Practical Considerations for Independent Student Learning: A Research Synthesis

Justin Sung

iCanStudy ™ Proprietary Limited

Author note

Justin Sung is the co-founder and head of learning at iCanStudy ™ Proprietary Limited, which sells evidence-based studying techniques. Findings from this report influence the design, development and presentation of products. This work is the intellectual property of iCanStudy Pty Limited and is protected under copyright laws.

Research Synthesis

Students engage in various activities in an attempt to increase their knowledge processing and retention. The process of acquiring knowledge is called learning. This research synthesis will review the literature on memory; encoding; retrieval; inquiry-based learning; and notetaking. Through examining the state of the science in these fields, we will derive the practical implications, gaps in research and potential hypotheses for a consilient approach to study technique development. Notably, this synthesis will take a practice-oriented view with emphasis on research implications for individual students, rather than for educational institutions.

Memory and Cognitive Load

Some sensory information we consume is retained while some is forgotten. If we can determine why we forget some information while retaining others, we can theoretically calibrate all subsequent studying techniques to these principles. Improving memory with a reduced reliance on repetition would help make studying less tedious, monotonous and time-consuming. Furthermore, the meta-cognitive knowledge of effective learning can be anticipated to improve student empowerment and facilitate student autonomy. This section will examine the nature of memory and establish some guidelines for creating and evaluating study techniques. We will draw significantly on the contributions of Professor John Sweller, the pioneer of cognitive load theory (CLT).

Human Cognitive Architecture and Cognitive Load Theory

CLT leverages evolutionary theory to create instructional procedures based on human cognitive architecture (Sweller, 2011). It expands on the work by Geary (2008), who categorised knowledge into biologically primary and secondary. Biologically primary knowledge is trans-generational knowledge, including skills such as listening to speech or recognising faces (Geary & Geary, 2007). This information may be genetically transferred and can be acquired easily without explicit instruction. However, in some cases, students may need to be instructed that certain transferable cognitive skills (known as generic-cognitive skills), such as metacognition, can be applied in a domain-specific context. Still, it is thought that the generic-cognitive skill itself does not need to be taught (Youssef-Shalala et al., 2014).

On the other hand, biologically secondary knowledge is acquired through instruction. For example, almost all domain-specific subject material learned in formal education is secondary knowledge. The characteristics and processes behind acquiring biologically primary knowledge are different from those for acquiring biologically secondary knowledge. As an important example, the current known limitations of human working memory are only relevant to acquiring biologically secondary knowledge and skills (Sweller, 2011). As explained by Sweller (2016b), once successfully acquired, this biologically secondary knowledge is stored in our long-term memory via elements of human cognitive architecture, which appear to follow several empirically discovered principles. More recent explanations of CLT emphasise the evolutionary basis of psychology, which assists in generating new hypotheses and explaining observations (Sweller, 2016a); however, it will not be discussed in detail in this report as its pragmatic implications are less immediately relevant.

The Information Store Principle

The information store principle dictates that a large volume of information is stored in aspects of our cognitive architecture through long-term memory. Though ontologically simple, it is important to understand that practical differences in knowledge and skill are influenced by differential knowledge in long-term memory as well as working memory capacity. For example, an expert with a large body of long-term memory in a domain is likely to overwhelm the advantage of a beginner with superior working memory through a higher proportion of automatic pattern recognition (system 2) thinking, compared to the beginner's reliance on slower problem-solving and cognitive load intensive (system 1) thinking

(Kahneman, 2011). On the other hand, two individuals with similar levels of knowledge could have different levels of skill due to differences in their respective working memories. This dynamic relationship between working memory capacity and prior long-term knowledge has been demonstrated in many studies spanning a wide range of domains and disciplines (Chiesi et al., 1979; Egan & Schwartz, 1979; Ericsson & Charness, 1994; Jeffries et al., 1981; Meinz & Hambrick, 2010; Simon & Gilmartin, 1973; Sweller & Cooper, 1985).

The Borrowing and Reorganising Principle

Most biologically secondary information is taken or "borrowed" from the long-term memories of others. While the act of taking information from others is biologically primary, the techniques we might use to facilitate this, such as notetaking, are biologically secondary. Once information has been borrowed, it is then reorganised. Information that conforms with previously held memories becomes enhanced, while information that does not align is diminished (Bartlett & Bartlett, 1995). Whenever information is recalled, new information is combined with existing information in our long-term memory, undergoing a constant process of recombination and reorganisation. Learning techniques that facilitate this borrowing and reorganisation seem effective, exemplified by the worked examples instruction technique outlined below (Renkl, 2014).

The Randomness as Genesis Principle

The borrowing and reorganising principle accounts for the majority of how biologically secondary information is acquired, whereas the randomness as genesis principle accounts for how this information was initially created. Drawing parallels with the concept of random genetic mutation in the process of natural selection, human problem solving is theorised to follow two potential pathways (Sweller, 2016b). First, if a problem is recognised via the patterns stored in long-term memory, the appropriate pattern can be retrieved and applied in the situation, similar to a template. Second, if no such pattern exists, a generate and test approach is followed, commonly as part of a means-end problem-solving strategy (Newell & Simon, 1972). In this strategy, possible moves are chosen and tested for effectiveness. If the tested solution reduces the difference between the current problem state and the goal state, it is accepted and stored in long-term memory.

There are strong parallels between the process implied by the randomness as genesis principle and the prolific experiential learning cycle by Kolb (2014). Interestingly, there are also many conceptual parallels with the popular inquiry-based learning approach (Khalaf, 2018), which Sweller (2021) himself admonished due to an apparent lack of both theoretical and empirical support. We will later reconcile some of these differences by explaining how these two approaches may not be mutually exclusive. In our experience, with novel modifications to the implementation procedure, these strategies can augment each other.

The Narrow Limits of Change Principle

The inherent limitations of the working memory mean that it is not viable to process massive volumes of novel information at once. This restricts the number of combinations of testing possibilities that occur during randomness as genesis. Though Sweller (2016b) describes this dynamic as the working memory protecting long-term knowledge structures from unmitigated, potentially harmful changes, there is no clear evidence to suggest that this evolved as a primarily protective, naturally selected feature. Instead, it may simply be that inherent limitations of the human working memory prohibit large volumes of information from being processed, with protection being a positive side effect. However, understanding the origin of this dynamic is functionally irrelevant, and it must simply be acknowledged that dramatic changes in knowledge structures are not possible due to a cap set by the working memory. This limitation has potentially significant implications when considering the formation of new techniques, which will be discussed later.

The Environmental Organising and Linking Principle

This principle describes the characteristics of memory when retrieving information that has already been encoded into long-term memory. While the narrow limits of change principle prevent massive amounts of data to be encoded at once, there are no clear limits to the amount of data that can be retrieved and utilised. As a result, the human brain appears capable of retrieving remarkable quantities of information and manipulating this information for a diverse range of applications, depending on the environmental signal and trigger that necessitated the retrieval in the first place (Sweller, 2016b). We hypothesise that working memory limitations during encoding can be functionally overcome through frequent cycles of encoding and retrieval, leveraging off the ability to mobilise vast volumes of information. This then can facilitate the encoding of novel information in a positive feedback cycle. This hypothesis will be developed further with our studying system.

Application of Cognitive Load Theory

Sweller (2016b) posits that extraneous cognitive load must be reduced through explicit instruction to overcome working memory limitations. This position is strongly supported by research on effective instructional techniques. For example, in a review by Renkl (2014), the usage of multiple worked examples was significantly superior in efficacy and time-efficiency compared to observational learning or analogical problem-solving. In the worked examples model of learning, skills are initially acquired through reviewing multiple problems with worked solutions. It is hypothesised to reduce the extraneous and learning-irrelevant cognitive load, thereby improving the learner’s ability for schema construction (Renkl, 2014). This model follows the skills acquisition phases, initially outlined by VanLehn (1996), of (a) principle encoding; (b) learning to solve problems and repairing knowledge gaps; and (c) automation. However, explicit instruction from only the instructor is insufficient for effective learning.

As Renkl (2014) notes, the effectiveness of using worked examples for learning depends on the learners' own ability to elaborate and rationalise the solutions to themselves. Students who can self-explain, either spontaneously or through prompting, were found to be superior at solving novel problems (Hilbert & Renkl, 2009; Renkl, 1997). Furthermore, spontaneous self-explanations may also be superior to prompted self-explanations (Hilbert & Renkl, 2009; Schworm & Renkl, 2006). The factors that influence a students' ability to spontaneous self-explain are countless and frankly not clearly mapped in the current research landscape. A dominant consideration is the balance between extraneous and intrinsic cognitive load, which we have briefly discussed so far. This optimisation of intrinsic load is a threshold concept that will guide all techniques discussed later.

