Working Memory, Cognitive Load & Instructional Design
Contents

Articles

Working Memory

Working memory 1
Short-term memory 14
Schema (psychology) 18
Baddeley's model of working memory 22

Cognitive Load

Cognitive load 26
Chunking (psychology) 30
The Magical Number Seven, Plus or Minus Two 33
Problem solving 36

Instructional Design

Instructional design 44
Multimedia learning 50
Split attention effect 52
Worked-example effect 54
Expertise reversal effect 56

References

Article Sources and Contributors 57
Image Sources, Licenses and Contributors 58

Article Licenses

License 59
Working Memory

Working memory

Working memory is the ability to actively hold information in the mind needed to do complex tasks such as reasoning, comprehension and learning. Working memory tasks are those that require the goal-oriented active monitoring or manipulation of information or behaviors in the face of interfering processes and distractions. The cognitive processes involved include the executive and attention control of short-term memory which provide for the interim integration, processing, disposal, and retrieval of information. Working memory is a theoretical concept central both to cognitive psychology and neuroscience.

Theories exist both regarding the theoretical structure of working memory and the role of specific parts of the brain involved in working memory. Research identifies the frontal cortex, parietal cortex, anterior cingulate, and parts of the basal ganglia as crucial. The neural basis of working memory has been derived from lesion experiments in animals and functional imaging upon humans. A study at the University of Stirling found that people with good working memories tend to be happy and more successful in their lives.[1]

History

The term "working memory" was coined by Miller, Galanter, and Pribram,[2] [3] and was used in the 1960s in the context of theories that likened the mind to a computer. Atkinson and Shiffrin (1968)[4] also used this term, "working memory" (p. 92) to describe their "short-term store." What we now call working memory was referred to as a "short-term store" or short-term memory, primary memory, immediate memory, operant memory, or provisional memory.[5] Short-term memory is the ability to remember information over a brief period of time (in the order of seconds). Most theorists today use the concept of working memory to replace or include the older concept of short-term memory, thereby marking a stronger emphasis on the notion of manipulation of information instead of passive maintenance.

The earliest mention of experiments on the neural basis of working memory can be traced back to over 100 years ago, when Hitzig and Ferrier described ablation experiments of the prefrontal cortex (PFC), they concluded that the frontal cortex was important for cognitive rather than sensory processes.[6] In 1935 and 1936, Carlyle Jacobsen and colleagues were the first to show the deleterious effect of prefrontal ablation on delayed response.[6] [7]

Theories

There have been numerous models proposed regarding how working memory functions, both anatomically and cognitively. Of those, three that are well known are summarized below.

Baddeley and Hitch

Baddeley and Hitch (1974)[8] introduced and made popular the multicomponent model of working memory. This theory proposes that two "slave systems" are responsible for short-term maintenance of information, and a "central executive" is responsible for the supervision of information integration and for coordinating the slave systems. One slave system, the phonological loop, stores phonological information (i.e., the sound of language) and prevents its decay by continuously articulating its contents, thereby refreshing the information in a rehearsal loop. It can, for example, maintain a seven-digit telephone number for as long as one repeats the number to oneself again and again. The other slave system, the visuo-spatial sketch pad, stores visual and spatial information. It can be used, for
example, for constructing and manipulating visual images, and for the representation of mental maps. The sketch pad can be further broken down into a visual subsystem (dealing with, for instance, shape, colour, and texture), and a spatial subsystem (dealing with location). The central executive (see executive system) is, among other things, responsible for directing attention to relevant information, suppressing irrelevant information and inappropriate actions, and for coordinating cognitive processes when more than one task must be done at the same time.

Baddeley (2000) extended the model by adding a fourth component, the episodic buffer, which holds representations that integrate phonological, visual, and spatial information, and possibly information not covered by the slave systems (e.g., semantic information, musical information). The component is episodic because it is assumed to bind information into a unitary episodic representation. The episodic buffer resembles Tulving's concept of episodic memory, but it differs in that the episodic buffer is a temporary store.

**Cowan**

Cowan\(^9\) \(^10\) regards working memory not as a separate system, but as a part of long-term memory. Representations in working memory are a subset of the representations in long-term memory. Working memory is organized into two embedded levels. The first level consists of long-term memory representations that are activated. There can be many of these, there is no limit to activation of representations in long-term memory. The second level is called the focus of attention. The focus is regarded as capacity limited and holds up to four of the activated representations.

Oberauer\(^11\) has extended the Cowan model by adding a third component, a more narrow focus of attention that holds only one chunk at a time. The one-element focus is embedded in the four-element focus and serves to select a single chunk for processing. For example, you can hold four digits in mind at the same time in Cowan's "focus of attention". Now imagine that you wish to perform some process on each of these digits, for example, adding the number two to each digit. Separate processing is required for each digit, as most individuals can not perform several mathematical processes in parallel. Oberauer's attentional component selects one of the digits for processing, and then shifts the attentional focus to the next digit, continuing until all of the digits have been processed.

**Ericsson and Kintsch**

Ericsson and Kintsch (1995) have argued that we use skilled memory in most everyday tasks. Tasks such as reading, for instance, require to maintain in memory much more than seven chunks - with a capacity of only seven chunks our working memory would be full after a few sentences, and we would never be able to understand the complex relations between thoughts expressed in a novel or a scientific text. We accomplish this by storing most of what we read in long-term memory, linking them together through retrieval structures. We need to hold only a few concepts in working memory, which serve as cues to retrieve everything associated to them by the retrieval structures. Anders Ericsson and Walter Kintsch refer to this set of processes as "long-term working memory". Retrieval structures vary according to the domain of expertise, yet as suggested by Gobet\(^12\) they can be categorized in three typologies: generic retrieval structures, domain knowledge retrieval structures and the episodic text structures. The first corresponds to Ericsson and Kintsch's 'classic' retrieval structure and the second to the elaborated memory structure. The first kind of structure is developed deliberately and is arbitrary (e.g. the method of loci), the second one is similar to patterns and schemas and the last one takes place exclusively during text comprehension. Concerning this last typology, Kintsch, Patel and Ericsson\(^13\) consider that every reader is able to form an episodic text structure during text comprehension, if the text is well written and if the content is familiar.
Capacity

Working memory is generally considered to have limited capacity. The earliest quantification of the capacity limit associated with short-term memory was the "magical number seven" introduced by Miller (1956).[14] He noticed that the memory span of young adults was around seven elements, called chunks, regardless whether the elements were digits, letters, words, or other units. Later research revealed that span does depend on the category of chunks used (e.g., span is around seven for digits, around six for letters, and around five for words), and even on features of the chunks within a category. For instance, span is lower for long words than for short words. In general, memory span for verbal contents (digits, letters, words, etc.) strongly depends on the time it takes to speak the contents aloud, and on the lexical status of the contents (i.e., whether the contents are words known to the person or not).[15] Several other factors also affect a person's measured span, and therefore it is difficult to pin down the capacity of short-term or working memory to a number of chunks. Nonetheless, Cowan (2001)[16] has proposed that working memory has a capacity of about four chunks in young adults (and fewer in children and old adults).

Whereas most adults can repeat about seven digits in correct order, some individuals have shown impressive enlargements of their digit span – up to 80 digits. This feat is possible by extensive training on an encoding strategy by which the digits in a list are grouped (usually in groups of three to five) and these groups are encoded as a single unit (a chunk). To do so one must be able to recognize the groups as some known string of digits. One person studied by K. Anders Ericsson and his colleagues, for example, used his extensive knowledge of racing times from the history of sports. Several such chunks can then be combined into a higher-order chunk, thereby forming a hierarchy of chunks. In this way, only a small number of chunks at the highest level of the hierarchy must be retained in working memory. At retrieval, the chunks are unpacked again. That is, the chunks in working memory act as retrieval cues that point to the digits that they contain. It is important to note that practicing memory skills such as these does not expand working memory capacity proper. This can be shown by using different materials - the person who could recall 80 digits was not exceptional when it came to recalling words.

Measures and correlates

Working memory capacity can be tested by a variety of tasks. A commonly used measure is a dual-task paradigm combining a memory span measure with a concurrent processing task, sometimes referred to as "complex span". Daneman and Carpenter invented the first version of this kind of task, the "reading span", in 1980.[17] Subjects read a number of sentences (usually between 2 and 6) and try to remember the last word of each sentence. At the end of the list of sentences, they repeat back the words in their correct order. Other tasks that don't have this dual-task nature have also been shown to be good measures of working memory capacity.[18] The question of what features a task must have to qualify as a good measure of working memory capacity is a topic of ongoing research.

Measures of working-memory capacity are strongly related to performance in other complex cognitive tasks such as reading comprehension, problem solving, and with any measures of the intelligence quotient.[19] Some researchers have argued[20] that working memory capacity reflects the efficiency of executive functions, most notably the ability to maintain a few task-relevant representations in the face of distracting irrelevant information. The tasks seem to reflect individual differences in ability to focus and maintain attention, particularly when other events are serving to capture attention. These effects seem to be a function of frontal brain areas.[21]

Others have argued that the capacity of working memory is better characterized as the ability to mentally form relations between elements, or to grasp relations in given information. This idea has been advanced, among others, by Graeme Halford, who illustrated it by our limited ability to understand statistical interactions between variables.[22] These authors asked people to compare written statements about the relations between several variables to graphs illustrating the same or a different relation, as in the following sentence: "If the cake is from France, then it has more sugar if it is made with chocolate than if it is made with cream, but if the cake is from Italy, then it has more sugar if it is made with cream than if it is made of chocolate." This statement describes a relation between three variables (country, ingredient, and amount of sugar), which is the maximum most individuals can understand. The
capacity limit apparent here is obviously not a memory limit (all relevant information can be seen continuously) but a limit on how many relationships are discerned simultaneously.

**Experimental studies of working memory capacity**

**Different approaches**

There are several hypotheses about the nature of the capacity limit. One is that there is a limited pool of cognitive resources needed to keep representations active and thereby available for processing, and for carrying out processes. Another hypothesis is that memory traces in working memory decay within a few seconds, unless refreshed through rehearsal, and because the speed of rehearsal is limited, we can maintain only a limited amount of information. Yet another idea is that representations held in working memory capacity interfere with each other.

