Improving student success in chemistry through cognitive science

JudithAnn R. Hartman[1] · Eric A. Nelson[2] · Paul A. Kirschner[3,4]

Abstract

Chemistry educator Alex H. Johnstone is perhaps best known for his insight that chemistry is best explained using macroscopic, submicroscopic, and symbolic perspectives. But in his writings, he stressed a broader thesis, namely that teaching should be guided by scientific research on how the brain learns: cognitive science. Since Johnstone's retirement, science's understanding of learning has progressed rapidly. A surprising discovery has been when solving chemistry problems of any complexity, reasoning does not work: students must apply very-well-memorized facts and algorithms. Following Johnstone's advice, we review recent discoveries of cognitive science research. Instructional strategies are recommended that cognitive studies have shown help students learn chemistry.

Keywords General chemistry · Working memory · Long-term memory · Overlearning · Cognitive science

Introduction

Alex H. Johnstone (1930–2017) is perhaps the most highly regarded chemistry educator of the past century. Dr. Johnstone taught and conducted research at the University of Glasgow from 1969 until his retirement in 2006 (Reid 2019). Known for his insight that chemistry teaching should help students connect macroscopic, submicroscopic, and symbolic perspectives, the "Johnstone triplet" was only a part of his fundamental belief that educators should know the science of how the brain learns.

Following his retirement, he received the *ACS Award for Achievement in Research for the Teaching and Learning of Chemistry*. In his award address, he noted that in seeking

---

✉ JudithAnn R. Hartman
  judithann.hartman@gmail.com

[1] Department of Chemistry (Emerita), United States Naval Academy, Annapolis, MD 21042, USA

[2] Fairfax County Public Schools (Retired), Fairfax, VA 22030, USA

[3] Faculty of Educational Sciences, Open University of the Netherlands, Heerlen, The Netherlands

[4] Expertise Centre Effective Learning, Thomas More University of Applied Sciences, Antwerp, Belgium

success in first-year college chemistry, a gateway to science majors, many students did not "get there." He observed that in chemistry education,

> [W]hat may have been missing was an understanding of how students learn…. Chemists require theories and models to inform their research …. My own research has led me to accept, develop, and use a model based on information processing. (2010)

The model he had adopted is diagrammed in Fig. 1.

This 'information processing (IP) model' had first been proposed by cognitive scientists Atkinson and Shiffrin (1968) and extended by Baddeley and Hitch (1974, 1999) to explain how the brain learns.

Johnstone wrote that during his career, he had applied the IP model to pursue a goal of all chemistry educators, to identify "problems students have with learning chemistry and indicate ways to remove the learning obstacles". Since his retirement, cognitive scientists have continued to refine the IP model. In recent decades, substantial progress has been made in understanding the model and its implications for education. An example is a 2021 diagram by cognitive scientist Stephen Chew (2021) which added "choke points" and 'pitfalls' in learning, with explanations of how students can work around each (Fig. 2).

During his career, Johnstone worked at the leading edge of learning research. Many of his understandings have proven to be correct. At other points, science's understanding has changed. But the fundamentals of Johnstone's adopted model, updated by recent discoveries, "indicate ways to remove the learning obstacles".

With help from new technologies (Tang and Pienta 2012; Delazer et al. 2004; Dhond et al. 2003), cognitive experts have measured substantial strengths and stringent limitations of the brain during problem solving (Geary et al. 2008). As a result of this progress, chemistry education has become a multi-disciplinary science: combining knowledge of molecular behavior with science described by Johnstone as "understanding of how students learn." (2010).

The findings of this new science offer an opportunity to transform chemistry education from opinion-based theories to science-informed best practices. There is substantial
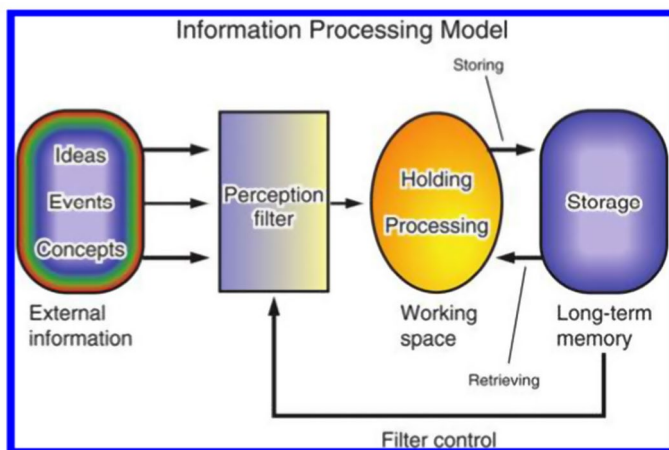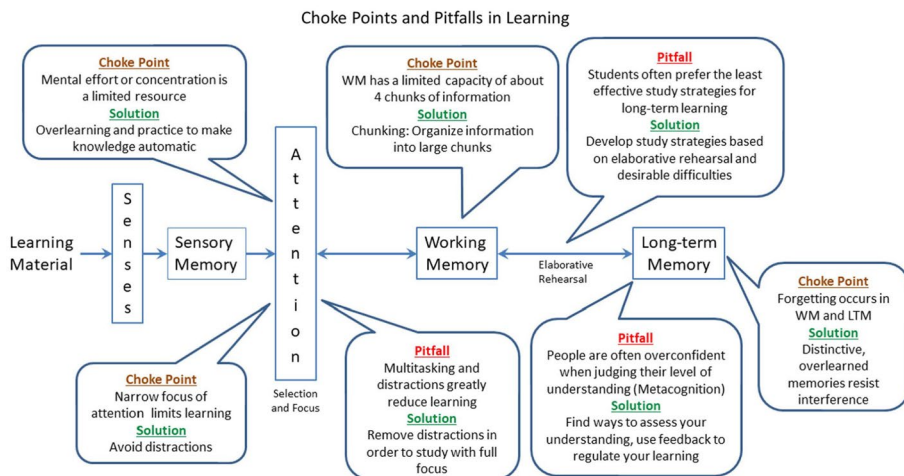

Fig. 1 The cognitive science model for information processing. From Johnstone (2010)

Choke Points and Pitfalls in Learning

evidence that when instructors design instruction based on the science of both their discipline *and* how the brain learns, students reach higher levels of achievement (Rosenshine 2012; Willis et al. 2021; Zhang et al. 2021; Gulacar et al. 2022).

Our paper compares Johnstone's understandings to the current scientific consensus on how the brain solves problems. Where possible, we cite research summaries written by cognitive scientists for educators that limit technical terminology but reference extensive peer-reviewed studies. Our goal is to assist faculty in designing experiments to better align instruction with the science of learning.

Cognitive science and learning

Cognitive science is the study of how the brain thinks and learns. Contributing disciplines include neuroscience, evolutionary biology, and cognitive and educational psychology. In chemistry, we ask students to solve problems. Within cognitive science, a sub-discipline focuses on how the brain solves problems and learns to do so.

Types of problems

By our DNA, humans are 'programmed' to solve some types of problems automatically (Geary 2002). We take our first breath automatically in response to appropriate stimulus and learn to communicate with our parents from day 1. But not all knowledge needed for survival is programmed. *Learning* evolved to adapt individuals to varied environments. To learn is to move knowledge gained through experience into the brain's *long-term memory (LTM)* (Dehaene 2021, p 25).

In a 1991 paper Johnstone asked, "Why is science difficult to learn?" Over the next decade, evolutionary psychologists discovered part of the answer.