In actual learning practice, numerous variables interact with each other, finely modifying the cognitive load at any given time. Some of these combinations produce predictable effects that have been defined and named. For example, the split-attention effect describes the increase in extraneous cognitive load when a learner splits their attention between multiple sources of information that must be integrated (Sweller, 2011). The redundancy effect describes when elements in the learning material are unnecessary to encode information, and extraneous load is required to process and distinguish necessary from unnecessary information (Sweller, 2011). Sweller (2016b) notes that most cognitive load effects, in practice, result from improper instruction creating extraneous load.

In contrast, other cognitive load effects are secondary to the optimisation of intrinsic cognitive load. For example, the variability effect describes when intrinsic cognitive load is enhanced through variable elements in the learning material or experience, improving learning outcomes (Paas & Van Merriënboer, 1994; Sweller, 2011). This phenomenon shares striking similarities to the highly researched and established practice of interleaving for skills development, which also has similar empirical findings (Taylor & Rohrer, 2010). Thus, variability effect on cognitive load may partially explain the efficacy of interleaving. Similarly, the generation effect occurs when learners have greater test performance by generating their own responses rather than being given instructional guidance (Chen et al., 2015). Where many elements are interacting with each other in the learning material, guidance can be beneficial to help manage the high cognitive load, while it can be harmful when working memory load is light (Sweller, 2016b). This reduction of cognitive overload seems to disproportionately affect students with a low level of prior knowledge, while those with higher prior knowledge are harmed by reducing intrinsic cognitive load (Ayres, 2006). Students engaging in generative learning are exposed to confusion and discomfort inherent to higher levels of cognitive load. Consequently, they can report a preference for non-generative instructional guidance (King, 1992; Wittrock, 1989), although their results may be significantly worse without generative techniques (Ritchie & Volkl, 2000; Wittrock, 2010). This is explained by the illusion of fluency, whereby individuals tend to overstimate the depth of their knowledge (Carey, 2015) and the belief that mastery over something has been achieved when it has not (Lang, 2016). The Dunning-Kruger effect (Dunning, 2011) is also likely to influence the general lack of metacognition by most students or even educators.

Of particular practical importance is the expertise reversal effect. This effect describes the reduction or even reversal of an instructional procedure’s efficacy when the learner has a higher level of expertise. While high element interactivity can increase intrinsic load in novices and create superior learning outcomes, for example with worked examples, the level of interactivity between elements reduces as an individual gains expertise, causing an effect reduction and eventual reversal (Kalyuga et al., 2001). This effect occurs because novices may perceive information as isolated individual elements, imposing a high cognitive load on the working memory. On the other hand, experts can see multiple elements as a simplified, single element, reducing working memory load. Thus, high element interactivity, which induces a high working memory load, can be reduced by greater domain expertise.

As such, a snowball effect is hypothesised for learning, where greater domain expertise increases the ability to process larger volumes of information. These observations of working memory are directly mirrored and extended with research on chunking theory, which describes a method of functionally bypassing working memory limitations through recoding of information into larger, related units, called “chunks” (Gobet & Clarkson, 2004; Gobet et al., 2001; Thalmann et al., 2019). Related to the expertise reversal effect and element interactivity effect is the isolated elements effect, which states that if information causes cognitive overload, it may need to be isolated and broken up, then reconstituted later from individual, isolated elements (Sweller, 2016a). We will modify this sequence in a novel way to reduce the amount of unnecessary isolation in our study system, while facilitating the snowball effect.

The transient information effect is relevant for modern formal learning. This effect describes the transient nature of information consumption in presentations or lectures. When presented in written form, complex information can be recoded more carefully, and cognitive load can be reduced to more optimum levels (Sweller, 2016a). However, this is not often in the learner's control, and therefore students should engage in preparatory strategies to reduce cognitive load, even without comprehensive written material being available. In our study system, students can functionally bypass the limitations of the transient information effect by increasing their expertise through semantic priming.

Additional Considerations for Cognitive Load Theory in Practice

There are a number of important findings to help guide the application of CLT. Firstly, several studies have demonstrated that total cognitive load can be reliably measured through subjective rating scales (Paas et al., 2003), with a high level of sensitivity for small differences in cognitive load (Paas et al., 2005; Van Gog et al., 2012). Secondly, there is no substantial evidence to suggest that working memory capacity cannot be trained, though it is an underlying assumption behind most cognitive load research (Sweller, 2020). This assumption seems unlikely given the research on neuroplasticity indicating that the human brain can undergo an astonishing level of adaptation (Bruel-Jungerman et al., 2007; Sagi et al., 2012). Finally, visual encoding appears to be less restricted than verbal encoding, with cognitive resources required to encode remaining stable, even if the pace of information is increased beyond optimal verbal encoding limits (Lang et al., 1999). We will leverage these effects in our studying system.

Retrieval, Active Recall and Spacing

Retrieval practice refers to any activity requiring learners to recall previously encoded information from their long-term memory. There is little doubt that retrieval is superior to restudying, as demonstrated by multiple meta-analyses (Adesope et al., 2017; Rowland, 2014). Moreover, decades of research have shown that the spacing of retrieval episodes improves knowledge retention (Latimier et al., 2021). However, the effects of spacing and the limits of its effectiveness are far from clear. In this review, I will briefly summarise some of the salient findings of spaced retrieval practice while emphasising the substantial limitations when making practical recommendations on real-world effectiveness. Note that a deeper understanding of the many potential theoretical explanations for the spacing effect does not offer a practical advantage at this time, so this discussion will be omitted from this review. Based on recent meta-analyses, the following statements about spaced retrieval are empirically supported.

- Retrieval practice is more beneficial when cognitive load is higher and more demanding during retrieval (Adesope et al., 2017).

- Gaining correct answer feedback on high-confidence correct answers can be an inefficient use of time for students (Hays et al., 2010).
- Students' predictions of recall performance do not correlate with actual performance (Karpicke & Roediger, 2008).
- Receiving delayed feedback on answers may be more beneficial for long-term memory than delayed, but only when the delayed feedback is thoroughly examined (Butler et al., 2007). In general, receiving feedback at all only has a slightly greater effect compared to no feedback (Adesope et al., 2017), and retrieval seems to be beneficial, regardless of whether it was correct or not (Carneiro et al., 2021).
- Retrieval with testing is more effective than restudying by a moderate effect size of g = 0.51, which "is arguably the most accurate indicator of the benefits of retrieval practice" (Adesope et al., 2017).
- Students commonly use rereading as a strategy (Karpicke et al., 2009) which is ineffective at improving test outcomes or retention (Callender & McDaniel, 2009). However, using rereading increases students' false sense of mastery and overconfidence (Koriat & Bjork, 2005).
- Both adolescent and adult learners are overwhelmingly unable to identify effective revision techniques. They are much more likely to identify easier revision techniques as more effective (Birnbaum et al., 2013; Kornell et al., 2010; Zulkiply & Burt, 2013), even when directly informed on which techniques are objectively superior (Logan et al., 2012; Simon & Bjork, 2001).
- Expanding spacing intervals does not seem to clearly improve learning outcomes compared to fixed spacing schedules (Latimier et al., 2021); however, the research is heavily conflicted, and conclusions cannot be made.

- Some studies suggest that when knowledge must be retained for longer, longer spacing intervals may be associated with better test performance (Cepeda et al., 2009; Greving & Richter, 2018).
- The positive effects of spacing seem to be present across age groups, domains, knowledge levels and working memory capacities (Adesope et al., 2017; Latimier et al., 2021).
- Retrieval seems to enhance learning new and novel information, even after interpolated testing stops (Chan et al., 2020). This is attributed to the forward testing effect (Chan et al., 2018; Kliegl & Bäuml, 2021).

Limitations of Current Research

Several limitations to the current state of science on spacing impose considerable caveats to its application. Most critically, the vast majority of studies examining the effect of spacing are laboratory-based (approximately 89%) compared to classroom-based (Adesope et al., 2017). Classroom studies have much higher variability in study time, extrinsic motivation and multiple interference variables (Roediger & Karpicke, 2006). Fewer studies still examine the effect for students using spaced repetition techniques independent of class (Adesope et al., 2017). As Latimier et al. (2021) remark in their meta-analysis of spaced retrieval practice, “diversity of experimental settings (particular stimuli, test types, population) was limited, making it impossible to fully address the moderating effects of these factors”.

