Optimizing Human Learning

Final Draft and Deliverable

Jacob Puthipiroj

Abstract

The benefits of spaced repetition over other more common study techniques is a long understood fact in academia. However, spaced repetition is still relatively obscure, for a number of reasons, including lack of knowledge on part of students of better methods, lack of incentives on part of educators to change their teaching style, and the inherent constraints on part of institutions and administrators to change the system.

Spaced Repetition is thus more commonly used as a self-studying technique, but when fully harnessed, can have extremely impressive results. This capstone attempts to create a space repetition solution that can be easily hackable and modified to fit customized user needs, in the form of an addon for Anki, an existing spaced repetition software. Unfortunately, we find Anki's algorithm inadequate for a number of reasons. We also create materials to educate people on how best to use Anki, because some of the UI is unintuitive. We achieve some success: in total, 132 people have downloaded the addon.

Contents

1 Background

Introduction to the Science of Learning

The ability to learn is an invaluable skill. Moreover, it is an invaluable metaskill: optimizing how we learn allows for more efficient acquisition of other skills.

Because it is a skill, the ability to learn can be improved. How learning can be better facilitated in real-world settings is explored in the field of the science of learning: an interdisciplinary field at the crossroads of cognitive neuroscience, educational psychology, and computer science. In general, this field suggests that most common learning techniques among students are suboptimal.

For example, a meta-analysis by Dunlosky et al [9], a cornerstone study in the science of learning, evaluates the effectiveness of 10 learning techniques. Common techniques, such as summarizing and highlighting text, were found to be generally ineffective for learning material, and comparable to a simple read-through of a textbook. Other techniques, such as distributed practice and practice testing, were found to be highly effective ways of learning and retaining information, but are less known. To date, one of the most powerful tools discovered in the science of learning is spaced repetition.

Spaced Repetition

Spaced repetition combines three key findings: the testing effect, the distribution effect, and the lag effect, into an efficient learning method.

The Testing Effect

The testing effect is an observation that learning techniques that involve active information retrieval, such as using flashcards as a form of studying [17], are generally more effective than passive re-reading. Here, 'testing' refers not only to formal exams but more broadly the attempt to retrieve information from memory. In fact, Allen [1] found that memories formed as a result of testing once are comparable in strength to those formed by 'studying' 5 times.

Testing is found to boost learning in multiple ways: memory for the tested information is improved [16], the rate of forgetting is slowed [5], learning is potentially transferred to new situations [4], and learner metacognition itself is improved [28].

A noteworthy study by Roediger & Karpicke [25] involves multiple experiments comparing the efficacy of testing with rereading over varying periods of time for learning Swahili vocabulary. Reading comprehension was greater among the rereading group over a period of 5 minutes. However, the self-testing group performed better over a period of 2 days and 1 week. Students wrongly believed that rereading was superior to testing for all intervals and rarely tested themselves. Furthermore, students almost never tested items already learned.

The Spacing Effect

The spacing effect refers not to a particular type of learning technique, but rather, the schedule of learning sessions. The spacing effect suggests that distributed practice, spreading learning out over time, is more effective than cramming, also called 'massed practice'[32]. Each learning session could involve restudying material, retrieving

information from memory, or practicing skills. Combining the testing effect with the spacing effect, using appropriately scheduled learning sessions to test retrieval (distributed practice testing), is found to be better than distributed study alone [5].

In a large meta-analytical study on distributed practice testing, Cepeda et al [6] review 184 articles with 317 experiments. The usual observation is that the massed practice group initially outperforms the distributed practice group, but the distributed practice group later outperforms the massed practice group over the span of days or even hours.

The spacing effect is a heavily generalizable finding. The spacing effect applies across various domains (e.g. new lists of words, perceptual-motor tasks, spatial tasks), age groups (infancy [12], childhood [33], adulthood [13], and the elderly [17]), and even species (rats [14], pigeons [24], flies [19], bees [20] and sea slugs [30]).

The Lag Effect

The lag effect answers the question of how to space each learning episode in distributed practice testing for optimal retention. Long lags between episodes are more effective than short lags. While the typical intervals between learning sessions can be expected to be on the order of days or weeks, most studies have used relatively short intervals of less than a day [6].

There is some debate about whether the intervals between each learning session should expand or remain constant in length. Landauer and Bjork [18] were the first to compare the efficacy of various schedules of distributed practice. Amongst the schedules tested, the expanding schedule yielded the highest recall followed by equal-interval, contracting, and massed practice, respectively. Furthermore, in a practical setting of attempting to learn foreign vocabulary over four weeks, Kang, Lindsey, and Mozer [15] found that an expanding schedule of distributed practice was superior. The idea of an expanding schedule is intuitively appealing and has been influential in the development of spaced repetition.

When the three effects are put together, we obtain an 'expanding distributed practice' schedule, more commonly known as 'spaced repetition'.

Spaced Repetition Software

Spaced repetition is one of the most powerful methods of learning discovered and has been successfully incorporated into noteworthy learning syllabi. The Pimsleur Language Programs, for example, uses a spaced repetition algorithm to present material, with the following intervals between learning sessions: 5 seconds, 25 seconds, 2 minutes, 10 minutes, 1 hour, 5 hours, 1 day, 5 days, 25 days, 4 months, and 2 years [22]. The optimal intervals for each person are dependent upon a number of factors, such as the base learning rate, time of day, difficulty of material etcetera.

In recent years, spaced repetition software and online platforms such as Anki and Duolingo have become increasingly popular, often replacing physical flashcards. The promise of these software is a more personalized, and thus more effective, spaced repetition algorithms. However, most of spaced repetition algorithms today remain simple rule-based heuristics with hard-coded parameters, which are unable to fulfill this promise. Adaptive, data-driven algorithms with provable guarantees have been largely missing until very recently [31].

Spaced Repetition Today

There is very little research on the effectiveness of spaced repetition when incorporated into a formal classroom setting. However, there have been cases of individuals achieving significant success using spaced repetition. For example, Ali Abdaal, a medical school student at Cambridge University, ranked 1st for his exams by using a spaced repetition-based essay memorization framework. More broadly, a large online community of medical school students has formed around the usage of Anki in learning large amounts of factual material. The reason this works for medical school students is because of their highly standardized memorization based curriculum, and the reason they use Anki is because its open source nature allows for flashcards to be easily shared across users.

Similar success has also been observed in non-traditional contexts. For example, Roger Craig set multiple records on the quiz show Jeopardy! by using a spaced repetition software (Anki) to memorize chunks of 200,000 past questions [2]. Another Jeopardy! winner, Arthur Chu, also used Anki in order to train.

Despite the promises of spaced repetition, a key problem observed in the science of learning is the inability to convince students to try what are considered better methods. As briefly hinted at in Roediger and Karpicke [25], students generally were not aware that testing and distributed practice were empirically better for recall. In a more recent survey, Morehead, Rhodes & Delozier [7] found that the majority of college students were aware the spaced study benefits learning, yet students still reported frequently massing their study before an exam [29]. Kang proposed that there were probably two broad factors that work against students' greater use of spaced practice: first, the lack of forethought or planning required to space out one's study and the discipline needed to follow through on the plan, and second, the 'subjective sense of fluency often engendered by passed practice, which can mislead students in feeling large gains in learning' (these gains, of course, do not last) [16].

A similar problem arises on the institutional level: while findings in spaced repetition have been known for a long time, they have never been properly incorporated into modern education. The spacing effect is 'a case study in the failure of how to apply the results of psychological research' [8]. According to Vash, "education policy setters know perfectly well that [spaced practice] works better [than massed practice]. They don't care. It isn't tidy. It doesn't let teachers teach a unit and dust off their hands quickly with a nice sense of 'Well, that's done.' " [34]. Again, Kang cites two reasons for this: first, instructors who often resort to familiar methods and techniques in teaching as opposed to research in the field, and second, the conventional instructional practice, which typically favors massed practice. [16]

This Capstone

Despite its wide-ranging applications, spaced repetition is still relatively obscure. Given the potential of the learning technique, this capstone aims to bring spaced repetition to a wider audience.

2 Design Thinking

Background and Context

The turn of the early 20th century saw a period of massive industrialization in what is now called the second industrial revolution, or the technological revolution. Advancements in manufacturing and production technology enabled the widespread adoption of new technologies, as well as the unprecedented movement of people and ideas. Scientific thinking allowed for innovations in long-established disciplines, such as chemistry, engineering and physics, and in some cases gave rise to entirely new fields altogether. The development of railroads, for example, led to creation of modern business management, as it required organized labor, capital, and expertise at a more complex level than anything yet experienced.

Frederick Winslow Taylor, one of the first management consultants, developed a comprehensive theory of management known as 'scientific management'. This was revolutionary for its time: detailed decomposition and documentation of an optimal production method, for example, allowed for easier automation of manufacturing and production processes, which eventually freed up large portions of the workforce for the service sector. Yet scientific management was almost immediately criticized for turning the working into an 'automaton' who performed only monotonous and unfulfilling work. By the 1930s, scientific management had become largely obsolete, though remnants of its body of knowledge remain in modern operations management and logistics.

Organizational management theory has since developed other trends. The late 1980s, for example, saw the rise of total quality management (TQM), which drew on statistics to monitor and control processes. The 1990s saw the rise of lean manufacturing, which emphasized the elimination of waste. And the 2000s saw significant interest in 'design thinking', with emphasis on a more flexible, human-centered approach towards innovation and management. Drawing not from the sciences but rather architecture and art, design thinking marked a distinct change in management philosophy.

What is Design Thinking?

While the concept of design is perhaps as old as history, the term "Design thinking" can be traced back to Herbert Simon's *The Sciences of The Artificial*, where it is defined as "changing existing circumstances into preferred ones" [27]. Free from the rigidity of previous structures, the growth of design thinking as a field has been closely related to the design and consultancy firm IDEO, and highly influential in business for organizational management, user experience, and user interfaces. Generally, design thinking is regarded both as a process, as well as a mindset for the effective execution of said process.