Instinctive primary learning

During 'window periods of human development' (also known as 'sensitive periods'), for *limited* topics, the brain evolved to instinctively, automatically, and seemingly effortlessly store knowledge gained from experience in LTM (Pinker 2007; Geary 2002; Geary and Berch 2016). As one example, creating speech is cognitively incredibly complex, but simply by exposure, children become fluent in speaking the language they hear spoken around them. During speech, we apply complex rules with minimal conscious knowledge of what those rules are (Pinker 2007).

Children also learn automatically in *limited* additional topics, including facial recognition, conventions of social relationships (evolutionary psychologists term these 'folk psychology'), and a practical 'folk physics' and 'folk biology' of how things work (Geary 2002; Geary and Berch 2016). Over thousands of generations, these drives evolved to promote survival in difficult primitive environments. Instinctive and automatic learning is termed evolutionarily *primary*. During childhood, extensive primary knowledge is automatically stored in LTM.

Secondary learning

Because reading, writing, and mathematics did not assist in prehistoric survival, their learning did not evolve to be automatic. Non-instinctive learning can be achieved but nearly always requires effort, attention, rehearsal, and practice at recall (Geary 2002; Geary and Berch 2016). Learning that requires effort is termed evolutionarily *secondary*.

Among cognitive scientists who study secondary problem-solving, a 'problem' is broadly defined as "cognitive work that involves a moderate challenge" (Willingham 2009a). The purpose of schools is to structure secondary learning: to teach citizens to solve problems found in modern society that we do not learn to solve automatically (Sweller 2008). When the window period for learning a primary topic closes, it becomes secondary. For example, after about age 12, gaining fluency in a new language nearly always requires effort (Pinker 2007, pp. PS17–18).

Types of problems

Recurrent and non-recurrent problems

Secondary learning can be divided into problems we learn to solve automatically after exerting the necessary effort, and those we do not. Cognitive experts term the former *recurrent* if they are familiar and encountered often. Recurrent skills are "performed as rule-based processes after the training; routine and sometimes fully automatic aspects of behavior". (Van Merriënboer and Kirschner 2007) Keeping your car in its lane becomes automatic and often unconscious with practice (Clark 2006; Dehaene 2021, pp. 160–61). Most problems assigned in introductory chemistry are *non-recurrent*—not yet solvable automatically.

Well-structured problems

Cognitive experts divide non-recurrent problems into two types. *Well-structured problems* have a specific goal, initial state, constraints, and precise correct answers that can be found by step-wise procedures. Scientific calculations are one example. All other types of unfamiliar problems can be categorized as *ill-structured*, including those for which correct answers are debatable or not known (Jonassen 1997).

Bennett (2008) determined that "over ninety percent" of chemistry examination questions at universities in England were well-structured. In widely-adopted U.S. textbooks, a similar portion of 'end-of-chapter' problems are well-structured.

This focus is determined by student goals. For example, in the U.S., for each college-graduating chemistry major about 14 students graduate in biology, health care, or engineering majors (Trapani and Hale 2019). Cognitive experts view chemistry, physics, and mathematics as overlapping subsets of the words and symbols invented by science to explain the physical universe. To learn the language, first-year chemistry is expected in part to teach strategies to solve well-structured problems encountered across the sciences which science knows how to solve precisely. Lecture sections grade primarily on how well students solve well-structured problems.

Because the work of scientists can impact public safety, students in 'science-major general chemistry' are taught to solve well-structured problems by applying proven procedures (algorithms) rather than less-reliable 'heuristics' that may involve speculation.

Scope

For the remainder of this article, except where noted, we limit our scope to questions of how students solve well-structured problems of the type assigned in the lecture component of college 'science-major' general chemistry and in college and high school courses preparing students for general chemistry.

Cognitive architecture

In broad outline, the IP model diagrammed by Johnstone above continues to be applied by cognitive scientists to explain problem solving and learning. The model has three major components: the *perception filter (attention)*, working space (or *working memory—WM*), and *long-term memory* (*LTM*) (Fig. 1).

Perception

Johnstone (2010) wrote that at every conscious moment, "we are victims of … a torrent of sensory stimuli," but "we have a filtration system that enables us to … focus upon what we consider to matter." (Fig. 1) The perception filter, he wrote, "must be driven by what we already know" to let limited items through. In stoichiometry, a student might read to first answer "what unit and formula are wanted?" The strategy previously stored in LTM (learned) will determine which data receive attention.

Johnstone then described the "information we admit through the filter" entering "working memory" (2010).

WM and attention

WM holds what we are conscious of at the moment. Johnstone described working memory as the.

> limited working space in which conscious thought takes place, bringing together new information admitted through the filter and information retrieved from long-term memory. There they interacted, looking for linkages between old and new knowledge (i.e., making sense of the new) … This working space has two functions: to hold information temporarily and to process it. (2010)

This summary is consistent with the current cognitive science explanation of working memory's functions during problem solving (Chew 2021; Dehaene 2021, pp. 160–61).

In recent descriptions, the 'perception filter" is often labeled 'attention' to emphasize a distracting noise or image can shift attention and send extraneous data into WM. Distracting data may bump stored problem data out of WM, which tends to cause confusion (Alloway and Gathercole 2006). From 'attention' we draw our first instructional implication of the IP model. Students may benefit from advice to study in a library or similar location where attention is on learning—and cell phones are *off*.

Long-term memory and neurons

Since Johnstone's retirement, the detail of science's description of LTM has been updated, in part based on neuroscience's improving ability to observe the brain during learning.

LTM is where we store what we have learned in networks of specialized cells termed *neurons*. Each neuron can form connections with and share information among thousands of other neurons. In neuronal networks, information is *encoded*: stored as representations that can be recalled to solve problems.

A neuron can *fire*, meaning it can create an electrical impulse which can transmit stored information to other neurons. Surrounding a neuron's central cell body are thousands of small fibers that carry incoming electrical signals to the central body and thousands of fibers that carry outgoing signals (Dehaene 2021, p. 10).

Synapses are structures at narrow gaps between wires of two neurons. When an outgoing signal reaches a synapse, molecules termed neurotransmitters, such as serotonin and dopamine, can be released and cross the gap. This can cause the fiber of the adjacent neuron to send an electrical signal to its cell body, which may cause the neuron to fire. In neuroscience, the fibers and synapses that can carry signals are termed the brain's *wiring*. Via wiring, each neuron can connect and exchange signals with thousands of other neurons (Dehaene 2021, pp. 86–89).

During learning, changes are made in the brain's wiring that enable knowledge to be stored. The human brain contains over 80 billion neurons, each with about 10,000 synaptic connection whose strength can vary. This gives human LTM an enormous potential capacity to store learning (Dehaene 2021, p. 10).

Memory storage in chunks

How does LTM store learning? Let's begin with an overview. LTM comes "pre-wired" to be able to break images and sounds into small elements LTM can encode (store). Among those elements, new connections in LTM can be made that result in new learning. (Anderson and Lebiere 1998) For example, a child learns to recognize a certain lines and curves as the number 5. During this learning, the brain creates a "wired connection" storing the fact those encoded lines, curves, and arrangement represent the symbol 5. (Dehaene 2021, pp. 6, 86, Furst 2020).

The storage of these relationships is said to create an LTM *chunk*. Cognitive science defines a *chunk* as a collection of connected knowledge elements that has meaning. (Anderson 1996; Willingham 2006; Dehaene 2021, p. 6). In chemistry problems, data are typically words, numbers, or symbols previously stored in LTM as a small chunk.

