Information-Processing Theory for Classroom Teachers*

Brian Beitzel

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 3.0

Abstract
This module is an introduction to the information-processing model, written for classroom teachers.
Also included are a summary of cognitive load theory and several principles for effective learning.

Information-processing theory is a psychological theory about how we process and learn information.
Clearly, this is a topic that is at the core of the everyday work of a classroom teacher, so let’s spend some
time exploring this theory and how it applies in the classroom.

1 Human Cognitive Architecture
The phrase human cognitive architecture is just a fancy academic way of referring to the areas of the human
brain involved in thinking. Don’t be dazzled by this term—it means little more than what I’ve just told you.

But now we’re going to explore the details of human cognitive architecture and show why this is such an
important topic for classroom teachers to understand.

2 Thinking Wasn’t Always Fashionable
Before we discuss cognitive architecture we should first say that it used to be the case that few scholars
wished to speculate about how the mind thinks. Researchers known as “behaviorists” preferred to talk only
about observable aspects of learning—in other words, what was put into the system (e.g., teachers’ questions)
and what came out of it (e.g., students’ responses). In fact, there was fierce resistance among these folks
to use terms such as “think” because there could be no direct observation of thinking; therefore, any claim
about thinking must necessarily be restricted to conjecture and was thus off-limits. A few of these folks are
still around today, but most of them have been converted to a new way of—dare I say it?—thinking.

3 Information Processing
Long ago and far away, in the late 1960s and throughout the 1970s, researchers became increasingly dissatis-
fied with the behaviorist explanations of learning and began to work on some new models explaining how
people learn. Most famously, Richard Atkinson and Richard Shiffrin (1968) ([1]) proposed a cognitive model
descriving how the mind processes information. This model remains popular even today, so we will take a
close look at it now.

*Version 1.1: Jan 25, 2012 9:04 am +0000
†http://creativecommons.org/licenses/by/3.0/

http://cnx.org/content/m42774/1.1/
Although somewhat oversimplified when compared to more recent work in this field, Atkinson and Shiffrin’s model has become known simply as “the information-processing model.” The basic notion of this model is that it tracks the flow of information as new knowledge moves from the entry point toward permanent storage within the information-processing system. The model proposes three storage compartments (see Figure 1), known as “stores,” which hold information at various points during processing.

![Figure 1: Atkinson and Shiffrin’s (1968) information-processing model. Note that short-term memory is now more commonly known as “working memory.”](http://cnx.org/content/m42774/1.1/)

### 3.1 Sensory Memory

The first store is known as **sensory memory**. This is the entry point for all information coming into the system. Specifically, the kinds of information that sensory memory processes are signals from the five senses: sight, hearing, taste, smell, and touch. Because these senses are always up and running, they are continuously delivering new data to the sensory memory (even during sleep). Take a moment to close your eyes and notice the information from your other four senses that you were unaware of when you began reading this paragraph (e.g., whether your chair is cushioned or hard, whether your neck feels warm or cool, etc.).

#### 3.1.1 Capacity and Duration

Although sensory memory can hold quite a lot of information, it cannot keep any of this information for very long due to the constant inflow of new data. Estimates of duration vary somewhat, but most agree that information cannot be kept active in sensory memory for more than a few (e.g., 3-5) seconds.

We cannot possibly process all of the data that sensory memory intakes. Therefore, we must select those sensory data that are relevant to whatever task we are currently undertaking—and ignore the rest. For the most part, we do this without being very aware of it.

One bothersome aspect of sensory memory is that it collects some sensory data that we wish we could ignore. Have you ever tried to concentrate, perhaps on school assignments, but felt distracted by the goings-on around you? That is a classic example of having sensory data that you felt compelled to process when it didn’t meaningfully benefit you.
Imagine Pierre, a student in a busy classroom where a teacher is giving a group of students directions for an assignment. Pierre is trying to concentrate on the teacher’s instructions, but some other students are creating a distraction with a butterfly display on the other side of the room. The problem here, from a cognitive perspective, is that Pierre cannot effectively process both the actions of his classmates and the teacher’s directions; he must choose whether to pay attention to the distraction or to his teacher. All of this information is contained in sensory memory, but not all of it can be processed in working memory, for reasons we discuss next.