As an emerging discipline, the design thinking process has multiple formulations, proposed by several different authors. Many formulations are cyclical, and most incorporate some element of nonlinearity. Dunne and Martin, for example, propose a 4-step circular process of induction, abduction (ideation), deduction, and testing [10]. Tim Brown, executive chair at IDEO, suggests instead a circular process of inspiration, ideation, implementation, and impact [3]. Walozsek, conducting an analysis of design thinking process formulations by various authors, identifies the following common themes [35]:

- Understanding the problem, usually via empathy.
- Observing Users
- Interpreting Results
- Generating Ideas

- Building Prototypes and Experimenting
- Testing, Implementing, and Improving the Design

Why is Design Thinking Relevant to This Capstone?

Design thinking is relevant as a solution-based approach to so-called 'wicked problems'. Coined by Horst Rittel in the 1960s, 'wicked problems' refers to problems which are ill-formulated, contain confusing information, and where many clients and decision-makers have opposite opinions [23]. In this capstone, the problem is ill-formulated, as the ideal solution would require a great number of people to change their mindsets and their behavior. As difficult as it is for spaced repetition to be incorporated into the routine of someone who is familiar with its benefits, it is harder to convince a student who does not have the relevant background knowledge or the discipline, and harder still to convince administrators and educators to incorporate it into their existing curricula, if indeed the goal is to modernize education and pedagogy for the 21st century.

As previously mentioned, despite the well-known benefits of spaced repetition in the learning sciences, spaced repetition is still relatively obscure in the general population, and few people manage to use spaced repetition solutions consistently for their own learning goals. Here, design thinking is especially relevant, as it encompasses a healthy middle ground of problem-solving. While it fully acknowledges the analytics, science, and rationale of why spaced repetition can potentially work, it places much-needed emphasis on user experience and feedback to obtain practical, rather than purely theoretical results.

Design thinking as a field is furthermore likely to produce good results, given its successful track record in the fields of education, marketing, and software development: all three of which are relevant to this capstone.

Inspiration

I first heard about spaced repetition in 2012, while looking through blog posts by Tim Ferriss. A particular post, "How to Learn Any Language in Record Time and Never Forget It", explored a method of second language acquisition, which relied heavily on Anki –a spaced repetition software– for building a large vocabulary. Although I was immediately interested, I remember quickly abandoning Anki. I remember being confused by the number of customizable settings available, and being discouraged by how the number of cards I had to review everyday seemed to endlessly increase.

I returned to spaced repetition after high school. During my gap year in Germany, I took German lessons, but finding the pacing to be too slow, I supplemented my learning with other tools: Anki, Duolingo, and online video games in German. In the beginning, I used Anki intensively, learning 200 new German words a day (which took 3-5 hours). After 2 weeks, I was frustrated by the amount of words I had to review every day, and again stopped. Within a month of starting my German lessons, I took my B1 German proficiency test, scoring over 90 %. Two months later, I would again score over 90% on a B2 German proficiency test, and by 6 months, I had passed my C1 examination, which qualified me for studying at the university level in Germany.

My own experiences in using spaced repetition to learn German convinced me personally of the efficacy of spaced repetition in real-world settings. According to the Common European Framework of Reference for Languages, an estimated 350-650 cumulative hours could be expected for German proficiency at the B1 level. Yet I would estimate that I had put in a maximum of 100 hours total in the first month. As good as an intensive, personalized spaced

repetition schedule was for learning vocabulary was, compared to in-person lessons, I believed that there was much room for improvement.

Frame Your Design Challenge

The 'Frame Your Design Challenge' is one of the tools of the IDEO design kit, that allows users to organize how to think about the solution, and clarify where to push the design in moments of ambiguity. The following include the questions and answers to the challenge:

- What is the problem you're trying to solve?
  - Improve existing spaced repetition solutions for students.
- Try to frame it as a design question
  - How might we improve existing spaced repetition solutions?
- What's the impact you're trying to have?
  - Provide an online customized learning schedule for individuals.
- What are some possible solutions to your problem?
  - Create an addon to Anki
- Finally, write some of the context and constraints that you are facing
  - The problem should directly address existing solutions, such as Anki, and their complement or pose it as a competitor.
  - The problem requires a solution in the form of an app, ideally be customizable to the level of the user
- Does your original question need a tweak? Try again.
  - How might we improve Anki's algorithm via an addon?

Now with more clarity for the direction of the design, I attempt to gather insights from potential users to identify if their needs, and if they would indeed be met by an Anki addon. I conducted interviews both online, in the form of online surveys as well as in-person interviews at Imperial College London about note taking and learning techniques for memorizing information. Participants universally indicated a need for a systematic method of reviewing and retaining information.

Users

The hypothetical ideal user for Anki is **a disciplined student who needs to memorize large amounts of information to be preserved in their long term memory and who is comfortable with technological solutions.** Because of large number of people who would fit this definition, any number of decompositions of the potential user base is possible, such as career path, educational institution, etc. However because I am aiming to develop a solution that is universally and freely accessible (i.e, online), I categorize users by channel of outreach. I believe this particular segmentation is optimal, since each channel/website caters to a distinct small community of students, and allows for the comparing and contrasting of niche extreme users to see differences in needs. The results reveal relatively different needs among the four communities, as follows:

1. (Survey) **The Anki and Medical School Anki subreddit** hosts a group of students who would like to efficiently prepare for medical school exams, such as USMLE. Almost all participants from this segment used Anki, Usually with a few add-ons.

2. (Survey) **HackerNews** users are are generally interested in computer science and entrepreneurship, with crowdsourced content for readers being categorized as 'anything that gratifies one's intellectual curiosity'. Less than 10% of participants used spaced repetition software such as Anki or Mnemosyne, although most had documents for a list of webpages stored somewhere they regularly referred to.
3. (Survey) **The Art of Memory Forum** users are a group of memory training enthusiasts who either have recreational (e.g. learn how to memorize the order of a deck of cards) or professional (e.g. learn how to memorize an entire speech to be delivered). Participants indicated varied learning-related needs, such as a system to efficiently memorize mnemonics, and a significant minority routinely used spaced repetition software.
4. (Interview) **Engineering students from Imperial College London** indicated a strong need for memorizing information for their exams in in-person interviews. Of the 6 that were interviewed, 4 had some sort of note-taking system (1 in particular used Anki - something they found from a viral YouTube video), and the other mainly relied on reading textbooks multiple times to remember information such as formulas.

As a result of the decomposition, it is clear that each group possesses different optimal needs, and showed that some communities, such as those of the HackerNews readers and engineering students may be poor choices for achieving product-market fit. Instead, the subredditors and memory enthusiasts, who were not necessarily technical experts but were highly interested in the practical aspect of a technical solution, and for whom the need was highly relevant to their professional or personal success, would make good potential users.

Interviews were conducted with friends of friends of mine at Imperial College London, a local university, where I came with a set of prepared questions. I would conduct the interview either in their dorms or the common room, and sometimes if they seemed comfortable I would ask to see their notes. Creating questions for both the interview and survey took time, and I referred to guidelines on questionnaire design from the Pew Research Center for help. The section on question wording was particularly salient, and I would reword my question multiple times and go over them with my roommate to avoid biases as much as possible. An example of this was 'to what extent are you satisfied or unsatisfied with your current learning system?' which was initially posed as 'Are you satisfied with your current note taking system?' The quantifiers were added to emphasize neutrality of the question, and implicitly let interviewees know that expressing dissatisfaction was perfectly acceptable.

Another designthinking tool used during the interview was the use of appropriate body language, as well as setting, in building empathy. Before the interview began, I would ensure a familiar, comfortable setting for the interviewee, establish rapport by spending some time beforehand, before starting the interview. I would then sit next to them or beside them (whichever felt most comfortable and natural), and then go through the questions and nod while they respond. After all of the prewritten questions were over, I would then ask followup questions about ones which received an interesting verbal or nonverbal response. One of the participants, for example, seemed to undergo a realization that her system of reading the textbook multiple times didn't work in university as well as it did in high school, and seemed rather embarrassed while responding. I would naturally laugh along with her, and then return to the question and ask about what differences she noticed and where the problem lied, such as the nature or difficulty of her university curriculum.

Needs

Participants from each of the four communities indicated varied needs. One of the survey questions - a Likert scale for 'how satisfied are you with your current learning system', was also posed to the interviewees in order to quantify the importance of satisfying the need to each community.

1. The Anki subredditors' principal indicated need generally leaned towards finding quality educational content to review and learn from, rather than a better learning system or framework. The modal answer for the Likert scale was a 4, with an average of 4.5. Many participants expressed gratitude for the existence of a spaced repetition schedule in their routine, and some even said they would likely fail medical school without such a system. Generally, participants indicated a strong interest in further improvements to Anki, such as having add-ons.
2. HackerNews readers were seemingly more enthusiastic, with the modal response on the Likert scale being a 2. However, when asked to elaborate on their need for a better learning system, most mentioned a centralized location for storing knowledge to refer to when looking up documentation, suggesting that a high-maintenance solution like Anki may be overkill. Overall, participants indicated their need was not met, but didn't seem to heavily suffer if the need was unaddressed.
3. The memory enthusiasts showed a bimodal pattern on the Likert scale, with 2 and 4 getting the most responses. Because of the familiarity of participants with various learning and memory techniques, I suspect these participants would be able to resort to other techniques besides spaced repetition for their particular use-case of memory feats. However, those who used Anki demonstrated strong interest in improvements, and some did not know about existing add-ons.
4. Interviewees from Imperial College London gave a modal answer of 3 on the Likert scale. When Anki was mentioned, participants questioned the amount of effort required in place of notes, and 3 mentioned the importance of doing past papers rather than rote learning in their curriculum, although all interviewees expressed interest in trying Anki for memorizing formulas.

As a result of the decomposition, it is clear that each group possesses different optimal needs, and showed that some communities, such as those of the HackerNews readers and engineering students may be poor choices for achieving product-market fit. Instead, the subredditors and memory enthusiasts, who were not necessarily technical experts but were highly interested in the practical aspect of a technical solution, and for whom the need was highly relevant to their professional or personal success, would make good potential users.

