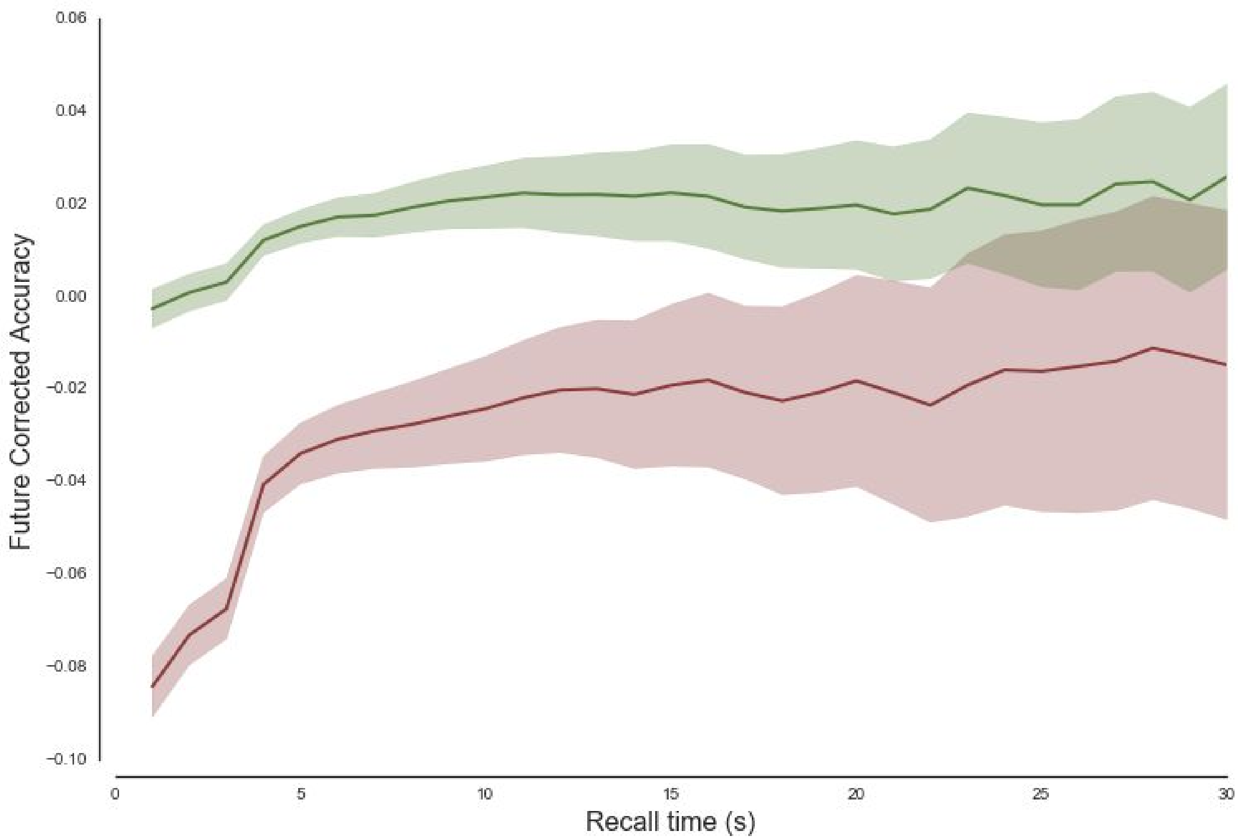

Figure 25: Future review accuracy (corrected for activation) as a function of recall time on the first review. Shown separately for items answered correctly (green; upper line) or incorrectly (red; lower line) on the first review following the recall screen. Whether the response was correct or incorrect, spending longer on the recall screen before answering improved review accuracy on later reviews for the same item.