There are several forms of interference discussed by theorists. One of the oldest ideas is that new items simply replace older ones in working memory. Another form of interference is retrieval competition. For example, when the task is to remember a list of 7 words in their order, we need to start recall with the first word. While trying to retrieve the first word, the second word, which is represented in close proximity, is accidentally retrieved as well, and the two compete for being recalled. Errors in serial recall tasks are often confusions of neighboring items on a memory list (so-called transpositions), showing that retrieval competition plays a role in limiting our ability to recall lists in order, and probably also in other working memory tasks. A third form of interference assumed by some authors is feature overwriting. The idea is that each word, digit, or other item in working memory is represented as a bundle of features, and when two items share some features, one of them steals the features from the other. The more items are held in working memory, and the more their features overlap, the more each of them will be degraded by the loss of some features.

**Time-based resource sharing model**

The theory most successful so far in explaining experimental data on the interaction of maintenance and processing in working memory is the "time-based resource sharing model". This theory assumes that representations in working memory decay unless they are refreshed. Refreshing them requires an attentional mechanism that is also needed for any concurrent processing task. When there are small time intervals in which the processing task does not require attention, this time can be used to refresh memory traces. The theory therefore predicts that the amount of forgetting depends on the temporal density of attentional demands of the processing task - this density is called "cognitive load". The cognitive load depends on two variables, the rate at which the processing task requires individual steps to be carried out, and the duration of each step. For example, if the processing task consists of adding digits, then having to add another digit every half second places a higher cognitive load on the system than having to add another digit every two seconds. Adding larger digits takes more time than adding smaller digits, and therefore cognitive load is higher when larger digits must be added. In a series of experiments, Barrouillet and colleagues have shown that memory for lists of letters depends on cognitive load, but not on the number of processing steps (a finding that is difficult to explain by an interference hypothesis) and not on the total time of processing (a finding difficult to explain by a simple decay hypothesis). One difficulty for the time-based resource-sharing model, however, is that the similarity between memory materials and materials processed also affects memory accuracy.
Limitations

None of these hypotheses can explain the experimental data entirely. The resource hypothesis, for example, was meant to explain the trade-off between maintenance and processing: The more information must be maintained in working memory, the slower and more error prone concurrent processes become, and with a higher demand on concurrent processing memory suffers. This trade-off has been investigated by tasks like the reading-span task described above. It has been found that the amount of trade-off depends on the similarity of the information to be remembered and the information to be processed. For example, remembering numbers while processing spatial information, or remembering spatial information while processing numbers, impair each other much less than when material of the same kind must be remembered and processed.\(^{28}\) Also, remembering words and processing digits, or remembering digits and processing words, is easier than remembering and processing materials of the same category.\(^{29}\) These findings are also difficult to explain for the decay hypothesis, because decay of memory representations should depend only on how long the processing task delays rehearsal or recall, not on the content of the processing task. A further problem for the decay hypothesis comes from experiments in which the recall of a list of letters was delayed, either by instructing participants to recall at a slower pace, or by instructing them to say an irrelevant word once or three times in between recall of each letter. Delaying recall had virtually no effect on recall accuracy.\(^{30}\) \(^{31}\) The Interference theory seems to fare best with explaining why the similarity between memory contents and the contents of concurrent processing tasks affects how much they impair each other. More similar materials are more likely to be confused, leading to retrieval competition, and they have more overlapping features, leading to more feature overwriting. One experiment directly manipulated the amount of overlap of phonological features between words to be remembered and other words to be processed.\(^{32}\) Those to-be-remembered words that had a high degree of overlap with the processed words were recalled worse, lending some support to the idea of interference through feature overwriting.

Development

The capacity of working memory increases gradually over childhood\(^{33}\) and declines gradually in old age.\(^{34}\)

Childhood

Measures of performance on tests of working memory increase continuously between early childhood and adolescence, while the structure of correlations between different tests remains largely constant.\(^{33}\) Thus, the development of working memory can be described as quantitative growth rather than a qualitative change. Starting with work in the Neo-Piagetian tradition,\(^{35}\) \(^{36}\) theorists have argued that the growth of working-memory capacity is a major driving force of cognitive development. This hypothesis has received substantial empirical support from studies showing that the capacity of working memory is a strong predictor of cognitive abilities in childhood.\(^{37}\) Particularly strong evidence for a role of working memory for development comes from a longitudinal study showing that working-memory capacity at one age predicts reasoning ability at a later age.\(^{38}\) Studies in the Neo-Piagetian tradition have added to this picture by analyzing the complexity of cognitive tasks in terms of the number of items or relations that have to be considered simultaneously for a solution. Across a broad range of tasks, children manage task versions of the same level of complexity at about the same age, consistent with the view that working memory capacity limits the complexity they can handle at a given age.\(^{39}\)

Aging

Working memory is among the cognitive functions most sensitive to decline in old age.\(^{40}\) \(^{41}\) Several explanations have been offered for this decline in psychology. One is the processing speed theory of cognitive aging by Tim Salthouse.\(^{42}\) Based on the finding of general slowing of cognitive processes as we grow older, Salthouse argues that slower processing leaves more time for working-memory contents to decay, thus reducing effective capacity. However, the decline of working-memory capacity cannot be entirely attributed to slowing because capacity declines
more in old age than speed. Another proposal is the inhibition hypothesis advanced by Lynn Hasher and Rose Zacks. This theory assumes a general deficit in old age in the ability to inhibit irrelevant, or no-longer relevant, information. Therefore, working memory tends to be cluttered with irrelevant contents that reduce the effective capacity for relevant content. The assumption of an inhibition deficit in old age has received much empirical support but so far it is not clear whether the decline in inhibitory ability fully explains the decline of working-memory capacity. An explanation on the neural level of the decline of working memory and other cognitive functions in old age has been proposed by West. He argued that working memory depends to a large degree on the PFC, which deteriorates more than other brain regions as we grow old.

**Training**

One theory of attention-deficit hyperactivity disorder states that ADHD can lead to deficits in working memory. Studies suggest that working memory can be improved by training in ADHD patients through computerized programs. This random controlled study has found that a period of working memory training increases a range of cognitive abilities and increases IQ test scores. Consequently, this study supports previous findings suggesting that working memory underlies general intelligence. Another study of the same group has shown that, after training, measured brain activity related to working memory increased in the prefrontal cortex, an area that many researchers have associated with working memory functions. It has been shown that working memory training leads to measurable density changes for cortical dopamine neuroreceptors in test persons.

A controversial study has shown that training with a working memory task (the dual n-back task) improves performance on a very specific fluid intelligence test in healthy young adults. The study's conclusion that improving or augmenting the brain's working memory ability increases fluid intelligence is backed by some and questioned by others. The study was replicated in 2010.

In Torkel Klingberg’s 2009 book *The Overflowing Brain*, he proposes that working memory is enhanced through exposure to excess neural activation. The brain map of an individual, he argues, can be altered by this activation to create a larger area of the brain activated by a particular type of sensory experience. An example would be that in learning to play guitar, the area activated by sensory impressions of the instrument is larger in the brain of a player than it is in a nonplayer.

There is evidence that optimal working memory performance links to the neural ability to focus attention on task-relevant information and ignore distractions, and that practice-related improvement in working memory is due to increasing these abilities.

Working memory performance may also be increased by high intensity exercise. A study was conducted with both sedentary and active females 18–25 years old in which the effects of short-term exercise to exhaustion on working memory was measured. While the working memory of the subjects decreased during and immediately after the exercise bouts, it was shown that the subjects' working memory had an increase following recovery.

However a recent review paper has called into question much of the "success" of working memory training studies. Shipstead et al. (2010) point out that working memory training studies are plagued with poor experimental design. The majority of training studies utilize a no-contact control group making it impossible to determine whether any benefit of training is due to actual improvement or a Hawthorne effect.
Working memory in the brain

Genetics
Little is known of the genetics of working memory. It is heritable, and, at a component level, one candidate gene has been proposed, namely ROBO1 for the phonological loop function of working memory.

Physiology and Psychopharmacology
The first insights into the neuronal and neurotransmitter basis of working memory came from animal research. The work of Jacobsen[60] and Fulton in the 1930s first showed that lesions to the PFC impaired spatial working memory performance in monkeys. The later work of Fuster[61] recorded the electrical activity of neurons in the PFC of monkeys while they were doing a delayed matching task. In that task, the monkey sees how the experimenter places a bit of food under one of two identical looking cups. A shutter is then lowered for a variable delay period, screening off the cups from the monkey's view. After the delay, the shutter opens and the monkey is allowed to retrieve the food from under the cups. Successful retrieval in the first attempt — something the animal can achieve after some training on the task — requires holding the location of the food in memory over the delay period. Fuster found neurons in the PFC that fired mostly during the delay period, suggesting that they were involved in representing the food location while it was invisible. Later research has shown similar delay-active neurons also in the posterior parietal cortex, the thalamus, the caudate, and the globus pallidus.[62] The work of Goldman-Rakic and others showed that principal sulcal, dorsolateral PFC interconnects with all of these brain regions, and that neuronal microcircuits within PFC are able to maintain information in working memory through recurrent excitatory glutamate networks of pyramidal cells that continue to fire throughout the delay period.[63] These circuits are tuned by lateral inhibition from GABAergic interneurons.[64] The neuromodulatory arousal systems markedly alter PFC working memory function; for example, either too little or too much dopamine or norepinephrine impairs PFC network firing[65] and working memory performance.[66]

Localization
Localization of brain functions in humans has become much easier with the advent of brain imaging methods (PET and fMRI). This research has confirmed that areas in the PFC are involved in working memory functions. During the 1990s much debate has centered on the different functions of the ventrolateral (i.e., lower areas) and the dorsolateral (higher) areas of the PFC. One view was that the dorsolateral areas are responsible for spatial working memory and the ventrolateral areas for non-spatial working memory. Another view proposed a functional distinction, arguing that ventrolateral areas are mostly involved in pure maintenance of information, whereas dorsolateral areas are more involved in tasks requiring some processing of the memorized material. The debate is not entirely resolved but most of the evidence supports the functional distinction.[67]