3 Technical Implementation

Background on Modeling Human Memory

Spaced repetition, and the science of learning in general, is often credited to Hermann Ebbinghaus, a German psychologist. Prior to Ebbinghaus, research into learning and memory centered mainly around philosophical observations and speculation. When Ebbinghaus's experiments on memory were first published in 1885, they had an immediate effect on the field of memory research. Among other things, Ebbinghaus is known for the discovery of the spacing effect, the lag effect, the 'savings effect', the 'learning curve' and the 'forgetting curve'. These terms are explained in detail below.

In his seminal 1885 paper 'On Memory' [11], Ebbinghaus attempted to memorize lists of 'trigrams': 3-letter nonsense syllables, such as 'WID' and 'HOF'. Trigrams were chosen as learning materials, (rather than, say, poems), as they were free of previous associations. Ebbinghaus would look at each syllable for 0.4 seconds, pause for 15 seconds, and then go through the list again, until he could correctly recite the entire list at speed. He would then wait until he could no longer recall any of the items in the list, and relearn the entire list again. Generally, the list would be memorized faster the second time. Ebbinghaus theorized that the difference between time taken to learn the list on the first and second passes were due to previously learned memories in his subconscious. Although the information could no longer be consciously assessed, the subconscious brain would retain information in what he termed 'savings'.

Learning and Forgetting Curves

Ebbinghaus would develop his theory of memory savings more concretely in his formulation of the 'learning curve', and its less well-known complement, the 'forgetting curve'.

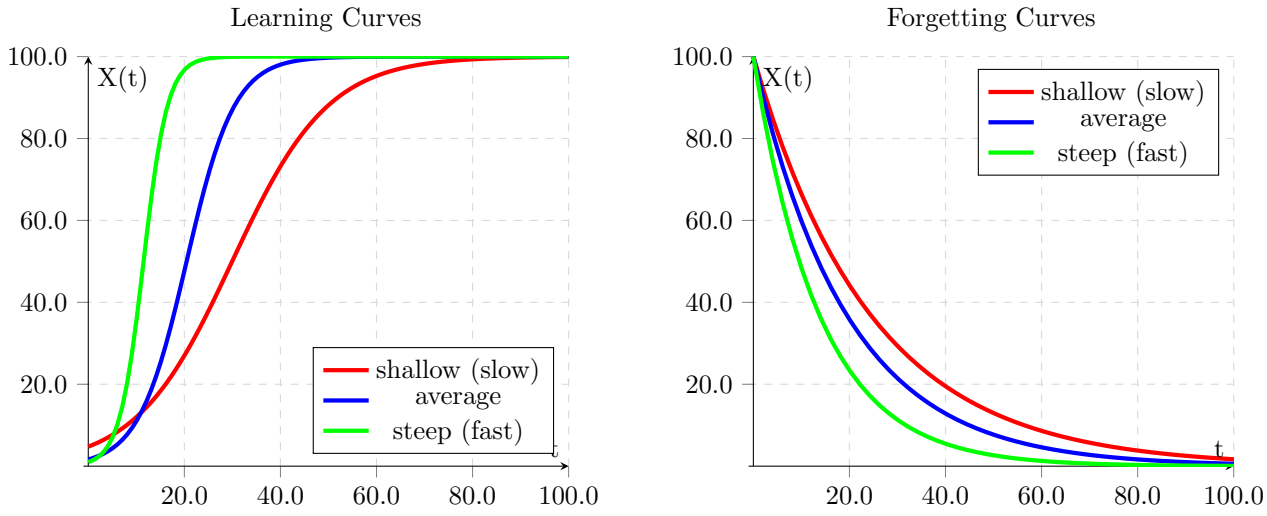


Figure 1: These two figures indicate the differences between shallow learning and forgetting curves in red, and steep learning and forgetting curves in green.

In his paper, Ebbinghaus referred to the learning curve as representing the increase in time to learn a list of trigrams

as the length of the list increased. More broadly, the learning curve - which has since seen growing use in other fields - represents how learning gradually increases over time. Learning curves are often modeled as sigmoidal functions, or 'S-curves', as it captures three commonly observed behaviors in mastering material: a slow initial improvement, then a rapid increase, and a leveling off as one plateaus. Exponential curves, or power law curves, are also sometimes used to approximate rapid learning without any initial difficulty. In his paper, Ebbinghaus assumed an exponential curve, though as will be later discussed, a power law was also considered.

A 'steep learning curve', as depicted in green, describes the progression of a skill that is quickly acquired over a relatively short period of time (Oxford Dictionary, American Dictionary, Merriam Webster). In colloquial use, a 'steep learning curve' has the opposite meaning of its original definition. In its scientific use, to say a subject has a 'steep learning curve', would mean that learning is rapidly gained in the initial period, i.e. is not difficult.

Conversely, forgetting curves describe how memories decay over time in the absence of reviews. A 'steep forgetting curve' would characterize a difficult skill, in which information is quickly lost, while a 'shallow forgetting curve' would characterize easily memorized information, with slow decay of retention. Ebbinghaus posited that each time a learning session occurs, 'savings' accumulates, and the forgetting curve after subsequent reviews become more shallow. This would also imply that the optimal intervals between learning sessions expand in length.

Limitations and Replication

By modern standards, Ebbinghaus' research had a few methodological issues, though the main consideration is the use of a single-subject design. Single-subject designs are not unusual in memory research: in fact, they often provide the advantage of having no inter-subject variability. However, variability on the level of the individual would still require hundreds of trials to be minimized to acceptable levels - Ebbinghaus' own estimates of the forgetting curve is based on seven months of arduous self-experimenting, and would often consume up to three hour-long sessions per day.

The main concern regarding single-subject designs is the lack of generalizability. In the absence of further research, it is unclear if the shape of the forgetting curve remains similar across different individuals. It could be possible that Ebbinghaus' own memory was exceptional in some way. For his own memory savings function, Ebbinghaus fit his observations with the curve:

$$Q(t) = \frac{1.84}{\log t^{1}.25 + 1.84}$$

'Savings' is specifically defined as the relative amount of time saved on the second learning trial as a given the first trial, and is analogous to retention rate. For example, a savings of 100% indicates all items from the first trial were remembered on the second, and a savings of 75% would indicate the relearning missed items required only 25% as long as in the first learning session. In modern formulations, we would say Ebbinghaus had assumed an exponential decay model, following an equation of the form:

$$R = e^{-t/s}$$

where R is the retrievability of memory (a measure of how easy it is to retrieve a piece of information from memory), s is the stability of memory (which determines how fast R falls over time in the absence of training, testing, or

other recall), and t is time.

Interestingly, in his first manuscript on 1880 (Before the 1885 submission), Ebbinghaus had in fact used a different equation, of

$$Q(t) = (1 + 0.51t)^{-0.099}$$

Recent developments indicate that a power-law function may better approximate the shape of the forgetting curve. Wickelgren a further distinction that an exponential function described forgetting from short-term memory but that a power function described forgetting from long-term memory [36]. A power law function has the form:

$$R = t^{-1/s}$$

In 2015, Murre and Dros replicated Ebbinghaus' experiment on J. Dros (also one of the coauthors) [21]. Instead of running 12-45 replications per time interval, Dros ran 10 replications per time. Unlike Ebbinghaus, who experimented at fixed times of the day (in the morning at 10am, sometimes additionally at 12pm and further at 7pm), Dros would review at convenient but inconsistent times of day. This is likely inconsequential: In another previous study by Heller et al (1991), time of day effects were not replicated. A further difference was the material used: while Ebbinghaus used 2300 trigrams formed from the German alphabet, Murre and Dros used 2100 trigrams formed from the phonotactics of Dutch, and removed syllables which too much meaning in Dutch to ensure uniform difficulty of the stimuli.

When using modern statistical techniques, computer-optimized curves to fit the observations of the Murre and Dros study were very similar to the one that Ebbinghaus calculated manually. Specifically, the parameter values for the power function were 0.523 and 0.101 (compared to Ebbinghaus' 0.51 and 0.099), and for the logarithmic/exponential function was 1.8 and 1.21 (compared to Ebbinghaus' 1.85 and 1.25). A more advanced Memory Chain Model (MCM) equation was also proposed, which incorporates more parameters, and which achieved a higher R^2 and lower AIC, but the overall conclusion was that Ebbinghaus' original study was sound in its methodology.

Evaluating Modern Spaced Repetition Algorithms

With the shape and the importance of the forgetting curve established, we attempt to evaluate a few noteworthy spaced repetition algorithms, either currently deployed or developed in academia for possible implementation in an Anki addon. We attempt also to evaluate the following algorithms in light of practical considerations: for example, most Anki users review their cards at most once per day. Thus, once a card has been learned and intervals have stretched beyond a day, granularity to the level finer than days are not practically significant.

Anki's Existing Algorithm

Anki uses the SM2 algorithm, originally developed by Polish software developer Piotr Wozniak for SuperMemo, another spaced repetition software in the 1980s. Anki has not improved its algorithm since, and upgrading to the latest member of the SuperMemo algorithm is not possible, as they are proprietary.

Anki's implementation of the SM2 algorithm does not offer binary quizzes. Instead of pass or fail, users can indicate the result of a flashcard they reviewed as: 'failed', 'passed - hard, 'passed', 'passed - easy'. The length of corresponding next intervals are as follows:

For failed reviews, where i is the length of the previous interval, and $m_0 = 0$ by default;

$$i_1 = m_0 * i$$

For hard reviews, where d is the delay (in days) between the due date and the review date;

$$i_2 = \max(i + 1, (i + d/4) * 1.2 * m))$$

For ok reviews;

$$i_3 = \max(i_2 + 1, (i + d/2) * (f/1000) * m)$$

For easy reviews, where $m_4 = 1.3$ by default;

$$i_4 = \max(i_3 + 1, (i + d) * (f/1000) * m * m_4)$$

Where f is the 'factor' of each card (referred to in the Anki sourcecode as `ease`, which is a number with a maximum of 1300 which gets incremented by 150 - 200 every review, depending on the type. f is simultaneously adjusted along with the interval at each step, as follows:
If a review is failed;

$$f' = \max(1300, f - 200)$$

If a review is hard;

$$f' = \max(1300, f - 150)$$

If a review is ok;

$$f' = f$$

If a review is easy;

$$f' = \max(1300, f + 150)$$

As mentioned previously, this is the default option, and is the main spaced repetition solution available online. Anki has a large number of users, and though the project of Anki has gotten a litlte out of hand (history of Anki).