The improvement in later trials with greater recall time is dramatic, and consistent with a learning advantage for more effortful retrieval. Strikingly, the learning advantage for recall effort exists independently of the outcome of the first review - whether the initial review ends up being correctly or incorrectly answered, **users still performed better on *later* reviews if they spent at least several seconds considering the question** before attempting to answer. Consistent with the learning science literature, promoting effortful retrieval by separating out the recall screen from the answer screen appears to improve learning outcomes in Cerego.

External Research and Case Studies

In this section we summarize a peer-reviewed third-party assessment of Cerego's impact on learning in a high-school environment [*Homer & Plass, 2015*]. We also highlight two example case studies demonstrating the use of Cerego in different learning environments: Dental school students at New York University, and a large-scale online MOOC on Jazz appreciation.

K-12 Peer-reviewed Study

Homer & Plass (*2015*) conducted an effectiveness study of Cerego for five high school subjects. Participants came from 7 schools in a network of charter schools in greater Miami. The primary goal was to determine if students who used the Cerego system had better learning outcomes compared to students in a business as usual control group. The results show the effectiveness of Cerego for learning, summarized below.

In total, 126 classes/sections across 7 schools with 4,111 students participated in the study. Classrooms in participating schools were randomly assigned to treatment conditions, either using
Cerego or not using Cerego ("business as usual" control). Since the use of any other learning software was continued, this means that Cerego + other software was compared to other software alone (i.e., without Cerego).

The participating teachers, with the support of instructional designers and curricular experts, designed Cerego courses for five subjects: Algebra, Biology, Civics, Geometry, and US History. These courses were designed to implement the applicable local standards (Next Generation Sunshine State Standards and Common Core Standards) on these subjects and to be complementary to classroom instruction.

Primary data collected included knowledge pre-tests and knowledge post-tests for each subject using regular end-of-term tests. The pre- and post-tests were the tests administered by schools as part of their regular educational assessments rather than tests specifically designed for this study.

Within the group using Cerego, the authors found that **the amount of time spent learning in Cerego was associated with significant increases in post-test score** (*Figure 26*).

The relation between post-test score (percent correct on end-of-term test) for students in all subjects who were assigned to the Cerego group was examined in relation to total time (in hours) spent using Cerego. A significant positive correlation was found between post-test scores and time spent using Cerego, $R^2 (934) = .045, p < .001$, indicating that students who spent more time with Cerego also had significantly higher post-test scores. The relation between time spent in Cerego and post-test score was also examined separately for each subject area in which minimum use time was met (i.e., Algebra, Biology and Civics). For each of these subject areas, significant correlations ($p < .05$) were found: Algebra, $R^2 (295) = .263, n = 295$; Biology, $R^2 (175) = .135$; and Civics, $R^2 = .051$.