Brain imaging has also revealed that working memory functions are by far not limited to the PFC. A review of numerous studies[68] shows areas of activation during working memory tasks scattered over a large part of the cortex. There is a tendency for spatial tasks to recruit more right-hemisphere areas, and for verbal and object working memory to recruit more left-hemisphere areas. The activation during verbal working memory tasks can be broken down into one component reflecting maintenance, in the left posterior parietal cortex, and a component reflecting subvocal rehearsal, in the left frontal cortex (Broca's area, known to be involved in speech production).[69] There is an emerging consensus that most working memory tasks recruit a network of PFC and parietal areas. A study has shown that during a working memory task the connectivity between these areas increases.[70] Another study has demonstrated that these areas are necessary for working memory, and not simply activated accidentally during working memory tasks, by temporarily blocking them through transcranial magnetic stimulation (TMS), thereby producing an impairment in task performance.[71]
A current debate concerns the function of these brain areas. The PFC has been found to be active in a variety of tasks that require executive functions.[21] This has led some researchers to argue that the role of PFC in working memory is in controlling attention, selecting strategies, and manipulating information in working memory, but not in maintenance of information. The maintenance function is attributed to more posterior areas of the brain, including the parietal cortex.[72] [73] Other authors interpret the activity in parietal cortex as reflecting executive functions, because the same area is also activated in other tasks requiring executive attention but no memory.[74]

Working memory has been suggested to involve two processes with different neuroanatomical locations in the frontal and parietal lobes.[75] First, a selection operation that retrieves the most relevant item, and second an updating operation that changes the focus of attention made upon it. Updating the attentional focus has been found to involve the transient activation in the caudal superior frontal sulcus and posterior parietal cortex. While increasing demands on selection selectively changes activation in the rostral superior frontal sulcus and posterior cingulate/precuneus.[75]

Articulating the differential function of brain regions involved in working memory is dependent on task able to distinguish these functions.[76] Most brain imaging studies of working memory have used recognition tasks such as delayed recognition of one or several stimuli, or the n-back task, in which each new stimulus in a long series must be compared to the one presented n steps back in the series. The advantage of recognition tasks is that they require minimal movement (just pressing one of two keys), making fixation of the head in the scanner easier. Experimental research and research on individual differences in working memory, however, has used largely recall tasks (e.g., the reading span task, see below). It is not clear to what degree recognition and recall tasks reflect the same processes and the same capacity limitations.

A few brain imaging studies have been conducted with the reading span task or related tasks. Increased activation during these tasks was found in the PFC and, in several studies, also in the anterior cingulate cortex (ACC). People performing better on the task showed larger increase of activation in these areas, and their activation was correlated more over time, suggesting that their neural activity in these two areas was better coordinated, possibly due to stronger connectivity.[77] [78]

**Effects of stress**

Working memory is impaired by acute and chronic psychological stress. This phenomenon was first discovered in animal studies by Arnsten and colleagues,[79] who have shown that stress-induced catecholamine release in PFC rapidly decreases PFC neuronal firing and impairs working memory performance through feedforward, intracellular signaling pathways.[80] Exposure to chronic stress leads to more profound working memory deficits and additional architectural changes in PFC, including dendritic atrophy and spine loss,[81] which can be prevented by inhibition of protein kinase C signaling.[82] fMRI research has extended this research to humans, and confirms that reduced working memory caused by acute stress links to reduced activation of the PFC, and stress increased levels of catecholamines.[83] Imaging studies of medical students undergoing stressful exams have also shown weakened PFC functional connectivity, consistent with the animal studies.[84] The marked effects of stress on PFC structure and function may help to explain how stress can cause or exacerbate mental illness.
Working memory

Neural maintenance

Much has been learned over the last two decades on where in the brain working memory functions are carried out. Much less is known on how the brain accomplishes short-term maintenance and goal-directed manipulation of information. The persistent firing of certain neurons in the delay period of working memory tasks shows that the brain has a mechanism of keeping representations active without external input.

Keeping representations active, however, is not enough if the task demands maintaining more than one chunk of information. In addition, the components and features of each chunk must be bound together to prevent them from being mixed up. For example, if a red triangle and a green square must be remembered at the same time, one must make sure that "red" is bound to "triangle" and "green" is bound to "square". One way of establishing such bindings is by having the neurons that represent features of the same chunk fire in synchrony, and those that represent features belonging to different chunks fire out of sync. In the example, neurons representing redness would fire in synchrony with neurons representing the triangular shape, but out of sync with those representing the square shape. So far, there is no direct evidence that working memory uses this binding mechanism, and other mechanisms have been proposed as well. It has been speculated that synchronous firing of neurons involved in working memory oscillate with frequencies in the theta band (4 to 8 Hz). Indeed, the power of theta frequency in the EEG increases with working memory load, and oscillations in the theta band measured over different parts of the skull become more coordinated when the person tries to remember the binding between two components of information.

One modern approach to explain the working memory in the brain is Prefrontal Cortex Basal Ganglia Working Memory (PBWM).

Learning

There is now extensive evidence that working memory is linked to key learning outcomes in literacy and numeracy. A longitudinal study confirmed that a child's working memory at 5 years old is a better predictor of academic success than IQ. Without appropriate intervention, these children lag behind their peers. A recent study of 37 school-age children with significant learning disabilities has shown that working memory capacity at baseline measurement, but not IQ, predicts learning outcomes two years later. This suggests that working memory impairments are associated with low learning outcomes and constitute a high risk factor for educational under achievement for children. In children with learning disabilities such as dyslexia, ADHD, and developmental coordination disorder, a similar pattern is evident.

In a classroom, common characteristics of working memory impairment include a failure to remember instructions and an inability to complete learning activities. Without early diagnosis, working memory impairment negatively impacts a child's performance throughout their scholastic career.

However, strategies that target the specific strengths and weaknesses of the student's working memory profile are available for educators.

Attention

Research suggests a close link between the working memory capacities of a person and their ability to control the information from the environment that they can selectively enhance or ignore. Such attention allows for example for the voluntarily shifting in regard to goals of a person's information processing to spatial locations or objects rather than ones that capture their attention due to their sensory saliency (such as an ambulance siren). The goal directing of attention is driven by "top-down" signals from the PFC that bias processing in posterior cortical areas and saliency capture by "bottom-up" control from subcortical structures and the primary sensory cortices.
ability to override sensory capture of attention differs greatly between individuals and this difference closely links to their working memory capacity. The greater a person’s working memory capacity, the greater their ability to resist sensory capture. The limited ability to override attentional capture is likely to result in the unnecessary storage of information in working memory, suggesting not only that having a poor working memory affects attention but that it can also limit the capacity of working memory even further.

Research

Today there are hundreds of research laboratories around the world studying various aspects of working memory. There are numerous applications of working memory in the field, such as using working memory capacity to explain intelligence, success at emotion regulation, and improving teaching methods, furthering the understanding of autism spectrum disorders, ADHD, motor dyspraxia, and creating artificial intelligence based on the human brain.

References

Working memory


Working memory


[58] PMID 19587320.


[76] M. Coltheart. (2006). What has functional neuroimaging told us about the mind (so far)?

Working memory


Working memory


**External links**

A working memory web forum (http://www.workingmemory.lancs.ac.uk) exists for researchers to post reference information for relevant journal articles and to discuss stimuli, test resources, etc.

**Short-term memory**

**Short-term memory** (or "primary" or "active memory") is the capacity for holding a small amount of information in mind in an active, readily available state for a short period of time. The duration of short-term memory (when rehearsal or active maintenance is prevented) is believed to be in the order of seconds. A commonly-cited capacity is 7 ± 2 elements. In contrast, long-term memory indefinitely stores a seemingly unlimited amount of information.

Short-term memory should be distinguished from working memory which refers to structures and processes used for temporarily storing and manipulating information (see more details below).

**Existence of a separate store**

The idea of a division of memory into short term and long term dates back to the 19th century. A classical model of memory developed in the 1960s, now known to be flawed,\cite{1} assumed that all memories pass from a short-term to a long-term store after a small period of time. This model is referred to as the "modal model" and has been most famously detailed by Shiffrin.\cite{2} The exact mechanisms by which this transfer takes place, whether all or only some memories are retained permanently, and indeed the existence of a genuine distinction between the two stores, remain controversial topics among experts.

One form of evidence, cited in favor of the separate existence of a short-term store comes from anterograde amnesia, the inability to learn new facts and episodes. Patients with this form of amnesia, have intact ability to retain small amounts of information over short time scales (up to 30 seconds) but are dramatically impaired in their ability to form longer-term memories (a famous example is patient HM). This is interpreted as showing that the short-term store is spared from amnesia and other brain diseases.

Other evidence comes from experimental studies showing that some manipulations (e.g., a distractor task, such as repeatedly subtracting a single-digit number from a larger number following learning) impair memory for the 3 to 5 most recently learned words of a list (presumably still held in short-term memory), while leaving recall for words from earlier in the list (presumably stored in long-term memory) unaffected; other manipulations (e.g., semantic similarity of the words) affect only memory for earlier list words,\cite{3} but do not affect memory for the last few words in a list. These results show that different factors affect short term recall (disruption of rehearsal) and long-term recall (semantic similarity). Together, these findings show that long-term memory and short-term memory can vary independently of each other.

Not all researchers agree that short-term and long-term memory are separate systems. Some theorists propose that memory is unitary over all time scales, from milliseconds to years.\cite{4} Support for the unitary memory hypothesis comes from the fact that it has been difficult to demarcate a clear boundary between short-term and long-term memory. For instance, Tarnow shows that the recall probability vs. latency curve is a straight line from 6 to 600 seconds (ten minutes), with the probability of failure to recall only saturating after 600 seconds.\cite{5} If there were really two different memory stores operating in this time frame, one could expect a discontinuity in this curve. Other research has shown that the detailed pattern of recall errors looks remarkably similar for recall of a list immediately...
after learning (presumably from short-term memory) and recall after 24 hours (necessarily from long-term memory). [6]

**Biological basis**

It is proposed by Tarnow that short term memory involves the firing of neurons which depletes the Readily Releasable Pool (RRP) of neurotransmitter vesicles at presynaptic terminals. [7] The pattern of depleted presynaptic terminals represents the long term memory trace and the depletion itself is the short term memory. After the firing has slowed down, endocytosis causes short term memory to decay. If the endocytosis is allowed to finish (the memory is not activated again), the pattern of exhausted postsynaptic terminals becomes invisible and the short term memory disappears. The long term memory remains as the metastable pattern of the neuronal excitations.