The following algorithms are considered as alternatives to Anki's algorithm.

Half-Life Regression

A recent spaced repetition algorithm that has seen practical use is half-life regression, developed by the team at Duolingo. Duolingo is a platform for second-language acquisition, where students develop skills in translating lexemes from one language to another.

The objective function to be minimized is

$$l(\langle p, \Delta, x\rangle; \Theta) = (p - 2^{-\frac{\Delta}{2^{\Theta x}}})^2 + \alpha(\frac{-\Delta}{\log_2 p} - 2^{\Theta x})^2 + \lambda||\Theta||_2^2$$

This is a convex function, and can be solved using stochastic gradient descent [26]. The team at Duolingo reported a reduction in the error of predicting recall rates by over 45%, and improved daily student engagement by 12%. However, because of the requirement for each lexeme to be known beforehand in training the data, it would be infeasible for implemnetation in Anki, where users study on their own.

MEMORIZE

Tabibian [31] considers the optimal reviewing schedule as a stochastic optimization problem, with the objective function to be minimized being:

$$\min \mathbb{E}\left[\int_0^T ((1-m(t))^2 + \frac{1}{2}qu^2(t))dt\right] \quad s.t. \quad u(t) \geq 0, \forall t \in (0,T]$$

Where $(1-m(t))^2$ is the probability of forgetting and $\frac{1}{2}qu^2(t)$ is the review rate. In particular, q is the tuning parameter that sets the tradeoff between the two objectives. The solution is obtained by reformulating the objective function as a partial differential equation, and deriving it using the Hamilton-Jacobi-Bellman equation, as:

$$u(t) = q^{-1/2}(1-m(t))$$

Where $u(t)$ is the probability of forgetting at any given point in time. So with perfect recall, $m(t) = 1 \implies u(t) = 0$, no reviews should be scheduled, and at zero recall, $m(t) = 0 \implies u(t) = q^{-1/2}$, and learner must review the item following an exponential decay curve $q^{-1/2}$. The rate at which the item should be reviewed is proportional to how close the learner is to forgetting the item. This is intuitive and relatively easy to implement.

The solution Tabibian provides is 'provably optimal'; it can be shown the algorithm cannot be improved upon via tools from optimal control theory. However, these proofs of optimality may be unconvincing to some potential users, as they practically assume, for example, that cards that are due will be reviewed exactly when they are scheduled. Furthermore, the stochastic nature of the algorithm may result in due times that are unintuitive and unappealing.

Ebisu

Ebisu models the results of each test as a Bernoulli, pass-fail attempt. For each piece of information, the results are therefore a Bernoulli random variable:

$$x_t \sim \text{Bernoulli}(p_t)$$

The conjugate prior to the Bernoulli is the Beta distribution

$$p_t \sim \text{Beta}(\alpha_t, \beta_t)$$

Conjugate priors have nice properties in that once we update (i.e. observe new results), our belief about the recall probability is updated. The resultant, new estimation follows another Beta distribution, with the following arguments:

$$p_t \sim \text{Beta}(\alpha_t + x_t, \beta_t + 1 - x_t)$$

In other words, since x_t is 1 for a correct answer and $x_t = 0$ when not, the updated parameters are $\text{Beta}(\alpha+1, \beta)$ for a correctly answered quiz, and $\text{Beta}(\alpha, \beta+1)$ for an incorrectly answered one.

In order to find the probability of recall at some point in the future, $t_2 = t$, we find that

$$p_{t_2} = p_{\delta t} = 2^{-\delta t/h} = (2^{-t/h})^\delta = (p_t)^\delta$$

When a random variable following a Beta distribution becomes raised to a positive power, it becomes non-Beta distributed. Following a manipulation of variables (provided in the appendix), this variable instead is distributed to the generalized Beta of the first kind.

$$p_t^\delta \sim \text{GB1}(p; 1/\delta, \alpha; \beta)$$

In short, a flashcard begins with the memory model as (α, β, t). After a quiz is administered, it becomes $(\mu(\frac{\mu(1-\mu)}{\sigma^2-1}), (1-\mu)(\frac{\mu(1-\mu)}{\sigma^2-1}), t')$, where $\mu = m_1$ and $\sigma^2 = m_2 - \mu^2$, where

$$m_n = \frac{\sum_{i=0}^{n-k} \binom{n-k}{i}(-1)^i * B(\alpha + (i+k)\delta) + N\delta(\frac{t'}{t_2}), \beta)}{\sum_{i=0}^{n-k} \binom{n-k}{i}(-1)^i * B(\alpha + (i+k)\delta, \beta)}$$

Choosing an Algorithm to Implement: Considerations

When time came to choose between the various algorithms, Ebisu was a clear winner, for a number of reasons.

Firstly, it was not possible to implement anything from the SuperMemo family of algorithms beyond SM2, as they are proprietary. I was discouraged from choosing Half-Life Regression - it was unclear to me if developing an on-line learning algorithm (one in which information was encoded beforehand) was possible. Although I was impressed by MEMORIZE, I was unsure if the addition of stochastically optimal results were necessary, as there are some limitations in the model (e.g. it does not account for the fact that humans sleep).

Additionally, there were also constraints on the side of Anki. Firstly, Anki only allows for minute-precise intervals up until 3 days, at which point it allows only for day-precise intervals. While this may seem a problem that should be worked around by solving, I would argue that incorporating this algorithm is not an issue, as users are unlikely to review more than once a day, and therefore require more granularity than once a day anyways. Second, Anki does not allow for the support of a large number of packages that would be used for machine learning and statistics (NumPy and SciPy). Thus if there were any packages beyond the base Anki that had to be used, I would have to create custom functions myself, heavily increasing the workload.

As it turns out, Ebisu was the perfect algorithm, as it achieved a reasonably good approximation, and is one of the most practical solutions used among quiz scheduling. To achieve for anything higher, such as Ebisu, would be considered unnecessary. Ebisu was also not reliant on many packages that wouldn't be allowed in Anki, such as NumPy, SciPy, and Pandas. There was a reliance on the function `logsumexp`, which I could construct myself using the `math` package.

Coding Implementation

I coded the Anki addon for Ebisu with help from Ahmed Fasih and Arthur Milchior. Fasih is an electrical engineer and the author of Ebisu, while Milchior is a top Anki contributor and the author of over 30 Anki addons. The Anki addon is attached in the accompanying file. In order to install it, users must have the latest Anki version (2.1.21).]

I met Milchior in person. Unbeknownst to me, Arthur and I were connected through our shared interest in the rationalist community, and I heard about Milchior first through a mutual friend. Milchior just so happened to be giving a talk in the Bay Area on Anki, which I attended. Afterwards, I asked him for tips and help for creating the addon. This was invaluable, as so much of the Anki backend code had been spaghettified over the past 12 years.

4 Growth Hacking

Introduction to the Anki Ecosystem

As a flashcard app, Anki appeals to learners who need to remember large amounts of factual information in a relatively short time, and does so based on the principles of the science of learning, particularly spaced repetition. By timing the spacing of each repetition, and customizing it to the pattern of reviews for each user, the addon takes full advantage of the 'spacing effect', and minimizes active review ('study') time. The addon, Ebisu, is a significant improvement over the existing algorithm, which does not customize review times individually.

There are a large number of people within the Anki ecosystem. At the time of writing, Anki is the third most-downloaded app in the Apple Store for Education. There are over a million downloads on the Android App Store. The largest subset of Anki users is by far medical school students, who need to memorize large amounts of technical, non-intuitive facts. Through this lens, it is clear that Anki users are self-motivated, capable of following even complex instructions, and will be comfortable with (appropriately) technical communication. As a result, giving a thorough yet easily digestible explanation would be convincing to this audience.

Purpose

There are 3 main relevant stakeholders:

- Developers and academics whose work has made the product possible.
- The existing Anki community, including the creator, Damien Elmes.
- New potential users of the addon.

Each of the stakeholders have various interests in mind. Developers would like their product to be used by a large number of people in order to improve education practices globally. The needs of the existing Anki community are more immediate; they would like a better user experience for their product and reduce the total time studying while keeping switching costs minimal. As for the potential users of the addon, their needs include recognizing that the product is relevant to them, and that they are the target demographic.

As a result of this breakdown, success in this project is measured via the lens of creating and demonstrating impact. The only other serious alternative, of generating revenue, would require proprietarizing the addon, and misalign incentives in such a way that certain potential users (e.g. schoolchildren who don't yet have access to online payments, people who cannot afford a formal education) would be unable to access my product.[1] This would be ultimately detrimental to the developers and academics, whose impact of work would be slowed, and could be divisive within the Anki community: while some members would happily pay for a worthwhile addon, others may not want to and may even resort to piracy.

[1] #utility: I explain why choosing to gear this project towards creating revenue, rather than impact would be potentially counter-productive when considering the impact to developers (whose work will reach less people), the current anki community (who may react negatively and even pirate the product anyways), and newer potential users (some of whom may not be able to access the product). Instead, keeping the addon free and open-source will benefit all stakeholders instead of arguably just me.

Measuring Impact

The original goals of this project includes getting 50 signups on the addon beta (while it was still under development), and getting 100 downloads once the addon was ready. Currently, the number of people who have signed up have included 132 people. The first goal was achieved by November 2019, and the second was achieved by March 2020.

Growth Hacks

Three growth hacks were run in total in an attempt to promote Ebisu and its prerequisite base, Anki. The hacks took place between September 29th, 2019 and December 21, 2019.