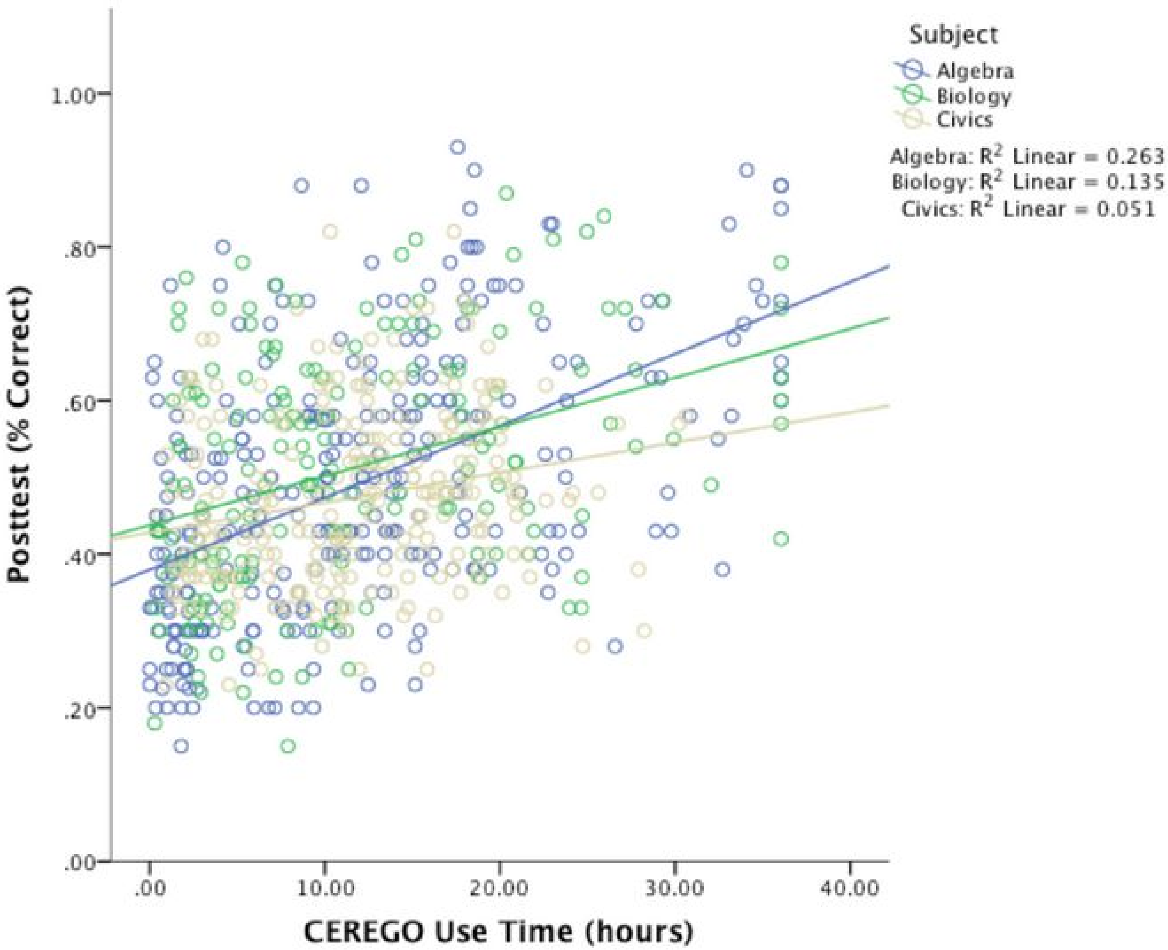

Figure 26: Relationship between time spent using Cerego and Post-test performance for students in the Cerego group.

The authors also found that **users who spent at least 1 hour per week learning in Cerego showed greater improvement between pre-test and post-test** than did users who did not use Cerego, or used it for less than the minimum time of 1hr per week. (*Figure 27*).

The authors examined whether or not adding Cerego to the classroom significantly improved learning outcomes for students. Using log data, students in the Cerego group were divided into two groups based on the "minimum use" criterion of 13 hours of total use time (one hour per week of assigned Cerego use). Learning outcomes were then compared for 1) Control Group (n = 739); 2) Cerego-assigned, but not meeting minimum requirements (n = 705); and 3) Cerego-assigned, meeting minimum requirements (n = 197). An ANCOVA was then conducted with post-test score as the dependent variable, condition (control, Cerego assigned – minimum time not met, and Cerego – minimum time met) and subject (Algebra, Biology, and Civics) as between subject variables, and pre-test score as a covariate. Results indicated a significant main effect of condition, $F(2, 1633) = 151.65$, $p < 0.001$. Planned comparisons indicated that **controlling for pre-test scores, the Cerego group meeting minimum use time requirements scored significantly higher** than the other two groups.