As learning is applied to solve new problems, small knowledge chunks become wired into increasingly larger chunks. The symbol and sound and spelling of 5 become linked into a larger chunk. As more complex problems are solved, this *chunking* combines growing networks of smaller chunks into rich and well-organized *conceptual frameworks* for topics (each termed a *schema*, plural *schemata*) (Taber 2013, Dehaene 2021, pp. 220–225, Kalyuga et al. 2003).

Cues prompt recall of chunks

How do students solve chemistry problems? A summary would be: problem data moved into WM that has been previously been stored within larger LTM chunks can locate useful relationships within those chunks. Those relationships can be recalled to convert problem data to the answer.

As the brain focuses attention on a problem, input collected by the senses enters WM. The brain uses the elements of data entering WM as a *cue* to search LTM for a matching chunk (Willingham 2008; Dehaene 2021 p 90). When a match for a cue such as '5' is found, LTM neurons holding the match *activate* (fire), sending a signal to other neurons to which they are connected (chunked). Depending on the characteristics of the signal, neurons in these connected chunks may activate and fire. Neuroscientists have observed the working of WM during problem solving as "the vigorous firing of many neurons" primarily in the brain's cortex (Dehaene 2021, p. 90, 160).

These *activated chunks* are said to be *recallable* by WM) (Willingham 2006; Furst 2018; Anderson et al. 2004). In WM, unique problem data *plus* activated, recallable chunks stored by previous learning can converted and integrated, step by step, to reach the problem goal (Alloway and Gathercole 2006; Willingham 2009a; Dehaene 2021, pp. 159–60). An example would be 'units to moles to moles to units' conversions in stoichiometry.

In brief, that is science's current description of how the brain solves problems. The steps are more detailed but similar to Johnstone's 2010 explanation of the operation of the IP model.

What makes learning science difficult?

Asking "Why is science difficult to learn?" Johnstone prepared a list of chemistry topics students found difficult. Research by his group found that for "questions of increasing complexity … a point is reached after which most students fail. It comes somewhere between five and six pieces of information and operations." (1997).

Working memory limits

Seeking an explanation, Johnstone found research by Harvard psychologist George Miller which Johnstown summarized, "working memory space can hold seven plus or minus two pieces of information … if no processing is required…. [If] holding and processing…. values are nearer five plus or minus two." (2010).

Johnstone's determination of WM capacity essentially replicated Miller's. Johnstone was among the first educators to consider the implications of this 'working memory limit' for student problem solving. He explained, "working space is of limited capacity… If there is too much to hold, there is not enough space for processing." (1997).

Johnstone also noted Miller's finding of a way to work around the WM limit: "This process is called chunking and it is this that enables us to use the limited working space efficiently." (1997).

How does chunking circumvent WM limits? Try this brief experiment.

Find a pencil and paper. Mentally read three times: 6.021023
Look away and write the value 100 times larger.
Mentally read three times: 4.850279
Hide the number and write the value 10 times larger.

Which problem was easier? For a chemist, the first number may contain smaller chunks (numbers) previously memorized together that can be recalled as a larger chunk (i.e., Avogadro's number). With fewer chunks in WM slots, remembering is easier and more room is available in WM for processing (De Groot, 1946; Chase and Simon 1973}.

Johnstone theorized chemistry was difficult because "the quantity of information that needs to be manipulated, … [was a] major source of *overload* of working memory." (2010) He noted Miller's observation that for experts, whose memory was organized in large chunks, WM limits could be circumvented (1980). But Johnstone worried, given an average WM capacity of five items, "Even the 'simplest' mole calculations require more than five manipulation steps for a novice…." and as courses progressed, "the complexity increased, leaving many students intellectually stranded." (2010).

Though WM is limited, there is good news. Science has discovered two additional strategies, in addition to chunking, to circumvent WM limits.

Automaticity and overlearning

Daniel Willingham is a cognitive scientist who has made a special effort to disseminate findings of cognitive research to educators. In a 2004 article, he advised the "lack of space in working memory is a fundamental bottleneck of human cognition" but described *two* recently discovered strategies, in support of chunking; to work-around WM limits: automaticity in recall of facts and automaticity in the application of algorithms. Both can be achieved by student effort termed *overlearning:* repeated practice to perfection in recalling information (Geary 1994; Willingham 2004).

In 2006, Kirschner, Sweller, and Clark summarized that WM is "very limited in duration and capacity" when processing information not quickly recallable from LTM, but "when dealing with previously learned information stored in long-term memory, these limitations disappear." In a 2008 report on well-structured problem solving, leading cognitive scientists Geary, Berch, Boykin, Embretson, Reyna, and Siegler advised:

There are several ways to improve the functional capacity of working memory. The most central of these is the achievement of automaticity, that is, the fast, implicit, and automatic retrieval of a fact or a procedure from long-term memory. (pp. 4–5)

'Achieving automaticity in retrieval' is also termed *automatization* or *automating recall*. Example: How much is 6 times 7? Quickly answering means the multiplication has been automated. Geary et al. add:

In support of complex problem solving, arithmetic [fundamental] facts and fundamental algorithms should be thoroughly mastered, and indeed, over-learned, rather than merely learned to a moderate degree of proficiency. (pp. 4-6)

To reliably solve multi-step problems in a topic, students must overlearn (*very* well memorize) its fundamental facts and algorithms.

WM'S strengths and limitations

These findings answer Johnstone's fundamental question: how can instructors "remove the learning obstacles?" Chunking works around WM limits –- if needed chunks are made recallable "with automaticity" and overlearned over time. The requirement for automaticity is explained by the measured characteristics of WM.

Limited slots in WM

WM can be described as composed of *slots*. During problem solving, these slots must hold novel (new, problem-specific) chunks of data, including the goal, initial data, answers found at middle steps, and needed relationships that must be 'looked up' or calculated (Willingham 2006; Alloway and Gathercole 2006; Luck and Vogel 2013). Johnstone wrote,

[W]orking space reaches a maximum about the age of 16 …. It seems that it cannot expand beyond that limit, but that we can learn to use it more efficiently in topics … in which we have some expertise… (2010)

Science's current description of WM is similar. If data are words, numbers, or symbols, adults typically have only 3–5 slots that can hold a novel data chunk (Cowan 2001, 2010). If data are supplied from multiple senses (such as both visual and auditory), a few additional slots may be available, but the number of novel slots remains quite limited (Paivio 2014).

On average, WM capacity roughly doubles between age 5 and 12, plateaus in adulthood, and declines in the elderly. In individuals, maximum adult WM capacity is resistant to change (Cowan 2001, 2010). WM is also limited in *duration*. When information is being processed, each WM slot can retain a novel chunk for less than 30 s (Peterson and Peterson 1959).

Unlimited room when automated

WM also has strengths. Information stored in LTM is "kept directly accessible by means of retrieval cues." (Ericsson and Kintsch 1995) This means for activated chunks, WM limits are circumvented. A data cue and the LTM chunks it activates are treated by WM as

one large chunk. Components within the chunk can be accessed by WM to convert data to reach the problem goal.

Capacity and 'bump out'

Johnstone found, "Students began to fail when working space was overloaded." (1997) Recent cognitive studies add detail. At problem steps, if a needed relationship has not been memorized, novel WM will then need slots for both the data cue *and* its needed looked-up relationship. In a complex problem, the limited slots for novel data are likely already full. Trying to store a non-recallable chunk in WM is then said to cause *overload* (Willingham 2004; Gathercole and Alloway 2004; Furst 2018).