The brand image Ebisu has is one of technical efficiency and friendly user interface. Amongst the Anki community, an often voiced complain is the confusing number of customizable settings, which heavily steepens the learning curve. Users may for example, manually tweak the number of cards they see per day, the initial review interval, and a plethora of more confusing settings, such as the 'starting ease', 'interval modifier', 'easy bonus', 'leech threshold', and so on. While there are a number of blog posts which detail some recommended settings, the default is not customized for each user, and there is no official manual. MEMORIZE fixes these issues by overwriting the UI, as well as the backend scheduling as to offer personally optimized review schedules.

Growth Hack 1

The first growth hack attempts to reach a large number of users via an outbound social media marketing campaign. Because the addon was still in development, this campaign creates a smokescreen MVP, in the form of a beta waitlist in order to confirm customer needs before the product is even created. The social media platforms chosen in order to make a post on were divided into two groups: targeted, and non-targeted, as follows:

- Targeted platforms: Art of Memory Forums, r/medicalschoolanki, r/anki
- Non-targeted platforms: Youtube, Quora, HackerNews

Here, the meta-segmentation considers that some websites encapsulate communities who may, as a whole, be considered potential users of Anki and Ebisu. Conversely, non-targeted platforms, such as Quora, contain massive amounts of users, of whom only a small fraction can be expected to become users in the early stage. The small probability of virality is offset by the massive returns in potential customers, and hence is a good platform for growth hacking as well, though content must be tailored differently and assume a more general audience. In communications with targeted platforms, word of mouth is expected spread far more efficiently, given the small size of the network. Any message specifically relevant to the homogeneous interests of the group can gain traction much more quickly than a similar message would on larger, diverse platforms like YouTube.

An example here is The Art of Memory Forums, a large online community of memory enthusiasts. Memory hobbyists, as a generalization, are obsessed with efficiency, readily embrace innovative and disruptive technologies, and are highly open to trying new 'memory hacks'. Posts directed at HackerNews could, therefore, focus on the story of Arthur Chu, a quiz show (Jeopardy!) contestant who built his knowledge base by using Anki. A link could be included to lead to a blog post going more in depth on how to download and use Anki, along with possible addons to consider enhancing the experience. Naturally, one of the addons is Ebisu.

A non-targeted community, on the other hand, are the users of HackerNews. While most readers of HackerNews are generally technically savvy (posts usually concern advances in computer science and entrepreneurship), it is unreasonable, especially because of its size, to expect that a significant fraction of HackerNews will be using Anki in the near term. However, communications can still be made relevant, by presenting the recent advances in spaced repetition, culminating in a comparison of Ebisu, Half-Life Regression, and MEMORIZE, along with a download link for Ebisu.

Aside from considerations to the group identity of users in each segment, the likeliest users to benefit from the product - what von Hippel (1986) calls 'lead users' – currently face needs that may soon be general in the marketplace, but are months or years ahead of a solution. These 'lead users' follow a specific psychographic profile, being more educated, prosperous and risk-oriented, and communications can therefore be technical and jargonistic within reason.

Overall, 4 different posts were written, each with a link to download Anki and a link to sign-up for the MEMORIZE algorithm beta waitlist. The second link, had a URL that was specific for the website the post was made on, in order to infer the difference in channels from the various websites. In some of the segments, posts would link to a lengthier article here, which gave a brief guide on current advances in spaced repetition and Anki, along with preliminary addons (including Anki) to further personalize Anki usage. This acted as a form of content marketing, which would be highly beneficial to users along the sales funnel. Many new Anki users are overwhelmed by the amount of customizations initially available to them, thus a guide that segues into MEMORIZE could help position the brand of the addon as one created by experts.

The posts were published within the same day on Sunday the 29th of September, 2019 and left for one week, with results collected on Saturday the 21st of December, 2019. Results at the end of the period were as follows:

User Segment	**Website(s)**	**Message Content**	**Sign-ups**
Memory Enthusiasts	*Art of Memory Forums*	Advances in Spaced Repetition Story of Arthur Chu Link to Blog Post Link to existing Trivia decks	*5*
Medical Students	*r/medicalschoolanki*	Reasons Anki's system is inadequate Link to existing USMLE decks	*34*
Programmers	*GitHub*	Link to project GitHub Installation and Setup Guide Link to Optimality Proof	1
Anki Users	*r/anki*	Reasons Anki's system is inadequate Link to Optimality Proof	*24*
General Audiences	*YouTube, Quora, HackerNews*	Link to Blog Post, which details: Advances in Spaced Repetition Anki Installation and Setup Guide Useful Addons inc. MEMORIZE	*12*

Table 1: The first four rows are considered targeted segments, and the last is non-targeted.

The number of signups in the last column can be used to obtain a chi-squared statistic about the un-evenness of

distributions. The Chi-squared statistic is $\chi = 27.02$, with a corresponding p-value of $p = 0.00002 < 0.05$. Hence we conclude that there is a significant difference between the 5 groups, and on deeper analysis, that specifically medical students and anki users are better websites to target than that of memory enthusiasts, general audiences, and GitHub users. A second test can be run by aggregating the targeted users versus non-targeted users, which gives a Chi-squared statistic of $\chi = 20.14$, with a corresponding p-value of $p =< 0.00001$. Hence targeted communications are much more effective than targeted ones, and further growth hacks should aim for targeted messages.

Qualitative data was also obtained from each of the websites by looking at the comments as a reply to each post. Generally, the opinions of each group were:

1. The Anki (r/anki and r/medicalschoolanki) subredditors' principal indicated need generally leaned towards finding quality educational content to review and learn from, rather than a better learning system or framework. Generally, participants indicated a strong interest in further improvements to Anki, such as having add-ons.
2. GitHub users and HackerNews readers, when asked to elaborate on their need for a better learning system, mentioned a centralized location for storing knowledge to refer to when looking up documentation, suggesting that a high-maintenance solution like Anki may be overkill. Overall, participants indicated their need was not met, but didn't seem to heavily suffer if the need was unaddressed.
3. The memory enthusiasts, from The Art of Memory Forums, were generally familiar with various learning and memory techniques, and perhaps can resort to other techniques besides spaced repetition for their particular use-case of memory feats. However, those who used Anki demonstrated strong interest in improvements, and some did not know about existing add-ons.

Customer Journey Map

Below is a journey map indicating the journey of a hypothetical medical student, John, who eventually ends up paying for the product given. On the left are the names of the stages, and in **bold** are stages which are considered significant 'moments of truth' that require significant work on part of the customer and which must be satisfied in order for the journey to progress to the desired direction (of purchase).

There are several key deviations that could be made for other hypothetical customers in order to get a broad range of likely scenarios. Another potential customer, for example, may not be as lucky as John, who found the solution to a problem he didn't even know he had while trying to solve the problem of initially setting up Anki. Another user who has used Anki over a period of years would be much less likely to search for 'best study tips' on YouTube and be more likely search from the Anki Subreddit, and in turn content posted on Reddit tailored for people like him could be more detailed and use more jargon in order to convince him that there is an existing problem with the algorithm and that our product (the solution) is a worthwhile one.

Stage	Description
Need created	John is a 22-year-old medical student who has recently been overwhelmed by the amount of information he is forced to remember in his course. Trying to find a more scientific solution to the problem of inefficient solution, John opens up YouTube and searches for 'best medicine studying tips'
Interest	The first reccomended video is by a medical student, Ali Abdaal. Ali runs a YouTube channel documenting his progress to the top scorer in Cambridge University medicine by using a spaced repetition app called Anki.
Engagement	John is intrigued and navigates to the Anki website via a google search, and downloads it himself.
Frustration	John opens up Anki, and tries to use it for a while, but soon realizes that there is a steep learning curve ahead of him.
Further Learning	John types into Google, "Best settings for Anki", and finds a guide on a small company blog showing common settings for fast Anki learning. The blog usefully guides him through the various settings and strategies to get him started.
Learning	At the bottom of the blog post, an additional option as mentioned: downloading the 'free trial' of an add-on that promises to fix several Anki settings. John is interested, and clicks on.
Download	John is brought to a page where it is briefly summarized what the addon does, along with a download button. Being persuaded by the article, and figuring out that he has nothing to lose, John downloads the add-on and installs it.
Purchase	A month later, John is very happy with his usage of Anki. A window pops up, telling him that his free trial is over, and asks that he pay a one-time $4.99 payment to keep using the product. John obliges.

Growth Hack 2

The success among targeted websites in the first growth hack compared to non-targeted social media platforms, suggests a further attempt to grow users among targeted communities would be wise. In this growth hack, influencer marketing is attempted by reaching out to influential YouTubers, as their access to a large audience could be mutually beneficial in propagating Ebisu through a network. Similar to in Growth Hack 1, where multiple websites were categorized by being targeted or non-targeted, and segmented further, a mix of targeted and non-targeted YouTube channels were considered.

An obvious example of an influencer in a targeted segment includes Glutanimate, a well-known and respected addon developer in the Anki community. Glutanimate has a YouTube channel where multiple addons are used and reviewed, including some which he has developed himself. Another similar YouTuber is AnKing, who primarily posts about addons. In both cases, a significant majority of their viewers can be seriously considered potential users.

Other YouTube channels, such as Mike Boyd, or Thomas Frank, have a much more general (but significantly larger) audience. Boyd's content primarily concerns himself attempting to learn various skills (e.g. axe-throwing, learning a new language, making a speech) on video, while Frank's includes general tips on studying and exam preparation. Both have used and promoted Anki on their channel, but because of their broad audience, only a small fraction of users have a strong need to require something like Anki (more so in Frank's case than in Boyd's).

In both scenarios, the approach was the same: reaching out to each of the identified influencers via email, twitter,

instagram, and any other channel available in order to encourage a response. The communications were tailored to each of the influencers in order to appeal to their perceived values and relationship with audience.