**Relationship with working memory**

The relationship between short-term memory and working memory is described differently by various theories, but it is generally acknowledged that the two concepts are distinct. Working memory is a theoretical framework that refers to structures and processes used for temporarily storing and manipulating information. As such, working memory might also be referred to as working attention. Short-term memory generally refers, in a theory-neutral manner, to the short-term storage of information, and it does not entail the manipulation or organization of material held in memory. Thus while there are short-term memory components to working memory models, the concept of short-term memory is distinct from these more hypothetical concepts. Within Baddeley's influential 1986 model of working memory there are two short-term storage mechanisms: the phonological loop and the visuospatial sketchpad. Most of the research referred to here involves the phonological loop, because most of the work done on short-term memory has used verbal material. In recent years, however, there has been a surge in research on visual short term memory, [8] and also increasing work on spatial short term memory. [9]

**Duration of short-term memory**

The limited duration of short-term memory immediately suggests that its contents spontaneously decays over time. The decay assumption is part of many theories of short-term memory, most notably Baddeley’s model of working memory. The decay assumption is usually paired with the idea of rapid covert rehearsal: In order to overcome the limitation of short-term memory, and retain information for longer, information must be periodically repeated, or rehearsed — either by articulating it out loud, or by mentally simulating such articulation. In this way, the information will re-enter the short-term store and be retained for a further period.

Several researchers, however, dispute that spontaneous decay plays any significant role in forgetting over the short term, [10] [11] and the evidence is far from conclusive. [12]

Authors doubting that decay causes forgetting from short-term memory often offer as an alternative some form of interference: When several elements (such as digits, words, or pictures) are held in short term memory simultaneously, their representations compete with each other for recall, or degrade each other. Thereby, new content gradually pushes out older content, unless the older content is actively protected against interference by rehearsal or by directing attention to it. [13]

**Capacity of short-term memory**

Whatever the cause or causes of forgetting over the short term may be, there is consensus that it severely limits the amount of new information that we can retain over brief periods of time. This limit is referred to as the finite capacity of short-term memory. The capacity of short-term memory is often called memory span, in reference to a common procedure of measuring it. In a memory span test, the experimenter presents lists of items (e.g. digits or words) of increasing length. An individual's span is determined as the longest list length that he or she can recall.
correctly in the given order on at least half of all trials.

In an early and highly influential article, The Magical Number Seven, Plus or Minus Two, the psychologist George Miller suggested that human short-term memory has a forward memory span of approximately seven items plus or minus two and that that was well known at the time (it seems to go back to the 19th century researcher Wundt). More recent research has shown that this “magical number seven” is roughly accurate for college students recalling lists of digits, but memory span varies widely with populations tested and with material used. For example, the ability to recall words in order depends on a number of characteristics of these words: fewer words can be recalled when the words have longer spoken duration; this is known as the word-length effect or when their speech sounds are similar to each other; this is called the phonological similarity effect. More words can be recalled when the words are highly familiar or occur frequently in the language. Recall performance is also better when all of the words in a list are taken from a single semantic category (such as sports) than when the words are taken from different categories. According to the available evidence, the best overall estimate of short-term memory is about four pieces or “chunks” of information. In free recall it has been shown, to the contrary, that there is no such “quantized” limit, rather it is a function of memory decaying with time.

**Chunking**

Chunking is the process with which we can expand our ability to remember things in the short term. Chunking is also a process by which a person organizes material into meaningful groups. Although the average person may only retain about four different units in short-term memory, chunking can greatly increase a person’s recall capacity. For example, in recalling a phone number, the person could chunk the digits into three groups: first, the area code (such as 215), then a three-digit chunk (123) and lastly a four-digit chunk (4567). This method of remembering phone numbers is far more effective than attempting to remember a string of 10 digits.

Practice and the usage of existing information in long-term memory can lead to additional improvements in one’s ability to use chunking. In one testing session, an American cross-country runner was able to recall a string of 79 digits after hearing them only once by chunking them into different running times (e.g. the first four numbers were 1518, a three-mile time.)

**Factors affecting short term memory**

It is very difficult to demonstrate the exact capacity of STM because it will vary depending on the nature of the material to be recalled. There is currently no way of defining the basic unit of information to be stored in the STM store. It is also possible that STM is not the store described by Atkinson and Shiffrin. In that case, the task of defining the task of STM becomes even more difficult.

However, capacity of STM can be affected by the following: Influence of long-term memory, Reading aloud, Pronunciation time and Individual differences.
References

Notes
[1] Tarnow
[16] Conrad & Hull, 1964
[18] Poirier & Saint-Aubin, 1995

Bibliography
• Tarnow, Eugen (2007) Response probability and latency: a straight line, an operational definition of meaning and the structure of short term memory (http://cogprints.org/6209/)

External links
• Short term memory in educational psychology (http://moodle.ed.uiuc.edu/wiked/index.php/Memory_short_term)
• An ERP Index of Visual Short Term Memory Capacity (http://develintel.blogspot.com/2006/02/neuroindices-of-memory-capacity.html)

Schema (psychology)

A schema (pl. schemata or schemas), in psychology and cognitive science, describes any of several concepts including:
• An organized pattern of thought or behavior.
• A structured cluster of pre-conceived ideas.
• A mental structure that represents some aspect of the world.
• A specific knowledge structure or cognitive representation of the self.
• A mental framework centering on a specific theme, that helps us to organize social information.
• Structures that organize our knowledge and assumptions about something and are used for interpreting and processing information.

A schema for oneself is called a "self schema". Schemata for other people are called "person schemata". Schemata for roles or occupations are called "role schemata", and schemata for events or situations are called "event schemata" (or scripts).

Schemata influence our attention, as we are more likely to notice things that fit into our schema. If something contradicts our schema, it may be encoded or interpreted as an exception or as unique. Thus, schemata are prone to distortion. They influence what we look for in a situation. They have a tendency to remain unchanged, even in the face of contradictory information. We are inclined to place people who do not fit our schema in a "special" or "different" category, rather than to consider the possibility that our schema may be faulty. As a result of schemata, we might act in such a way that actually causes our expectations to come true.

The concept of schemata was initially introduced into psychology and education through the work of the British psychologist, Sir Frederic Bartlett (1886–1969). This learning theory views organized knowledge as an elaborate network of abstract mental structures that represent one's understanding of the world. Schema theory was developed
by the educational psychologist R. C. Anderson. The term schema was used by Jean Piaget in 1926, so it was not an entirely new concept. Anderson, however, expanded the meaning.\footnote{2}

People use schemata to organize current knowledge and provide a framework for future understanding. Examples of schemata include Rubric (academic), social schemas, stereotypes, social roles, scripts, worldviews, and archetypes. In Piaget's theory of development, children adopt a series of schemata to understand the world.

**History of schema theory**

Early developments of the idea in psychology emerged with the Gestalt psychologists and Piaget. However, it is with the work of Sir Frederic Bartlett\footnote{3} (himself drawing on the term as used by the neurologist Sir Henry Head) that the term came to be used in its modern sense. Bartlett's work was neglected in America during the behaviouristic era until its wholesale recapitulation in Ulric Neisser's massively influential Cognitive Psychology (1967).\footnote{4} Neisser's work led to the ubiquity of the term in psychology, and its extension to other disciplines, notably the cognitive and computational sciences. Since that time, many other terms have been used as well, including "frame", "scene", and "script".

**Thought using schemata**

Schemata are an effective tool for understanding the world. Through the use of schemata, most everyday situations do not require effortful processing—automatic processing is all that is required. People can quickly organize new perceptions into schemata and act effectively without effort. For example, most people have a stairway schema and can apply it to climb staircases they've never seen before.

However, schemata can influence and hamper the uptake of new information (proactive interference), such as when existing stereotypes, giving rise to limited or biased discourses and expectations (prejudices), may lead an individual to "see" or "remember" something that has not happened because it is more believable in terms of his/her schema. For example, if a well-dressed businessman draws a knife on a vagrant, the schemata of onlookers may (and often do) lead them to "remember" the vagrant pulling the knife. Such distortion of memory has been demonstrated. (See Background research below.)

Schemata are interrelated and multiple conflicting schemata can be applied to the same information. Schemata are generally thought to have a level of activation, which can spread among related schemata. Which schema is selected can depend on factors such as current activation, accessibility, and priming.

*Accessibility* is how easily a schema comes to mind, and is determined by personal experience and expertise. This can be used as a cognitive shortcut; it allows the most common explanation to be chosen for new information.

With priming, a brief imperceptible stimulus temporarily provides enough activation to a schema so that it is used for subsequent ambiguous information. Although this may suggest the possibility of subliminal messages, the effect of priming is so fleeting that it is difficult to detect outside laboratory conditions. Furthermore, the mere exposure effect—which requires consciousness of the stimuli—is far more effective than priming.
Background research

Sufferers of Korsakov’s syndrome are unable to form new memories, and must approach every situation as if they had just seen it for the first time. Many sufferers adapt by continually forcing their world into barely-applicable schemata, often to the point of incoherence and self-contradiction.

The original concept of schemata is linked with that of reconstructive memory as proposed and demonstrated in a series of experiments by Bartlett (1932). By presenting participants with information that was unfamiliar to their cultural backgrounds and expectations and then monitoring how they recalled these different items of information (stories, etc.), Bartlett was able to establish that individuals’ existing schemata and stereotypes influence not only how they interpret “schema-foreign” new information but also how they recall the information over time. One of his most famous investigations involved asking participants to read a Native American folk tale, “The War of the Ghosts”, and recall it several times up to a year later. All the participants transformed the details of the story in such a way that it reflected their cultural norms and expectations, i.e. in line with their schemata. The factors that influenced their recall were:

- Omission of information that was considered irrelevant to a participant;
- Transformation of some of the details, or of the order in which events, etc., were recalled; a shift of focus and emphasis in terms of what was considered the most important aspects of the tale;
- Rationalization: details and aspects of the tale that would not make sense would be “padded out” and explained in an attempt to render them comprehensible to the individual in question;
- Cultural shifts: the content and the style of the story were altered in order to appear more coherent and appropriate in terms of the cultural background of the participant.