1. Application Activity
How many times have you seen a penny? Would you be able to recognize a penny if you saw one? Go to [http://go.edpsych.net/cents](http://go.edpsych.net/cents) and see if you can identify which one is the real penny. Explain, based on the information-processing model, why you (or someone else) might have difficulty with this task.

3.2 Working Memory
Let’s clear something up before we get ourselves too involved talking about working memory. Atkinson and Shiffrin originally called this store “short-term memory” (to contrast with long-term memory), but modern researchers use the term “working memory” instead. These two terms have some rather subtle distinctions (which cognitive scholars care deeply about), but for our purposes the differences are negligible. Thus, in this discussion, we will prefer the more common term “working memory.”

Working memory is where the real business of thinking takes place. This is where your students will process the content of your carefully crafted lessons as well as your instructions for how to complete their assignments—oh yes, and your warnings regarding proper decorum in the classroom. This is where rocket scientists do their thing, eventually accomplishing moon landings, sending spacecraft to land with precision on other planets millions of miles away, and the like. Now you can see that working memory is a space to be respected (please remove your hat, if you are wearing one).

3.2.1 Capacity
The pity is, in spite of all of its capabilities, working memory is a very small place. Well before Atkinson and Shiffrin developed their information-processing model, George Miller (1956) ([2]) discovered that most individuals have approximately seven cognitive “slots” available to be filled with information at any given time and that this number varies by about two slots across the population, yielding the now-popular estimate of “seven plus or minus two” elements available in working memory to hold all the information one wishes to cram in.

Take a moment now to think how you might remember a telephone number if you had to look it up and then walk over to a phone a short distance away to dial that number. If it is a local number, it has seven digits. As long as your working memory has seven slots available, you should be good to go. But what if someone delivers some surprising news to you halfway through your walk to the phone? It is unlikely that you will remember the phone number because you will be using some of your working-memory capacity to process the news you have just received. The point is, working memory is just too small for us to do everything we would like to be capable of doing.

This limited capacity has profound implications for teaching and learning. Let us now consider how the students in your classroom are affected. If you provide them with complicated instructions for an assignment, there is likely minimal space remaining in working memory for them to comprehend the content. Likewise, if

[http://go.edpsych.net/cents](http://go.edpsych.net/cents)
the pace of your instruction is too fast, with lots of information conveyed at a quick pace, there will be little chance for your students to sufficiently process the information, and maybe even not enough opportunity for them to write it down for later study. A disciplined, controlled pace of presentation is essential if meaningful learning is to occur.

2: Demonstration Activity
You might be surprised how quickly information escapes from your working memory. Go to http://go.edpsych.net/wm\(^2\) and see how well you score!

3.2.2 Maintenance and Elaborative Rehearsal
You might be wondering exactly what happens while information is being processed in working memory. What does it mean for information to be “processed”? We typically refer to processing in working memory with the term \textit{rehearsal}. There are two principal types of rehearsal: maintenance and elaborative. \textit{Maintenance rehearsal} is what you were likely doing as you walked from the telephone book to the nearest phone across the room—i.e., repeating the information over and over to yourself in order to keep it “active.” This is by far the easiest type of rehearsal but it is also the least effective. How well will you remember that phone number two hours from now?

The second type of rehearsal is \textit{elaborative rehearsal}. When one uses elaborative rehearsal, one connects the new information with previously learned information; this integration of old and new information has a dramatic impact on the memorability of the new information. Let’s go back to that phone number. Imagine that you recognize the last four digits to be the same as the house number where you lived for your entire childhood. Now is the phone number easier to remember? Of course it is. The integration of prior knowledge (the house number) with new information (the phone number) improves the memorability of the phone number.