This lack of evidence in real-world academic settings is relevant from a student’s perspective, as the student is not in control of how their class is facilitated. Thus, even if spaced testing is effective in classroom settings, this does not help a student in a class that does not facilitate spaced testing. In addition, laboratory testing and even classroom studies do not sufficiently account for the range of other academic pressures a student must navigate, including the presence of multiple subjects with varied methods of instruction or time spent on homework, which may in itself be harmful to the student (Fernández-Alonso et al., 2017).

Furthermore, most studies do not test retention after a testing retrieval episode longer than one day, with longer time-delay studies measuring still less than one week (Adesope et al., 2017). Studies measuring the effect of spaced retrieval across weeks in realistic educational settings are exceedingly rare across several decades of studies (Carpenter, 2017). The studies that have measured across longer intervals do sufficiently demonstrate some form of long-term benefit (Bahrick et al., 1993; Bahrick & Hall, 2005; Rawson & Dunlosky, 2013), but the character of this benefit is still unclear. For example, Smith and Scarf (2017) reviewed studies of spaced retrieval across exclusively longer time scales for language learning. They found patterns not present in short-term data, such as the lack of benefit of spacing to help learn words or grammar among adults. To my knowledge, no studies have examined the effect of spaced retrieval in a setting that matches all of the following conditions, despite these conditions being representative of reality for nearly all secondary and tertiary students:

- adolescents or young adults;
- realistic educational settings;
- realistic assessments;
- spacing intervals across weeks;
- multiple simultaneous subjects.

In addition, although spaced retrieval is undoubtedly effective, the magnitude of this effect may be exaggerated by popular media. The vast majority of studies compare the usage of spaced retrieval to no spaced retrieval or non-testing study activity such as rereading (Adesope et al., 2017). Even when spaced retrieval is shown to be significantly effective, individual results range widely, with results for some individuals who use spaced retrieval being lower than those who do not use it (Adesope et al., 2017). The size of beneficial effects are mostly moderate. When publication bias or moderating factors are accounted for, there are very few large ($g \geq 0.8$) or very large effect sizes ($g \geq 1.3$), even in laboratory and controlled classroom studies. Larger effect sizes are often countered by significant heterogeneity between studies, sometimes showing a small effect for the same outcome (Latimier et al., 2021). Ultimately, the moderate tendency of effect sizes suggests that it is unlikely that an individual will receive a dramatic improvement in any learning outcome using primarily spaced retrieval. The wide variability in results has not been explained or reconciled (Latimier et al., 2021), indicating that while spaced retrieval is effective at a population level, moderating factors have considerable influence on individual results.

One of the most significant moderating factors may be cognitive load. However, despite the relevance of cognitive load theory on memory, relatively few studies compare the efficacy of spaced retrieval against working memory load. In a Chinese laboratory study of 1032 university students, C. Yang et al. (2020) demonstrated that spaced testing has the greatest effect on those with low working memory capacity compared to those with high working memory capacity. This finding is consistent with those of a prior study by Agarwal et al. (2017) in a laboratory study of 166 students from Washington University. Notably, the forward testing effect in isolation has been shown to be effective, independent of the working memory capacity (Pastötter & Frings, 2019); however, this finding does not consider spacing. These observations are theoretically logical as students with lower working memory capacity encode less into their long-term memory, increasing the relative proportion of information that would be forgotten without adequate spaced retrieval.

Conclusion

Some recommendations regarding spacing are reasonably safe to make at an institutional level. However, the research is far from being able to extrapolate the findings into precise recommendations for secondary and tertiary students, the majority of whom do not have active facilitation of spaced testing during class. We may even be decades away from having this level of research due to the exponentially increasing difficulty in controlling for multiple interference effects in more real-world settings. Nevertheless, we can reasonably conclude that incorporating some element of spaced retrieval is highly likely to have some positive effect. This effect is most likely to be moderate, though it may be reduced for those with higher working memory capacity. While potentially longer spacing intervals may benefit test performance, specific guidelines on exact scheduling cannot be made. If a learner is already using spaced retrieval, there is certainly no evidence basis to suggest that relying more on spaced retrieval would have a positive real-world benefit.

In my experience, the heavy reliance on spaced repetition for learning modern curriculum under modern assessment criteria is often associated with more isolated and superficial processing and a tendency for rote learning. The offloading of information into superficially processed notes or flashcards seem to reduce cognitive load during encoding episodes. Due to the inherently repetitive nature of spaced retrieval and under the pressure of multiple subjects and limited time, I have observed an inverted-U effect, whereby results and mental health are negatively impacted at high levels of spaced repetition. This effect may be unmasked only for those studying more challenging material, higher total volumes of material, learning at a faster pace, or aiming for higher test scores. Therefore, I hypothesise that (a) spaced retrieval is optimally effective when augmenting a studying system that primarily optimises intrinsic cognitive load and (b) spaced retrieval has a negative effect on test outcomes and mental health when encoding and recoding processes are below an intrinsic load threshold.

Inquiry-Based Learning

Fundamentals of Inquiry-Based Learning

First proposed over 60 years ago (Bruner, 1961), inquiry-based learning (IBL) (previously and interchangeably called discovery learning, constructivist learning and problem-based learning) is a method of learning based around a now outdated understanding of cognitive architecture (Sweller, 2021). The theory centers around the premise that humans are inherently problem-solving creatures and capitalising on the brain's problem-solving cognitive habits leads to superior learning outcomes. Inquiry-based learning was one of the first non-traditional approaches to learning that disrupted the long-held pedagogoical models that are still dominant in the education space (Khalaf, 2018). To first understand the fundamentals of IBL, and the subsequent flaws and practical implications, we must first understand the defining features of it compared to traditional learning.

Traditional Learning

While definitions vary, traditional learning is typically seen as having the following characteristics.

- Divided into two stages: encoding and decoding, followed by a term examination to evaluate student performance outcomes (Hall, 2002; Johnson, 1991).
- The teacher talks for most of the time, usually with whole class participation, rather than individual or group activities (Rashty, 2003).
- The teaching is driven by a fixed curriculum, independent of the learner's gradual development of knowledge (Rashty, 2003).
- Lessons are dictated by the teacher on the underlying assumption that the teacher knows what is best for the student (Austin et al., 2001).

Traditional learning faces criticism for encouraging superficial learning and increased memorisation (Biggs, 1996), which creates future setbacks for students, especially for practical science and problem-solving (Entwistle & Tait, 1995). Traditional learning is thought to fail students in facilitating depth of knowledge mastery (Khalaf, 2018) with some outright stating that traditional learning is no longer effective in the educational field (Kiraly, 2005). IBL is described to overcome the problems of teacher-centric models (Barrow, 2006). Unfortunately, the implementation and outcomes of IBL are extremely varied (Khalaf, 2018) with low consistency between studies (Rönnebeck et al., 2016). This also makes an agreed-upon list of characteristics challenging, but the following seem to be typical core features of IBL models.

- IBL tends to involve problem-solving or testing of a hypothesis with evaluation of the findings (Pedaste et al., 2012; Pedaste & Sarapuu, 2006).
- Learners formulate explanations based on evidence to do with a subject and then communicate this in some form (Dewey, 1933).

In practice, these core features are adapted in countless ways with more specific models and frameworks. The implementation of IBL varies depending on culture, institution, age group and domain, to name a few (Khalaf, 2018).

Problems With Inquiry-Based Learning

At present, the contructivist school of thought, favouring inquiry-based learning, and the cognitivist school, favouring cognitive load theory, are almost mutually exclusive. Sweller (2021) directly states that “based on both theory and data, there is little justification for the current emphasis on inquiry learning”, in an executive summary titled “Why Inquiry-Based Approaches Harm Students’ Learning”. Although various studies have demonstrated improvements in student motivation, depth of learning and student engagement through using IBL (Khalaf, 2018), these results are often inconsistent and conflicted. To date, a clear benefit of IBL has not been empirically demonstrated (Sweller, 2021), despite gaining widespread popularity and adoption (Sundberg et al., 2005). However, the vast majority of studies investigating IBL are focused on institutional implementation.

There are many documented barriers for institutional implementation of IBL. Some of the most significant and prevalent barriers revolve around the extensive level of teacher training to implement IBL with any success (Dorier & Maab, 2012), causing considerable variability depending on the teacher or facilitator. For example, in a recent meta-analysis, IBL approaches were found to improve motivation and engagement for science teaching, but the need for strong teacher training was noted as a major factor (Areepattamannil et al., 2020). Though these barriers are largely still present and methods of consistently overcoming them have not been reported, this review is focused on non-institutional, individual implications of research. The following are limitations of IBL that have been suggested that may impact an individual student attempting to independently apply IBL into their own practice.