For example, one of the influencers was Andrew Ng. Ng is adjunct professor of machine learning at Stanford, business executive, investor, as well as online education pioneer, having cofounded Coursera and published multiple online courses for free. When reaching out to Ng, communications focused on his interests of machine learning and open online education, and Ebisu was presented as a tool that could integrate pedagogy and technology and help him achieve his goal of 'democratizing deep learning' (as he puts it). It was made clear that there was no monetary investment required in this free software, and that a mention on his social media about Ebisu would be mutually beneficial. If successful, Ebisu could easily propagate through the multiple networks of which he is a key influencer: academics (both in the field of education, as well as machine learning) could observe and critique some of the assumptions behind Ebisu, and his students on Coursera would be encouraged to use Ebisu in order to efficiently learn from their online courses. Moreover, as one of the world's most famous computer scientists and as one of Time magazine's 100 Most Influential People, this could allow spaced repetition software to 'go mainstream' and be significantly used worldwide.

	Influencers
No Reply	Thomas Frank, Barbara Oakley, Mike Boyd
Replied, Not Interested	Alex Mullen, Andrew Ng
Replied, Interested	Med School Insiders Glutanimate Ali Abdaal

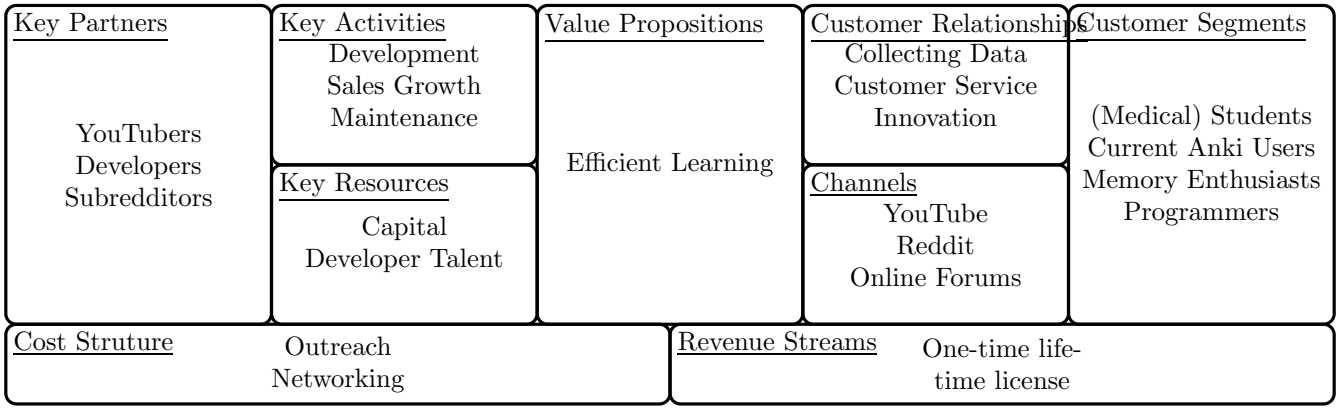
In general, there were no responses from influencers with a large following (such as Mike Boyd And Thomas Frank), and interest was essentially inversely proportional to their reach. Others, such as Andrew Ng, responded, but did not feel comfortable giving me a shout out on their platform. The influencers with which there was any form of success were those who already used Anki regularly and recognized the significance of an optimized learning scheduling, such as Med School Insiders, Glutanimate, and Ali Abdaal. While they commanded relatively smaller audiences, although capturing interest from even a fraction of their following would be more than enough to reach the growth hacking goal.

Ali Abdaal and Glutanimate were interested, but neither were willing to immediately give a shout out to Ebisu on their platform. Instead, an agreement was made: they would experiment and try out my addon in their own routines, and then perhaps give me a shoutout if Ebisu was deemed sufficiently better than the existing algorithm. In communications with the group of 'replied, not interested' influencers, a mistake was potentially made when making a seemingly unbalanced ask to promote Ebisu to their followers without a clear emphasis on quality. Among among those who responded positively, this was made clear: in Abdaal's case, MEMORIZE was presented as a project developed by a fellow student who found a complementary learning technique, and in Glutanimate's case, as a fellow developer who created a tool that could later be integrated in base Anki.

Unfortunately, none of the YouTubers has as yet given Ebisu a shoutout. However, 3 are currently evaluating the

product, and may soon be ready to spread Ebisu to a larger network. While feedback on the addon was agreed to have come before the submission of the final project, the influencers have been slow to reply. As a result, there were no additional users from this growth hack.

Business Model Canvas



Key Partners

We have several Key Partners, with the main ones being YouTubers, developers, and Subredditors. YouTubers increase awareness of our brand and we even explicitly promote our product on their channel. This may or may not include a sponsorship deal with the firm. Developers refers to open-source developers to may create derivative works of our addons (as would be legal under open-source software distribution laws), who we can also commission to develop further addons from us. Members of the Anki subreddit are also considered partners, though much more loosely, as they provide a easy way to engage with customers and bounce ideas off them. Some negotiation will be required to get influencers on board, but since they too have predictable motivations –such as the growth and revenues from their channels– which may be aligned with ours, a mutually beneficial negotiation, whether in the form of a personalized promo-code or free usage of the addon (partners are also potential customers), should be possible. In some cases, partners may not even be interested in monetary compensation, as their primary goal is to promo more efficient education, no matter the form.

Key Activities

I identified 4 main activities: the development, sales, growth, and maintenance of the product. The development will take around 2-4 weeks, while sales, growth, and maintenance are considered perpetual activities. The development of the sales platform (a one-time payment to PayPal), will also occur along with the development, but has a further deadline due to the initial free-trial month. Most importantly, growth refers to outreach activities, such as creating and maintaining relationships with partners and influencers, but also making content to spread the word about the addon, as well as monitoring competitors and the industry in general which may offer obstacles to growth. Maintenance refers to customer support and takes a more passive role, as the whole Anki acts as a crowdsourced customer support agent.

Key Resources

The company is designed to run based only on myself, and those constraints come with it a set of advantages and disadvantages. I am much more flexible than a larger company can be, and I can (and am more or less forced to) operate based on small scale experiments to validate what does and does not work. My extensive involvement with every part of the business will give me a broader vision of the future of the company, and thus key resources involve my own time and capital invested in the business. Realistically, there are business functions I may not be able to immediately grasp, such as the development of the purchasing platform (and corresponding databases), and I can contract some developer talent to help out.

Value Proposition

At the broadest level, the company is trying to provide a means for more efficient learning for the public. For each of the different customer segments, this has different implications: memory enthusiast will be able to remember more factoids more quickly; a medical student will spent less time revising and get better grades on his exams; a programmer, having the relevant documentation in his head, will be better at his job. The company, while for-profit, is also deliberately aiming to maximize a positive impact in the world by improving human education.
At some level, this is considered ethically problematic: the monetization of a good that would universally benefit mankind. Yet because of the open-source nature of the base software, overt attempts to license the product is not possible (and would actually be unethical). The primary aim is to spread awareness of the product, which, even if pirated and distributed for free to public, would provide a social good. Conversely, not attempting to monetize this product would be wastefully leaving money on the table, and imply no attempt to recoup on the labor and capital invested into the project.

Customer Relationships

Many customer relationships will be created and significant effort will be made to maintain them. Partially, this is due to significant overlap between partners and customers: Some partners, e.g. YouTube influencers will also be customers, while some customers e.g. beta testers will also be partners in providing feedback about flaws and improvements to the product. Maintaining relationships includes, but is not limited to answering emails, fixing bugs, and participating actively in community discussions on reddit and other internet forums.

Channels

Channels used will primarily be online in nature: While I expect to be engaging with potential customers in the real world, such as tech conferences where I will talk about my add-on online (such as one organized with London Python in mid-late December), this is not seen as a viable long-term growth strategy (these talks will be recorded and uploaded to YouTube in order to maximize reach). YouTube in particular is used as a channel for the joint purposes of growth and sales; both educating the public on exactly what the addon does and how it affects the Anki algorithm and a quick link to download. Furthermore, because of its large community, ease of use, and availability to get videos directly linked with those of partners/influencers, YouTube makes for an ideal platform. It is also key to embrace the nuances of each channel and medium to provide effective, engaging communication. Videos, for example, allow the insertion of memes to take a more casual tone, the inclusion of a friendly voice and 'demo'-like walkthroughs, and moving graphics and data visualizations. Blog posts, on the other hand can allow for a more lengthy, authoritative tone, copyable code can be neatly displayed for analysis, and equations can be typsetted.

Customer Segments

As identified in previous assignments, the core customer segments include, in order of importance, medical school students, memory enthusiasts, and computer programmers. Previous research showed that each group possesses different optimal needs, and showed that some communities, such as those of the HackerNews readers, may face problems with product-market fit. Specifically, this decomposition of customers in to segments were preferred over others, as they neatly separate users in terms of academic/professional/amateur users, include successful extreme users in all three fields, can be reached through roughly mutually exclusive channels, and possess their own specific lingo.

Cost Structure

Major drivers of cost are to the key activities of growth and maintenance. As potentially more users use the product and more bugs are solved, code maintenance costs will go down. Since the product is software, there are no variable costs, but fixed costs will grow with revenues (but at a much slower rate), and more money will have to be invested into activities like outreach to potential partners and networking.

Revenue Streams

The base app –Anki– is currently free and open-source, however our add-on is not, and relies on a one-month free trial followed by a one-time purchase of $4.99. Additionally, promo-codes are also applicable (e.g. a 20% discount for those following a YouTube channel, or a 100% discount for a private, limited corporate social responsibility campaign). Customers pay on the app, via PayPal. A subscription model, rather than lifetime model, could also, and potentially bring in more money, but this has to be considered in further detail.

Growth Hack 3

The third growth hack is the final iteration of the first two, and attempts to incorporate prior lessons. It was observed in the first growth hack that targeted communities are much more lucrative (in terms of new user signups) than non-targeted social media marketing, where a larger (but much shallower) net is cast. In the second growth hack, contact was made with key influencers in a targeted vs non-targeted audience, and while there were no new downloads, it was established that smaller influencers were more likely to advertise MEMORIZE than larger channels, and again, influencers of targeted communities were more cooperative. In the third growth, a talk was given at London Python about Anki and MEMORIZE, acting as a form of content marketing. Instead of reaching out to an influencer, the talk was given in order to establish a mini-influencer status, and grow MEMORIZE to a small but attentive audience, who could ask questions and receive answers in real-time.