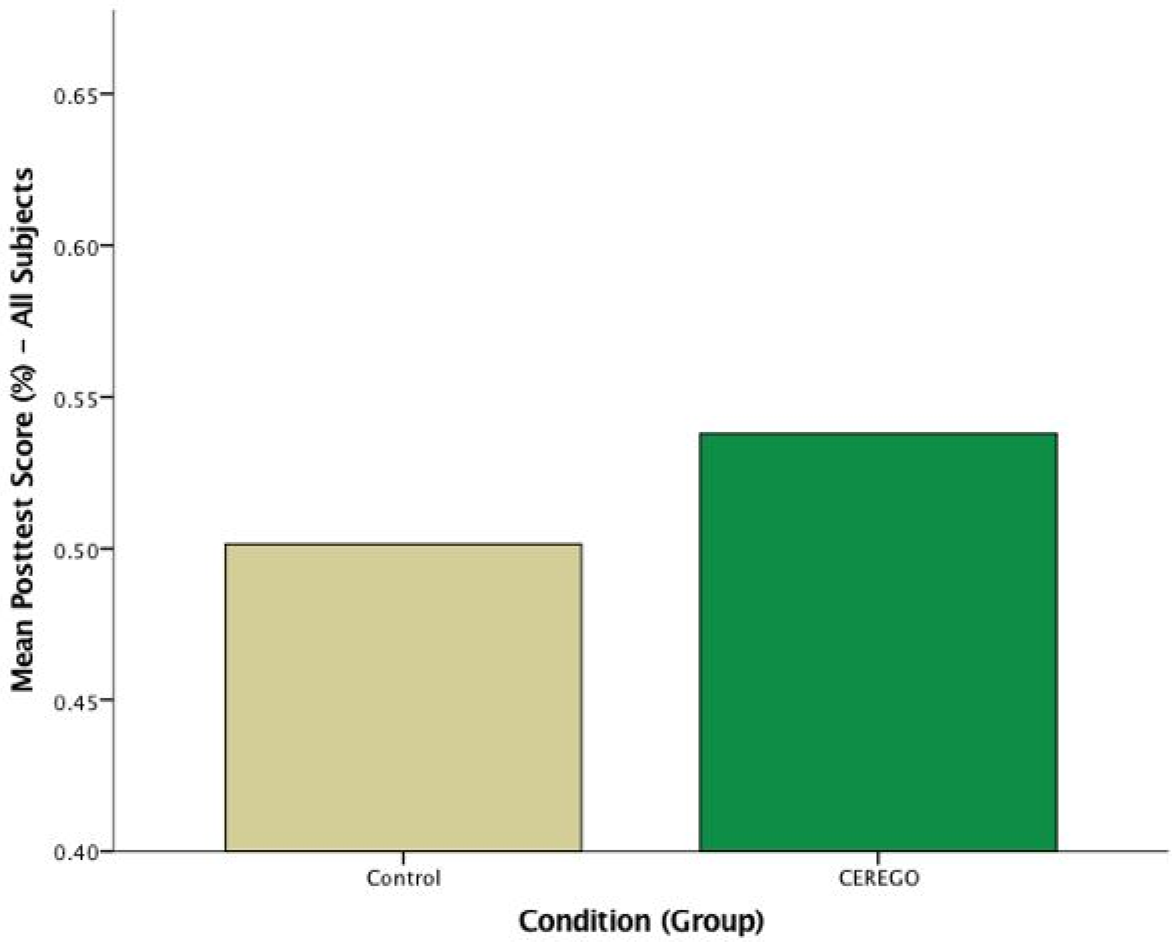

Figure 27: Post-test scores for users assigned Cerego, and controls. Users assigned to the Cerego group who engaged with Cerego for the minimum time of 1 hour per week showed significantly greater improvement during the course than both controls and lighter users, as measured by post-test scores (controlling for pre-test scores).

The full study can be found in *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2015* (pp. 869-878).

Case Studies

Cerego is being used in partnership with EdX and the University of Texas at Austin to provide a 10-week massive online open course about Jazz music. Using Cerego to increase engagement and retention of the course material:

- More than twice the proportion of initial signups **completed the course successfully** than with a typical MOOC (12% v 5%).

- Of the completing students, 1408 (59%) scored above **90%**, including 773 (32%) who achieved **perfect** scores.

- Consistent with the findings from [*Homer & Plass, 2015*] that greater Cerego use improved learning outcomes, high achieving students spent **more time studying in Cerego** per week (66 minutes) than the average student who completed the course (51 minutes).