- Students may lack sufficient intrinsic motivation to engage in correct IBL processes (Krajcik et al., 1994). This has been predominantly reported for middle school aged students.
- Students may lack the skills to properly engage in IBL, either in investigating problems sufficiently or synthesising appropriate discussions or explanations (Edelson et al., 1999; Krajcik et al., 1998).

Consilient Hypothesis

Inquiry-based learning faces significant challenges to its theoretical and empirical basis. While traditional learning is certainly antiquated, IBL does not seem to be a strongly evidence-based direction forward at this time. However almost all of the criticism for IBL is from its institutional applications. Constantinou et al. (2018) remarks that IBL approaches have been misconstrued as narrowly a framework for teaching, instead of the broader encouragement of learner engagement through observation, analysis and problem-solving. We hypothesise that elements of IBL can be used to facilitate optimal intrinsic cognitive load for students through our studying system. Our preliminary results show great promise to this bridging approach.

Impact of Notetaking Strategies

Primarily, notetaking is seen as serving two benefits (Di Vesta & Gray, 1972). Firstly, they serve as a storage of information for future reference. Secondly, notetaking induces various levels of deeper processing and knowledge encoding (Kiewra, 1989). It has been empirically demonstrated that variations in notetaking style and technique, including longhand vs typed forms, significantly influence the level of encoding and prominence of undesirable adverse effects (Peper & Mayer, 1978; Peters, 1972). The same techniques can also vary in effect between individuals and conditions (Mueller & Oppenheimer, 2014; Peverly et al., 2007).

How Notetaking Affects Performance

Many students are told by their teachers and lecturers to take notes during class. However, the method of notetaking and the cognitive process that occurs before the pen touches paper or the keyboard is pressed can have dramatic effects.

A classic study by Peters (1972) found that notetaking during lectures was correlated with worse test performance. Whether the information was presented in written form, spoken slowly, or spoken quickly did not change the overall negative effect of notetaking on test scores. A later study by Peper and Mayer (1978) found the opposite, whereby notetaking was correlated to higher test performance regardless of presented modality. Furthermore, a positive effect of notetaking was demonstrated by Schoen (2012), with laptop typed notetaking being the superior form. Notably, each of these studies had low sample sizes, low statistical power, and the chance of an interference effect, as noted in an integrative review of notetaking by Jansen et al. (2017).

Individual differences also have significant effects on notetaking. Students receive a greater benefit from notetaking when their cognitive ability and working memory scores are higher (Berliner, 1971; Kiewra & Benton, 1988; Kiewra et al., 1987; Peverly et al., 2007), with those at lower levels of performance potentially not benefiting at all (Berliner, 1971; Peper & Mayer, 1978). Research on the effect of individual variance regarding the ability to tolerate high cognitive load and create mental models has not been sufficiently studied. However, existing evidence suggests a significant influence of these on the success of notetaking (Bui & McDaniel, 2015). In other words, notetaking skills used by top students may be ineffective for other students who lack the fundamental cognitive processes to benefit from them. In these cases, it is reasonable to suggest that these fundamental processes should be trained first.

Typed vs Longhand Notetaking

Studies on notetaking for informationally complex topics, such as lectures, where the student is exposed to a high rate of information transfer, show that typed notetaking is superior to longhand notetaking for short-term memory (Bui et al., 2013; Schoen, 2012). Presumably, this is due to typing being faster than longhand notetaking. However, this conclusion is not so straightforward. Famously, Mueller and Oppenheimer (2014) found longhand notetaking to be superior to typed notetaking. An extension of this study by Morehead et al. (2019) concluded that there was no significant difference between longhand or typed notetaking on memory performance for both short- and long-term. Further still, a systematic review by H. H. Yang et al. (2020) on the effect of digital notetaking in the classroom found insufficient evidence on the superiority of digital or longhand notetaking with several theories and frameworks to support either modality. The authors identified a range of economic, software and hardware limitations to digital notetaking and a distinct lack of empirical studies to show superiority of one form over another. Presently, the question of which form of notetaking is superior remains unanswered. However, many seemingly contradictory findings can potentially be reconciled by considering cognitive load as a significant influencer on encoding and memory, which will be discussed later.

Structuring Notes

Some students believe that writing lots of notes is superior to writing fewer notes. This advice is anecdotally echoed by teachers, with some teachers even enforcing students to write copious notes during and after class. Unfortunately, this indiscriminate advice is more likely to be detrimental than beneficial.

While some studies show that transcribed, verbatim-style notetaking is superior to structured and more processed notetaking for short-term memory, this effect is reversed with delayed testing (Bui et al., 2013). Similarly, most other studies have found that structured notes with clear outlines and greater organisation are superior (Bretzing & Kulhavy, 1979; Bui & McDaniel, 2015; Kauffman et al., 2011; Peverly et al., 2013). This effect seemed to vary in differing modalities of information delivery (spoken vs written), type of organisation and outline, and content difficulty. The effect on memory may be diminished or disappear entirely when students have more time to revise their notes (Katayama & Robinson, 2000).

Ultimately, empirical research has not thoroughly examined the effect of different types of structures, quality of structures, learning modalities, and content difficulties. Thus, the current state of evidence is far from sufficient to conclude superiority.

Note Content

In the systematic review by Jansen et al. (2017), higher note quality was significantly correlated with higher knowledge retention and test performance. Conversely, verbatim notes were correlated with lower retention due to reduced processing and encoding of information.

Controversially, due to the nature of measuring note quality, most studies use potentially inaccurate proxy measures of note quality, such as the number of factual statements in a set of notes. Confounding effects such as usage of non-verbal notetaking techniques and spatial arrangement is not accounted for, nor is there much consideration of multiple intervention interference effects. Research on the impact that content of notes relative to other processes that occur in the studying system, such as prior knowledge, level of semantic priming, or usage of information beyond immediate recall have not been sufficiently evaluated. It should be particularly noted that the superior level of recall found when information was included in notes was consistently around 40 to 50% (Aiken et al., 1975; Einstein et al., 1985; Peper & Mayer, 1986), which is much higher than the 6 to 12% found with information that was not included in notes, but still indicating a considerable degree of knowledge decay. The effect of notetaking techniques on knowledge decay in its ability to improve encoding has previously been shown; however, there are no established techniques or guidelines on how to write notes to reduce the rate of forgetting. Hypotheses based on existing empirical findings are discussed later.

Importantly, there is a general lack of modern studies examining the effect of note content and quality on retention. The absence of more recent studies is likely to be important given the considerable shift in curriculums, assessment styles and student climates over the last 40 years (Schleicher, 2018), with some indication that current assessments can even be harmful for students (Mayes & Howell, 2018).

A Unifying Theory

So far, the evidence on notetaking is sparse, empirically lacking, often contradictory, seemingly highly susceptible to interference effects and in many cases, simply outdated or statistically underpowered. Naturally, this makes it difficult to form a clear picture.

Impressively, many of these findings can be unified with cognitive load theory (Plass et al., 2010). To an extent, individuals who can handle higher levels of cognitive load can encode more and therefore improve their memory and subsequent test performance, while supra-optimal cognitive load causes performance to suffer. As Jansen et al. (2017) state, by analysing the cognitive load capacity of an individual as well as the level of cognitive load induced by a task, "we can make more fine-grained predictions about when notetaking improves performance on memory tests".

Current evidence strongly supports this with similar conclusions across an enormous range of studies in the domain of cognitive load theory (Hilbert & Renkl, 2009; Olive & Barbier, 2017; Renkl, 2014; Svinicki, 2017; Sweller, 2016b). One interesting study by Casteleyn et al. (2013) investigated the effect of graphics and multimedia in presentations on learning outcomes. Across 155 university students, usage of multimedia in teaching presentations and learning material did not create a significant difference in cognitive load or actual knowledge gain. Notably, participants subjectively preferred presentations with more graphics though it had no objective impact. This observation mirrors the illusion of fluency previously discussed in that student preferences are not an accurate predictor of actual efficacy. It also supports the theory that encoding and the facilitation of cognitive load are not easy to measure or control externally if the student lacks the independent skill to engage in the proper encoding processes.

