Report on Learning: A Practical and Learner-Centric Perspective

Justin Sung

iCanStudy ™ Proprietary Limited

Author note

Justin Sung is the co-founder and head of learning at iCanStudy ™ Proprietary Limited, which teaches, promotes, services, and commercialises evidence-based studying techniques. Findings from this report influence the design, development, delivery, and presentation of educational products, including courses, programs, and workshops. This work is the intellectual property of iCanStudy Pty Limited and is protected under copyright laws.

Table of Contents

Disclaimer

The primary objective of this report is to provide a broad starting point for discussion, rather than a deep understanding of all factors. The secondary objective is to rationalise the iCanStudy™ program in relation to these factors. As such, this document is a relatively superficial and brief summary of multiple areas of research pertaining to learning skills. All of these factors are considered at the core of the iCanStudy Program; however, this document is not a comprehensive account of all factors and considerations. Due to the complexity of learning science, only the most important ideas have been expressed, while dozens of sub-concepts and nuances have been omitted for conciseness. Many of the concepts included in this report are only discussed to a surface level of detail with nuanced discussion excluded, especially with regard to the intricate interactions between higher-order thinking, cognitive load theory, and self-motivation. Notably, this document entirely does not address any of the following domains, which are of equal importance within the full program:

- Growth mindsets and the implications on behaviour change, learner-centric self-development, propensity for long-term academic success, and other sequelae.
- Mood and affective disorders, including internalising or externalising emotional behavioural disorder, that can affect the learning process.
- Habit formation and frameworks for behaviour change, specifically those surrounding environment- and antecedent-focused methodologies that have a lower likelihood of creating burnout cycles or willpower exhaustion.
- Goal setting and effective planning, including reverse goal setting.
- Decision-making, including frameworks for positional decision-making and expected value predictions for wicked multivariate environments.

- Time and task management, including frameworks for prioritisation and implications of scheduling, time-blocking, focus, and net productivity.
- Focus and attention management, including the management of distractions or procrastination triggers.

All of the above can have significant impacts on learning or the self-management necessary to engage in effective learning under realistic and predictable pressures. Therefore, they are all taught and guided through in considerable depth in an integrated fashion throughout the iCanStudy program. However, as their influences on the learning process are indirect, they will not be discussed in this report. A high-level overview of some of the key considerations in effective learning has been mapped in figure 1. A high-resolution downloadable PDF version is available for download at iCanStudy.com.

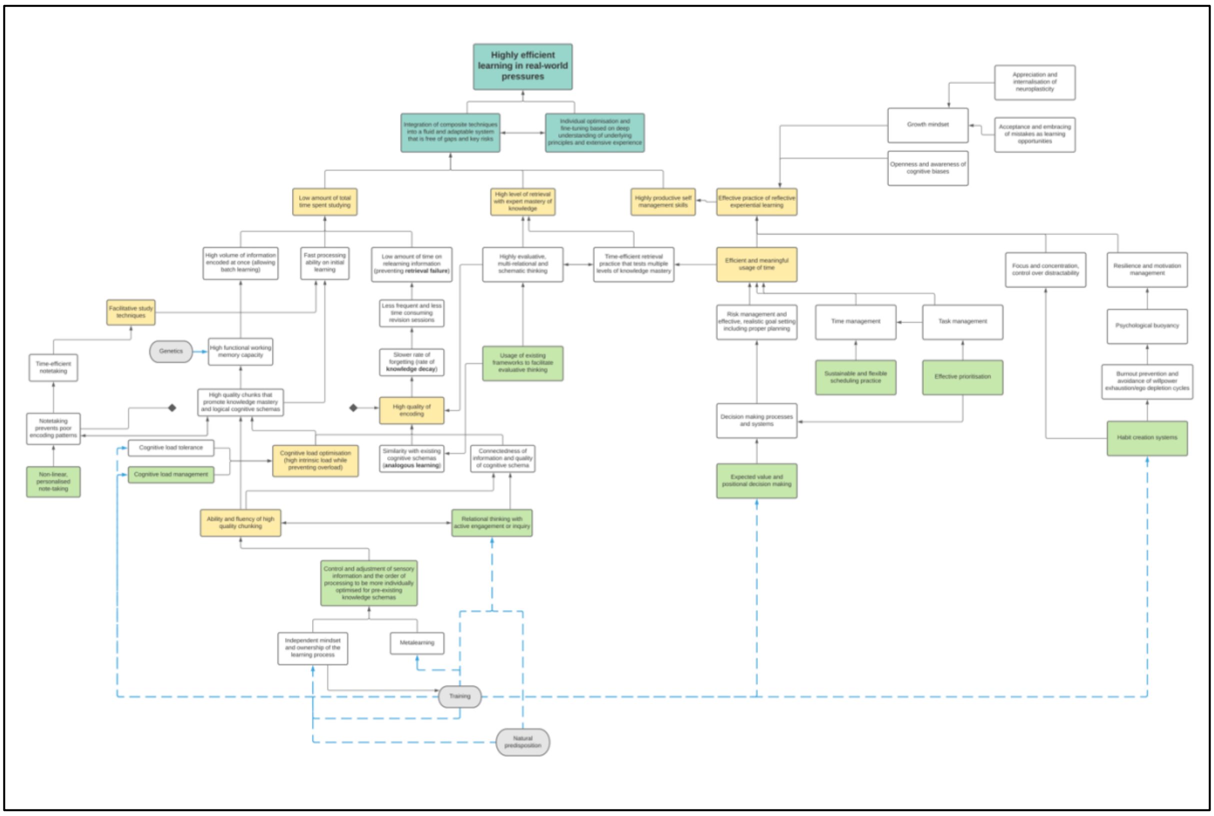

Figure 1: High-level map of learning considerations

Background to the Report

Students engage in various activities in an attempt to increase their knowledge processing and retention. The process of acquiring knowledge is called “learning”. Growing evidence suggests that the conventional approaches to studying, particularly in academic settings, is inefficient at meeting real-world demands. This is likely due to an extensive research-practice gap in the educational field spanning multiple decades of delay between new research and changes to mainstream practice (Neal et al., 2015), the highly multi-variate nature of learning (Baars et al., 2020), and the inherent difficulty with measuring and observing the learning process to progress research smoothly (Schnaubert & Schneider, 2022). Figure 2 represents the conceptualisation of this delay. A high-resolution PDF is downloadable at iCanStudy.com.

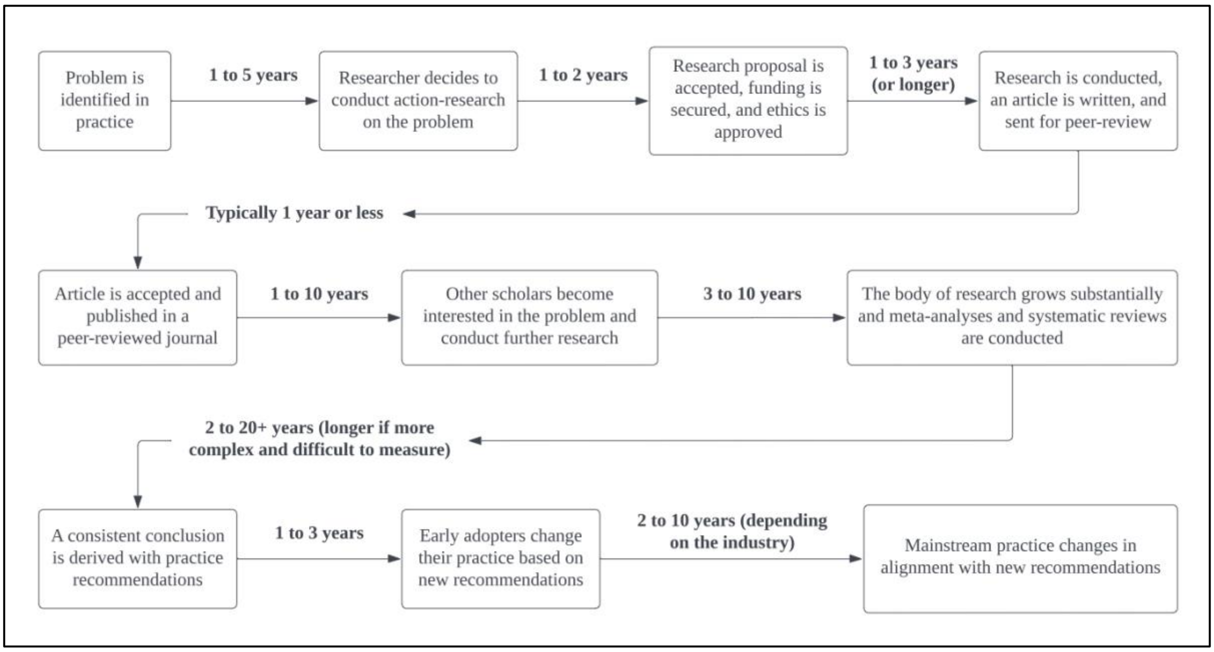

Figure 2: Practice-research-practice gap. In learning science, the total delay between problem and mainstream action is approximately 30 to 50 years.