What happens in overload? If 'phosphate ion' is supplied as data, it needs one slot. But if its multi-component *formula* is needed and must be looked up, storage for transfer of the formula's components to paper will also require WM slots. If WM slots are already full, either the formula chunks will not store, or they store by 'bumping out' problem data previously stored in slots. If the 'bumped out' data chunk is needed at later steps of processing, confusion results, as when trying to remember 4.850279 during processing.

In contrast, if the name to formula relationship is in a well-memorized chunk in LTM, only one component needs a novel WM slot and overload is less likely.

Speed and 'time out'

During problem solving, speed is also important. Automated information is recalled instantly. Finding an answer from a table, calculator, or the internet takes time. Because information in WM is stored for less than 30 s during processing, during a search, other problem data being held in novel slots tend to 'time out' and be lost. But if a relationship is automated, information 'timed out" can be quickly restored by cued recall (Ericsson and Kintsch 1995; Kirschner et al. 2006; Willingham 2008).

The impact of limits

When prior production of chunks by students has been limited, Alloway and Gathercole advise,

> The capacity of working memory is limited, and the imposition of either excess storage or processing demands in the course of an on-going cognitive activity will lead to catastrophic loss of information from this temporary memory system. (2006)

Among cognitive scientists, WM limits and their impact are not contested. Neuroscience and cognitive psychology may vary in terminology, but all current scientific descriptions of problem solving include space in WM that is stringently limited for novel data and non-recallable relationships but essentially unlimited for relationships quickly recallable.

Three work-arounds

Johnstone summarized,

> Chunking usually depends upon some recognizable conceptual framework that enables us to draw on old, or systematize new, material. For an experienced chemist, the recognition… bases …are related provides a helpful chunking device. (1997)

Since Johnstone's retirement, cognitive research has confirmed that to efficiently learn a new and well-structured topic, to circumvent WM limits, students must take three steps.

1. *Facts* that are fundamental must be made recallable 'with *automaticity'* early in the study of a new topic – by *rehearsal* then *distributed retrieval practice*.
2. Algorithms solving topic problems by applying recallable facts must be *automated* using *interleaved* and *distributed practice* that solves problems in a variety of distinctive contexts.
3. To *chunk* new knowledge into a robust and long-lasting conceptual framework, initially automated facts and algorithms must be *overlearned* by practice in applications over days, then re-visited in weeks, then months.

Richard Clark summarizes, "We appear to have innate, unconscious routines for automating all behavior that is perceived as successful and repeated over time." (2006) Stanislas Dehaene, awarded the 2014 Brain Prize in neuroscience, explains the basis in the brain for the automaticity work-around:

> Automatization mechanisms 'compile' the operations we use regularly into more efficient routines. They transfer them to other brain circuits, outside our conscious awareness. As long as a mental operation …has not yet been automated by overlearning, it … prevents us from focusing on anything else. (2021, pp. 222–23)

Each new topic in general chemistry includes new vocabulary and other fundamentals. To build conceptual frameworks efficiently, fundamentals must first be automated.

Information in automated circuits can be applied with minimum use of the WM where conscious knowledge is held. To work around WM limits, in any structured activity (including sports and music), a goal is to automate facts and procedures needed frequently. After automation, steps are effortless.

Automate factual recall

As a priority, what must be automated? Facts and procedures needed most often. Let's begin with facts.

Recalling facts

A fact (also termed 'declarative knowledge') is composed of two or more related chunks of knowledge. A new term is defined with terms previously memorized. If 'mole' is unfamiliar but types of small particles and exponential notation have both been well learned, the factual definition '1 mol of particles $= 6.02 \times 10^{23}$ particles' can be understood.

Facts can be definitions of words, symbols, and abbreviations. In introductory courses, pico-, phosphate, potential energy, proton, photon, P, *P*, Pb, *pH, pKa,* and kPa are likely to be vocabulary of an unfamiliar foreign language. When new terms are

needed for a new topic, definitions must be automated early in study as a foundation for understanding.

Facts can be rules. $10^6/10^{-3} = 10^9$. Cation is pronounced as cat ion. Facts can be mathematical relationships. $pH = - \log[H^+]$. Individual facts can be automated into larger chunks with conditions and context chunks attached. For ideal gases: $PV = nRT$. For first-order kinetics: $\ln(\text{fraction remaining}) = - kt$.

When listening to lecture, reading a text, or solving problems, if the definition of a new terms have *not* been stored in automated circuits, 'information overload ' can occur. But after automating the meaning of new technical vocabulary, students can listen to lecture and read detailed textbooks with improved comprehension.

Johnstone and Kellett wrote in 1980, "If the pupil can 'chunk' the information [s]he may have sufficient 'spare capacity' to operate upon the information with some hope of success." We know now what in 1980 they did not. Chunking can provide more than "hope" for help in solving complex problems—if recall of new fundamentals has first been automated.

Maintenance rehearsal

At the start of a new topic, needed new facts can be learned quickly. The steps are: the instructor supplies a limited list of new and prerequisite facts for the topic, students practice *rehearsal* and *retrieval practice* of those fundamentals for a few days, and a brief announced quiz encourages assignment completion. Without this instructor guidance, novice learners tend to have difficulty identifying what is most important in a detailed reference text.

Factual learning begins with *maintenance rehearsal* (or simply *rehearsal*). Repeatedly reciting an unfamiliar phone number—or an unfamiliar chemistry definition—is the first step in making it recallable. To speed learning, maintenance rehearsal should involve as many senses as possible: hearing, seeing, saying, thinking, and writing the fact repeated for several days.

When learning to speak, the brain of a child automatically moves frequently overheard phrases into LTM. After about age 12, learning phrases in a new language, including chemistry, becomes more difficult, but it remains achievable with effort and practice (Pinker 2007, pp. PS17–18). Flashcards can help. If a student writes 'phosphate' on an index card and $PO_4{}^{3-}$ on the other side, then finds a place to whisper without distracting others: "phosphate is (flip) P O 4 3 minus," recital practice for several days in both directions automates recall of both phrases.

Elaborative rehearsal

Thinking about meaning is *elaborative rehearsal*. Here, associations are made between the new information and what you already know. This can involve organizing or reorganizing, thinking of examples, creating an image in your head, or applying a mnemonic device. For information which is complex or likely to be encountered in specific contexts, elaborative rehearsal is generally more efficient than maintenance rehearsal at making new information retrievable (Bransford 2000; Willingham 2004).

While both types of rehearsal should be practiced, when learning *simple* facts, maintenance rehearsal is usually sufficient.

Retrieval practice

Retrieval practice divides rehearsal into a question and an answer, then requires the answer be recalled. Retrieval practice creates a *desirable difficulty* (Bjork and Bjork, 2019). When effort is made to recall information from LTM, it strengthens the wired connection within the rehearsed chunk. The effort can include self-testing such as flashcard use, teachers giving short, no-stakes quizzes, writing tables of relationships (such as metric-prefix definitions) from memory, recalling mnemonics (RoyGBiv), sequence recitation (methane, ethane, …), and answering clicker questions (Agarwal et al. 2013; Carpenter and Agarwal 2019).

Flashcards are especially useful for simple relationships such as vocabulary definitions. If possible, *each* of the two parts of a flashcard relationship should be practiced as visual cues, spoken cues, and for written practice: Seeing one side, write the other. Writing words teaches spelling. Writing chemical symbols teaches case, subscripts, and superscripts. With sufficient rehearsal and retrieval practice, answers are moved into automated circuits in LTM, working around WM limits. That's the goal.

Retrieval practice applies the *testing effect*: recall is strengthened more by testing, including low- or no-stakes or self-testing, than by highlighting or re-reading (Dunlosky 2013; Brown et al. 2014, Deans for Impact 2015).