3.2.3 Duration
How long does information hang around in working memory? If it is being rehearsed, information will be active until rehearsal of that information ceases. But if no particular processing strategy (e.g., maintenance or elaborative rehearsal) is being applied to the information, it will vanish from working memory within 5-20 seconds. The reason for that broad estimate is that the information could die a slower death if no new information is imported into working memory to consume whatever limited space is available. Neglected information will not survive long in a busy working memory.

As a teacher, it is tempting to conclude that your students will understand and remember simple information (like a brief fact or quick directions) that you have just told them. However, consider the situation in which a student, Karla, is still trying to understand a previously explained concept. Karla’s working memory is operating at full capacity with attempts to process earlier information and thus cannot successfully deal with the simple instruction or fact that you have now stated. Working memory is altogether too limited to thoroughly process rapidly delivered information.

3.2.4 Attention
The concept of \textit{attention} is one’s focus on a given portion of all possible stimuli. This is also the layman’s understanding of the term “attention,” so you are already familiar with this idea. Whatever you are thinking about (i.e., whatever is currently in working memory) is what you are paying \textit{attention} to. We sometimes use

\(^2\)http://go.edpsych.net/wm
the phrase “selective attention” to indicate that we must select a limited amount of information to process, and ignore the remainder of the incoming information streams.

3: Demonstration Activity
Notice how selective attention is necessary to focus on the target voice and number in this activity: http://go.edpsych.net/selattn

3.2.5 Practice and Automaticity
The longer a piece of information is effectively processed (e.g., through elaborative rehearsal), the more we understand it and the more likely we will be to remember it at a later time. In layman’s terms, this is called practice. In the classroom, first-graders will need to practice their reading skills more than sixth-graders because the sixth-graders have “put in their time” already and have spent a considerable number of hours practicing their reading to the point where it is now automatic. When a skill (such as reading) has been automatized, it requires fewer working-memory resources and thus consumes less space in working memory; this has the benefit of freeing up the remaining space in working memory for other thoughts. For example, how burdensome is it for you to figure out how to pronounce the word “conundrum” compared to the time it would take a first-grader? Because you can easily process this word, you can simultaneously consider the ideas “conundrum” and “Aunt Mary’s wallet is missing from her purse, and we didn’t see anyone enter or exit the room.” A first-grader would be capable of comparing these ideas but would require much more time to arrive at a complete understanding of the intersection between these two ideas than you would need, because you have already automatized much of the requisite processing.

In summary, then, practice speeds up processing because it automatizes critical skills.

3.3 Long-Term Memory
Long-term memory is just what it sounds like: an area that stores information permanently. To arrive in long-term memory, information must have been sufficiently processed in working memory. Stated another way, working memory is the exclusive route to long-term memory. How does information become “sufficiently processed” in working memory? By considering both the amount of time spent and the quality of processing encountered there. We will discuss different qualities of information processing in Section 5 (Principles of Effective Learning). For now, keep in mind that the amount of time one spends thinking about a topic (e.g., preparing for an exam) does not necessarily predict one’s memory for that material at a future point in time.

3.3.1 Duration and Capacity
As far as we know, information is maintained in long-term memory indefinitely; there are no known expiration dates here. Additionally, there is no known limit to the amount of knowledge that can be stored in long-term memory. No one can credibly make the excuse that they don’t have room to store any more information!

Now perhaps you can begin to see why it is important for teachers to understand human cognitive architecture. Without fully appreciating the capabilities and limitations of the information-processing system, teachers could easily have unrealistic expectations for their students—and that is not good for anybody.