Moreover, the majority of the available research targets educators and instructional design, leaving individual learners disempowered to independently optimise their learning. As evidence-based practitioners with extensive experience with learners of all levels, the iCanStudy team has sought to reduce the gaps between research, practice, and individual learners, while aiming to facilitate the development of critical, learner-centric research. Subsequently iCanStudy have developed a novel, first-principles system of learning based on the consilient synthesis of latest research. This report will outline the literature on some of the most pertinent elements of the learning process and the relevant modifications or considerations that comprise the iCanStudy cognitive retraining program. By examining the state of the science in the broad field of learning science, we will derive the practical implications, gaps in research and potential hypotheses for learner-centric strategy development and instruction. Notably, this synthesis will take a practice-oriented view with emphasis on research implications for individual students, rather than for educational institutions.

High-level Conceptualisation of Learning in Practice

The iCanStudy™ program conceptualises learning based on an adapted Atkinson-Shiffrin multi-store model of memory (Atkinson & Shiffrin, 1968). This model of memory, and many like it, have received substantial valid criticism, especially regarding the rigid structural model of memory, the broad categorisation of long-term memory, the proposed characteristics of long-term memory, the separation of sensory memory, and the oversimplification of the model (Chechile & Ehrensbeck, 1983; Izawa, 1999; Malmberg et al., 2019). While we do not agree with many of the tenants of the model initially proposed by Atkinson and Shiffrin (1968), the model serves as a convenient tool for establishing some basic conceptualisations about learning. Figure 3 shows our adapted model of learning, which has been modified to either eliminate or rework the most contentious elements. For comprehensiveness, we have also presented basic principles of knowledge decay (i.e. the forgetting of information that has previously attempted to be learned) and the forgetting curve (i.e. the rate at which knowledge decay occurs) in the same diagram, famously proposed by Ebbinghaus (1885).

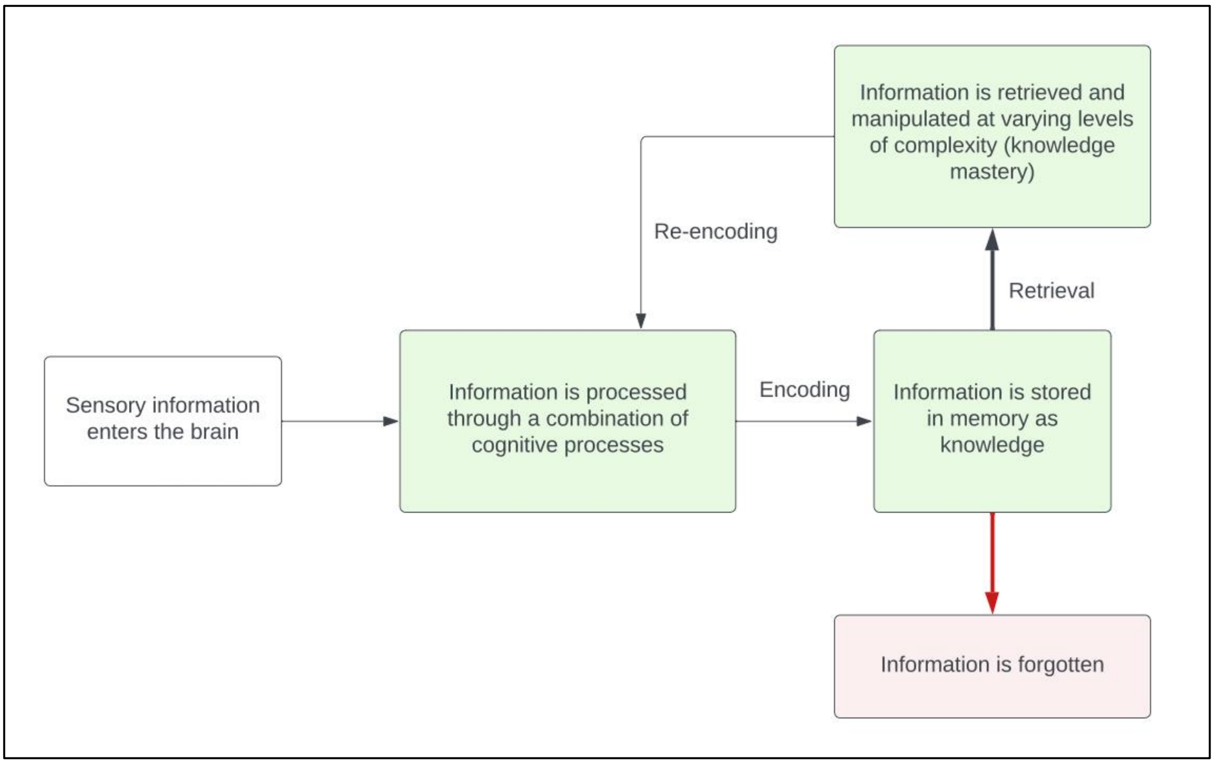

Figure 3: High-level conceptualisation of the learning process

In our model, sensory information enters the brain. This information is then processed to be stored in memory through a process called "encoding". Information stored in memory is then retrieved or forgotten, where the act of retrieving can facilitate re-encoding. The model also involves the following assumptions, theories, and hypotheses, which are not evident in the diagram alone:

- The desirable pathway for learning is for information to be encoded into memory and retrieved for varying levels of knowledge mastery.
- Learners are generally unable to significantly modify the raw sensory information that enters the brain.
- The way in which information is processed is directly related to the cognitive processes that are activated during the learning event.

- Knowledge must be retrieved to be useful, and it must be retrieved at different levels of complexity (knowledge mastery).
- The ability to retain and retrieve information at one level does not necessarily translate to the same ability at other levels.
- The way in which information is processed by the brain directly affects how strongly the information is encoded in the memory, as well as the level of knowledge mastery it can be retrieved at. Subsequently, different cognitive processes or combinations of processes result in different rates of knowledge decay.
- The way in which information is retrieved directly affects how effective the re-encoding process is at strengthening the encoding of information in the memory and the level of mastery it can be retrieved at.
- Differences in procedural, conditional, and declarative knowledge are not represented in this model (although they are covered comprehensively in the iCanStudy™ program).

In practical contexts, we propose that a high rate of knowledge decay (i.e. a proportionally high volume of information that is forgotten) results in less efficient learning due to the additional time investment to relearn the forgotten information. In addition, high levels of knowledge mastery are required to meet assessment criteria. Therefore, processes that do not result in higher levels of mastery are also less efficient due to the additional time investment to retrieve at the required knowledge level. This report will explore some of the nuances of this learning process, as well as research to support our conceptual model. We will focus on key elements that we have identified as practically significant for individual learners.

Introduction to Memory and Cognitive Load

Some sensory information is retained while some is forgotten. Improving memory with a reduced reliance on repetition would help make studying less tedious, monotonous and time-consuming. Furthermore, the metacognitive knowledge of effective learning can be anticipated to empower students and facilitate their autonomy. This section will examine the nature of memory from a cognitivist school of thought and establish some guidelines for creating and evaluating study techniques.

Human Cognitive Architecture and Cognitive Load Theory

Cognitive load theory (CLT) leverages evolutionary theory to create instructional procedures based on human cognitive architecture (Sweller, 2011). Originating as a grounded theory of learning, CLT sought to unify the many fragmented and theoretically uncorrelated empirical findings of learning under a cognitivist school of thought. It theorises that human information processing and memory operate through the mobilisation of fixed cognitive resources. The investment of cognitive resources, broadly referred to as “cognitive load”, is perceived as mental effort, which can be defined as the “cognitive capacity that is actually allocated to accommodate the demands imposed by the task” (Paas et al., 2003). Over the years, multiple types and subtypes of cognitive load have been conceptualised with some debate; however, the predominant two types that are more agreed upon are intrinsic load (cognitive load associated with learning) versus extraneous load (all other cognitive load that does not benefit the learning process). An extended discussion of the different types of cognitive load is of limited practical utility for learners, and so will be omitted in this report. In sum, growing evidence around CLT demonstrates that the strategic and optimal investment of cognitive resources and the maintenance of cognitive load within certain parameters can produce more efficient learning effects.

CLT expands on the work by Geary (2008), who categorised knowledge into biologically primary and secondary. Biologically primary knowledge is transgenerational knowledge, including skills such as listening to speech and 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. Notably, it is traditionally thought that generic-cognitive skills do not need to be taught and may not be coachable (Youssef-Shalala et al., 2014), a finding that iCanStudy disputes based on our own evaluations.