Verbatim versus gist learning

Retrieval practice is especially important in chemistry because knowledge in the physical sciences is precise. Geary et al. (2008) note that *verbatim* (precise) facts "are encoded separately from *gist*" (summary) memories, and verbatim information "often requires more effort to learn than the gist…. Verbatim recall … requires a great deal of time, effort, and practice." For example, everyone has a summary, operational, gist understanding of temperature, but "temperature is a measure of the average kinetic energy of particles" requires overlearning to remember long-term.

Why retrieval works

Dehaene suggests we advise students:

> Active engagement followed by error feedback maximizes learning…. Using flashcards, for each card, try to remember the answer (prediction) before checking it by turning it to the other side (error feedback). (2021, p.186)

Repeated engagement, prediction, and feedback signals the brain to move the relationships among chunks being processed in WM into automated circuits in LTM.

Willingham's best-known adage may be, "memory is the residue of thought." (Willingham 2009c) We tend to remember what we think about, especially if we think about it often. For example, during a lecture or reading assignment that includes many new terms, an occasional pause for 'clicker questions' can move short definitions from WM into initial LTM storage. (Trafton 2017) This can free space in WM for additional new information, assisting for a day or two with comprehension of speech or reading, and tends to assist longer-term if retrieval is repeated (Willingham 2006).

Overlearn with spaced practice

Recall practiced several times in one long sitting (such as the night before a test) but *not* in shorter sittings on multiple days is termed *massed practice* or *cramming*. Crammed knowledge can help for the next day or two (increasing *performance*) but tends to be quickly forgotten (it is not *learned*; there is no change in LTM) (Kirschner et al. 2006). After several days without re-study, to regain recall of crammed knowledge tends to require a repeat of intensive and time-consuming study.

The *spacing effect* is the improved recall that results from retrieval practiced over multiple days. Repeated practice over several days, combining the *testing* and *spacing* effects, is termed *distributed practice* (Willingham 2002a, 2015; Carpenter and Agarwal 2019). For a day or two after massed practice, learning can be used to solve problems because *retrieval strength* is high. However, *storage strength* is low (Bjork and Bjork 2019). Distributed practice promotes both storage in and retrieval of knowledge from LTM.

Repeated practice to *perfection* is *overlearning*. Willingham advises, "Practice makes perfect—but only if you practice beyond the point of perfection… Regular, ongoing review … past the point of mastery" is necessary to move knowledge into automated circuits. (Willingham 2004).

Overlearning distributed over weeks and months is termed *spaced overlearning.* Neuroscientists advise that when studying new facts and procedures, "To keep the information in memory as long as possible…. start with rehearsals every day, then review the information after a week, a month, then a year." (Dehaene 2021, p. 219) For maximum *long*-term retention, the goal in study for a science career, spaced overlearning is required.

Must everything be well-memorized? No. Willingham suggests what should be overlearned are "core skills and knowledge that will be used again and again." (2004) Geary et al. (2008) write that *fundamental* facts and procedures should be overlearned. If fundamentals are not thoroughly memorized at the start of a topic, cognitive experts predict new learning will not be as efficient and effective as needed for the rigor and pace of science-major science courses (Kirschner et al. 2006; Willingham 2009b).

Making learning stick

On a cumulative final examination, if facts, algorithms, and concepts are remembered, they were successfully overlearned. But three months later, some knowledge will no longer be recallable. In six months, more will be forgotten (Ebbinghaus 1885). Is that a problem? Cognitive experts say no.

Johnstone expressed the concern that "rote-learned material …. is easily lost" from LTM (2010). Cognitive studies have found, for material previously *overlearned*, forgetting may occur, but if needed in a higher-level course, quick recall can be restored with far less re-study. Cognitive experts call this ability to refresh overlearned memory the 'savings in relearning' effect (Willingham 2002a, 2015). After the quick review, the necessary foundation of prerequisites is accessible to expand conceptual frameworks by problem solving.

Spaced overlearning also tends to 'flatten the forgetting curve,' meaning what has been learned tends to be better remembered for longer periods of time (Ebbinghaus 1885). Spacing day or longer time gaps between retrieval practice sessions lead to some forgetting, creating a desirable difficulty: The increased mental effort required for retrieval (i.e., the difficulty) promotes longer lasting cued recall (i.e., which is desirable). 'Forgetting then remembering' strengthens both storage and retrieval (Bjork and Bjork 2019).

Rote memorization

Johnstone's concern, shared possibly by other educators, was that "students can imagine that learning chemistry is a rote process." (1997) Though a new chemistry topic must get well beyond learning initial vocabulary, in introductory courses, new topic fundamentals will tie up WM slots *until* they are well-memorized. After fundamentals are automated, solving problems wires the new fundamentals to prior learning, speeding construction of the robust cognitive schemata needed for deeper conceptual understanding. Neuroscience educator Efrat Furst summarizes, "Time is better spent at teaching the basics than trying to teach the new without it." (2018).

Willingham advises much of what is deprecated as "rote learning" is actually "inflexible knowledge" with a narrow meaning. He advises, "What turns the inflexible knowledge of a beginning student into the flexible knowledge of an expert seems to be a lot more knowledge, more examples, and more practice." (2002b).

Incidental memorization

Instead of by initial retrieval practice, can students learn new vocabulary as they solve problems? Yes, but it is slow and frustrating. In a problem with unfamiliar terms, WM overloads quickly. If a student uses problems to learn what new terms mean, given WM limits, solving the problem will be less likely. Success is motivating; repeated failure is not.

Automate algorithms by practice

Johnstone expressed concern that, because WM capacity was limited, the multiple steps of complex calculations would leave "many students intellectually stranded." (2010) Cognitive science research has recently shown that even for complex, many-step problems, structured algorithms can work around the WM bottleneck.

Algorithms

An *algorithm* (also termed a *well-structured* or *fixed procedure)* is a 'recipe' that solves a type of complex problem in a sequence of steps (Willingham 2009a; Van Merriënboer and Kirschner 2007). Examples of algorithms include sequences remembered by mnemonics (RICE, ICE, BCA tables), solubility schemes, the algorithms of arithmetic and algebra, and the steps of a worked example.

For a specific type of problem, a useful algorithm is one that has empirically proven to successfully convert data to an answer, breaking a problem into a sequence of steps such that at each step, WM does not overload (Geary et al. 2008; Willingham 2009a, 2009b). Cognitive studies have found, "With mastery… algorithms can be executed automatically and without need for explicit recall and representation of each problem-solving step in working memory." (Geary et al. 2008, pp. 4–32) Practice moves algorithms into automated circuits that minimize the need to for WM processing.

When trying to solve problems with complex data and/or multiple steps, reasoning strategies that do not rely on algorithms nearly always fail. Trying to reason without an algorithm quickly overloads WM (Kirschner et al. 2006; Alloway and Gathercole 2006; Geary et al. 2008).

For each problem type, many algorithms will likely work. However, if multiple algorithms are practiced, recall of different steps tends to 'interfere' with each other. For this reason, cognitive experts suggest instructors identify and teach one 'best' algorithm for each problem type (Anderson and Neely 1996; Dewar et al. 2007). The most useful algorithms will be widely applicable, rely on fundamental concepts, apply fundamental factual relationships, and be easy to remember.

Algorithms require implicit retrieval

Avoiding WM limits requires "the fast, *implicit*, and automatic retrieval of a fact or a *procedure*" from LTM (Geary et al. 2008). *Implicit retrieval* can be described as *intuitive* or *tacit* recall, which may not include a conscious ability to identify a particular chunk or why it was recalled (Clark 2006). Implicit means the student must be able to look at problem data and intuitively, fluently decide which automated facts and algorithms to apply.