Interestingly, Jansen et al. (2017) identify five types of cognitive load that is induced by notetaking:

1. Comprehending the lecture material
2. Identifying key points
3. Linking the material to prior knowledge and prior notes
4. Paraphrasing or summarising

5. Transforming to written form

Given the inconclusive state of note quality and nuances of note structuring, I would note that to allow these five types the highest chance for accuracy, "written form" should be interpreted broadly as the documentation of ideas, which may not be limited to verbal expression.

In summary, Jansen et al. (2017) posit that notetaking that creates sufficient, tolerable cognitive load, while preventing cognitive overload produces the highest level of performance. This cleanly explains many of the discrepancies in findings from empirical studies so far. However, validating this hypothesis is incredibly challenging due to the difficulty of directly measuring cognitive load, analysing individual tolerances for cognitive load, and the effect of different test types, test timings, and complexity of manipulation (simple fact recall vs synthesis or evaluative manipulation of knowledge).

A Potential Technique

Based on the available evidence, the best method of notetaking depends on the desired outcome. Techniques for students with a high tolerance for cognitive load would be vastly different to a student whose cognitive load capacity must be trained and extended.

Objectives

As our goal is to allow as many students as possible to study to a level of efficiency, the following objectives should be set:

- Increase a student's capacity for cognitive load
- Improve a student's ability to self-regulate cognitive load for techniques
- Equip the student with a technique that creates an optimum level of cognitive load

Attributes

Furthermore, to be as empirically supported as possible, the technique should prioritise the following attributes:

- A high level of clear organisation and processing
- Time-efficiency for use in varying rates of information transfer (e.g. fast lectures vs written sources)
- Preferably leveraging the benefits and convenience of digital notetaking while mitigating the distractions and potential for reducing cognitive load below optimum.

To my knowledge, there is no notetaking system that has been designed and subsequently instructed with all of these considerations optimised.

Synthesis and Proposed System

So far, we have examined the nature of human memory with emphasis on human cognitive architecture and the role of cognitive load theory; the benefits and limitations of spaced retrieval practice; the current state of science regarding inquiry-based learning; and the relationship between notetaking and cognitive load. In this section, we will bridge these domains together into a single studying system that leverages off the strengths of each, while mitigating their limitations. We describe a novel and precise combination of theories and frameworks, a brief history of the system's development, and preliminary findings on effectiveness.

Our studying system is named the Bear Hunter System (BHS) in reference to the evolutionary theory of human cognition which is shared in fundamental principle between IBL and CLT. The system is divided into discrete steps with each step designed to facilitate a specific series of cognitive processes. Students are introduced to each subsequent step progressively, once sufficient mastery has been obtained for the prior step. Though the stages of learner development loosely follow the previously mentioned stages of skills acquisition set by VanLehn (1996), to my knowledge, there is no current system that consistently navigates secondary or tertiary students through these stages in real-world settings, nor any system for any group that incorporates IBL principles, chunking, best-practice notetaking and intrinsic cognitive load optimisation.

By following these steps, the students engage in the following sequence of activities.

1. Identification of key words and terminology for the topic to be studied.
2. Usage of inquiry to drive spontaneous self-explanations that lead towards the formation of main chunks. The focus of inquiry is split between (a) evaluating the relationship between concepts and (b) questions examining the functional or conceptual importance of a concept. We refer to this process of inquiry which is restricted to a select few generic questions as "restricted inquiry".
3. Usage of the same restricted inquiry processes to clearly identify relationships between concepts and chunks, with evaluation of relationships to assign strict relative priority.
4. Non-linear representation of these ideas and relationships using visualisation such as doodles to supplement verbal expression where appropriate.
5. Repeating of restricted inquiry with separation of isolated elements into separate methods for spaced retrieval or rote-memorisation.

The core features of the BHS involve all of the following:

- The usage of restricted inquiry to identify commonalities between concepts, resulting in chunks;
- Strict prioritisation of relationships using restricted inquiry;
- Multi-staged development of knowledge while learning material, based on the gradual development of domain-specific cognitive schemas;
- Subsequent representation of this schema through non-linear notetaking.

Theoretical Basis for the Bear Hunter System

The BHS promotes non-linear, recursive encoding of information with constant application of restricted inquiry. The sequential activation of clear steps (not outlined in full in this document) breaks up the process into meaningful subgoals. This improves the likelihood that the solution can be applied correctly in novel contexts, without learners being confused by the correct procedure for solving or encoding information (Renkl, 2014). Split-attention effect is reduced through the collection of key words and concepts into a single source via a defined step with low working memory load requirements. Restricted inquiry reduce redundancy effects and leverages the environmental organising and linking principle, while following this with chunking and non-linear notetaking facilitates generative learning and spontaneous self-explanations.

Information is learned in the order that is deemed most relevant by the student, based on their self-perceived areas of greatest curiosity which is informed by the schema that was created in the aim step. Though information is sometimes learned out of order in relation to the material, we argue that the information is learned in the correct order when viewed in relation to what is most likely to be effectively encoded at any given time. The hypothesis is that cognitive overload can be prevented by minimising element interactivity and unnecessary elements through self-identifying material that feels the most necessary, based on a basic means-end strategy whereby the answers to restricted inquiry are set as the goal outcome. The non-linear encoding of information creates a naturally multi-stage development of knowledge that is deeply processed, while the usage of restricted inquiry facilitates generative learning and activating of the forward testing effect with spaced retrieval between each stage of development.

Based on the expertise reversal effect, high element interactivity, which induces a high working memory load, can be reduced by greater domain expertise. Following the BHS procedure, domain expertise is increased through a path of least resistance, aligning with the narrow limits of change principle. As the restricted inquiry process constantly aims to create simplified chunks, we hypothesise that working memory limitations during encoding can be functionally overcome through frequent cycles of encoding and retrieval, leveraging off the ability to mobilise vast volumes of information to facilitate the encoding of novel information in a positive feedback cycle. We also hypothesise that this reduces the quantity of isolated elements which need to be memorised by rote, which can subsequently create unnecessary additional revision time during future spaced repetitions.

To my knowledge, the following aspects are novel:

- The usage of a restricted inquiry process to create chunks, which functionally bypasses working memory limits;
- Increase in intrinsic cognitive load and facilitation of explicit spontaneous self-explanation using restricted inquiry processes;
- Prevention of cognitive overload through reduction of unnecessary or isolated elements via learner-led, multi-staged, non-linear knowledge development;
- Non-linear mindmap-style notetaking primarily focused on the visual prioritisation of relationships between chunks formed through restricted inquiry.

Results From Trials

Coaching the ability for self-explanations has been found to be somewhat effective, though not explored through comprehensive, multi-factor programs (Bielaczyc et al., 1995; Stark et al., 2002). When looking at transferability across domains, current evidence has not found that teaching cognitive-generic skills improve performance in far transfer domains (Ritchie et al., 2015; Sweller, 2016a; Tricot & Sweller, 2014). Due to this current lack of evidence for the benefit of cognitive-generic skills in far transfer performance, Sweller

(2020) notes that “we need to look elsewhere for effective instructional procedures”. When students are supported to interrelate multiple different information sources, outcomes are better, though it is not yet known what procedure is best to support this (Renkl, 2014). Therefore, the effective results from our BHS are novel.

We taught the BHS to 797 students from predominantly Australian secondary schools between December 2020 to October 2021 through the iCanStudy™ online membership, involving online courses, group coaching through video call for two hours per month, brief feedback on work every month (seven minutes on average per student per month), and uncounted engagement points through an online Discord Community.

A baseline data survey was introduced after approximately the first 6 months. Results were recorded from 364 students on the iCanStudy course. Of these, 70.3% had used active recall and 60.9% had used spaced repetition, while only 8.0% had attempted to use cognitive load theory. The quality of execution before joining the course was not assessed, however it is reasonable to predict a very high level of procedural variation. The average self-reported studying efficiency was 43.3% with a median of 45% , and 59.3% reported moderate to significant procrastination. The average hours spent studying per week, excluding school time, was 24.1 hours with a median of 21 hours. The average self-reported retention after one week was 51.2% with a median of 50%. Fifty-seven percent of students typically received test scores of 85% or lower, while 17.6% typically received above 95%. The following table shows the results taken at intermittent checkpoints throughout the course. Note that this data is taken at each group in a between-groups analysis, rather than within-subject analysis. Unfortunately, due to a technical incident, some early course data of approximately 400 students was deleted which has not yet been recovered.