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 of 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.

The differentiation of biologically primary and secondary knowledge is relevant for educators and course creators considering optimal instructional design. However, for individual learners, these categorisations are relatively unimportant as they do not significantly impact the operational process of learning. Instead, attention should be geared towards the principles of CLT. These principles can be activated through deliberate processes and ultimately chained into an integrative, cognitively optimised learning system. The following is an incomplete list of some of the most pertinent cognitive principles identified within CLT.

The Information Store Principle

The information store principle states 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 differences in knowledge and skill are influenced by differences in prior long-term knowledge 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 1) thinking, compared to the beginner's reliance on slower problem-solving and cognitive load intensive (system 2) 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 prior long-term knowledge and working memory capacity 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 note-taking, are biologically secondary. Once information has been acquired, it is then reorganised. Pieces of information that conform with previously held memories become enhanced, while ones that do not align are 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 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 a 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 significant 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 the apparently mutually exclusive concepts of CLT and inquiry-based learning. 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 ability to test possibilities through randomness as genesis by limiting the number of combinations that can be processed in a time period. 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 it 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 is more pragmatic to plainly acknowledge 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 prevents massive amounts of data from being 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 signals and triggers 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 the brain's ability to mobilise vast volumes of information, which may facilitate the encoding of novel information in a positive feedback cycle. This hypothesis will be developed further with our studying system.

The Self-Regulation Principle

Self-regulation is a recent area of growing research in CLT. It focuses on how learners are able to manage and optimise their cognitive load (with or without educator facilitation). Effective self-regulation has been shown to protect against dramatically changing learning environments and pressures, such as during COVID-19 lockdowns (Hadwin et al., 2022), enhance the achievement and significance of learning events from current techniques (Bjork et al., 2013; Oudman et al., 2022), as well as overcome barriers and learning obstacles that are inherent when learning through cognitively suboptimal instruction (Roodenrys et al., 2012). Unfortunately, there are several crucial challenges to developing self-regulatory practices. Not only is self-regulation an extremely multifactorial skill that is intricately influenced by learners' metacognition (Schnaubert & Schneider, 2022), but the very concept of training learners to become proficient at it is in its infancy (de Bruin et al., 2020; McDaniel & Einstein, 2020). To the best of our knowledge, our system is the first program through which generic-cognitive skills of self-regulation practices have been consistently trained in a cross-domain setting, while producing significant positive results in learner outcomes. Practical challenges with teaching effective self-regulation will be discussed later.

Orders of Learning

The activities and processes involved in cognitively optimum learning closely resemble those of higher-order thinking skills (HOTS), which sit within a broader field of higher-order learning (HOL). HOTS and HOL are not considered as concepts or principles within CLT; however, we have noticed a practical advantage to understanding these terms to appreciate the full scope of learning efficiency, as will be described in this section.

Levels of thinking have previously been categorised by learning theorists. The most prominent categorisation is Bloom's revised taxonomy (Conklin, 2005), which categorises levels of knowledge mastery into six stages, commonly clustered further into two main divisions: lower-order thinking and higher-order thinking. The goal of studying requires both divisions; however, the latter is more emphasised as learners progress through education. In professional contexts, the need for lower-order learning is minimal.

Lower-order learning is reinforced by the requirement to memorise concepts, definitions, terminologies, and facts which serve as a foundation for higher levels of knowledge. Common strategies are highlighting and writing summary notes, using flashcards, and implementing the “cover, copy, check” method. However, this form of learning tends to emphasise the isolation of ideas for the sake of surface memorisation. Counterproductively, a heavy focus on lower-order knowledge acquisition results in a very time-consuming and tedious experience of reading. With the vast workload at the university level, these strategies are often unviable or unsustainable.

Studying for higher-order learning requires learners to be focused on evaluating information with purpose and attention to context. Although theorists still debate the exact scope and definition of higher-order thinking (Lewis & Smith, 1993), a characteristic feature that tends to distinguish it from lower-order thinking is that learners do not view information in isolation; rather, they engage in a comparative process whereby they make judgements (e.g. explicit and implicit importance) and prioritisations on relationships between ideas (e.g. identifying which influences and relationships are more important in varying contexts). Learners must also apply ideas and consider the consequences of applying knowledge (e.g. the implications of changing actions or perspectives on other related factors). To engage in higher-order learning, crucial strategies that learners should use are comparing ideas against each other; identifying relationships; prioritising and making judgements on the nature and importance of the relationships; considering the purpose of information and how it can be applied; and focusing on creating clear and explicit networks of information rather than simply understanding isolated concepts.

Practical Implications 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 that this approach reduces 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 abilities to elaborate and rationalise the solutions for 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 be superior to prompted self-explanations (Hilbert & Renkl, 2009; Schworm & Renkl, 2006). The factors that influence a students' ability to spontaneously self-explain are countless and not clearly mapped in the current research landscape. A pertinent consideration is the balance between extraneous and intrinsic cognitive load, which we have briefly discussed so far. The 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.

Moreover, 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 with the highly researched and established practice of interleaving for skills development, which also has similar empirical findings (Taylor & Rohrer, 2010). Thus, the variability effect on cognitive load may partially explain the efficacy of interleaving. As another example, 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 overestimate the depth of their knowledge (Carey, 2015) and believe that mastery over something has been achieved when it has not (Lang, 2016). This illusion is congruent with the Dunning-Kruger effect (Dunning, 2011) which describes the general overconfidence of the under-informed.

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, the aforementioned 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, and 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 additional minor findings to help guide the application of CLT. Firstly, 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). Secondly, visual encoding appears to be less restricted than verbal encoding, with cognitive resources required to encode visual information remaining stable, even if the pace of information presentation is increased beyond optimal verbal encoding limits (Lang et al., 1999). The iCanStudy Program is designed with a hypothesis that functional working memory can be developed through effective learning processes, while leveraging the higher capacity for learning of visual encoding.

Practical Barriers to Effective Learning

Misinterpreted Effort

To achieve higher-order thinking, learners must invest more cognitive resources (Afflerbach et al., 2015) and exert greater mental effort (Leppink & Pérez-Fuster, 2019). Although this process is beneficial for learning based on CLT (Sweller, 2016a), a commonly observed barrier is learners' counterproductive response to increased effort—a phenomenon called the "misinterpreted-effort hypothesis" (Kirk-Johnson et al., 2019) whereby learners tend to view increased effort as poor learning and therefore use less effective strategies (Baars et al., 2020; Carpenter et al., 2020; Groep, 2021). This adverse response to effort causes learners to opt for less effective learning techniques which ultimately increase the difficulty for them to perform at their desired level. As a result, despite investing time and effort into studying, learners consistently make decisions about their usage of learning strategies that are detrimental to their learning goals (Roodenrys et al., 2012).

The reason for this response is complex. It may be due to inadequate metacognition with regard to learning processes (i.e. meta-learning) or because of educational assessments (e.g. predictable keywords on marking rubrics that can be superficially memorised and inserted) that allow for mimicry of expertise rather than genuine mastery (Didau, 2015; Schnaubert & Schneider, 2022; Zaidi et al., 2018). Furthermore, learners are often disinclined to synthesise ideas and build knowledge mastery when they do not have much time.

Practical Consequences of the Misinterpreted-Effort Hypothesis

The consequences of the misinterpreted-effort hypothesis are dire. Modern curriculums are increasing in content and assessment volume (Schleicher, 2018) while competitive academic degrees such as medicine and dentistry are becoming even more competitive (Bound et al., 2009). Subsequently, the average hours spent on studying is also increasing (Dominguez & Novak, 2019). Research on youth mental health shows a global trend of increasing mood and affective mental health problems such as depression, burnout, and anxiety, with one of the most consistently dominant stressors being academic pressure (Australian Government Department of Health, 2015; Erskine et al., 2017; Kalberg et al., 2011; Lee & Larson, 2000; Mayes & Howell, 2018; Merikangas et al., 2010).

To manage these challenges, students may place a high value on using time-efficient learning strategies that help them perform well in these competitive assessments, even if they unknowingly make decisions that reduce their performance. For example, the usage of flashcards with spaced retrieval is a commonly used strategy, with some studies reporting over 80% of students utilising it (Wissman et al., 2012). However, despite students investing their limited time in using this strategy, there is insufficient research to indicate that this strategy is effective at helping students achieve high grades for most subjects when assessed at the varying levels of knowledge mastery required in a typical assessment (Carpenter et al., 2020). This topic of flashcards and spaced retrieval strategies are discussed in depth later.