4 Cognitive Load Theory
The information-processing model has given rise to a theory of instructional design called cognitive load theory (Sweller & Chandler, 1994 ([4]); van Merrienboer & Sweller, 2005 ([3])). Because working memory is the principal player in the process of learning new information, cognitive load theory focuses exclusively on working memory. The gist of this theory is that there are distinct types of demands imposed upon working memory during learning: intrinsic, extraneous, and germane. We now examine each of these.
4.1 Intrinsic Cognitive Load

Intrinsic cognitive load represents the burden imposed on working memory by the inherent nature of the material. In other words, simple topics require very little processing capacity in working memory, and complex topics demand a large amount of space. For example, it requires considerably more focus to safely drive a semi truck through a rainstorm than to sign your name with a pen on paper. Driving the semi requires attention to many different information inputs (e.g., gauges, mirrors, windshield) and coordinating the requisite motor skills in response; all of this processing is conducted in working memory. Signing one’s name takes barely any attention at all (for adults) because it has been done thousands of times before. Thus, the effect of having practiced the skill reduces its intrinsic cognitive load.

But practice alone cannot reduce the intrinsic cognitive load of all tasks. The element interactivity (i.e., coordination among multiple aspects) inherent in some tasks cannot ultimately reduce the task to a trivial activity, even with extensive practice. If that were the case, we should all be capable of becoming skilled airline pilots or successful politicians.

For beginners learning an essential skill, element interactivity becomes problematic and must be temporarily reduced. When learning a language, one first learns the alphabet and then proceeds to acquire simple words or phrases—not complex prose. But one cannot be considered proficient in a language unless one can understand its complex prose. This is an example of element interactivity because understanding prose depends upon not only understanding its nouns, verbs, adverbs, etc., but also how each of them modifies or alters the meaning of other words nearby. Topics or skills that contain element interactivity must at first be oversimplified and then gradually built up to their full complexity before one can successfully deal with the intrinsic cognitive load.

4.2 Extraneous Cognitive Load

Extraneous cognitive load is the set of mental demands that are irrelevant to the current task. Recall the butterfly display that was distracting Pierre (Section 3.1.1 (Capacity and Duration))? The butterflies had nothing to do with the directions the teacher was giving Pierre’s group, yet Pierre was distracted by his classmates on the other side of the room. His classmates’ activity was extraneous cognitive load for Pierre because it was consuming his precious cognitive resources yet not providing any real benefit to him in the task of understanding his teacher’s directions.

It is critical to realize that these various forms of cognitive load are additive—that is, they each increase the amount of processing space that is active in working memory. For example, if the intrinsic load is already high, there is not much room for any extraneous load unless the learner decides (like Pierre) to reduce the processing of the intrinsic load and focus more on the extraneous load. Teachers should strive to reduce extraneous cognitive load in their classrooms because students are likely to sacrifice attention to important material and distract themselves with the extraneous stimuli.

Extraneous cognitive load is, for the most part, under the direct control of the teacher. Have you ever seen presentations that were decorated with graphics which were only tangentially related to the content? You probably found yourself sidetracked by the images and not paying sufficient attention to the material itself. Because working memory has such a limited capacity, we cannot afford to “clutter up” this valuable space with unproductive ideas that divert attention from more important content. As a teacher, you should make earnest efforts to avoid exposing students to extra “fluff” during learning activities.

4.3 germane Cognitive Load

Germane cognitive load has been explained in various ways. The explanation I prefer is the more traditional characterization that germane load represents increased demand upon working memory in the service of the learning goal. This can be explained more easily through an example. Most (if not all) languages have forms of expression that are not appropriate for all audiences. For example, in English one would not address the President of the United States in the same informal way as one would address a close friend (“How is your day going, Mr. President?” versus “Hey dude, whazzup?”). The meaning of the utterance expressed to these
two individuals may be the same, but the words and intonation are somewhat different. If an international student were learning English, it would be important for the language teacher to communicate not only the meaning of the words (intrinsic load) but also the contexts in which those words are appropriate (germane load). Learning the situations in which certain phrases are most appropriately used goes beyond intrinsic load but could hardly be considered extraneous if one’s purpose is to learn the language well.