The first event was organized at White Collar Factory, on Old St. in London at the offices of BGL Group in joint partnership with London Python. In total, there were 74 attendees, and the talk went on for approximately 30 minutes. The talk covered multiple aspects of spaced repetition, and the slide deck is available here.

In an attempt to be broadly accessible to the technical audience, the uses and potential of Anki were explored via the lens of influential figures, such as Arthur Chu, a quiz show contest who trained for Jeopardy! Using spaced repetition, and Ali Abdaal, who trained for his Cambridge University medical exams using Anki. The talk also included a personal account of how Anki was used by the speaker to learn German in 3 months, with the specific deck shared with the audience. Audience members were encouraged to download Anki, as well as give immediate

feedback on whether they would use Anki after the talk.

Taking advantage of the technically-comfortable audience, and the time afforded to the speaker, the talk included a brief mathematical primer on how human forgetting and learning is modeled. The speaker briefly walked through some of the basics of our scientific understanding of spaced repetition; including the testing effect (i.e. mock tests are more effective that reading a textbook), the spacing effect (i.e. spreading out a few mock tests in the weeks leading up to a test is more effective than cramming the night before), and the lag effect (i.e. exposure to a test increases the retention of information in memory and decreases the rate of forgetting, even if the test was incorrectly answered), and how all three effects give rise to the 'forgetting curve'. In retrospect, this was necessary for the audience to understand why the current implementation in Anki was inadequate; as it did not conform to what scientists know are the best ways to learn (despite often having been advertised as such).

Immediately following the talk was a Q&A session lasting around 25 minutes. There were questions from multiple previously neglected perspectives: A young lady asked if Anki could be used to aid her grandfather –who had alzheimer's– in storing basic facts in long-term memory. Another university student asked if they could use Anki in training for their pub quiz (an emphatic yes, although he would have to find an online question bank to scrape from if he didn't want to create the flashcards himself), which suggests that a community of quiz bowl enthusiasts could be a significant, previously neglected customer segment.

Based on the Q&A sessions, and the approximately 20 people who approached the speaker after the talk, the impression is that the utility of spaced repetition was successfully conveyed to the audience. Some audience members showed excitement and determination in learning more about how to use Anki and its addons. Others were upfront about probably not using Anki for various reasons: they were either not technically comfortable with adjusting some of the settings without supervision, did not have the time to create flashcards for them to review, or did not even have the time (20-30 minutes per day) to review their own flashcards. It became obvious that one of the hurdles of Anki was the habit formation required on part of users, and perhaps companies in similar spaces could offer insights (e.g. meditation apps or to-do lists which require consistent daily use to be effective in changing habits).

Analysis and Conclusion

From the 3 growth hacks, a total of 132 signups were made for the addon (122 from GH1, 0 from GH2, and 6 from GH3), which surpasses the originally set goal of 50 signups. The previously neglected user segment of quiz show enthusiasts could be pursued in an analogous fashion to the three previous growth hacks: making posts on various quiz and trivia shows, such as (r/quizbowl, r/trivia, r/universitychallenge). Rapid adoption by influential members in various quizzing communities at all levels (secondary school, undergraduate, professional etc.) could offer the ability to quickly propagate through multiple networks. A possible direction of future work could include training for, and winning a quiz bowl, and revealing afterwards that the quiz was trained for using Anki and Ebisu. This would be a clear and easily sharable way of demonstrating how the average person could practically use spaced repetition to achieve success in their daily lives.

5 Appendix

LOs: Universal

#planningarchitecture

Throughout this project, I developed timelines to guide important project goals and deadlines. I have never worked on a project this large before, and in order to work efficiently, I used systems, such as Mendeley and BibTex for my references, and a google doc to organize my writing into several chapters which I could work on individually. I had accountability groups - aside from my capstone group and my professor, who I did not want to disappoint, I made a verbal agreement with my roommates to keep each other accountable during the final weeks before the capstone deadline. This was essential as the global coronavirus pandemic had canceled some of my plans for the capstone: for example, I had planned to speak at SXSW.

#accountability

I demonstrated accountability throughout my growth hacks, all of which required significant amounts of self-driven work. I had to create my own websites, make posts on social media, connect with influencers (making sure my responses and replies were timely), and find systems to mitigate my own inertia. One of the systems I found helpful was the reliance on social pressure created by others, which forced me to be accountable. For example, for Growth Hack 3, I pushed myself to ask to speak at London Python and asked some of my classmates to attend the event, knowing that I would have to submit a quality piece of work in order to not disappoint them. I also found ways to integrate capstone with my academic journey, with some of the content in this capstone being highly related to B144 and B154. This pushed me to work towards my capstone even in order to succeed academically.

#qualitydeliverables

There were two quality deliverables; the Anki addon, and this capstone itself. The justification for the depth, scope and rigor of the addon is explained in section 3 under the Ebisu section, and the justification for the whole capstone itself is covered in section 2, which explains the approach of design thinking.

#connect

I reached out to multiple people during the course of this capstone, and in each case, I made sure put in in-depth research. When reaching out to Ahmed Fasih in order to ask for help on Ebisu, for example, I made sure to thoroughly read his documentation on GitHub and confirm my understanding that Ebisu would be appropriate for Anki. When reaching out to Arthur Milchior, I leveraged my own personal networks in the Bay Area rationalist community, listened to his concerns (e.g. making sure that the addon was 'reversible') and compensating for his time as a consultant. In reaching out to academics, I leveraged my own networks, such as Stephen Kosslyn, who I personally met from my time at Minerva, and professor Michael Mozer, who I connected with through a classmate.

#research

Research is demonstrated perhaps most clearly in section 1, where I explain the research behind spaced repetition, and breaking it down into the testing effect, the spacing effect, and the lag effect. I thoroughly explained the research behind all three, and demonstrated this again in section 3, where I looked at existing research and implementation of spaced repetition. Research was also used in a lesser extent in section 2, where I explained the background behind

design thinking, and in section 4, where each of the growth hacks were themselves a resource, and thus I had to execute the growth hack that would provide the greatest learning experience.

#feedback

Because some parts of the capstone, such as the design thinking tools and the growth hacks were covered in B144 and B154, I received real-time feedback from my classmates about my capstone which I incorporated. For example, in the beginning I had hoped to center my capstone around the addon, but my classmates made me realize that there was a lot more to write about the importance of using 'design thinking' as opposed to 'scientific thinking', because the latter had not yet produced practically useful results. Feedback from experts were also incorporated along the way - at first I had planned to make an addon implementing the MEMORIZE algorithm, but Ahmed Fasih convinced me that the supposed optimality was only relevant for a specific model of stochastic reviews over an infinite timeline that may not be practically relevant. Instead, I was encouraged to use Ebisu, and Milchior also helped to steer me in this direction by noting that Anki had limited support for outside packages, which would be necessary for an implementation of MEMORIZE.

#metrics

Metrics were used in the form of the number of downloads for the addon. While the addon was still in development, I used a waiting list that users could sign up for and be notified when the addon had been released, which allowed me to gauge interest. During each growth hack, I used metrics to evaluate my work.

LOs: Project Specific

#metrics (B154)

Metrics were thoroughly used in the three growth hacks. The KPI developed was that total number of unique email signups on the mailing list, and then the number of unique downloads. This was separated into 5 categories, tracked via different url links. Moreoverm, the 'unique' part was especially important to satisfy the assumption of the chi-square test that all of the categories were mutually exclusive.

#inspiration (B144)

I explain the details of a user need based on market research in section 2, where I explained my own experiences, framed my design challenge in the context of user needs, and then conducted research to identify where different groups of customers (e.g. Anki subredditors, memory enthusiasts, hackernews readers) might want to use my addon.

#implementation (B144)

I conduct user testing for feedback in section 4, where I center my research focus around the question, 'how might we make reviewing and testing rewarding instead of stressful?' and gathering insights via an online survey, and a case study of an extreme user. I create a customer journey map after growth hack 1 to empathize with the customer, and a business model canvas after growth hack 2 to see how my solution would interact with a potential customer.

#pythonimplementation (CS146)

In section 3, I demonstrate the derivation for a statistical model, Ebisu, and implement it in Python, using code comments to justify the steps in the addon implementation.

#models (CS111)

In section 3, I analyze the existing Anki algorithm, the half life regression model, the MEMORIZE model, and Ebisu. In each case, I describe the domain of validity of the model, and relate them to practical issues. For example, the threshold model, HLR, and Ebisu assume binary results for flashcards, while in Anki there are 4 difficulty buttons. While the latter would seem to be more 'realistic', it is much harder to solve mathematically. Furthermore, I mention how while MEMORIZE is stochastically 'optimal', it is only optimal within the model of people who are willing to review their cards at any time.

HCs: Foregrounded

#professionalism

Given the context of my capstone as an intersection of the relatively new fields of the learning sciences (itself a combination of computer science, educational psychology, and cognitive neuroscience), growth hacking (as a subfield of marketing), and design thinking, I use appropriate forms of speech by following academic conventions for citations and attributions (APA 7), while embracing the human-centered spirit of design thinking by conveying complex thoughts with minimal jargon. Furthermore, I demonstrate a deep grasp of nuanced conventions most evidently when citing 'non-traditional' sources, such as blogs and Reddit Posts.

#organization

I effectively organize my capstone by breaking down each part of my capstone into disjoint parts, while ensuring that each part leads naturally to the next. This is clearest in Chapter 1: Background, where I situated spaced repetition in the context of the science of learning, explained what spaced repetition is and why it works, and ending by re-situated the spaced repetition in the context of today's algorithms, and finally ending with a call to action.