The Time-Efficiency Trap of Lower-Order Learning

Many students express a tendency towards lower-order thinking because it allows them to cover content faster. However, this thinking does not take into account multiple other variables that dictate performance and learning success. Lower-order thinking is faster at a surface level but much more time-consuming to reach the higher-order levels demanded of learners in senior secondary school, university, and beyond. Evidence suggests that learners who struggle to engage in higher-order thinking find it difficult to engage with material at the depth of knowledge expected from their program (Zoller, 2016). From a time investment standpoint, learning also has a snowball effect whereby it is easier to learn about a topic when there is more prior knowledge about it (Ayres, 2006). Growing research suggests that higher-order thinking helps to build prior knowledge schemas more efficiently and deeply than lower-order thinking (Liu et al., 2021). Conversely, the isolated nature of lower-order thinking means that learners can struggle to become more efficient over time as they fail to efficiently build leverageable schemas for future learning. This may manifest in content overwhelm when a learner is faced with high reading loads, creating an entrenched cycle of high workload followed by inefficient learning strategies, resulting in higher academic pressures and lower confidence to use alternative strategies.

Overcoming the Misinterpreted-Effort Hypothesis

The erroneous monitoring judgements students make have been noted as a significant barrier to student learning improvement. Researchers have explored whether effort can be self-monitored alongside metacognitive training to help students make more beneficial decisions about their learning strategies (Kirk-Johnson et al., 2019). The "cue utilisation framework" is often used to describe learners' decision-making processes (Koriat, 1997). In this framework, the effort is seen as a cue that is monitored by the learner. The cue is processed and the learner makes a judgement. In the context of learners making decisions about their learning strategies, their "monitoring judgements" can be considered disadvantageous as learners decide not to use more effective techniques due to the cue of high effort being monitored. Interestingly, learners seem to have at least a partially accurate ability to monitor effort, despite having poor eventual judgement (Carpenter et al., 2020). As Bosch et al. (2021) remark in their investigation of the behaviours of successful students, students tend to use fewer learning strategies than they originally intended to and more work is necessary to help guide students towards adopting effective ones.

Measuring Effort and Cognitive Load

Several researchers have attempted to improve monitoring judgements through the usage of self-reports and metacognition. The earliest and potentially most common measure of cognitive load is the Paas scale (Paas, 1992) which is rapid to administer and continues to be used to this day. This scale has repeatedly shown that instructional design that optimises cognitive load increases test scores (Sweller, 2011). Some studies have even been so specific as to find that a continuous numerical measurement such as a 0 to 100% scale was more accurate and reliable for measuring cognitive load effects on complex problem solving than standard nine-point Likert scales (Ouwehand et al., 2021). However, the measurement of cognitive load and effort is an area of sparse research and scholars have noted a need for further investigation in this field (Baars et al., 2020; Händel et al., 2020).

Outside of CLT, researchers have attempted to more broadly measure metacognition and the effect it has on learning through other tools. One of the most prominent and well-researched measures is the Learning and Study Strategies Inventory (LASSI), initially developed by Weinstein et al. (1987). The latest LASSI 3rd Edition (Weinstein et al., 2016) features 60 questions across six subscales; it seeks to help students gauge their own usages of learning and study strategies, as well as educators implement useful interventions to provide targeted support for the students. The LASSI is based on Weinstein's (2017) Model of Strategic Learning (MSL), which categorises learning into skill, will, and self-regulation. The skill and will categories include three subscales each, while self-regulation includes four. The LASSI was the first measurement tool that was designed for diagnostic purposes and has seen widespread usage in educational institutions around the world. A recent rigorous meta-analysis of LASSI by Fong et al. (2021) found low to moderate positive correlations with LASSI scores to grades and GPA. While the magnitude of correlation was not high, the consistency of findings demonstrated that the LASSI had a high level of utility for broadly assessing student learning and usage of learning strategies. However, this questionnaire was developed to be an all-encompassing, broad measure of student learning strategy and does not directly correlate mental effort with monitoring judgements. Instead, students filling in the LASSI have already made monitoring judgements and their cognitive processes are measured after a learning strategy has already been used.

In later years, some researchers challenged the validity of self-reported measures of metacognition, including perceived cognitive load. The premise was that self-reported measures have been sometimes inaccurate (Hadwin et al., 2007), which has prompted research in more objective trace measurements such as keystrokes and mouse click patterns to make inferences on learning beliefs and behaviours (Boekaerts & Corno, 2005). In response, McCardle and Hadwin (2015) considered the measurement of self-regulation as an activity where learners' self-assessments were critical to the monitoring and decision-making process. They posited that inaccuracies in self-reporting are important to consider as part of the learners' self-regulatory processes. Supporting this view, a recent review of metacognition self-reporting concluded that self-reports offer valuable insights into the awareness and accuracy of students, reinforcing the position that self-reports are not only valid, but are a fundamentally indispensable method of measuring monitoring judgements (Craig et al., 2020).

Ultimately, when considering the domains of research presented in this report, there is an opportunity to consolidate existing research to better inform practice. Evidence thus far suggests that cognitive load optimised learning strategies are being selected against due to the misinterpreted-effort hypothesis. Promisingly, learner metacognition can improve self-regulatory behaviours to combat this negative effect, and perceived cognitive load may be measurable through self-reporting. However, there is currently no established practice that involves measuring students' cognitive load through self-assessment to try and improve self-regulatory behaviours and promote more beneficial monitoring judgements. Therefore, the iCanStudy program integrates a metacognitive training component to enhance learners' ability for cognitive load self-assessment. These components have been shown to be grossly effective based on current data; however, further research is needed to characterise the parameters of effectiveness and the consistency of results.

Practical Challenges of Utilising Effort as a Cue

Recent research indicates that while mental effort can be detected, there are likely to be numerous unexplored moderators that influence its effect on behaviour. One prominent moderator is motivation. Baars and Wijnia (2018) performed a study on 178 Dutch secondary school students to examine how motivational profiles of students, including their perceptions of mental effort, affected their ability to self-regulate. They hypothesised that students with higher motivation would demonstrate greater self-regulatory capabilities than those with low motivation. To assess this, students sat a pre-test on a topic that had not been covered in their curriculum, followed by a video teaching about self-regulated learning and a 0 to 100 self-reported assessment of their perceived confidence to be self-effective at studying the topic. Students were then shown a series of videos on the curriculum topic and given a post-test with problem-solving questions that were isomorphic from the pre-test. After each question, students rated themselves on perceived mental effort, performance, and made a choice on whether they wanted to re-study the relevant concepts. Upon completion, students completed a motivation questionnaire. On analysis of the results, the authors found that students with higher motivational profiles tended to have higher monitoring accuracy and learning outcomes than those with poor motivation. The correlations were low to moderate, though statistically significant as determined by their p-value. Therefore, the authors concluded that motivation has an influence on a student's ability to self-regulate their learning and the amount of effort they are willing to invest. This finding is intuitive and predictable, but important to consider when attempting to use effort itself as a cue for more productive self-regulation.

In an examination of the consistency of effort as a cue, Blissett et al. (2018) investigated if mental effort was used as a cue to make judgements about diagnostic certainty for medical doctors. They also investigated if perceived effort was correlated to the accuracy of the diagnoses. A total of 22 medical doctors from a single centre in Canada participated in the study. Their diagnostic certainties were self-reported on a scale from 0 to 100%, and their cognitive efforts were measured using the Paas scale (Paas, 1992). Diagnostic accuracy was coded binarily as either correct or incorrect. On analysis, the authors found a moderate negative correlation between mental effort and diagnostic accuracy, indicating that diagnoses made with higher effort were more likely to be incorrect. Their findings support both that effort can be consistently self-reported, and that effort is inversely correlated to perceptions of certainty and objective accuracy. However, the moderate level of magnitude in correlation reinforces the notion that other factors are likely to be influential, and the small sample size puts this study at high risk of confounding effects and random error.

Suitability of Current Methods of Effort Measurement

Instead of examining the utility of effort as a cue, some of the research focused on how appropriate the accepted methods are for measuring and analysing effort. Vangsness and Young (2021) hypothesised that the frequency at which learners were overtly asked to judge the difficulty of their learning could have an unnatural influence on their cue use and judgement accuracy. In their lab-based experiments, 100 participants were recruited in exchange for research credit, with 59 participants completing the experimental task. A multi-level analysis was used which allowed the usage of partial and incomplete data sets. The task consisted of identifying a target circle from an array of distractor items within a time limit. This process repeated for approximately one hour with difficulty increasing through more indistinguishable distractor items. On analysis, the authors found that more frequent overt prompts to make judgements on difficulty correlated to a higher tendency to utilise inaccurate and extraneous cues, as well as reduced performance in a standardized assessment. This study is the first and only of its kind to examine how the frequency of prompts affects cue utilisation. Although the study was limited in that their experiments did not feature control groups, isolated changes in variables, or task conditions that closely reflect broader authentic academic challenges, the findings suggest that the frequency of metacognitive self-reporting may influence natural cue utilisation behaviours. It is also possible that due to volunteer bias, the participants may be more invested in learning efficacy and have more positive motivational profiles. As Baars and Wijnia (2018) demonstrated, this could have an impact on the willingness to invest effort and therefore habitual tendencies of effort-based cue utilisation.