Bartlett’s work was crucially important in demonstrating that long-term memories are neither fixed nor immutable but are constantly being adjusted as our schemata evolve with experience. In a sense it supports the existentialist view that we construct our past and present in a constant process of narrative/discursive adjustment, and that much of what we “remember” is actually confabulated (adjusted and rationalized) narrative that allows us to think of our past as a continuous and coherent string of events, even though it is probable that large sections of our memory (both episodic and semantic) are irretrievable to our conscious memory at any given time.

Further work on the concept of schemata was conducted by Brewer and Treyens (1981) who demonstrated that the schema-driven expectation of the presence of an object was sometimes sufficient to trigger its erroneous recollection. An experiment was conducted where participants were requested to wait in a room identified as an academic’s study and were later asked about the room’s contents. A number of the participants recalled having seen books in the study whereas none were present. Brewer and Treyens concluded that the participants’ expectations that books are present in academics’ studies were enough to prevent their accurate recollection of the scenes.

Modification of schemata

New information that falls within an individual’s schema is easily remembered and incorporated into their worldview. However, when new information is perceived that does not fit a schema, many things can happen. The most common reaction is to simply ignore or quickly forget the new information. This can happen on a deep level—frequently an individual does not become conscious of or even perceive the new information. However, when the new information cannot be ignored, existing schemata must be changed.

Assimilation is the reuse of schemata to fit the new information. For example, when an unfamiliar dog is seen, a person will probably just assimilate it into their dog schema. However, if the dog behaves strangely, and in ways that don’t seem dog-like, there will be “accommodation” as a new schema is formed for that particular dog.
Self-schemata

Schemata about oneself are considered to be grounded in the present and based on past experiences. Memories, as mentioned, are framed in the light of one's self-conception. There are three major implications of self-schemata. First, information about oneself is processed faster and more efficiently, especially consistent information. Second, one retrieves and remembers information that is relevant to one's self-schema. Third, one will tend to resist information in the environment that is contradictory to one's self-schema. This is also related to self-verification.

Schema therapy

Schema therapy was founded by Dr Jeffrey Young, and represents a development of cognitive behaviour therapy (CBT) specifically for treating personality disorders. The therapy blends CBT with elements of Gestalt, object relations, constructivist and psychoanalytical therapies in order to treat the characterological difficulties which both constitute personality disorders and which underlie many of the chronic depressive or anxiety-involving symptoms which present in the clinic. CBT, Young felt, may be an effective treatment for presenting symptoms, but without the conceptual or clinical resources for tackling the underlying structures (maladaptive schemas) which consistently organise the patient's experience, the patient is likely to lapse back into unhelpful modes of relating to others and attempting to meet their needs. Early Maladaptive Schemas are described by Young as: broad and pervasive themes or patterns; made up of memories, feelings, sensations, and thoughts; regarding oneself and one's relationships with others; developing during childhood or adolescence; elaborated throughout life; and dysfunctional in that they lead to self-defeating behaviour. Examples include schemas of abandonment/instability, mistrust/abuse, emotional deprivation, and defectiveness/shame.

References


Further reading

• Bartlett, F.C. (1932), Remembering: An Experimental and Social Study. Cambridge: Cambridge University Press

External links

• Schema theory of learning (http://www.sil.org/lingualinks/literacy/ImplementALiteracyProgram/SchemaTheoryOfLearning.htm)
• Emotion Schemas - Psychic structures that shape our individual personalities, influence how we interact with others, experience our emotions, and interpret our reactions. (http://www.thereferentialprocess.org/theory/emotion-schemas)
Baddeley's model of working memory

Alan Baddeley and Graham Hitch proposed a **model of working memory** in 1974, in an attempt to describe a more accurate model of short-term memory.

Baddeley & Hitch proposed their tripartite working memory model as an alternative to the short-term store in Atkinson & Shiffrin's 'multi-store' memory model (1968). This model is later expanded upon by Baddeley and other co-workers and has become the dominant view in the field of working memory. However, alternative models are developing (see working memory) providing a different perspective on the working memory system.

The original model of Baddeley & Hitch was composed of three main components; the **central executive** which acts as supervisory system and controls the flow of information from and to its **slave systems**: the **phonological loop** and the **visuo-spatial sketchpad**. The slave systems are short-term storage systems dedicated to a content domain (verbal and visuo-spatial, respectively). In 2000 Baddeley added a third slave system to his model, the **episodic buffer**.

Baddeley & Hitch's argument for the distinction of two domain-specific slave systems in the older model was derived from experimental findings with dual-task paradigms. Performance of two simultaneous tasks requiring the use of two separate perceptual domains (i.e. a visual and a verbal task) is nearly as efficient as performance of the tasks individually. In contrast, when a person tries to carry out two tasks simultaneously that use the same perceptual domain, performance is less efficient than when performing the tasks individually.

**Components**

**Central executive**

The central executive is a flexible system responsible for the control and regulation of cognitive processes. It has the following functions:

- binding information from a number of sources into coherent episodes
- coordination of the slave systems
- shifting between tasks or retrieval strategies
- selective attention and inhibition

It can be thought of as a supervisory system that controls cognitive processes and intervenes when they go astray.

Using the dual-task paradigm, Baddeley and Erses have found, for instance, that patients with Alzheimer's dementia are impaired when performing multiple tasks simultaneously, even when the difficulty of the individual tasks is adapted to their abilities.\(^1\)

Recent research on executive functions suggests that the 'central' executive is not as central as conceived in the Baddeley & Hitch model. Rather, there seem to be separate executive functions that can vary largely independently between individuals and can be selectively impaired or spared by brain damage.\(^2\)
Baddeley’s model of working memory

**Phonological loop**

The phonological loop (or "articulatory loop") as a whole deals with sound or phonological information. It consists of two parts: a short-term phonological store with auditory memory traces that are subject to rapid decay and an articulatory rehearsal component (sometimes called the articulatory loop) that can revive the memory traces.

Any auditory verbal information is assumed to enter automatically into the phonological store. Visually presented language can be transformed into phonological code by silent articulation and thereby be encoded into the phonological store. This transformation is facilitated by the articulatory control process. The phonological store acts as an 'inner ear', remembering speech sounds in their temporal order, whilst the articulatory process acts as an 'inner voice' and repeats the series of words (or other speech elements) on a loop to prevent them from decaying. The phonological loop may play a key role in the acquisition of vocabulary, particularly in the early childhood years.\(^3\) It may also be vital for learning a second language.

Five main findings provide evidence for the phonological loop:

1. The effect of phonological similarity:
   Lists of words that sound similar are more difficult to remember than words that sound different. Semantic similarity (similarity of meaning) has comparatively little effect, supporting the assumption that verbal information is coded largely phonologically in working memory.\(^4\)

2. The effect of articulatory suppression:
   Memory for verbal material is impaired when people are asked to say something irrelevant aloud. This is assumed to block the articulatory rehearsal process, thereby leaving memory traces in the phonological loop to decay.\(^5\)

3. Transfer of information between codes:
   With visually presented items, adults usually name and sub-vocally rehearse them, so the information is transferred from a visual to an auditory code. Articulatory suppression prevents this transfer, and in that case the above mentioned effect of phonological similarity is erased for visually presented items.\(^6\)

4. Neuropsychological evidence:
   A defective phonological store explains the behavior of patients with a specific deficit in phonological short-term memory. Aphasic patients with dyspraxia are unable to set up the speech motor codes necessary for articulation, caused by a deficiency of the articulatory rehearsal process.\(^7\)

5. On the other hand, patients with dysarthria, whose speech problems are secondary, show a normal capacity for rehearsal. This suggests that it is the subvocal rehearsing that is crucial.\(^8\)

**Biology**

In terms of genetics, the gene ROBO1 has been associated with phonological buffer integrity or length \(^9\)

**Visuospatial sketchpad**

The visuospatial sketchpad is assumed to hold information about what we see. It is used in the temporary storage and manipulation of spatial and visual information, such as remembering shapes and colours, or the location or speed of objects in space. It is also involved in tasks which involve planning of spatial movements, like planning one's way through a complex building. The visuospatial sketchpad can be divided into separate visual, spatial and possibly...
kinaesthetic (movement) components. It is principally represented within the right hemisphere of the brain.\[10\]

**Logie’s elaboration of the visuospatial sketchpad**

Logie has proposed that the visuospatial sketchpad can be further subdivided into two components:
1. The visual cache, which stores information about form and color.
2. The inner scribe, which deals with spatial and movement information. It also rehearses information in the visual cache and transfers information to the central executive.\[11\]

Three main findings provide evidence for the distinction between visual and spatial parts of the visuospatial sketchpad:
1. There is less interference between visual and spatial tasks than between two visual tasks or two spatial tasks.\[12\]
2. Brain damage can influence one of the components without influencing the other.\[13\]
3. Results from brain-imaging show that working memory tasks with visual objects activate mostly areas in the left hemisphere, whereas tasks with spatial information activate more areas in the right hemisphere.\[14\]

**Episodic buffer**

In 2000 Baddeley added a fourth component to the model, called the ‘episodic buffer’. This component is a third slave system, dedicated to linking information across domains to form integrated units of visual, spatial, and verbal information with time sequencing (or chronological ordering), such as the memory of a story or a movie scene. The episodic buffer is also assumed to have links to long-term memory and semantical meaning.\[15\]

The main motivation for introducing this component was the observation that some (in particular, highly intelligent) patients with amnesia, who presumably have no ability to encode new information in long-term memory, nevertheless have good short-term recall of stories, recalling much more information than could be held in the phonological loop.\[16\]

**Validity of the model**

The strength of Baddeley’s model is its ability to integrate a large amount of findings from work on short-term and working memory. Additionally, the mechanisms of the slave systems, especially the phonological loop, has inspired a wealth of research in experimental psychology, neuropsychology, and cognitive neuroscience. However, criticisms have been raised, for instance of the phonological-loop component, because some details of the findings are not easily explained by the original Baddeley & Hitch model.\[17\]\[18\]

The episodic buffer is seen as a helpful addition to the model of working memory, but it has not been investigated extensively and its functions remain unclear.\[19\]

**References**

**Notes**


   b) Baddeley AD (November 1966). "Short-term memory for word sequences as a function of acoustic, semantic and formal similarity" (http://


Bibliography

Cognitive Load

Cognitive load

The term cognitive load is used in cognitive psychology to illustrate the load related to the executive control of working memory (WM).