	Stage (estimated weeks on course) [Number of respondents]		
	Technique Training (4 to 8) [251]	Basecamp (24 to 35) [76]	Camp Two (42 to 50) [40]
Students rating "somewhat less", "significantly less" or "minimal" procrastination	77.3%	97.3%	100%
Average self-reported retention after one week	70.7%	84.9%	86.9%
Self-reported improvement in efficiency of at least 30%	66.5%	96.1%	97.5%
Self-reported improvement in efficiency of at least double	27.9%	90.8%	90.0%
Students who sat a recent assessment who received over 90%	60.9%	71.2%	90%
Students who feel "significantly more" confident about achieving their academic goals	45.8%	71.1%	70%

Table 1: Progression of student metrics throughout the course

As the results show, there is a consistent improvement in procrastination, efficiency, test results and academic confidence. It should be remembered that students tend to overestimate their actual retention (Koriat & Bjork, 2005), however as all results are also self-reported, we can consider any relative change to be equally biased. In reality, improved metacognition while progressing through the course may reduce overestimation, however, this has not been verified.

Future plans

"Learning" is a complex cognitive phenomenon whereby sensory information is stored in our memory in such a way that this information can be retrieved, manipulated and applied. Although advancements in educational research in the last few decades have significantly elucidated how learning occurs, much of the research remains fragmented or isolated. Furthermore, much of the research does not take into consideration the pragmatic variables such as limited time available for learners or depth of knowledge mastery achieved by learning strategies.

However, there are several key challenges in consolidating and unifying the current isolated findings across learning strategies. Findings often do not measure against a standardised or empirically validated metric. The variation and contestability of outcome measures make meta-analyses and systematic reviews difficult to conduct or interpret. The generalisability of findings is limited to the scope of the metrics and strict, often lab-based and unrealistic, parameters of the study. Despite constant research on the contextual effectiveness of isolated learning strategies, we are facing diminishing returns on our ability to answer the pragmatic question of “which learning strategies are most efficient for a learner to achieve their learning goals?”

Due to a lack of existing research focus on time efficiency and level of knowledge mastery to research, as well no presently validated standardised metric for meta-analyses and comparative reviews, we will be conducting our own research on this field to help bridge the research-practice gap. We expect a series of publications between 2022 and 2025 addressing these issues, facilitated by Monash University. By producing such research, not only can individual learners be empowered to accurately assess their own learning strategies, but educators and researchers will be able to compare different techniques and strategies against each other

History of Technique Development

- In 2011 and 2012, I relied heavily on spaced retrieval and flashcards. The literature on effective learning was superficially reviewed and I began tutoring students independently.
- In 2013, I completed a certification in teritary teaching and learning, exposing me to the ideas of metacognition and threshold concepts. These ideas intrigued me as they offered perspectives on learning that I had previously not prioritised or strongly considered.
- Following this, I read more on the literature of threshold concepts and began to apply them to my own practice. During this time, I started noticing the issues with spaced repetition dependent studying techniques among my students. Combined with my stronger understanding of threshold concepts, I investigated methods to allow students to uncover threshold concepts more easily and quickly.
- In 2015 I had delved into the ideas of encoding and retrieval and the cognitive architectural basis for memory. Some basic techniques were derived to try to facilitate this, however effect was not consistent and coaching was very time-consuming and intensive. By the end of 2015, the idea of using inquiry-based learning to induce cognitive load was hypothesised.
- By 2017, with more experience working with students and experimenting on my own studies, some guidelines for chunking and how inquiry-based principles, not formally described in the literature, could be combined. By this point, component techniques had been taught to and evaluated across approximately 500 students.

- In 2018, I developed a "seek-and-receive" technique centred around inquiry-based chunking to optimise cognitive load. By the end of 2018, the problems of cognitive overload, intrinsic motivation of students and far transfer of skills was clearly apparent. At this point, the overwhelming majority of students were unable to master this skill to a level of superiority above their conventional techniques. By the end of 2018, approximately 250 additional students were evaluated on the seek-and-receive technique.
- By 2019, some of the nuanced effects of cognitive load were considered, as well as creating restrictions on inquiry. Stricter guidelines on chunking were developed to more explicitly instruct students on the correct procedure. Though this helped, still the results were inconsistent and many of my students were unable to master the technique. By the end of 2019, approximately 1,500 additional students were taught variations of the component techniques with results evaluated.
- In early 2020, I introduced the multi-stage progression and consolidating numerous separate techniques that I had been teaching, including the concept of non-linear learning based on inquiry. Guidelines on restricted inquiry were made stricter and the technique was reframed with chunking and relational processing as a priority. The technique was then tiered with progressions. This form of technique was far more consistent with most students able to achieve superior outcomes with intensive coaching. By mid-2020, approximately 500 additional students had been evaluated with the now consolidated technique with acceptably consistent results.
- In mid-2020, an online course for this new set of techniques was created. The majority of students who did not have the intrinsic motivation for an online

course were unable to master the technique, though results through coaching were now consistently high. Approximately 400 students went through this early stage course.

- In late 2020, iCanStudy™ was founded and a fully blended learning experience was designed to be a balance of coaching and online courses. Results from students were highly predictable and consistent. At this point, the decision was made for scalable commercialisation without the need for significant changes to the core techniques. Results from the first 700 students were evaluated with high levels of consistency in achieving results.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659-701.

Agarwal, P. K., Finley, J. R., Rose, N. S., & Roediger, H. L., 3rd. (2017, Jul). Benefits from retrieval practice are greater for students with lower working memory capacity. *Memory, 25*(6), 764-771. https://doi.org/10.1080/09658211.2016.1220579

Aiken, E. G., Thomas, G. S., & Shennum, W. A. (1975). Memory for a lecture: effects of notes, lecture rate, and informational density. *Journal of educational psychology, 67*(3), 439.

Areepattamannil, S., Cairns, D., & Dickson, M. (2020, 2020/08/17). Teacher-Directed Versus Inquiry-Based Science Instruction: Investigating Links to Adolescent Students' Science Dispositions Across 66 Countries. *Journal of Science Teacher Education, 31*(6), 675-704. https://doi.org/10.1080/1046560X.2020.1753309

Austin, K., Orcutt, S., & Rosso, J. (2001). How people learn: Introduction to learning theories. *The learning classroom: Theory into practice–a telecourse for teacher education and professional development*.

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20*(3), 287-298. https://doi.org/10.1002/acp.1245

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological science, 4*(5), 316-321.

Bahrick, H. P., & Hall, L. K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language, 52*(4), 566-577.

Barrow, L. H. (2006). A brief history of inquiry: From Dewey to standards. *Journal of Science Teacher Education, 17*(3), 265-278.

Bartlett, F. C., & Bartlett, F. C. (1995). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Berliner, D. C. (1971). Aptitude-Treatment Interactions in Two Studies of Learning from Lecture Instruction.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and instruction, 13*(2), 221-252.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education, 32*(3), 347-364.

Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & cognition, 41*(3), 392-402.

Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology, 4*(2), 145-153.

Bruel-Jungerman, E., Davis, S., & Laroche, S. (2007). Brain plasticity mechanisms and memory: a party of four. *The Neuroscientist, 13*(5), 492-505.

Bruner, J. S. (1961). The act of discovery. *Harvard educational review*.

Bui, D. C., & McDaniel, M. A. (2015). Enhancing learning during lecture note-taking using outlines and illustrative diagrams. *Journal of Applied Research in Memory and Cognition, 4*(2), 129-135.

Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of educational psychology, 105*(2), 299.

Butler, A. C., Karpicke, J. D., & Roediger, H. L., 3rd. (2007, Dec). The effect of type and timing of feedback on learning from multiple-choice tests. *J Exp Psychol Appl, 13*(4), 273-281. https://doi.org/10.1037/1076-898x.13.4.273

Callender, A. A., & McDaniel, M. A. (2009, 2009/01/01/). The limited benefits of rereading educational texts. *Contemporary Educational Psychology, 34*(1), 30-41. https://doi.org/10.1016/j.cedpsych.2008.07.001

Carey, B. (2015). *How we learn: The surprising truth about when, where, and why it happens*. Random House Trade Paperbacks.