Moreover, the conventional linear model of perceived effort has recently been challenged. Typically, mental effort is correlated with performance through a linear measurement of each variable. While these results are easy to understand, they are likely to be inaccurate if the relationships between variables are non-linear. In response, Leppink and Pérez-Fuster (2019) conducted a study reanalysing the data from four recent publications that had measured self-reported effort. Upon reanalysis, the authors found some evidence that mental effort may not be linearly correlated to workload, response time, and certainty. Most notably, a cubic relationship was found between mental effort and certainty, wherein the results suggest that there may be a reduced ability to distinguish high or low levels of certainty within a significant range of mental effort.

The non-linear relationship between effort and certainty is directly contradictory to the findings from previous linear models which suggested that individuals can accurately distinguish high and low levels of certainty within the full range of mental effort. With respect to this review, it calls into question the statistical validity of the findings in other studies, which did not specify the usage of a non-linear model to analyse the relationships between effort and other variables. Indeed, it is likely that the magnitude of correlation in the other studies may change if the analysis used a non-linear regression model. This is disconcerting in cases like Blissett et al. (2018) where they used the extremely popular Paas scale (Paas et al., 2003; Paas, 1992) to measure cognitive load, which has almost exclusively been used and validated against other variables with linear models.

Discussion of Key Findings Relating to Mental Effort Measurement

The relationship between mental effort and monitoring judgements is complex, especially with consideration of recent research. Prior reviews and meta-analyses on self-regulation and metacognition measurement suggest that mental effort is likely to be a reliable subcomponent of metacognition that can be self-assessed (Craig et al., 2020; Händel et al., 2020; van Gog et al., 2020). These findings are at least partially supported by Blissett et al. (2018) and Vangsness and Young (2021). However, it is likely that the relationship between mental effort and actual self-regulatory behaviour, including monitoring judgements, is non-linear and much more multi-factorial (Baars & Wijnia, 2018; Leppink & Pérez-Fuster, 2019). Despite thousands of papers published in just the last few years, there is almost no focus on the integration of cognitive load measurement and the effect it can have on improving learners' decision-making and monitoring judgements. This is surprising, given the clear emphasis for future research in this area stipulated by prior reviews and meta-analyses (Paas et al., 2003; Sweller, 2018).

Furthermore, even if mental effort monitoring produced predictable and consistent results, Leppink and Pérez-Fuster (2019) suggest that current linear models and measurement scales may be fundamentally unsuited to evaluate how effort correlates to other variables. Given the lack of wider research on alternative methods of cognitive load measurement outside of the popular Paas scale (Paas, 1992), the unsuitability of linear measurements could be problematic in that a suitable measurement tool may not be available for practical usage. Recently, the measurement of gaze-shift frequency in clinical ultrasound interpretation has been inversely correlated with Paas scores (Aldekhyl et al., 2018). The hypothesis behind gaze-shift frequency is that those with greater expertise may require less time looking at certain areas of information, leading to a more frequent shift in gaze (Sweller, 2018). Measurement of gaze-shift frequency may therefore provide an alternative method of cognitive load measurement, which may confer an added benefit of continuous measurement without overt prompting. This may additionally influence their natural cue utilisation behaviours as Vangsness and Young (2021) posited. However, this method of measurement is not easily accessible for most practitioners and does not give insight into the students' metacognition, which is an important variable in evaluating the self-regulatory process (McCardle & Hadwin, 2015). Further research on appropriate measurement tools for assessing effort and cognitive load is indicated.

In conclusion, while there is no empirical support to suggest the superiority of effort monitoring, there is no theoretical or empirical evidence of a significant negative consequence either, as long as the frequency of monitoring is not excessive (Vangsness & Young, 2021). It is unclear as to the exact frequency at which monitoring could be deemed excessive and it is unlikely that this will be an objectively consistent number for all individuals and contexts. In this regard, practitioners will have to take care and be observant to whether overt prompting for effort monitoring is disruptive to their learning process or not. Summarily, educators may wish to consider training students on responding more positively to mental effort and helping them identify high and low effort cues explicitly. This recommendation is theoretically justifiable with minimal risk of harm, despite a lack of empirical support at this time.

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 limitations of spacing and the parameters within which these beneficial effects can be produced are far from clear.

In this section, I will briefly summarise some of the salient findings of spaced retrieval practice while emphasising the substantial limitations of existing research when extrapolating on them to make broad practical recommendations, as is commonly done. 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 such discussion will be omitted from this review. Based on a review of recent research, 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 feedback that their answers are correct when students are highly confident in their answers can be an inefficient use of time for them (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 immediate feedback, 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, knowledge domains, knowledge levels, and working memory capacities (Adesope et al., 2017; Latimier et al., 2021).
- Retrieval seems to enhance learning new 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 notably, 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 research 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 academic subjects with varied methods of instruction or the 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 retrieval across longer intervals do sufficiently demonstrate some form of long-term benefit for learning (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 and grammar among adults. To our 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 effect sizes suggest 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.

By evaluating spaced retrieval through the lens of CLT, these variations can be explained. A growing number of studies have found that when cognitive load is outside of its optimal range due to the over- or under-activation of cognitive effects, performance improvements of spaced retrieval are not consistently observed (Carpenter et al., 2020; Chen et al., 2018a; Chen et al., 2018b; Sweller et al., 2019). For example, studies that utilise a method of retrieval that activates beneficial cognitive effects, such as the generation effect or testing effect, tend to demonstrate a more robust association with better performance outcomes (Schwieren et al., 2017), especially when the student generates self-explanations of the concepts during retrieval (Bielaczyc et al., 1995; Schworm & Renkl, 2006).

In accordance with CLT, the effectiveness of using worked examples for learning depends on the learners' own ability to elaborate and rationalise the solutions to themselves (Renkl, 2014). Students who can self-explain, either spontaneously or through prompting, were found to be superior at solving novel problems (Hilbert & Renkl, 2009; Renkl, 1997). Retrieval methods that encourage spontaneous self-explanations tend to be superior for retention compared to prompted self-explanations where a cue is provided before retrieval, which is congruent with existing research in CLT (Hilbert & Renkl, 2009; Schworm & Renkl, 2006).

More focused research on the relationship between spaced retrieval and CLT also show the parameters of spacing effectiveness. For example, Sweller (2016b) posits that extraneous and unhelpful cognitive load must be reduced to overcome working memory limitations. In a Chinese laboratory study of 1,032 university students, C. Yang et al. (2020) demonstrated that spaced testing has the greatest effect on those with low working memory capacities compared to those with high working memory capacities. 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 and 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 capacities encode less into their long-term memories, 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 our 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 seems to reduce cognitive load below optimal levels during encoding sessions. In the context of studying multiple subjects, limited time, and the need for higher-orders of learning in assessments, we have observed an inverted-U effect on the benefit of spaced retrieval, whereby results and mental health are negatively impacted at high levels of spaced retrieval, due to the inherently repetitive and time-consuming nature of this strategy. This effect may become apparent only for those studying more challenging material, higher total volumes of material, learning at a faster pace, or aiming for higher test scores. Therefore, we 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 on a now outdated understanding of cognitive architecture (Sweller, 2021). The theory centres around the premise that humans are inherently problem-solving creatures, and that 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 pedagogical 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 its defining features as compared to those of 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 students (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 scholars 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). While this makes a consistent list of characteristics challenging, 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 them 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 constructivist 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. To date, clear benefits of IBL have not been empirically demonstrated (Sweller, 2021), despite the concept gaining widespread popularity and adoption (Sundberg et al., 2005). However, the potential utility of IBL principles and approaches may be relatively unexplored as the vast majority of studies investigating IBL are focused on institutional implementation.

There are many documented barriers for the 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 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 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 to 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 ideas to engage in 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 strong, evidence-based direction forward at this time. However, almost all of the criticism for IBL is from its institutional applications. Constantinou et al. (2018) remark 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 Note-taking Strategies

Primarily, note-taking is seen as serving two benefits (Di Vesta & Gray, 1972). Firstly, they serve as a storage of information for future reference. Secondly, note-taking induces various levels of deeper processing and knowledge encoding (Kiewra, 1989). It has been empirically demonstrated that variations in note-taking style and technique, including longhand versus typed forms, significantly influence the level of encoding and quality of learning (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 Note-taking Affects Performance

Many students are told by their lecturers to take notes during class. However, the method of note-taking 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 note-taking during lectures was correlated with worse test performance. Whether the information was written, spoken slowly, or spoken quickly in presentation did not change the overall negative effect of note-taking on test scores. A later study by Peper and Mayer (1978) found the opposite: note-taking was correlated with higher test performance regardless of presentation modality. Furthermore, a positive effect of note-taking was demonstrated by Schoen (2012), with laptop-typed note-taking being the superior form. Notably, each of these studies had a low sample size, low statistical power, and the chance of an interference effect, as noted in an integrative review of note-taking by Jansen et al. (2017).