To help students gain algorithmic fluency, instructors should assign practice problems containing typical problem cues and distinctive problem contexts (Willingham 2003). Practice that processes context cues at the same time as facts and algorithms tends to connect all of those related chunks into a larger chunk in an accessible memory schema. The brain is then able to choose correct facts and algorithms to recall intuitively in a manner similar to assembling fluent speech, a task humans evolved to accomplish with ease (Geary et al. 2008; Pinker 2007). Anderson and Neely write,

> Retrieval cues can be anything from component of the desired memory to incidental concepts associated with that item during its processing.… Retrieving a target item [occurs] when the cues available at the time of recall are sufficiently related to that target to identify it uniquely in memory. (1996)

Interleaved practice

In problem sets, practicing one algorithm is termed *blocked practice*. Mixing problem types that require different algorithms in a random order is termed *interleaved practice*. Seemingly different problems may require the same algorithm while seemingly similar problems may require different algorithms. Practice that is interleaved helps students to discriminate among different problem types. With interleaving, solving is initially more difficult, but the difficulty is desirable. Long-term, students attain improved memory of which cues and contexts are paired with which algorithms (Bjork and Bjork 2019; Gulacar et al. 2022).

Johnstone observed, if "much information has to be held [in WM], little room remains for processing." (2010) However, if facts and algorithms have been automated, WM slots tend to remain open to store context cues. Students then learn a sense of which facts and algorithms to apply while solving *fewer* problems, improving study efficiency.

Concepts and reasoning

Concepts are fundamental principles that categorize and explain knowledge, often within a hierarchical structure. As defined by Geary et al., "conceptual knowledge refers to general knowledge and understanding stored in long-term memory." (2008)

Willingham writes, "conceptual knowledge refers to an understanding of meaning, … understanding *why*" something is true (2009b). Potential energy, conservation of energy, and conservation of matter are examples of concepts that organize knowledge components.

Facts before concepts

Cognitive science emphasizes that conceptual understanding is vitally important and needs to be taught (Siegler and Lortie-Forgues 2015). Geary et al. advise, "The cognitive processes that facilitate rote retention… such as repeated practice, can differ from the processes that facilitate transfer and long-term retention, such as conceptual understanding." (2008).

Learning concepts helps the brain efficiently consolidate knowledge in LTM by the deeper structure of its meaning. But among facts, procedures, and concepts, Willingham writes, "conceptual knowledge is the most difficult to acquire… A teacher cannot pour concepts directly into students' heads. Rather, new concepts must build on something students already know." (2009b) Furst summarizes, "New knowledge is built on the basis of the previous knowledge and they must be related by meaningful connections." (2019).

Willingham cites evidence that during the initial moving of information into memory, "the mind much prefers that new ideas be framed in concrete rather than abstract terms." (2002b) Johnstone's advice in chemistry was similar: "Begin with things that they will perceive as interesting and familiar so that there are already anchorages in their long term memory on which to attach the new knowledge," (2000) and "Concepts must be built from the macroscopic and gradually be enriched with submicroscopic and representational aspects." (2010).

If examples to illustrate concepts are simple, concrete, and familiar, and quantitative reasoning can be solved using automated facts and algorithms, explanations of concepts tend to avoid WM overload (Willingham 2002b; Geary et al. 2008).

Use concepts, not memory?

To solve problems, some in chemistry education have advocated using "online resources" to 'look up' rather than memorize factual knowledge (Pienta 2018). This assumption, that the brain could apply new information with the same facility it applies well-memorized information, has proven to be mistaken. Cognitive studies have found stringent limits apply when processing not-quickly-recallable information that do not apply when processing relationships quickly recallable from LTM.

Some instructional reform proposals have expressed concern that "students are often able to answer well-defined (i.e., close-ended) problems, without making use of conceptual understanding." (Cooper and Stowe, 2018) Widely implemented 'reform' curriculum for general chemistry topics has been based on "shifting away from a paradigm that emphasized algorithmic problem solving and content knowledge…" (Rushton 2014).

Students need to be taught concepts, but cognitive research has established that to reliably solve problems of any complexity, reasoning based on conceptual understanding that does not apply memorized algorithms and memorized content fundamentals is highly

unlikely to work. Without recallable algorithms and facts, WM overloads (Geary et al. 2008; Rosenshine 2012).

John R. Anderson, a leader in the study of information processing, counsels, "one fundamentally learns to solve problems by mimicking examples of solutions," such as those learned from worked examples (1996). Willingham advises that when students *seem* to solve complex problems based on "understanding" without a recalled procedure, research nearly always finds "understanding is remembering in disguise." (2009c, pp. 68–72).

Reasoning

When does generalized reasoning without an algorithm work to solve problems? Experts can reason in their discipline because of their vast storehouse of knowledge in LTM (De Groot 1946; Chase and Simon, 1973; Kirschner et al. 2006). For topics of primary (instinctive) learning, including speech and the 'folk' biology, physics, and psychology of daily life, the brain fills with knowledge automatically. In those areas, we all become experts and able to reason generally. But in secondary learning, including chemistry, moving knowledge into LTM is not automatic. Becoming an expert requires years of increasingly complex study. In chemistry, students can use general reasoning strategies to solve *only very* simple problems in which data and steps do not overload WM (Geary et al. 2008).

How to speed chunking

In 1980, Johnstone and Kellett wrote, "problem-solving ability is associated with students' ability to organize or 'chunk' the information …." Forty years later, chunking remains one of the three strategies identified by cognitive experts that circumvent WM limits. How can instructors help to speed the rate at which students wire chunks?

Chunking in detail

To solve problems, WM holds and processes information. This processing is also the first step in learning. Furst writes: "processing in working memory is … information's 'entry ticket'" to LTM (2018). During problem solving, if a step is "perceived as useful and successful" (Clark 2006), the chunks that are processed together in WM either tend to form new connections or their existing connections are reinforced in LTM (Trafton 2017). In the (simplified) formula of neuroscience: "Neurons that *fire* together, *wire* together." (Hebb 1949) Feedback can help to signal when a step has been successful (Dehaene 2021, p. 186).

If similar processing does *not* take place again over several days, new wiring tends to be lost (Willingham 2015; Dehaene 2021, p. 216). But if chunks are repeatedly processed together during the next several days, a record of chunks processed at the same time tends to become *consolidated*: organized and wired into *long*-term memory (Taber 2013, Trafton 2017; Dehaene 2021, pp. 221–235; Furst 2020).

Using advanced microscopy, neuroscientists have imaged the brain's learning *plasticity*: the growth and strengthening of synaptic connections among neurons in response to problem solving, as well as the loss of connections not repeatedly used in processing (Yang et al. 2014; Dehaene 2021, p. 87, 137). As new relationships are applied to subsequent

problems, connections among related components strengthen, so that if one component is a problem data cue, it more likely brings to mind (activates) others in the larger chunk.

Room for context cues

Johnstone noted that in WM, "if much information has to be held, little room remains for processing." (2010) Science's updated understanding is similar. To speed the construction of conceptual frameworks (schemata), as many slots as possible in WM should be kept open to store "context cues" during processing. The context in which a problem is solved helps to provide an intuitive, implicit, automated sense of which facts and procedures to recall out of the millions stored in LTM.

The bottleneck of WM limits means that learning must be gradual, step by problem-solving step. But by solving problems in a variety of distinctive contexts, cued recall can solve more problems with more success, speeding the rate of learning. Willingham notes "Knowledge is not only cumulative, it grows exponentially. Those with a rich base of factual knowledge find it easier to learn more—the rich get richer." (2006).