This article mainly deals with an aspect of cognitive load, specifically it mostly deals with the load on working memory during instruction.

Instruction may be aimed at teaching learners problem solving skills, thinking and reasoning skills (including perception, memory, language, etc.).[1] Many would agree that people learn better when they can build on what they already understand (known as a schema), but the more a person has to learn in a shorter amount of time, the more difficult it is to process that information in working memory. Consider the difference between having to study a subject in one's native language versus trying to study a subject in a foreign language. The cognitive load is much higher in the second instance because the brain must work to translate the language while simultaneously trying to understand the new information.

Another aspect of cognitive load theory involves understanding how many discrete units of information can be retained in short-term memory before information loss occurs. An example of this principle that seems to be commonly cited is the use of 7-digit phone numbers, based on the theory that most people can only retain seven "chunks" of information in their short-term memory.

Cognitive load theory

"Cognitive load theory has been designed to provide guidelines intended to assist in the presentation of information in a manner that encourages learner activities that optimize intellectual performance".[2] John Sweller's theory employs aspects of information processing theory to emphasize the inherent limitations of concurrent working memory load on learning during instruction. It makes use of schemas as the unit of analysis for the design of instructional materials.

The history of cognitive load theory

The history of cognitive load theory can be traced back to the beginning of Cognitive Science and the work of G.A. Miller. Miller was perhaps the first to suggest our working memory capacity was limited in his classic paper.[3] His experimental results suggested that humans are only able to hold seven plus or minus two digits of information in their short-term memory. Simon and Chase[4] used the term "chunk" to describe how people might organize information in their short-term memory, this chunking of memory components has also been described as schema construction.

John Sweller developed cognitive load theory (CLT) while studying problem solving.[1] While studying learners as they solved problems, he and his associates found that learners often use a problem solving strategy called means-ends analysis. He suggests problem solving by means-ends analysis requires a relatively large amount of cognitive processing capacity, which may not be devoted to schema construction. Instead of problem solving, Sweller suggests Instructional designers should limit cognitive load by designing instructional materials like worked-examples, or goal-free problems.

In the 1990s, cognitive load theory was applied in several contexts and the empirical results from these studies led to the demonstration of several learning effects: the completion-problem effect;[5] modality effect,[6] [7] split-attention
Cognitive load theory has broad implications for instructional design. This theory provides a general framework for instructional designers for it allows them to control the conditions of learning, within an environment, or more generally within most instructional materials. Specifically, it provides empirically-based guidelines that help instructional designers to decrease extraneous cognitive load during learning, and refocus learner's attention toward germane materials, increasing germane (schema related) cognitive load. This theory differentiates between three types of cognitive load: intrinsic cognitive load, germane cognitive load, and extraneous cognitive load.

Intrinsic

Intrinsic cognitive load is the inherent level of difficulty associated with instructional materials. The term was first used by Chandler and Sweller. According to them, all instruction has an inherent difficulty associated with it (e.g., the calculation of \(2 + 2\), versus solving a differential equation). This inherent difficulty may not be altered by an instructor. However many schemas may be broken into individual "subschemas" and taught in isolation, to be later brought back together and described as a combined whole.

Extraneous

Extraneous cognitive load is generated by the manner in which information is presented to learners and is under the control of instructional designers. This load can be attributed to the design of the instructional materials. Because there is a single, limited cognitive resource, using resources to process the extraneous load reduces the amount of resources available to process the intrinsic load and germane load (i.e., learning). Thus, especially when intrinsic and/or germane load is high (i.e., when a problem is difficult), materials should be designed so as to reduce the extraneous load.

An example of extraneous cognitive load occurs when there are two possible ways to describe a square to a student. A square is a figure and should be described using a figural medium. Certainly an instructor can describe a square in a verbal medium, but it takes just a second and far less effort to see what the instructor is talking about when a learner is shown a square, rather than having one described verbally. In this instance, the efficiency of the visual medium is preferred. This is because it does not unduly load the learner with unnecessary information. This unnecessary cognitive load is described as extraneous.

Chandler and Sweller introduced the concept of extraneous cognitive load. This article was written to report the results of six experiments that they conducted to investigate this working memory load. Many of these experiments involved materials demonstrating the split attention effect. They found that the format of instructional materials either promoted or limited learning. They proposed that differences in performance were due to higher levels of the cognitive load imposed by the format of instruction. "Extraneous cognitive load" is a term for this unnecessary (artificially induced) cognitive load.

Germane

Germane cognitive load is that load devoted to the processing, construction and automation of schemas. It was first described by Sweller, van Merrienboer and Paas in 1998. While intrinsic cognitive load is generally thought to be immutable (although techniques can be applied to manage complexity by segmenting and sequencing complex material), instructional designers can manipulate extraneous and germane load. It is suggested that they limit extraneous load and promote germane load.

Until the 1998 article by Sweller, van Merrienboer & Paas, cognitive load theory primarily concentrated on the reduction of extraneous cognitive load. With this article, cognitive load researchers began to seek ways of...
Cognitive load redesigning instruction to redirect what would be extraneous load, to now be focused toward schema construction (germane load). Thus it is very important for instructional designers to "reduce extraneous cognitive load and redirect learners' attention to cognitive processes that are directly relevant to the construction of schemas."[16]

**Measurement**

Paas and van Merriënboer[17] developed a construct (known as relative condition efficiency) which helps researchers measure perceived mental effort, an index of cognitive load. This construct provides a relatively simple means of comparing instructional conditions. It combines mental effort ratings with performance scores. Group mean z-scores are graphed and may be compared with a one-way Analysis of variance (ANOVA).

Paas and van Merriënboer used relative condition efficiency to compare three instructional conditions (worked examples, completion problems, and discovery practice). They found learners who studied worked examples were the most efficient, followed by those who used the problem completion strategy. Since this early study many other researchers have used this and other constructs to measure cognitive load as it relates to learning and instruction.[18] The ergonomic approach seeks a quantitative neurophysiological expression of cognitive load which can be measured using common instruments, for example using the heart rate-blood pressure product (RPP) as a measure of both cognitive and physical occupational workload.[19] They believe that it may be possible to use RPP measures to set limits on workloads and for establishing work allowance.

**Individual differences in processing capacity**

Evidence has been found that individuals systematically differ in their processing capacity.[20] [21] A series of experiments support the assumption that each individual has a fixed capacity for processing information, irrespective of the task in question, or more accurately, irrespective of the processes an individual uses in solving any given task. Tasks ranged from remembering simple lists, lists supplemented with a fixed constant and simple arithmetic. Identifying the processing capacity of individuals could be extremely useful in further adapting instruction (or predicting the behavior) of individuals. Accordingly, further research would clearly be desirable. First, it is essential to compute the memory load imposed by detailed analysis of the processes to be used. Second, it is essential to ensure that individual subjects are actually using those processes. The latter requires intensive pre-training.

**Effects of heavy cognitive load**

Some are:

- Error
- Fundamental Attribution Error

**References**

For those wishing to learn more about cognitive load theory, please consider reading these journals and special issues of those journals:

- *Educational Psychologist* vol. 43 (2008)
- *Learning and Instruction* vol. 12 (2002)

For ergonomics standards see:

- ISO 9241 Ergonomics of Human System Interaction
Further reading


External links

- Cognitive Load Theory – a brief introduction to Cognitive load theory (http://knol.google.com/k/david-lewis-phd/cognitive-load-theory/3k9oibq1mbt1c/2)
- video of John Sweller explaining Extraneous load (http://www.youtube.com/watch?v=RyuOU2RasRQ)

Chunking (psychology)

In cognitive psychology and mnemonics, **chunking** refers to a strategy for making more efficient use of short-term memory by recoding information. More generally, Herbert Simon has used the term **chunk** to indicate long-term memory structures that can be used as units of perception and meaning, and **chunking** as the learning mechanisms leading to the acquisition of these chunks.

Chunking means to organize items into familiar manageable units.

"Magic number seven"

The word **chunking** comes from a famous 1956 paper by George A. Miller, *The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information*. At a time when information theory was beginning to be applied in psychology, Miller observed that whereas some human cognitive tasks fit the model of a "channel capacity" characterized by a roughly constant capacity in bits, short-term memory did not. A variety of studies could be summarized by saying that short-term memory had a capacity of about "seven plus-or-minus two" chunks. Miller wrote that "With binary items the span is about nine and, although it drops to about five with monosyllabic English words, the difference is far less than the hypothesis of constant information would require. The span of immediate memory seems to be almost independent of the number of bits per chunk, at least over the range that has been examined to date." Miller acknowledged that "we are not very definite about what constitutes a chunk of information."

Miller noted that according to this theory, it should be possible to effectively increase short-term memory for low-information-content items by mentally recoding them into a smaller number of high-information-content items. "A man just beginning to learn radio-telegraphic code hears each dit and dah as a separate chunk. Soon he is able to organize these sounds into letters and then he can deal with the letters as chunks. Then the letters organize
themselves as words, which are still larger chunks, and he begins to hear whole phrases.” Thus, a telegrapher can effectively "remember" several dozen dits and dahs as a single phrase. Naive subjects can only remember about nine binary items, but Miller reports a 1954 experiment in which people were trained to listen to a string of binary digits and (in one case) mentally group them into groups of five, recode each group into a name (e.g. "twenty-one" for 10101), and remember the names. With sufficient drill, people found it possible to remember as many as forty binary digits. Miller wrote:

"It is a little dramatic to watch a person get 40 binary digits in a row and then repeat them back without error. However, if you think of this merely as a mnemonic trick for extending the memory span, you will miss the more important point that is implicit in nearly all such mnemonic devices. The point is that recoding is an extremely powerful weapon for increasing the amount of information that we can deal with."