Individual differences also have significant effects on note-taking. Students receive a greater benefit from note-taking when their cognitive abilities 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). 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 on the success of note-taking (Bui & McDaniel, 2015). In other words, note-taking 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 Versus Longhand Note-taking

Studies on taking notes for informationally complex topics, such as during lectures where students are exposed to a high rate of information transfer, show that typed note-taking is superior to longhand note-taking for short-term memory (Bui et al., 2013; Schoen, 2012). Presumably, this is due to typing being faster than longhand note-taking. However, this conclusion is not so straightforward. Famously, Mueller and Oppenheimer (2014) found longhand note-taking to be superior to typed note-taking. An extension of this study by Morehead et al. (2019) concluded that there was no significant difference between longhand and typed note-taking on memory performance for both the short- and long-term. Further still, a systematic review by H. H. Yang et al. (2020) on the effect of digital note-taking in the classroom found insufficient evidence on the superiority of digital or longhand note-taking with several theories and frameworks to support either modality. The authors identified a range of economic, software, and hardware limitations to digital note-taking and a distinct lack of empirical studies to show superiority of one form over another. Presently, the question of which form of note-taking 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 even enforcing students to write copious notes during and after class. Unfortunately, this generalised and grossly unsupported advice is more likely to be detrimental than beneficial.

While some studies show that transcribed, verbatim-style note-taking is superior to structured and more processed note-taking 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). The impact of note structure on learning quality seemed to vary in differing modalities of information delivery (spoken versus written), types of organisation and outline, and content difficulty. The effect on knowledge retention and retrieval accuracy 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 the superiority of one form of structure over another. Current research suggests that the most appropriate structure will differ depending on the other factors present in the learning event. In alignment with this view, we hypothesise that improving metacognition around note-taking structures and self-regulatory capacities around adapting the structure to fit the dynamics of the learning event results in positive effects on knowledge retention and level of mastery.

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 variables such as the usage of non-verbal note-taking techniques and spatial arrangement are not accounted for, nor is there much consideration of multiple intervention interference effects. Research on the impact that the actual content of notes has on learning, compared to other processes that occur in the studying system, such as prior knowledge, level of semantic priming, and usage of information beyond immediate recall, has 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). Although this percentage of retention is much higher than the 6 to 12% found with information that was not included in notes, it still indicates a considerable degree of knowledge decay. The effect of note-taking techniques on knowledge decay in their 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 pertinent given the considerable shift in curriculums, assessment styles, and student climates over the last 40 years (Schleicher, 2018). Furthermore, measuring the effectiveness of note-taking relative to proxy measures of learning such as assessment results is particularly contentious given the indication that current assessments can even be harmful to students (Mayes & Howell, 2018).

Explaining Note-taking with CLT

So far, the evidence on note-taking is sparse, lacking in sufficient empirical support, often contradictory, seemingly highly susceptible to interference effects, and in many cases, simply outdated or statistically underpowered. Naturally, these problems make it difficult to form a clear picture of the effects and factors that influence how note-taking can facilitate efficient learning practice.

Impressively, many of the disparate findings on note-taking can be unified with CLT (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 loads cause 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 note-taking improves performance on memory tests".

Current evidence strongly supports this conclusion with similar findings across an enormous range of studies in CLT (Hilbert & Renkl, 2009; Olive & Barbier, 2017; Renkl, 2014; Svinicki, 2017; Sweller, 2016b). For example, one study by Casteleyn et al. (2013) investigated the effect of graphics and multimedia in presentations on learning outcomes. Across 155 university students, the 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 and control externally if the student lacks the independent skill to engage in proper encoding processes.

To elucidate the function of cognitive load in note-taking, Jansen et al. (2017) identify five types of load that are typically induced by note-taking techniques:

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, we suggest 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 note-taking that creates sufficient, tolerable cognitive load while preventing cognitive overload produces the highest level of performance. This 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 moderating influences of different test types, test timings, and varying complexities of knowledge manipulation (e.g. simple fact recall versus multi-relational HOL).

A Potential System

Based on the available evidence, the best method of note-taking depends on the desired outcome. Techniques for students with high tolerance for cognitive load would be vastly different to students whose cognitive load capacities 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 versus written sources)
- Preferably leveraging the benefits and convenience of digital note-taking while mitigating distractions and the potential for reducing cognitive load below optimum.

To our knowledge, there is no note-taking system that has been designed and subsequently instructed with all of these considerations optimised.

The iCanStudy Program: A Novel Learning 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 note-taking 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.

We consider our program to be a cognitive retraining program to teach a fully integrated system of learning that is based on the latest research and refined through extensive testing. The cognitive retraining component refers to creating new cognitive habits and skills for processing information, monitoring judgements, and learning strategies. For the sake of developing self-sustainable learners, self-regulation, self-modification, and habit formation of core processes are necessary. To our best knowledge, the program is characteristically novel and unique for the following reasons:

- The program is more of a revolution than an evolution. It addresses learning from a deep understanding of the latest research, instead of iterating on existing norms that may be fundamentally misdirected. This makes our techniques significantly different from conventional studying techniques and also more efficient. We theorise that the emphasis on correct encoding and re-encoding, instead of heavy spacing and retrieval is a major reason for our approach's efficiency.
- The theoretical basis is consilient. This means the program draws on research across multiple domains that could influence either the learning outcome directly or the facilitation of self-sustainable learner development, including mindset, learning, productivity, mental health, blended learning, and behavioural analysis.
- The system is fully integrated and end-to-end. To our knowledge, it is the first fully-integrated system of techniques that have been tested in real-world settings to achieve a high level of consistency and predictability. All of its individual components and techniques are designed to complement, augment, or mitigate deficiencies of other techniques and components within the system.

- It is ahead of the research on learning science with respect to learner-centric study skills and self-regulation. We estimate that our learning system is approximately 50 years ahead of the research space, based on predicted future research volume and historical adoption trends. This is primarily because our own practical observations and rate of trialling were faster than the research being released, as we were unrestricted by the research-practice delays that are inherent to peer-reviewed publication of data. The latest research continually supports the methods we have developed while increasingly identifying problems with conventional techniques.
- The results are highly replicable. Our comprehensive training methodology and approach to technology usage allow for consistent and reliable large-scale training of the system to a high level of execution fidelity.
- The system is cross-domain and transferable. Our system is fundamentally concerned with generic-cognitive skills training and can be applied to any form of declarative memory and knowledge acquisition, making it beneficial for an enormous range of learning applications.

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 inquiry-based learning and cognitive load theory. 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 the best of our knowledge, there is no current system that consistently navigates secondary or tertiary students through these stages in real-world settings, or any system in general that incorporates IBL principles, chunking, best-practice note-taking, and intrinsic cognitive load optimisation.

By following the following steps, our students are taught to engage in the following sequence of activities with a high level of procedural fluency and fidelity.

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 domain-transferrable 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 illustrations 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 schemas through non-linear note-taking.

Theoretical Basis for the Bear Hunter System

The BHS promotes non-linear, recursive encoding of information with constant application of restricted inquiry. Through the sequential activation of explicit and well-defined steps (not outlined in full in this document), the entire learning process is divided into progressive sub-processes. This sequential and compartmentalised approach improves the likelihood that the knowledge generated can be applied correctly in novel contexts, without learners being confused by the correct procedure for solving or encoding information (Renkl, 2014). Each step of the process helps to support the activation of beneficial cognitive effects and HOTS, while reducing detrimental cognitive effects or extraneous load. The following examples of BHS elements provide a brief insight into some of the notable considerations:

- The split-attention effect is reduced through the collection of key terminology into a single source via a specific technique designed to reduce working memory load requirements.
- Restricted inquiry reduces redundancy effects and leverages the environmental organising and linking principle.
- Following restricted inquiry with chunking and non-linear note-taking helps to facilitate 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 an earlier 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.

- Non-linear encoding of information creates a recursive, multi-stage development of knowledge that is deeply processed.
- The usage of restricted inquiry facilitates generative learning and activation of the forward testing effect by incorporating spaced retrieval between each stage of knowledge mastery 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 layered approach of progressive detail and knowledge complexity. Learners are taught to engage in HOTS from an early stage of learning, progressively developing the more detailed lower-order knowledge retention. As HOTS across a broad topic creates change in general knowledge schemas, as opposed to lower-order thinking which is isolated in nature, this approach utilises similar total levels of cognitive load and abides by the narrow limits of change principle. We have found that this "reverse thinking" approach that prioritises HOL chronologically before lower-order thinking and knowledge mastery produces learning efficiency, ultimately allowing learners to attain higher-levels of knowledge mastery in a shorter time, without sacrificing lower-order knowledge competency.
- 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, utilising the brain's much less load-restricted ability to retrieve high volumes of knowledge. We also hypothesise that layered encoding and retrieval cycles following the reverse thinking

approach above reduces the quantity of isolated elements that need to be memorised by rote, which can subsequently reduce unnecessary additional revision time during future spaced repetitions.