Inquiry and discovery

Cognitive research strongly supports activities that engage students, but the timing of those activities markedly affects the learning that may result. Scientific studies have consistently found "the most effective teachers" employed "hands-on activities, but they always did the experiential activities after, not before, the basic material was learned." (Rosenshine 2012) Cognitive studies also have repeatedly shown that it is not a best practice to ask students to engage in "inquiry" to discover what experts struggled to learn (Kirschner et al, 2006; Mayer 2004). Dehaene summarizes: "Discovery learning methods are seductive ideas whose ineffectiveness, unfortunately, has repeatedly been demonstrated." (2021, p. 180).

Conclusion

Johnstone recommended that as instructors, we consider the implications of cognitive science research. Scientists who study the brain have found problem solving of any complexity must be based on remembering, not reasoning. Efficient learning builds on overlearned knowledge of fundamental facts and algorithms, gradually moving the student toward the knowledge and intuition of an expert.

Science's discovery of the necessity for spaced overlearning to achieve efficient long-term learning is a paradigm shift: a change in science's fundamental understanding of how the brain manages information. Accepting paradigm shifts can be difficult (Kuhn 1962).

In the United States and some other nations, many instructional reforms have been proposed, and some adopted, that de-emphasize memorization and emphasize reliance on conceptual understanding and non-algorithmic reasoning. Others de-emphasize explicit instruction in favor of inquiry or discovery learning. In crucial respects, those reforms oppose the strategies cognitive science has verified are necessary to learn chemistry. This may be contributing to the high failure rates seen in general chemistry.

If memorization can be reduced, we should try. But learning is a progression. Automating recall of new and pre-requisite fundamentals is an essential first step in creating a conceptual foundation for subsequent learning.

Johnstone's fundamental insight? Students benefit when instructors know both the science of molecular behavior and the science of learning. Cognitive studies have identified specific steps students can take to learn with more success, and how instructors can assist them in doing so. This scientific progress is potentially a great gift to students and our society.

Johnstone sought ways "that students will learn efficiently," but also "with understanding and enjoyment." (2010) Applying cognitive research, we can design more efficient learning during both lecture and study time. That efficiency opens opportunities during lecture for activities such as demonstrations that engage students in chemistry.

As recently witnessed, science can save millions of lives, but work at the front lines can require personal risk and sacrifice. Students who seek to pursue challenging science majors have courage. If we apply the science of learning in our instruction, can we better help our brave students as Johnstone wished, with efficiency, understanding, and enjoyment, pass through the chemistry gateway to scientific careers?

Declarations

Conflict of interest The authors declare the following competing interest(s): Eric Nelson has co-authored textbooks in chemistry.

References

Agarwal, P.K., Roediger, H.L., McDaniel, M.A., McDermont, K.B.: How to Use Retrieval Practice to Improve Learning. WUSL Institute of Education Sciences, St. Louis (2013). http://pdf.retrievalpractice.org/RetrievalPracticeGuide.pdf. Accessed 14 Feb. 2022

Alloway, T.P., Gathercole, S.E.: How does working memory work in the classroom? Educ. Res. Rev. **1**(4), 134–139 (2006)

Anderson, M.C., Neely, J.H.: Interference and inhibition in memory retrieval. In: Memory, pp. 237–313. Academic Press (1996)

Anderson, J.R.: ACT: A simple theory of complex cognition. Am. Psychol. **51**(4), 355–365 (1996)

Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychol. Rev. **111**(4), 1036–1060 (2004)

Anderson, J., Lebiere, C. (eds.): The atomic components of thought. Erlbaum Publishers, Mahwah, NJ (1998)

Atkinson, R., Shiffrin, R.: Human memory: a proposed system and its control processes. In: Spence, K., Spence, J. (eds.). The Psychology of Learning and Motivation, vol. 2. Academic Press, New York (1968)

Baddeley, A.D., Hitch, G.: Working memory. In: Psychology of Learning and Motivation, vol. 8, pp. 47–89. Academic Press (1974)

Baddeley, A.D.: Essentials of human memory. Psychology Press Ltd., Hove (1999)

Bennett, S.W.: Problem solving: can anybody do it? Chem. Ed. Res. Pract. **9**(1), 60–64 (2008)

Bjork, R.A., Bjork, E.L.: Forgetting as the friend of learning: Implications for teaching and self-regulated learning. Adv. Physiol. Educ. **43**, 164–167 (2019)

Bransford, J.: How People Learn: Brain, Mind, Experience, and School. National Academy Press, Washington DC (2000)

Brown, P., Roediger, H.L., McDaniel, M.A.: Make It Stick: the Science of Successful Learning. Harvard University Press, Cambridge, MA (2014)

Carpenter, S.K., Agarwal, P.K.: How to Use Spaced Retrieval Practice to Boost Learning. Iowa State University, Ames, IA (2019). http://pdf.retrievalpractice.org/SpacingGuide.pdf. Accessed 14 Feb. 2022

Chase, W.G., Simon, H.A.: Perception in chess. Cog. Psych. **4.1**, 55–81 (1973)

Chew, S.L.: An advance organizer for student learning: choke points and pitfalls in studying. Can. Psychol. **62**(4), 420–427 (2021). https://doi.org/10.1037/cap0000290

Clark, R.E.: Not knowing what we don't know: reframing the Importance of automated knowledge for educational research. In: Clareboutand, G., Elen, J. (eds.) Avoiding Simplicity, Confronting Complexity, pp. 3–14. Sense Publishers, Rotterdam (2006)

Clark, R.E., Sweller, J., Kirschner, P.A.: Putting students on the path to learning: the case for fully guided instruction. Am. Educ. **36**(1), 6–11 (2012)

Cooper, M., Stowe, R.: Chemistry education research—from personal empiricism to evidence, theory, and informed practice. Chem. Rev. **118**(12), 6053–6087 (2018). https://doi.org/10.1021/acs.chemrev.8b00020

Cowan, N.: The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behav. Brain Sci. **24**, 87–185 (2001)

Cowan, N.: The magical mystery four: how is working memory capacity limited, and why? Curr. Dir. in Psych. Sci. **19**(1), 51–57 (2010)

Deans for Impact.: The Science of Learning. Deans for Impact, Austin (2015). https://deansforimpact.org/wp-content/uploads/2016/12/The_Science_of_Learning.pdf Accessed 14 Feb. 2022

Dehaene, S.: How we learn. Penguin, New York (2021)

Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Locky, A., Trieb, T.: The acquisition of arithmetic knowledge—an fMRI study. Cortex **40**, 166–167 (2004)

Dewar, M., Cowan, N., Della Sala, S.: Forgetting due to retroactive interference. Cortex **43.5**, 616–634 (2007)

Dhond, R.P., Marinkovic, K., Dale, A.M., Witzel, T., Halgren, E.: Spatiotemporal maps of past-tense verb inflection. Neuroimage **19**(1), 91–100 (2003)

Dunlosky, J.: Strengthening the student toolbox: study strategies to boost learning. Am. Educ. **37**(3), 12–21 (2013)

Ebbinghaus, H.: Memory: A Contribution to Experimental Psychology. Ruger H.A., Bussenius C.E., translator. Columbia University, New York (1885/1913)

Ericsson, K.A., Kintsch, W.: Long-term working memory. Psych. Rev. **102.2**, 211–245 (1995)

Furst, E.: Learning in the brain (2018). https://sites.google.com/view/efratfurst/learning-in-the-brain. Accessed 14 Feb. 2022