- Surveyed students reported that using Cerego helped them to learn **faster** (74%) and retain their knowledge for **longer** (82%) than traditional study methods.

More details can be found at https://cerego.com/pdf/Edx_Infographic.pdf.

Cerego was also piloted by New York University College of Dentistry in summer 2015. Around 350 students taking review classes for Board exams in the fall were given the opportunity to supplement their reviews with Cerego content created by the professor.

- Students studied 424 review items in Cerego, replacing **96 hours** of class teaching time (reducing the professor's in-class teaching by **50%**).

- **100%** of students in the review class passed the subsequent Board exams.

- Aggregate Board exam results for the students were exceptionally high; **2.6 standard deviations** above the national average.

- A majority of surveyed students reported that using Cerego to review was worthwhile, by a margin of **58%** versus **16%**.

References

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, *21*(8), 627-635.

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General*, *108*(3), 296.

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective?. *Cognitive psychology*, *61*(3), 228-247.

Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123-144). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185-205). Cambridge, MA: MIT Press

Bjork. R.A. & Yan, V.X. (2014). The Increasing importance of learning how to learn. In M. McDaniel, R. Frey, S. Fitzpatrick, & H.L. Roediger (Eds), Integrating cognitive science with innovative teaching in STEM disciplines [E-reader version]. doi: http://dx.doi.org/10.7936/K7QN64NR

Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). How people learn: Brain, mind, experience, and school. (Expanded ed.). *Washington D.C.: National Academy Press*, ISBN 0309070368.

Broudy, H. S. (1977). Building a philosophy of education (2nd Edition). *R.E. Krieger Publishing Company,* ISBN 0882755439

Byrnes, J. P. (1996). Cognitive Development and Learning in Instructional Contexts. *Boston: Allyn and Bacon,* ISBN 0205159508

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

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T. & Pashler, H. (2008). Spacing Effects in Learning: A Temporal Ridgeline of Optimal Retention. *Psychological Science, 19*, 1095-1102.

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

Dempster, F. N., & Farris, R. (1990). The spacing effect: Research and practice. *Journal of Research & Development in Education, 23*, 97–101.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4-58.

Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Books on Demand.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology*, *15*(1), 1-38.

Homer, B. & Plass, J. (2015). Innovating randomized effectiveness trials: The case of an adaptive learning engine for E-Learning in high schools. In *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2015* (pp. 869-878). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Karpicke, J. D. & Roediger, H. L., III. (2008). The critical importance of retrieval for learning. *Science, 15*, 966–968.

Keppel, G. (1964). Facilitation in short-and long-term retention of paired associates following distributed practice in learning. *Journal of Verbal Learning and Verbal Behavior*, *3*(2), 91-111.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories is spacing the "enemy of induction"?. *Psychological science*, *19*(6), 585-592.

McDaniel, M. A.; Roediger, H. L. & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review, 14*, 200–206. doi:10.3758/bf03194052.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*(3), 159-163.

Peterson, S. E. (1992). The cognitive functions of underlining as a study technique. Reading Research and Instruction, 31, 49–56.

Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough?. *Journal of Experimental Psychology: General*, *140*(3), 283.

Rawson, K. A., Dunlosky, J., & Sciartelli, S. M. (2013). The power of successive relearning: Improving performance on course exams and long-term retention. *Educational Psychology Review, 25*, 523-548.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological science*, *17*(3), 249-255.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in cognitive sciences*, *15*(1), 20-27.

Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied, 17*, 382-395.

Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 907.

Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, *25*(5), 763-767.

Veenman, M. V. J., Wilhelm, P. & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14*, 89–109.

Wang, A. Y., Thomas, M. H. & Ouellette, J. A. (1992). Keyword mnemonic and retention of second-language vocabulary words. *Journal of Educational Psychology, 84*, 520-528.

Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society*, *15*(1), 41-44.