To our knowledge, the following aspects are novel and have not been described in either the research literature or presented commercially:

- The usage of a restricted inquiry process to create chunks, which functionally bypass 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 mind map-style note-taking primarily focused on the visual prioritisation of relationships between chunks formed through restricted inquiry.
- Reversing the priority of HOL and lower-order learning as chronological steps of a learning system.

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 research has not found that cognitive-generic skills improve learners' performance in far transfer domains (Ritchie et al., 2015; Sweller, 2016a; Tricot & Sweller, 2014). Due to this current lack of evidence, 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 and October 2021 through the iCanStudy™ online program. The program involves 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 numerous unmeasured engagement points through an online Discord Community.

A baseline data survey was introduced after approximately the first six 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 45.0%, 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%.

Table 1 shows the results of a survey taken at intermittent checkpoints throughout the program. Note that this data is taken at each group in a between-group analysis, rather than within-subject analysis. An improved method of data collection and technical infrastructure is currently under development to allow for more comprehensive metrics, including the capacity for within-subject and multi-variate analyses. Unfortunately, due to a technical incident, some early course data of approximately 400 students were deleted from our servers, which have not yet been recovered.

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. This has not been verified.

	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 and received over 90%	60.9%	71.2%	90.0%
Students who feel "significantly more" confident about achieving their academic goals	45.8%	71.1%	70.0%

Table 1: Progression of student metrics throughout the course

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 pragmatic variables such as the limited time available for learners and depth of knowledge mastery achieved by learning strategies. Research on learning science tends to focus on educators' perspectives, and many branches of educational research are yet to be well consolidated.

There are several key challenges in consolidating 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 knowledge mastery, as well as the lack of presently validated standardised metric for meta-analyses and comparative reviews, we will be producing reports on our results so that 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 System Development

This section will outline the chronological progression of system development. The record below is a reflective documentation of the history and rationale for our approaches, rather than an academic or technical record. It is written in response to a growing number of individuals who have expressed interest in the developmental history of iCanStudy™, considering the unconventional nature of our approaches. Numbers and dates are given as approximations based on best recollection but may differ slightly to actual figures.

Chronology of Events

- In 2011 and 2012, Justin Sung relied heavily on spaced retrieval and flashcards. The literature on effective learning was superficially reviewed and Justin began tutoring students independently.
- In 2013, Justin completed a certification in tertiary teaching and learning, exposing him to the ideas of metacognition and threshold concepts. These ideas intrigued him as they offered perspectives on learning that he had previously not prioritised or strongly considered.
- Following this, Justin read more on the literature of threshold concepts and began to apply them to his own practice. During this time, he started noticing the issues with spaced repetition–dependent studying techniques among his students. Combined with his stronger understanding of threshold concepts, he investigated methods to allow students to uncover threshold concepts more easily and quickly.
- In 2015, Justin 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, positive effects on learning efficiency were 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 his own studies, some guidelines were hypothesised for combining the usage of chunking and inquiry-based learning principles that had not been described in the literature. By this point, component techniques had been taught to and evaluated across approximately 500 students.

- In 2018, Justin 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, the intrinsic motivation of students, and problems with the far transfer of skills became clearly apparent. At this point, the overwhelming majority of students were unable to master the seek-and-receive technique 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, and restrictions were created 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 his 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, Justin hypothesised and eventually integrated the reverse thinking approach of layered, multi-stage learning progression and consolidated numerous separate techniques that he 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 the primary outcome priority. The system of techniques was then tiered with progressions of skill development to enhance the scalable and accurate training of this system to others. This form of development 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 system and teaching method with acceptably consistent results.

- In mid-2020, an online course for this system 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 as a scalable balance of coaching and online lessons. 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 or teaching methodology. Results from the first 700 students were evaluated with high levels of consistency in achieving results.

Current Focuses for Delivery Optimisation

By early 2022, over 4,000 students had joined the iCanStudy™ program. New problems with training consistency were observed among this larger population of diverse learners. More prominently, rushing, whereby learners progressed through learning techniques without sufficient practice, and selective learning, whereby learners opted not to learn or practice certain principles or techniques had been clearly identified as major moderating factors for success. Learners who demonstrated rushing and selective learning were up to nine times more likely to fail to develop competency with the system than those who did not.

Additional measures were introduced to optimise program outcomes and prevent rushing and selective learning through clear sign-posting; activity-based course progression; time-based lesson “dripping”; gamification and incentives for correct practice; new activities, quizzes, and lessons on correct progression and practice methods; and lesson reframing and restructuring to create looped pathways where rushing behaviour could be directed without detriment. As of August 2022, approximately 80% of learners are able to achieve system competency and attain predicted results. Further research and data collection is necessary to monitor progress long-term and across larger population sizes.

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.

Afflerbach, P., Cho, B.-Y., & Kim, J.-Y. (2015). Conceptualizing and assessing higher-order thinking in reading. *Theory Into Practice*, *54*(3), 203-212.

Agarwal, P. K., Finley, J. R., Rose, N. S., & Roediger, H. L., 3rd. (2017). 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.

Aldekhyl, S., Cavalcanti, R. B., & Naismith, L. M. (2018). Cognitive load predicts point-of-care ultrasound simulator performance. *Perspectives on medical education*, *7*(1), 23-32. https://doi.org/10.1007/s40037-017-0392-7

Areepattamannil, S., Cairns, D., & Dickson, M. (2020). 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

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Elsevier.

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*.

Australian Government Department of Health. (2015). *The mental health of children and adolescents: Report on the second Australian Child and Adolescent Survey of Mental Health and Wellbeing*. https://www.health.gov.au/resources/publications/the-mental-health-of-children-and-adolescents

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

Baars, M., & Wijnia, L. (2018). The relation between task-specific motivational profiles and training of self-regulated learning skills. *Learning and Individual Differences*, *64*, 125-137. https://doi.org/10.1016/j.lindif.2018.05.007

Baars, M., Wijnia, L., de Bruin, A., & Paas, F. (2020). The relation between students' effort and monitoring judgments during learning: a meta-analysis. *Educational Psychology Review*, *32*(4), 979-1002. https://doi.org/10.1007/s10648-020-09569-3

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.
Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual review of psychology*, *64*, 417-444. https://doi.org/10.1146/annurev-psych-113011-143823
Blissett, S., Sibbald, M., Kok, E., & van Merrienboer, J. (2018). Optimizing self-regulation of performance: is mental effort a cue? *Advances in Health Sciences Education*, *23*(5), 891-898. https://doi.org/10.1007/s10459-018-9838-x
Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, *54*(2), 199-231. https://doi.org/10.1111/j.1464-0597.2005.00205.x
Bosch, E., Seifried, E., & Spinath, B. (2021). What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement. *Learning and Individual Differences*, *91*, 102056. https://doi.org/10.1016/j.lindif.2021.102056
Bound, J., Hershbein, B., & Long, B. T. (2009). Playing the admissions game: Student reactions to increasing college competition. *Journal of Economic Perspectives*, *23*(4), 119-146.
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). 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). 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

Carpenter, S. K., Endres, T., & Hui, L. (2020). Students' use of retrieval in self-regulated learning: Implications for monitoring and regulating effortful learning experiences. *Educational Psychology Review*, *32*(4), 1029-1054. https://doi.org/10.1007/s10648-020-09562-w

Casteleyn, J., Mottart, A., & Valcke, M. (2013). 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). 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

Chechile, R. A., & Ehrensbeck, K. (1983). Long-term storage losses: A dilemma for multistore models. *The Journal of general psychology*, *109*(1), 15-30.

Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018a). Extending cognitive load theory to incorporate working memory resource depletion: evidence from the spacing effect. *Educational Psychology Review*, *30*(2), 483-501.

Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018b). Undesirable difficulty effects in the learning of high-element interactivity materials [Review]. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01483

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.

Conklin, J. (2005). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives complete edition. In: JSTOR.

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

Craig, K., Hale, D., Grainger, C., & Stewart, M. E. (2020). Evaluating metacognitive self-reports: Systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning*, *15*(2), 155-213. https://doi.org/10.1007/s11409-020-09222-y

de Bruin, A. B., Roelle, J., Carpenter, S. K., & Baars, M. (2020). Synthesizing cognitive load and self-regulation theory: A theoretical framework and research agenda. *Educational Psychology Review*, *32*(4), 903-915. https://doi.org/10.1007/s10648-020-09576-4

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.