Furst, E.: Understanding understanding (2019). https://sites.google.com/view/efratfurst/understanding-understanding. Accessed 14 Feb. 2022

Furst, E.: Reconsolidation (2020). https://sites.google.com/view/efratfurst/reconsolidation. Accessed 14 Feb 2022

Gathercole, S.E., Alloway, T.P.: Working memory and classroom learning. Dyslexia Rev. **15**, 4–9 (2004)

Geary, D.C.: Children's mathematical development: research and practical applications. Am. Psychol. Assoc. (1994)

Geary, D.C., Boykin, A.W., Embretson, S., Reyna, V., Siegler, R., Berch, D.B., Graban, J.: Chapter 4: the report of the task group on learning processes, pp. 4-2–4-10. U.S. Dept. of Education, Washington, DC (2008). https://files.eric.ed.gov/fulltext/ED502980.pdf. Accessed 14 Feb. 2022

Geary, D.C., Berch, D.B.: Evolution and children's cognitive and academic development. In: Evolutionary Perspectives on Child Development and Education, pp. 217–249. Springer, Cham (2016)

Geary, D.C.: Principles of evolutionary educational psychology. Learn. Indiv. Differ. **12**, 317–345 (2002)

De Groot, A.D.: Thought and Choice in Chess. The Hague, Netherlands, Mouton (1946/1965)

Gulacar, et al.: Benefits of desirable difficulties: comparing the influence of mixed practice to that of categorized sets of questions on students' problem-solving performance in chemistry. Chem. Ed. Res. Pract. (2022)

Hebb, D.: The Organization of Behavior. Wiley, New York (1949)

Johnstone, A.H.: Why is science difficult to learn? Things are seldom what they seem. J. Comput. Assist. Learn. **7**(2), 75–83 (1991)

Johnstone, A.H.: Chemistry teaching—science or alchemy? J. Chem. Educ. **74**(3), 262 (1997)

Johnstone, A.H.: Teaching of chemistry—logical or psychological? Chem. Ed. Res. Pract. **1**(1), 9–15 (2000)

Johnstone, A.H.: You can't get there from here. J. Chem. Educ. **87**(1), 22–29 (2010)

Johnstone, A.H., Kellett, N.C.: Learning difficulties in school science–towards a working hypothesis. Euro. J. Sci. Ed. **2**(2), 175–181 (1980)

Jonassen, D.H.: Instructional design models for well-structured and ill-structured problem-solving learning outcomes. Educ. Tech. Res. Dev. **45**, 65–94 (1997). https://doi.org/10.1007/BF02299613

Kalyuga, S., Ayres, P., Chandler, P., Sweller, J.: The Expertise reversal effect. Educ. Psychol. **38**(1), 23–31 (2003)

Kirschner, P.A., Sweller, J., Clark, R.E.: Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem based, experiential and inquiry-based teaching. Educ. Psychol. **41**, 75–86 (2006)

Kuhn, T.S.: The Structure of Scientific Revolutions. University of Chicago Press, Chicago (1962)

Luck, S.J., Vogel, E.K.: Visual working memory capacity: from psychophysics and neurobiology to individual differences. Trends Cogn. Sci. **17**(8), 391–400 (2013)

Mayer, R.: Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. Am. Psych. **59**, 14–19 (2004)

Van Merriënboer, J.J., Kirschner, P.A.: Ten Steps to Complex Learning. A Systematic Approach to Four Component Instructional Design. Lawrence Erlbaum Associates, Mahwah, NJ (2007)

Paivio, A.: Mind and its Evolution: A Dual Coding Theoretical Approach. Psychology Press (2014)

Peterson, L., Peterson, M.: Short-term retention of individual verbal items. J. Exp. Psychol. **58**, 193–198 (1959)

Pienta, N.J.: Hey, professor—something is wrong in the book. J. Chem. Ed. **95**(2), 187–188 (2018). https://doi.org/10.1021/acs.jchemed.8b00032

Pinker, S.A.: The Language Instinct: How the Mind Creates Language. HarperCollins, New York (2007/1994)

Reid, N.: A tribute to Professor Alex H Johnstone (1930–2017): His unique contribution to chemistry education research. Chem. Teacher Int. **1**(1), 20180016 (2019). https://doi.org/10.1515/cti-2018-0016

Rosenshine, B.: Principles of instruction: research-based strategies that all teachers should know. Am. Educ. **36**(1), 12–19 (2012)

Rushton, G.T.: Introducing the journal of chemical education's "Special Issue: advanced placement (AP) chemistry." J. Chem. Educ. **91**(9), 1273–1275 (2014). https://doi.org/10.1021/ed500476r

Siegler, R., Lortie-Forgues, H.: Conceptual knowledge of fraction arithmetic. J. Ed. Psych. **107**(3), 909 (2015)

Sweller, J.: Instructional implications of David C. Geary's Evolutionary Educational Psychology. Ed. Psychol. **43**, 214–216 (2008)

Taber, K.S.: Revisiting the chemistry triplet: drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. Chem. Ed. Res. Pract. **14**(2), 156–168 (2013)

Tang, H., Pienta, N.: Eye-tracking study of complexity in gas law problems. J Chem. Educ. **89**(8), 988–994 (2012). https://doi.org/10.1021/ed200644k

Trafton, A.: Neuroscientists identify brain circuit necessary for memory formation (2017). http://news.mit.edu/2017/neuroscientists-identify-brain-circuit-necessary-memory-formation-0406. Accessed 14 Feb 2022

Trapani, J., Hale, K.: Higher education in science and engineering. In: Science & Engineering Indicators 2020. NSB-2019-7. Table S2-6. National Science Foundation, Arlington, VA (2019)

Willingham, D.T.: Allocating student study time: Massed versus distributed practice. Am. Educ. **26**(2), 37–39 (2002a)

Willingham, D.T.: Inflexible knowledge: the first step to expertise. Am. Educ. **26**(4), 31–33 (2002b)

Willingham, D.T.: Students remember what they think about. Am. Educ. **27**(2), 37–41 (2003)

Willingham, D.T.: Practice makes perfect—but only if you practice beyond the point of perfection. Am. Educ. **28**(1), 31–33 (2004)

Willingham, D.T.: How knowledge helps. Am. Educ. **30**(1), 30–37 (2006)

Willingham, D.T.: What will improve a student's memory? Am. Educ. **32**(4), 17–25 (2008)

Willingham, D.T.: Why don't students like school? Am. Educ. **33**(2), 4–13 (2009a)

Willingham, D.T.: Is it true that some people just can't do math? Am. Educ. **33**(4), 14–19 (2009b)

Willingham, D.T.: Why Don't Students Like School: A Cognitive Scientist Answers Questions About How the Mind Works and What It Means for the Classroom. Wiley, New York (2009c)

Willingham, D.T.: Do students remember what they learn in school? Am. Educ. **39**(3), 33–38 (2015)

Willis, W.K., Williamson, V.M., Chuu, E., Dabney, A.R.: The relationship between a student's success in first-semester general chemistry and their mathematics fluency, profile, and performance on common questions. J. Sci. Ed. Tech. 1–5. (2021)

Yang, G., Lai, C.S., Cichon, J., Ma, L., Li, W., Gan, W.B.: Sleep promotes branch-specific formation of dendritic spines after learning. Science **344**(6188), 1173–1178 (2014)

Zhang L., Kirschner P.A., Cobern W.W., Sweller J.: There is an evidence crisis in science educational policy. Ed. Psych. Rev. 1–20 (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.