### Chunking in motor learning

Chunking is a flexible way of learning. Karl Lashley, in his classic paper on serial order (Lashley, 1951), argued that the sequential responses that appear to be organized in linear and flat fashion concealed an underlying hierarchical structure. This was demonstrated in motor control by Rosenbaum et al. (1983). Thus sequences can consist of sub-sequences and these can in turn consist of sub-sub-sequences. Hierarchical representations of sequences have an edge over linear representations. They combine efficient local action at low hierarchical levels while maintaining the guidance of an overall structure. While the representation of a linear sequence is simple from storage point of view, there can be potential problems during retrieval. For instance, if there is a break in the sequence chain, subsequent elements will become inaccessible. On the other hand, a hierarchical representation would have multiple levels of representation. A break in the link between lower level nodes does not render any part of the sequence inaccessible, since the control nodes (chunk nodes) at the higher level would still be able to facilitate access to the lower level nodes.

![Diagram of Chunking in Motor Learning](image-url)

Chunks in motor learning are identified by pauses between successive actions (Terrace, 2001). He also suggested that during the sequence performance stage (after learning), subjects download list items as chunks during pauses. Terrace also argued for an operational definition of chunks suggesting a distinction between the notions of input and output chunks from the ideas of short-term and long-term memory. Input chunks reflect the limitation of working memory during the encoding of new information, i.e., how new information is stored in long-term memory, and how it is retrieved during subsequent recall. Output chunks reflect the organization of over-learned motor programs that
are generated on-line in working memory. Sakai et al. (2003) showed that subjects spontaneously organize a sequence into a number of chunks across few sets, and that these chunks were distinct among subjects tested on the same sequence. Sakai et al. (2003) showed that performance of a shuffled sequence was poorer when the chunk patterns were disrupted than when the chunk patterns were preserved. Chunking patterns also seem to depend on the effectors used.

**Memory training systems**

The phenomenon of chunking as a memory mechanism can be observed in the way we group numbers and information in our day-to-day life. For example, when recalling a number such as 14101946, if we group the numbers as 14, 10 and 1946, we are creating a mnemonic for this number as a day, month and year. An illustration of the limited capacity of working memory as suggested by Miller can be seen from the following example: While recalling a mobile phone number such as 9849523450, we might break this into 98 495 234 50. Thus, instead of remembering 10 separate digits that is beyond the "seven plus-or-minus two", we are remembering 4 groups of numbers.

Various kinds of memory training systems and mnemonics include training and drill in specially-designed recoding or chunking schemes. Such systems existed before Miller's paper, but there was no convenient term to describe the general strategy. The term "chunking" is now often used in reference to these systems.

**Chunking as the learning of long-term memory structures**

This usage derives from Miller's (1956) idea of chunking as grouping, but the emphasis is now on long-term memory rather than on short-term memory. A chunk can then be defined as "a collection of elements having strong associations with one another, but weak associations with elements within other chunks" (Gobet et al., 2001, p. 236). Chase and Simon (1973), and later Gobet, Retschitzki and de Voogt (2004), showed that chunking could explain several phenomena linked to expertise in chess. Several successful computational models of learning and expertise have been developed using this idea, such as EPAM (Elementary Perceiver and Memorizer) and CHREST (Chunk Hierarchy and REtrieval STructures). Chunking has also been used with models of language acquisition.

**References**

External links

- The Magical Number Seven, Plus or Minus Two [1]: Full text of Miller's 1956 paper
- The Magical Number Seven, Plus or Minus Two [2]: Alternate text of Miller's 1956 paper

References


The Magical Number Seven, Plus or Minus Two

"The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information" is one of the most highly cited papers in psychology. [1] [2] [3] It was published in 1956 by the cognitive psychologist George A. Miller of Princeton University's Department of Psychology in Psychological Review. It supposedly argues that the number of objects an average human can hold in working memory is 7 ± 2. This is frequently referred to as Miller's Law (not to be confused with his theory of communication: Miller's Law).

Recent research has demonstrated that not only is the "law" based on a misinterpretation of Miller's paper, but that the correct number is probably around three or four. [4]

Miller's article

In his article, Miller discussed a coincidence between the limits of one-dimensional absolute judgment and the limits of short-term memory. In a one-dimensional absolute-judgment task, a person is presented with a number of stimuli that vary on one dimension (e.g., 10 different tones varying only in pitch) and responds to each stimulus with a corresponding response (learned before). Performance is nearly perfect up to 5 or 6 different stimuli but declines as the number of different stimuli is increased. The task can be described as one of information transmission: The input consists of one out of \( n \) possible stimuli, and the output consists of one out of \( n \) responses. The information contained in the input can be determined by the number of binary decisions that need to be made to arrive at the selected stimulus, and the same holds for the response. Therefore, people's maximum performance on one-dimensional absolute judgement can be characterized as an information channel capacity with approximately 2 to 3 bits of information, which corresponds to the ability to distinguish between 4 and 8 alternatives.

The second cognitive limitation Miller discusses is memory span. Memory span refers to the longest list of items (e.g., digits, letters, words) that a person can repeat back immediately after presentation in correct order on 50% of trials. Miller observed that memory span of young adults is approximately 7 items. He noticed that memory span is approximately the same for stimuli with vastly different amount of information - for instance, binary digits have 1 bit,
The Magical Number Seven, Plus or Minus Two

Miller concluded that memory span is not limited in terms of bits but rather in terms of chunks. A chunk is the largest meaningful unit in the presented material that the person recognizes - thus, it depends on the knowledge of the person what counts as a chunk. For instance, a word is a single chunk for a speaker of the language but breaks down into as many chunks as the word has letters for someone who is totally unfamiliar with the language.

Miller recognized that the correspondence between the limits of one-dimensional absolute judgment and of short-term memory span was only a coincidence, because only the first limit, not the second, can be characterized in information-theoretic terms (i.e., as a roughly constant number of bits). Therefore, there is nothing "magical" about the number 7, and Miller used the expression only ironically. Nevertheless, the idea of a "magical number 7" inspired much theorizing, rigorous and less rigorous, about the capacity limits of human cognition.

The "magical number 7" and working memory capacity

Later research on short-term memory and working memory revealed that memory span is not a constant even when measured in terms of a number of chunks. The number of chunks a human can recall immediately after presentation depends on the category of chunks used (e.g., span is around seven for digits, around six for letters, and around five for words), and even on features of the chunks within a category. For instance, span is lower for long words than it is for short words. In general, memory span for verbal contents (digits, letters, words, etc.) strongly depends on the time it takes to speak the contents aloud. Some researchers have therefore proposed that the limited capacity of short-term memory for verbal material is not a "magic number" but rather a "magic spell". Baddeley used this finding to postulate that one component of his model of working memory, the phonological loop, is capable of holding around 2 seconds of sound. However, the limit of short-term memory cannot easily be characterized as a constant "magic spell" either, because memory span depends also on other factors besides speaking duration. For instance, span depends on the lexical status of the contents (i.e., whether the contents are words known to the person or not). Several other factors also affect a person's measured span, and therefore it is difficult to pin down the capacity of short-term or working memory to a number of chunks. Nonetheless, Cowan (2001) proposed that working memory has a capacity of about four chunks in young adults (and less in children and older adults).

Tarnow (2010) finds that in a classic experiment typically argued as supporting a 4 item buffer by Murdock (1962). There is in fact no evidence for such and thus the "magical number", at least in the Murdock experiment, is 1.

Other cognitive numeric limits

Cowan also noted a number of other limits of cognition that point to a "magical number four", and different from Miller, he argued that this correspondence is no coincidence. One other process that seems to be limited at about four elements is subitizing, the rapid enumeration of small numbers of objects. When a number of objects is flashed briefly, their number can be determined very quickly, at a glance, when the number does not exceed the subitizing limit, which is about four objects. Larger numbers of objects must be counted, which is a slower process. The film Rain Man portrayed an autistic savant, who was able to rapidly determine the number of toothpicks from an entire box spilled on the floor, apparently subitizing a much larger number than four objects. A similar feat was informally observed by neuropsychologist Oliver Sacks and reported in his book The Man Who Mistook His Wife for a Hat. (Autistic expert Daniel Tammet has suggested that the children Sacks observed may have pre-counted the matches in the box.) Therefore one might suppose that this limit is an arbitrary limit imposed by our cognition rather than necessarily being a physical limit.
Notes


References


External links

• Single neuron binding properties and the magical number 7 - Michele Migliore,Gaspare Novara and Domenico Tegolo (http://www3.interscience.wiley.com/journal/121360014/abstract) Hippocampus - Volume 18 Issue 11, Pages 1122 - 1130 - Published Online: 4 Aug 2008
• Version of the paper with figures adapted for HTML and proofread and approved by Miller in 1997 (http://www.musanim.com/miller1956/)
• The 7±2 Urban Legend (http://www.knosof.co.uk/cbook/misart.pdf) (pdf file)
Problem solving

**Problem solving** is a mental process and is part of the larger problem process that includes problem finding and problem shaping. Considered the most complex of all intellectual functions, problem solving has been defined as higher-order cognitive process that requires the modulation and control of more routine or fundamental skills.\[1\] Problem solving occurs when an organism or an artificial intelligence system needs to move from a given state to a desired goal state.

**Overview**

The nature of human problem solving methods has been studied by psychologists over the past hundred years. There are several methods of studying problem solving, including: introspection, behaviorism, simulation, computer modeling and experiment.

Beginning with the early experimental work of the Gestaltists in Germany (e.g. Duncker, 1935 \[2\]), and continuing through the 1960s and early 1970s, research on problem solving typically conducted relatively simple, laboratory tasks (e.g. Duncker's "X-ray" problem; Ewert & Lambert's 1932 "disk" problem, later known as Tower of Hanoi) that appeared novel to participants (e.g. Mayer, 1992 \[3\]). Various reasons account for the choice of simple novel tasks: they had clearly defined optimal solutions, they were solvable within a relatively short time frame, researchers could trace participants' problem-solving steps, and so on. The researchers made the underlying assumption, of course, that simple tasks such as the Tower of Hanoi captured the main properties of "real world" problems, and that the cognitive processes underlying participants' attempts to solve simple problems were representative of the processes engaged in when solving "real world" problems. Thus researchers used simple problems for reasons of convenience, and thought generalizations to more complex problems would become possible. Perhaps the best-known and most impressive example of this line of research remains the work by Allen Newell and Herbert Simon \[4\].