Carneiro, P., Lapa, A., & Finn, B. (2021). Memory updating after retrieval: when new information is false or correct. *Memory*, 1-20. https://doi.org/10.1080/09658211.2021.1968438

Carpenter, S. K. (2017). 2.26 - Spacing Effects on Learning and Memory☆. In J. H. Byrne (Ed.), *Learning and Memory: A Comprehensive Reference (Second Edition)* (pp. 465-485). Academic Press. https://doi.org/10.1016/B978-0-12-809324-5.21054-7

Casteleyn, J., Mottart, A., & Valcke, M. (2013, 2013/10/01). The impact of graphic organisers on learning from presentations. *Technology, Pedagogy and Education, 22*(3), 283-301. https://doi.org/10.1080/1475939X.2013.784621

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing Distributed Practice. *Experimental Psychology, 56*(4), 236-246. https://doi.org/10.1027/1618-3169.56.4.236

Chan, J. C. K., Manley, K. D., & Ahn, D. (2020, 2020/12/01/). Does retrieval potentiate new learning when retrieval stops but new learning continues? *Journal of Memory and Language, 115*, 104150. https://doi.org/10.1016/j.jml.2020.104150

Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin, 144*(11), 1111-1146. https://doi.org/10.1037/bul0000166

Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of educational psychology, 107*(3), 689.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of verbal learning and verbal behavior, 18*(3), 257-273.

Constantinou, C. P., Tsivitanidou, O. E., & Rybska, E. (2018). What Is Inquiry-Based Science Teaching and Learning? In O. E. Tsivitanidou, P. Gray, E. Rybska, L. Louca, & C. P. Constantinou (Eds.), *Professional Development for Inquiry-Based Science Teaching and Learning* (pp. 1-23). Springer International Publishing. https://doi.org/10.1007/978-3-319-91406-0_1

Dewey, J. (1933). *A restatement of the relation of reflective thinking to the educative process*. DC Heath.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of educational psychology, 63*(1), 8.

Dorier, J., & Maab, K. (2012). The PRIMAS Project: Promoting inquiry-based learning (IBL) in mathematics and science education across Europe PRIMAS context analysis for the implementation of IBL: International Synthesis Report PRIMAS–Promoting Inquiry-Based Learning in Mathemati (Vol. 1). *Lokaliseret på: www. primasproject. eu/servlet/supportBinaryFiles*.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In *Advances in experimental social psychology* (Vol. 44, pp. 247-296). Elsevier.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences, 8*(3-4), 391-450.

Egan, D. E., & Schwartz, B. J. (1979). Chunking in recall of symbolic drawings. *Memory & cognition, 7*(2), 149-158.

Einstein, G. O., Morris, J., & Smith, S. (1985). Note-taking, individual differences, and memory for lecture information. *Journal of educational psychology, 77*(5), 522.

Entwistle, N., & Tait, H. (1995). Approaches to studying and perceptions of the learning environment across disciplines. *New directions for teaching and learning, 1995*(64), 93-103.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American psychologist, 49*(8), 725.

Fernández-Alonso, R., Álvarez-Díaz, M., Suárez-Álvarez, J., & Muñiz, J. (2017, 2017-March-07). Students' Achievement and Homework Assignment Strategies [Original Research]. *Frontiers in Psychology*, *8*(286). https://doi.org/10.3389/fpsyg.2017.00286

Geary, D. C. (2008). An evolutionarily informed education science. *Educational Psychologist*, *43*(4), 179-195.

Geary, D. C., & Geary, D. C. (2007). Educating the evolved mind. *Educating the evolved mind*, 1-99.

Gobet, F., & Clarkson, G. (2004, 2004/11/01). Chunks in expert memory: Evidence for the magical number four … or is it two? *Memory*, *12*(6), 732-747. https://doi.org/10.1080/09658210344000530

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001, 2001/06/01/). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236-243. https://doi.org/10.1016/S1364-6613(00)01662-4

Greving, S., & Richter, T. (2018, 2018-December-04). Examining the Testing Effect in University Teaching: Retrievability and Question Format Matter [Original Research]. *Frontiers in Psychology*, *9*(2412). https://doi.org/10.3389/fpsyg.2018.02412

Hall, T. (2002). *Differentiated instruction*. http://www.cast.org/publications/ncac/ncac_diffinstruc.html

Hays, M. J., Kornell, N., & Bjork, R. A. (2010, 2010/12/01). The costs and benefits of providing feedback during learning. *Psychonomic Bulletin & Review*, *17*(6), 797-801. https://doi.org/10.3758/PBR.17.6.797

Hilbert, T. S., & Renkl, A. (2009). Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. *Computers in Human Behavior*, *25*(2), 267-274.

Jansen, R. S., Lakens, D., & Ijsselsteijn, W. A. (2017, 2017/11/01/). An integrative review of the cognitive costs and benefits of note-taking. *Educational Research Review*, *22*, 223-233. https://doi.org/10.1016/j.edurev.2017.10.001

Jeffries, R., Turner, A. A., Polson, P. G., & Atwood, M. E. (1981). The processes involved in designing software. *Cognitive skills and their acquisition*, *255*, 283.

Johnson, D. W. (1991). *Cooperative Learning: Increasing College Faculty Instructional Productivity. ASHE-ERIC Higher Education Report No. 4, 1991*. ERIC.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of educational psychology*, *93*(3), 579.

Karpicke, J. D., Butler, A. C., & Roediger Iii, H. L. (2009, 2009/04/01). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*(4), 471-479. https://doi.org/10.1080/09658210802647009

Karpicke, J. D., & Roediger, H. L. (2008). The Critical Importance of Retrieval for Learning. *Science, 319*(5865), 966-968. https://doi.org/doi:10.1126/science.1152408

Katayama, A. D., & Robinson, D. H. (2000). Getting students "partially" involved in note-taking using graphic organizers. *The Journal of Experimental Education, 68*(2), 119-133.

Kauffman, D. F., Zhao, R., & Yang, Y.-S. (2011). Effects of online note taking formats and self-monitoring prompts on learning from online text: Using technology to enhance self-regulated learning. *Contemporary Educational Psychology, 36*(4), 313-322.

Khalaf, B. K. (2018). Traditional and Inquiry-Based Learning Pedagogy: A Systematic Critical Review. *International Journal of Instruction, 11*(4), 545-564.

Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review, 1*(2), 147-172.

Kiewra, K. A., & Benton, S. L. (1988). The relationship between information-processing ability and notetaking. *Contemporary Educational Psychology, 13*(1), 33-44.

Kiewra, K. A., Benton, S. L., & Lewis, L. B. (1987). Qualitative aspects of notetaking and their relationship with information-processing ability and academic achievement. *Journal of Instructional Psychology, 14*(3), 110.

King, A. (1992). Comparison of Self-Questioning, Summarizing, and Notetaking-Review as Strategies for Learning from Lectures. *American educational research journal, 29*(2), 303-323. https://doi.org/10.2307/1163370

Kiraly, D. (2005). Project-based learning: A case for situated translation. *Meta: journal des traducteurs/Meta: Translators' Journal, 50*(4), 1098-1111.

Kliegl, O., & Bäuml, K.-H. T. (2021, 2021/10/01/). When retrieval practice promotes new learning – The critical role of study material. *Journal of Memory and Language, 120*, 104253. https://doi.org/10.1016/j.jml.2021.104253

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.

Koriat, A., & Bjork, R. A. (2005, Mar). Illusions of competence in monitoring one's knowledge during study. *J Exp Psychol Learn Mem Cogn, 31*(2), 187-194. https://doi.org/10.1037/0278-7393.31.2.187

Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and aging, 25*(2), 498.

Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., Fredricks, J., & Soloway, E. (1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences, 7*(3-4), 313-350.

Krajcik, J. S., Blumenfeld, P. C., Marx, R. W., & Soloway, E. (1994). A collaborative model for helping middle grade science teachers learn project-based instruction. *The elementary school journal*, *94*(5), 483-497.

Lang, A., Potter, R. F., & Bolls, P. D. (1999, 1999/06/01). Something for Nothing: Is Visual Encoding Automatic? *Media Psychology*, *1*(2), 145-163. https://doi.org/10.1207/s1532785xmep0102_4

Lang, J. M. (2016). Small changes in teaching: The last 5 minutes of class. *The chronicle of higher education*.

Latimier, A., Peyre, H., & Ramus, F. (2021, 2021/09/01). A Meta-Analytic Review of the Benefit of Spacing out Retrieval Practice Episodes on Retention. *Educational Psychology Review*, *33*(3), 959-987. https://doi.org/10.1007/s10648-020-09572-8

Logan, J. M., Castel, A. D., Haber, S., & Viehman, E. J. (2012). Metacognition and the spacing effect: the role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal. *Metacognition and Learning*, *7*(3), 175-195.

Mayes, E., & Howell, A. (2018, 2018/10/03). The (hidden) injuries of NAPLAN: two standardised test events and the making of 'at risk' student subjects. *International Journal of Inclusive Education*, *22*(10), 1108-1123. https://doi.org/10.1080/13603116.2017.1415383

Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological science*, *21*(7), 914-919.