It goes without saying that beginning learners should not be exposed to germane cognitive load; the intrinsic load for many tasks is of sufficient complexity that beginners cannot handle any additional processing burdens. However, as learning proceeds and the intrinsic load becomes more and more automatized, teachers can add aspects of additional complexity that enhance students’ understanding of the material in a germane way.

5 Principles of Effective Learning

We now turn to a few empirical principles, derived from decades of research, that are known to improve learning. These principles will not all apply to every learning situation; however, each of them has been sufficiently demonstrated through carefully controlled scientific studies to merit mentioning them here.

The overarching goal here is to select processing strategies that will increase the likelihood of a learner recalling new information at a later point in time.

5.1 Activate Prior Knowledge

One of the most important cognitive principles for a teacher to keep in mind is the importance of relating information from long-term memory to information newly entering the system. Recall our discussion of elaborative rehearsal earlier, in which I indicated that making a connection to prior knowledge is a superior learning method to simply repeating information over and over without altering it.

Any good lesson-plan format begins the class with some form of prior-knowledge activation. It might be a reminder or a brief review of what was studied in the previous day’s lesson, or it could be a question similar to, “Have you ever had a problem you couldn’t solve?” The purpose of this phase of the lesson is to activate prior knowledge—i.e., bring long-term memories back into working memory—so that new knowledge can be mingled with old with the result of more solid understanding of the new (and perhaps even the old) information.

4: Demonstration Activity

Remember that having prior knowledge is not good enough; that knowledge needs to be activated in order to make use of it. Go to http://go.edpsych.net/implicit and notice that you will identify previously seen words more quickly than “new” words (which you have prior knowledge of, but have not recently been activated).

5.2 Organization

This is one principle that applies to a rather restricted set of instructional situations, but it is so powerful that it deserves mention here. In contexts where there is a list of items to commit to memory, the task of memorizing the list will be much easier if the items are grouped together (i.e., organized) in a meaningful way. This also works as a basic memory strategy in everyday life—think about your latest visit to the grocery store and imagine remembering a rather random assortment of items versus grouping the dairy items together, the produce items together, etc.

4http://go.edpsych.net/implicit
5.3 Deep Processing

It is easy to become convinced that if a student spends, say, twenty hours reviewing for an exam, that student should be expected to excel on the exam. However, cognitive studies show that it is not specifically the time one spends studying that matters most; what one does during that time matters even more.

Consider, for example, the all-too-common exam-preparation strategy of using flash cards. Students often take terms from the textbook or class discussions, write them down on flash cards, and then rehearse what is written down until the flash cards are memorized. Such a student will walk into the exam confident that the material has been thoroughly mastered. The problem with this approach to studying is that the student has only done “surface-level processing” of the material, rather than “deep” processing. It is surface-level because the student has memorized terms and definitions rather than truly understanding the meaning and applications of those concepts.

Deep processing happens when one uses elaborative rehearsal to connect a concept to other concepts that are already known or are being learned. For example, one could write a summary of a concept in one’s own words to check for comprehension. Another approach to facilitate deep processing is to think of examples of the newly learned concept from one’s own life. One could even make up fictitious examples of the concept if no examples come to mind from one’s past experience.

The point is, learning that comes from surface-level processing is not durable. One does not remember the content of flash cards for very long after the exam. But spending the same amount of time (or even less time) meaningfully engaged with the to-be-learned ideas can result in learning that could last for a lifetime.

5.4 Distributed Practice

There is one final principle for effective learning that must be mentioned here. To be the most effective learner, one should “space” or “distribute” one’s studying over a period of time. Attempting to cram a lot of learning into one or two concentrated study sessions rarely works. Research cannot prescribe the specific number or length of study sessions required to maximize learning—there are too many variables to account for (e.g., one’s prior knowledge of the topic, one’s knowledge of related topics, the quality of one’s study strategies, etc.). But the benefits of distributing one’s study sessions over a period of time are well documented in the research literature.

References


http://cnx.org/content/m42774/1.1/