#evidencebased

Evidence is used thoroughly in this capstone in order to justify arguments. In section 1, evidence was used in order to show that each individual component of spaced repetition (the testing effect, the spacing effect, and the lag effect), was supported by the scientific literature, as well as the efficacy of spaced repetition as a whole (e.g the example of Ali Abdaal, Robert Craig). Evidence was also used to show that despite the clear benefits of spaced repetition, which are known to students and instructors alike, there was a widespread failure to apply spaced repetition. In section 3, evidence was used to determine which algorithm was best to implement. This evidence comes both empirically, in surveys of people asked, as well as mathematically, with a proof of optimality of the MEMORIZE algorithm for stochastic reviewing.

#sourcequality

Closely linked to evidencebased, I attempt to find and use the most relevant, current, accurate, authoritative, and purposely sources. In section 1, which serves as a background and Literature Review, I structure the development of my arguments around as Dunlosky et al [9], Cepeda et al [6], and Roedig and Karpicke [25], which are highly influential sources with over thousands of citations. When sources may have had issues, I evaluated its quality: in Ebbinghaus [11], I noted some of the disadvantages, such as the its incomplete and single-subject nature, and contrasted with its advantages, such as how it modernized the field, by using modern statistical analysis tools and attempting the scientific method, and even how it was replicated in modern times [21].

#responsibility

#responsibility was touched on explicitly in the text in section 3 and section 4, when I had to be responsible and self-directed in conducting market research and growth hacks. I explained how responsibility was important in section 3: technical implementation, in able to coordinate the work of both Arthur Milchior and Ahmed Fasih in constructing a mini-team in order to create the Anki Add-on for Ebisu.

#scienceoflearning

This HC is foregrounded, and applied thoroughly in section 1 and section 3.

HCs: Project Specfic

#designthinking

Design thinking was used throughout the capstone, including the framing of the capstone itself in the 'frame your design challenge', in the customer journey map, in the creation of the business model canvas, and in the inspiration, ideation, and implementation stages leading up to and including the growth hacks. Justification for why design thinking is relevant to the capstone, compared to 'scientific thinking' is given in section 2.

#constraints

Constraint satisfaction is clearly used in the choosing of which algorithm to implement in Anki in section 3, where Ebisu was chosen over the other algorithms because of its easily tractable nature, the ability to perform on-line learning on the level of the individual (discounting half life regression), the requirement for Anki to use few external pacakages (discounting MEMORIZE), and legal considerations (discounting supermemo algorithms).

#studyreplication

This was used in the analysis of the Murre and Dros study, which replicated Ebbinghaus' seminal self-experiments. I explain why even though Ebbinghaus' original study was influential, the fact that it was a self-experiment and had only a single subject could make it biased. To establish if it was biased or not, a replication was necessary, and Murre and Dros' replication empirically vindicated Ebbinghaus' original data. I also made sure to note that there were some deviations from Ebbinghaus' original procedures, but not to the extent to invalidate the replication.

#emotionaliq

I used my own self knowledge of the emotions that potential users of Anki would go through, drawing from my own experiences to identify the emotions of curiosity, excitement, disillusionment, and despair (akin to a hype cycle), and apply this in the customer journey map in section 3.

#audience

I explain the perspectives of my target audience in the various growth hacks. For example, I demonstrate understanding of the memory enthusiasts, and tailor my message towards them by mentioning Arthur Chu, while I tailor my message to medical school students by offering links to USMLE decks which they can easily download and start using.

#multipleagents

In section 2 and section 4, I explore the decomposition of the customer/user segment into the three categories of 'medical school students', 'memory enthusiasts', and 'programmers', and –in the customer journey map– show the dimensions along which there are separations as well as similarities.

#ethicalconflicts

In the business model canvas in section 4, I explored the question of if it was moral to charge money for a service that should be easily available to everyone. While the addon is ultimately free, I considered the possibility of a CSR campaign where people who cannot afford the product can simply be given a lifetime license, though this process would still involve an element of exclusivity in practice.

#interviewsurvey

In section 2, I explained the underlying concepts of using an online survey, explained the results and how they served only as a basis for asking and answering further questions. The survey was clearly effective as it provided new questions and even solutions. I also demonstrated how I designed my survey questions to be as unbiased as possible, such as the switch from 'which addons do you use?' to 'if you use any addons, which? And if not, why not?'

#purpose

In section 2, and especially during the 'frame your design challenge', I explain my my underlying goals and values of aiming to provide spaced repetition tools to students around the world in the form of an addon, with the assumption that they are indeed more effective than most types of 'learning' conducted in modern institutional settings.

References

[1] Gordon A Allen, William A Mahler, and WK Estes. Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 8(4):463–470, 1969.

[2] Stephen Baker. *Final Jeopardy: Man vs. machine and the quest to know everything.* Houghton Mifflin Co., 2011.

[3] Tim Brown et al. Design thinking. *Harvard business review*, 86(6):84, 2008.

[4] Andrew C Butler. Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5):1118, 2010.

[5] Shana K Carpenter, Harold Pashler, and Nicholas J Cepeda. Using tests to enhance 8th grade students' retention of us history facts. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 23(6):760–771, 2009.

[6] Nicholas J Cepeda, Harold Pashler, Edward Vul, John T Wixted, and Doug Rohrer. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological bulletin*, 132(3):354, 2006.

[7] Sarah J DeLozier and Matthew G Rhodes. Flipped classrooms: a review of key ideas and recommendations for practice. *Educational Psychology Review*, 29(1):141–151, 2017.

[8] Frank N Dempster. The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, 43(8):627, 1988.

[9] John Dunlosky, Katherine A Rawson, Elizabeth J Marsh, Mitchell J Nathan, and Daniel T Willingham. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1):4–58, 2013.

[10] David Dunne and Roger Martin. Design thinking and how it will change management education: An interview and discussion. *Academy of Management Learning & Education*, 5(4):512–523, 2006.

[11] Hermann Ebbinghaus. *Über das gedächtnis: untersuchungen zur experimentellen psychologie.* Duncker & Humblot, 1885.

[12] Lissa Galluccio and Carolyn Rovee-Collier. Nonuniform effects of reinstatement within the time window. *Learning and Motivation*, 37(1):1–17, 2006.

[13] Arthur M Glenberg. Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, 7(2):95–112, 1979.

[14] Sheena A Josselyn, Chanjun Shi, William A Carlezon, Rachael L Neve, Eric J Nestler, and Michael Davis. Long-term memory is facilitated by camp response element-binding protein overexpression in the amygdala. *Journal of Neuroscience*, 21(7):2404–2412, 2001.

[15] Sean HK Kang, Robert V Lindsey, Michael C Mozer, and Harold Pashler. Retrieval practice over the long term: Should spacing be expanding or equal-interval? *Psychonomic bulletin & review*, 21(6):1544–1550, 2014.

[16] Sean HK Kang, Kathleen B McDermott, and Henry L Roediger III. Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19(4-5):528–558, 2007.

[17] Nate Kornell, Alan D Castel, Teal S Eich, and Robert A Bjork. Spacing as the friend of both memory and induction in young and older adults. *Psychology and aging*, 25(2):498, 2010.

[18] TK Landauer, RA Bjork, MM Gruneberg, PE Morris, and RN Sykes. Practical aspects of memory. *M. Gruneberg, P. Morris, & R. Sykes (Eds.)*, pages 625–632, 1978.

[19] Carla Margulies, Tim Tully, and Josh Dubnau. Deconstructing memory in drosophila. *Current biology*, 15(17):R700–R713, 2005.

[20] Randolf Menzel, Gisela Manz, Rebecca Menzel, and Uwe Greggers. Massed and spaced learning in honeybees: the role of cs, us, the intertrial interval, and the test interval. *Learning & Memory*, 8(4):198–208, 2001.

[21] Jaap MJ Murre and Joeri Dros. Replication and analysis of ebbinghaus' forgetting curve. *PloS one*, 10(7), 2015.

[22] Paul Pimsleur. A memory schedule. *The Modern Language Journal*, 51(2):73–75, 1967.

[23] Horst WJ Rittel and Melvin M Webber. Dilemmas in a general theory of planning. *Policy sciences*, 4(2):155–169, 1973.

[24] William A Roberts. Short-term memory in the pigeon: Effects of repetition and spacing. *Journal of Experimental Psychology*, 94(1):74, 1972.

[25] Henry L Roediger III and Jeffrey D Karpicke. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science*, 17(3):249–255, 2006.

[26] Burr Settles and Brendan Meeder. A trainable spaced repetition model for language learning. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1848–1858, 2016.

[27] Herbert A Simon. *The sciences of the artificial*. MIT press, 2019.

[28] Nicholas C Soderstrom and Robert A Bjork. Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, 73:99–115, 2014.

[29] Jonathan A Susser and Jennifer McCabe. From the lab to the dorm room: Metacognitive awareness and use of spaced study. *Instructional Science*, 41(2):345–363, 2013.

[30] Michael A Sutton, Jasmine Ide, Sarah E Masters, and Thomas J Carew. Interaction between amount and pattern of training in the induction of intermediate-and long-term memory for sensitization in aplysia. *Learning & Memory*, 9(1):29–40, 2002.

[31] Behzad Tabibian, Utkarsh Upadhyay, Abir De, Ali Zarezade, Bernhard Schölkopf, and Manuel Gomez-Rodriguez. Enhancing human learning via spaced repetition optimization. *Proceedings of the National Academy of Sciences*, 116(10):3988–3993, 2019.

[32] Will Thalheimer. Spacing learning events over time: What the research says. *Retrieved March*, 21:2007, 2006.

[33] Thomas C Toppino. The spacing effect in preschool children's free recall of pictures and words. *Bulletin of the Psychonomic Society*, 31(1):27–30, 1993.

[34] Carolyn L Vash. " the spacing effect: A case study in the failure to apply the results of psychological research": Comment. 1989.

[35] Gerd Waloszek. Introduction to design thinking. *SAP AG, SAP User Experience*, 2012.

[36] John T Wixted, Shana K Carpenter, et al. The wickelgren power law and the ebbinghaus savings function. *Psychological Science*, 18(2):133, 2007.