Morehead, K., Dunlosky, J., & Rawson, K. A. (2019, 2019/09/01). How Much Mightier Is the Pen than the Keyboard for Note-Taking? A Replication and Extension of Mueller and Oppenheimer (2014). *Educational Psychology Review*, *31*(3), 753-780. https://doi.org/10.1007/s10648-019-09468-2

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological science*, *25*(6), 1159-1168.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104). Prentice-hall Englewood Cliffs, NJ.

Olive, T., & Barbier, M.-L. (2017). Processing time and cognitive effort of longhand note taking when reading and summarizing a structured or linear text. *Written Communication*, *34*(2), 224-246.

Paas, F., Tuovinen, J., Van Merrienboer, J. J. G., & Darabi, A. (2005, 09/01). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development*, *53*, 25-34. https://doi.org/10.1007/BF02504795

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003, 2003/03/01). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63-71. https://doi.org/10.1207/S15326985EP3801_8

Paas, F. G., & Van Merriënboer, J. J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of educational psychology*, *86*(1), 122.

Pastötter, B., & Frings, C. (2019). The Forward Testing Effect is Reliable and Independent of Learners' Working Memory Capacity. *Journal of cognition*, *2*(1), 37-37. https://doi.org/10.5334/joc.82

Pedaste, M., Mäeots, M., Leijen, Ä., & Sarapuu, S. (2012). Improving students' inquiry skills through reflection and self-regulation scaffolds. *Technology, Instruction, Cognition and Learning*, *9*(1-2), 81-95.

Pedaste, M., & Sarapuu, T. (2006). Developing an effective support system for inquiry learning in a web-based environment. *Journal of computer assisted learning*, *22*(1), 47-62.

Peper, R. J., & Mayer, R. E. (1978). Note taking as a generative activity. *Journal of educational psychology*, *70*(4), 514.

Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of educational psychology*, *78*(1), 34.

Peters, D. L. (1972). Effects of note taking and rate of presentation on short-term objective test performance. *Journal of educational psychology*, *63*(3), 276.

Peverly, S. T., Ramaswamy, V., Brown, C., Sumowski, J., Alidoost, M., & Garner, J. (2007). What predicts skill in lecture note taking? *Journal of educational psychology*, *99*(1), 167.

Peverly, S. T., Vekaria, P. C., Reddington, L. A., Sumowski, J. F., Johnson, K. R., & Ramsay, C. M. (2013). The relationship of handwriting speed, working memory, language comprehension and outlines to lecture note-taking and test-taking among college students. *Applied Cognitive Psychology*, *27*(1), 115-126.

Plass, J. L., Moreno, R., & Brünken, R. (2010). Cognitive load theory.

Rashty, D. (2003). Traditional learning vs. elearning. *Retrieved May*, *10*, 2011.

Rawson, K. A., & Dunlosky, J. (2013). Relearning attenuates the benefits and costs of spacing. *Journal of Experimental Psychology: General*, *142*(4), 1113.

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

Renkl, A. (2014). Toward an Instructionally Oriented Theory of Example-Based Learning. *Cognitive Science*, *38*(1), 1-37. https://doi.org/10.1111/cogs.12086

Ritchie, D., & Volkl, C. (2000). Effectiveness of Two Generative Learning Strategies in the Science Classroom. *School Science and Mathematics, 100*(2), 83-89. https://doi.org/10.1111/j.1949-8594.2000.tb17240.x

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental psychology, 51*(5), 573.

Roediger, H. L., & Karpicke, J. D. (2006, 2006/09/01). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science, 1*(3), 181-210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Rönnebeck, S., Bernholt, S., & Ropohl, M. (2016, 2016/07/02). Searching for a common ground – A literature review of empirical research on scientific inquiry activities. *Studies in Science Education, 52*(2), 161-197. https://doi.org/10.1080/03057267.2016.1206351

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432.

Sagi, Y., Tavor, I., Hofstetter, S., Tzur-Moryosef, S., Blumenfeld-Katzir, T., & Assaf, Y. (2012, 2012/03/22/). Learning in the Fast Lane: New Insights into Neuroplasticity. *Neuron, 73*(6), 1195-1203. https://doi.org/https://doi.org/10.1016/j.neuron.2012.01.025

Schleicher, A. (2018, 2018/03/01). Educating Learners for Their Future, Not Our Past. *ECNU Review of Education, 1*(1), 58-75. https://doi.org/10.30926/ecnuroe2018010104

Schoen, I. (2012). Effects of method and context of note-taking on memory: handwriting versus typing in lecture and textbook-reading contexts.

Schworm, S., & Renkl, A. (2006, 2006/05/01/). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education, 46*(4), 426-445. https://doi.org/10.1016/j.compedu.2004.08.011

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

Simon, H. A., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology, 5*(1), 29-46.

Smith, C. D., & Scarf, D. (2017). Spacing Repetitions Over Long Timescales: A Review and a Reconsolidation Explanation. *Frontiers in Psychology, 8*, 962-962. https://doi.org/10.3389/fpsyg.2017.00962

Stark, R., Mandl, H., Gruber, H., & Renkl, A. (2002). Conditions and effects of example elaboration. *Learning and instruction, 12*(1), 39-60.

Sundberg, M. D., Armstrong, J. E., & Wischusen, E. W. (2005). A reappraisal of the status of introductory biology laboratory education in US colleges & universities. *The American Biology Teacher*, *67*(9), 525-529.

Svinicki, M. (2017). *Supporting the Cognitive Skills Behind Note-Taking* (1057-2880). (The National Teaching & Learning Forum, Issue.

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37-76). Elsevier.

Sweller, J. (2016a). Cognitive Load Theory, Evolutionary Educational Psychology, and Instructional Design. In D. C. Geary & D. B. Berch (Eds.), *Evolutionary Perspectives on Child Development and Education* (pp. 291-306). Springer International Publishing. https://doi.org/10.1007/978-3-319-29986-0_12

Sweller, J. (2016b, 2016/12/01/). Working Memory, Long-term Memory, and Instructional Design. *Journal of Applied Research in Memory and Cognition, 5*(4), 360-367. https://doi.org/10.1016/j.jarmac.2015.12.002

Sweller, J. (2020, 2020/02/01). Cognitive load theory and educational technology. *Educational Technology Research and Development, 68*(1), 1-16. https://doi.org/10.1007/s11423-019-09701-3

Sweller, J. (2021). *Why Inquiry-based Approaches Harm Students' Learning*. The Centre For Independent Studies. https://www.cis.org.au/publications/analysis-papers/why-inquiry-based-approaches-harm-students-learning/

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and instruction, 2*(1), 59-89.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology, 24*(6), 837-848.

Thalmann, M., Souza, A. S., & Oberauer, K. (2019, Jan). How does chunking help working memory? *J Exp Psychol Learn Mem Cogn, 45*(1), 37-55. https://doi.org/10.1037/xlm0000578

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review, 26*(2), 265-283.

Van Gog, T., Kirschner, F., Kester, L., & Paas, F. (2012). Timing and frequency of mental effort measurement: Evidence in favour of repeated measures. *Applied Cognitive Psychology, 26*(6), 833-839.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual review of psychology, 47*(1), 513-539.

Wittrock, M. C. (1989, 1989/09/01). Generative Processes of Comprehension. *Educational Psychologist, 24*(4), 345-376. https://doi.org/10.1207/s15326985ep2404_2

Wittrock, M. C. (2010, 2010/01/21). Learning as a Generative Process. *Educational Psychologist*, *45*(1), 40-45. https://doi.org/10.1080/00461520903433554

Yang, C., Sun, B., Potts, R., Yu, R., Luo, L., & Shanks, D. (2020, 06/04). Do working memory capacity and test anxiety modulate the beneficial effects of testing on new learning? *Journal of Experimental Psychology: Applied*, *26*. https://doi.org/10.1037/xap0000278

Yang, H. H., Shi, Y., Yang, H., & Pu, Q. (2020). *The Impacts of Digital Note-Taking on Classroom Instruction: A Literature Review* (978-981-33-4594-2). L.-K. Lee, L. H. U, F. L. Wang, S. K. S. Cheung, O. Au, & K. C. Li.

Youssef-Shalala, A., Ayres, P., Schubert, C., & Sweller, J. (2014). Using a general problem-solving strategy to promote transfer. *Journal of Experimental Psychology: Applied*, *20*(3), 215.

Zulkiply, N., & Burt, J. S. (2013). The exemplar interleaving effect in inductive learning: Moderation by the difficulty of category discriminations. *Memory & cognition*, *41*(1), 16-27.