Didau, D. (2015). *What if everything you knew about education was wrong?* Crown House Publishing.

Dominguez, A., & Novak, H. (2019). *National Survey of Student Engagement (NSSE) 2019 Results* (Institutional Research, Planning and Effectiveness, Issue.
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.
Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.
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.
Erskine, H. E., Baxter, A. J., Patton, G., Moffitt, T. E., Patel, V., Whiteford, H. A., & Scott, J. G. (2017). The global coverage of prevalence data for mental disorders in children and adolescents. *Epidemiology and Psychiatric Sciences*, *26*(4), 395-402. https://doi.org/10.1017/S2045796015001158
Fernández-Alonso, R., Álvarez-Díaz, M., Suárez-Álvarez, J., & Muñiz, J. (2017). Students' Achievement and Homework Assignment Strategies [Original Research]. *Frontiers in Psychology*, *8*(286). https://doi.org/10.3389/fpsyg.2017.00286
Fong, C. J., Krou, M. R., Johnston-Ashton, K., Hoff, M. A., Lin, S., & Gonzales, C. (2021). LASSI's great adventure: A meta-analysis of the Learning and Study Strategies Inventory and academic outcomes. *Educational Research Review*, *34*, 100407. https://doi.org/10.1016/j.edurev.2021.100407
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). 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). 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). 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
Groep, L. (2021). *The relationship between mental effort and judgement of learning on students' perceived study effectiveness and their decisions on the use of future study strategies*

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, *2*(2), 107-124. https://doi.org/10.1007/s11409-007-9016-7

Hadwin, A. F., Sukhawathanakul, P., Rostampour, R., & Bahena-Olivares, L. M. (2022). Do self-regulated learning practices and intervention mitigate the impact of academic challenges and COVID-19 distress on academic performance during online learning? [Original Research]. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.813529

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

Händel, M., Harder, B., & Dresel, M. (2020). Enhanced monitoring accuracy and test performance: Incremental effects of judgment training over and above repeated testing. *Learning and instruction*, *65*, 101245. https://doi.org/10.1016/j.learninstruc.2019.101245

Hays, M. J., Kornell, N., & Bjork, R. A. (2010). 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.

Izawa, C. (1999). *On human memory: Evolution, progress, and reflections on the 30th anniversary of the Atkinson-Shiffrin model*. Psychology Press.

Jansen, R. S., Lakens, D., & Ijsselsteijn, W. A. (2017). 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.

Kalberg, J. R., Lane, K. L., Driscoll, S., & Wehby, J. (2011). Systematic Screening for Emotional and Behavioral Disorders at the High School Level: A Formidable and Necessary Task. *Remedial and Special Education*, *32*(6), 506-520. https://doi.org/10.1177/0741932510362508

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). 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.
Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. H. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive Psychology*, *115*, 101237. https://doi.org/10.1016/j.cogpsych.2019.101237
Kliegl, O., & Bäuml, K.-H. T. (2021). 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. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349. https://doi.org/10.1037/0096-3445.126.4.349
Koriat, A., & Bjork, R. A. (2005). 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). 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). 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
Lee, M., & Larson, R. (2000). The Korean 'examination hell': Long hours of studying, distress, and depression. *Journal of Youth and Adolescence*, *29*(2), 249-271.
Leppink, J., & Pérez-Fuster, P. (2019). Mental effort, workload, time on task, and certainty: Beyond linear models. *Educational Psychology Review*, *31*(2), 421-438. https://doi.org/10.1007/s10648-018-09460-2
Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory Into Practice*, *32*(3), 131-137.

Liu, J., Ma, Y., Sun, X., Zhu, Z., & Xu, Y. (2021). A systematic review of higher-order thinking by visualizing its structure through HistCite and CiteSpace software. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-021-00614-5

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.

Malmberg, K. J., Raaijmakers, J. G., & Shiffrin, R. M. (2019). 50 years of research sparked by Atkinson and Shiffrin (1968). *Memory & cognition*, *47*(4), 561-574.

Mayes, E., & Howell, A. (2018). 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

McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning*, *10*(1), 43-75. https://doi.org/10.1007/s11409-014-9132-0

McDaniel, M. A., & Einstein, G. O. (2020). Training learning strategies to promote self-regulation and transfer: The knowledge, belief, commitment, and planning framework. *Perspectives on Psychological Science*, *15*(6), 1363-1381. https://doi.org/10.1177/1745691620920723

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.

Merikangas, K. R., He, J.-p., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime Prevalence of Mental Disorders in U.S. Adolescents: Results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(10), 980-989. https://doi.org/10.1016/j.jaac.2010.05.017

Morehead, K., Dunlosky, J., & Rawson, K. A. (2019). 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.

Neal, J. W., Neal, Z. P., Kornbluh, M., Mills, K. J., & Lawlor, J. A. (2015). Brokering the Research–Practice Gap: A typology. *American Journal of Community Psychology*, *56*(3), 422-435. https://doi.org/10.1007/s10464-015-9745-8

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.

Oudman, S., van de Pol, J., & van Gog, T. (2022). Effects of self-scoring their math problem solutions on primary school students' monitoring and regulation. *Metacognition and Learning*, *17*(1), 213-239. https://doi.org/10.1007/s11409-021-09281-9

Ouwehand, K., Koroef, A. v. d., Wong, J., & Paas, F. (2021). Measuring cognitive load: Are there more valid alternatives to Likert rating scales. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.702616

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). 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. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *Journal of educational psychology*, *84*(4), 429.
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). 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). 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

Roodenrys, K., Agostinho, S., Roodenrys, S., & Chandler, P. (2012). Managing one's own cognitive load when evidence of split attention is present. *Applied Cognitive Psychology*, *26*(6), 878-886. https://doi.org/10.1002/acp.2889
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). Learning in the Fast Lane: New Insights into Neuroplasticity. *Neuron*, *73*(6), 1195-1203. https://doi.org/10.1016/j.neuron.2012.01.025
Schleicher, A. (2018). Educating Learners for Their Future, Not Our Past. *ECNU Review of Education*, *1*(1), 58-75. https://doi.org/10.30926/ecnuroe2018010104
Schnaubert, L., & Schneider, S. (2022). Analysing the relationship between mental load or mental effort and metacomprehension under different conditions of multimedia design [Original Research]. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.648319
Schoen, I. (2012). Effects of method and context of note-taking on memory: handwriting versus typing in lecture and textbook-reading contexts.
Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching*, *16*(2), 179-196.
Schworm, S., & Renkl, A. (2006). 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). 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. (2018). Measuring cognitive load. *Perspectives on medical education*, *7*(1), 1-2. https://doi.org/10.1007/s40037-017-0395-4
Sweller, J. (2020). 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.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261-292. https://doi.org/10.1007/s10648-019-09465-5

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). 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., Hoogerheide, V., & van Harsel, M. (2020). The role of mental effort in fostering self-regulated learning with problem-solving tasks. *Educational Psychology Review*, *32*(4), 1055-1072. https://doi.org/10.1007/s10648-020-09544-y

Vangsness, L., & Young, M. E. (2021). More isn't always better: when metacognitive prompts are misleading. *Metacognition and Learning*, *16*(1), 135-156. https://doi.org/10.1007/s11409-020-09241-9

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

Weinstein, C. (2007). Model of strategic learning. *Clearwater, FL: HandH Publishing*.

Weinstein, C. E., Palmer, D., & Schulte, A. C. (1987). Learning and study strategies inventory (LASSI). *Clearwater, FL: H & H Publishing*.

Weinstein, C. E., Palmer, D. R., & Acee, T. W. (2016). LASSI 3rd edition learning and study strategies inventory. In: Clearwater, FL: H & H Publishing.

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2012). How and when do students use flashcards? *Memory*, *20*(6), 568-579.

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

Wittrock, M. C. (2010). 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). 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.

Zaidi, N. L. B., Grob, K. L., Monrad, S. M., Kurtz, J. B., Tai, A., Ahmed, A. Z., Gruppen, L. D., & Santen, S. A. (2018). Pushing critical thinking skills with multiple-choice questions: does bloom's taxonomy work? *Academic Medicine*, *93*(6). https://journals.lww.com/academicmedicine/Fulltext/2018/06000/Pushing_Critical_Thinking_Skills_With.28.aspx

Zoller, U. (2016). From algorithmic science teaching to “know” to research-based transformative inter-transdisciplinary learning to “think”: Problem solving in the STES/STEM and sustainability contexts. In N. Papadouris, A. Hadjigeorgiou, & C. P. Constantinou (Eds.), *Insights from research in science teaching and learning: Selected papers from the ESERA 2013 conference* (pp. 153-168). Springer International Publishing. https://doi.org/10.1007/978-3-319-20074-3_11

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.