Simple laboratory-based tasks can be useful in explicating the steps of logic and reasoning that underlie problem solving; however, they omit the complexity and emotional valence of "real-world" problems. In clinical psychology, researchers have focused on the role of emotions in problem solving (D'Zurilla & Goldfried, 1971; D'Zurilla & Nezu, 1982), demonstrating that poor emotional control can disrupt focus on the target task and impede problem resolution (Rath, Langenhahn, Simon, Sherr, & Diller, 2004). In this conceptualization, human problem solving consists of two related processes: problem orientation, the motivational/attitudinal/affective approach to problematic situations and problem-solving skills, the actual cognitive-behavioral steps, which, if successfully implemented, lead to effective problem resolution. Working with individuals with frontal lobe injuries, neuropsychologists have discovered that deficits in emotional control and reasoning can be remediated, improving the capacity of injured persons to resolve everyday problems successfully (Rath, Simon, Langenhahn, Sherr, & Diller, 2003).

**Europe**

In Europe, two main approaches have surfaced, one initiated by Donald Broadbent (1977; see Berry & Broadbent, 1995) in the United Kingdom and the other one by Dietrich Dörner (1975, 1985; see Dörner & Wearing, 1995) in Germany. The two approaches have in common an emphasis on relatively complex, semantically rich, computerized laboratory tasks, constructed to resemble real-life problems. The approaches differ somewhat in their theoretical goals and methodology, however. The tradition initiated by Broadbent emphasizes the distinction between cognitive problem-solving processes that operate under awareness versus outside of awareness, and typically employs mathematically well-defined computerized systems. The tradition initiated by Dörner, on the other hand, has an interest in the interplay of the cognitive, motivational, and social components of problem solving, and utilizes very complex computerized scenarios that contain up to 2,000 highly interconnected variables (e.g., Dörner, Kreuzig,

To sum up, researchers' realization that problem-solving processes differ across knowledge domains and across levels of expertise (e.g. Sternberg, 1995) and that, consequently, findings obtained in the laboratory cannot necessarily generalize to problem-solving situations outside the laboratory, has during the past two decades led to an emphasis on real-world problem solving. This emphasis has been expressed quite differently in North America and Europe, however. Whereas North American research has typically concentrated on studying problem solving in separate, natural knowledge domains, much of the European research has focused on novel, complex problems, and has been performed with computerized scenarios (see Funke, 1991, for an overview).

**USA and Canada**

In North America, initiated by the work of Herbert Simon on learning by doing in semantically rich domains (e.g. Anzai & Simon, 1979; Bhaskar & Simon, 1977), researchers began to investigate problem solving separately in different natural knowledge domains – such as physics, writing, or chess playing – thus relinquishing their attempts to extract a global theory of problem solving (e.g. Sternberg & Frensch, 1991). Instead, these researchers have frequently focused on the development of problem solving within a certain domain, that is on the development of expertise (e.g. Anderson, Boyle & Reiser, 1985; Chase & Simon, 1973; Chi, Feltovich & Glaser, 1981).

Areas that have attracted rather intensive attention in North America include such diverse fields as:

- Problem Solving (Kepner & Tregoe, 1958)
- Reading (Stanovich & Cunningham, 1991)
- Writing (Bryson, Bereiter, Scardamalia & Joram, 1991)
- Calculation (Sokol & McCloskey, 1991)
- Political decision making (Voss, Wolfe, Lawrence & Engle, 1991)
- Problem Solving for Business (Cornell, 2010)
- Managerial problem solving (Wagner, 1991)
- Lawyers' reasoning (Amsel, Langer & Loutzenhiser, 1991)
- Mechanical problem solving (Hegarty, 1991)
- Problem solving in electronics (Lesgold & Lajoie, 1991)
- Computer skills (Kay, 1991)
- Game playing (Frensch & Sternberg, 1991)
- Personal problem solving (Heppner & Krauskopf, 1987)
- Mathematical problem solving (Polya, 1945; Schoenfeld, 1985)
- Social problem solving (D'Zurilla & Goldfreid, 1971; D'Zurilla & Nezu, 1982)

**Characteristics of difficult problems**

As elucidated by Dietrich Dörner and later expanded upon by Joachim Funke, difficult problems have some typical characteristics that can be summarized as follows:

- Intransparency (lack of clarity of the situation)
  - commencement opacity
  - continuation opacity
- Polytely (multiple goals)
  - inexpressiveness
  - opposition
  - transience
Problem solving

- Complexity (large numbers of items, interrelations and decisions)
  - enumerability
  - connectivity (hierarchy relation, communication relation, allocation relation)
  - heterogeneity
- Dynamics (time considerations)
  - temporal constraints
  - temporal sensitivity
  - phase effects
  - dynamic unpredictability

The resolution of difficult problems requires a direct attack on each of these characteristics that are encountered.

In reform mathematics, greater emphasis is placed on problem solving relative to basic skills, where basic operations can be done with calculators. However some "problems" may actually have standard solutions taught in higher grades. For example, kindergarteners could be asked how many fingers are there on all the gloves of 3 children, which can be solved with multiplication.\(^5\)

**Problem-solving techniques**

- Abstraction: solving the problem in a model of the system before applying it to the real system
- Analogy: using a solution that solved an analogous problem
- Brainstorming: (especially among groups of people) suggesting a large number of solutions or ideas and combining and developing them until an optimum is found
- Divide and conquer: breaking down a large, complex problem into smaller, solvable problems
- Hypothesis testing: assuming a possible explanation to the problem and trying to prove (or, in some contexts, disprove) the assumption
- Lateral thinking: approaching solutions indirectly and creatively
- Means-ends analysis: choosing an action at each step to move closer to the goal
- Method of focal objects: synthesizing seemingly non-matching characteristics of different objects into something new
- Morphological analysis: assessing the output and interactions of an entire system
- Reduction: transforming the problem into another problem for which solutions exist
- Research: employing existing ideas or adapting existing solutions to similar problems
- Root cause analysis: eliminating the cause of the problem
- Trial-and-error: testing possible solutions until the right one is found

"A solution, to be a solution, must share some of the problems characteristics." Richard L. Kempe
Problem-solving methodologies

- Eight Disciplines Problem Solving
- 5Φ (IAPIE)
- GROW model
- How to solve it
- Kepner-Tregoe
- Southbeach Notation
- PDCA
- RPR Problem Diagnosis
- TRIZ (Teoriya Resheniya Izobretatelskikh Zadatch, “theory of solving inventor's problems”)
- WebKaizen

Example applications

Problem solving is of crucial importance in engineering when products or processes fail, so corrective action can be taken to prevent further failures. Perhaps of more value, problem solving can be applied to a product or process prior to an actual fail event i.e. a potential problem can be predicted, analyzed and mitigation applied so the problem never actually occurs. Techniques like Failure Mode Effects Analysis can be used to proactively reduce the likelihood of problems occurring. Forensic engineering is an important technique of failure analysis which involves tracing product defects and flaws. Corrective action can then be taken to prevent further failures.

Notes


References


External links

• Computer Skills for Information Problem-Solving: Learning and Teaching Technology in Context (http://wwwERICdigests.org/1996-4/skills.htm)

• Problem solving-Elementary level (http://moodle.ed.uiuc.edu/wiked/index.php/Problem_solving-Elementary_level)

• CROP (Communities Resolving Our Problems) (http://ceap.wcu.edu/houghton/Learner/basicidea.html)

• The Altshuller Institute for TRIZ Studies, Worcester, MA (http://www.aitriz.org)
Instructional Design

Instructional design

Instructional Design (also called Instructional Systems Design (ISD)) is the practice of maximizing the effectiveness, efficiency and appeal of instruction and other learning experiences. The process consists broadly of determining the current state and needs of the learner, defining the end goal of instruction, and creating some "intervention" to assist in the transition. Ideally the process is informed by pedagogically (process of teaching) and andragogically (adult learning) tested theories of learning and may take place in student-only, teacher-led or community-based settings. The outcome of this instruction may be directly observable and scientifically measured or completely hidden and assumed. There are many instructional design models but many are based on the ADDIE model with the five phases: 1) analysis, 2) design, 3) development, 4) implementation, and 5) evaluation. As a field, instructional design is historically and traditionally rooted in cognitive and behavioral psychology.

History

Much of the foundation of the field of instructional design was laid in World War II, when the U.S. military faced the need to rapidly train large numbers of people to perform complex technical tasks, from field-stripping a carbine to navigating across the ocean to building a bomber—see "Training Within Industry (TWI)"). Drawing on the research and theories of B.F. Skinner on operant conditioning, training programs focused on observable behaviors. Tasks were broken down into subtasks, and each subtask treated as a separate learning goal. Training was designed to reward correct performance and remediate incorrect performance. Mastery was assumed to be possible for every learner, given enough repetition and feedback. After the war, the success of the wartime training model was replicated in business and industrial training, and to a lesser extent in the primary and secondary classroom. The approach is still common in the U.S. military.[1]

In 1956, a committee led by Benjamin Bloom published an influential taxonomy of what he termed the three domains of learning: Cognitive (what one knows or thinks), Psychomotor (what one does, physically) and Affective (what one feels, or what attitudes one has). These taxonomies still influence the design of instruction.[2]

During the latter half of the 20th century, learning theories began to be influenced by the growth of digital computers.

In the 1970s, many instructional design theorists began to adopt an information-processing-based approach to the design of instruction. David Merrill for instance developed Component Display Theory (CDT),[3] which concentrates on the means of presenting instructional materials (presentation techniques).[4]

Later in the 1980s and throughout the 1990s cognitive load theory began to find empirical support for a variety of presentation techniques.[5]
**Cognitive load theory and the design of instruction**

Cognitive load theory developed out of several empirical studies of learners, as they interacted with instructional materials.[6] Sweller and his associates began to measure the effects of working memory load, and found that the format of instructional materials has a direct effect on the performance of the learners using those materials.[7] [8] [9]

While the media debates of the 1990s focused on the influences of media on learning, cognitive load effects were being documented in several journals. Rather than attempting to substantiate the use of media, these cognitive load learning effects provided an empirical basis for the use of instructional strategies. Mayer asked the instructional design community to reassess the media debate, to refocus their attention on what was most important: learning.[10]

By the mid- to late-1990s, Sweller and his associates had discovered several learning effects related to cognitive load and the design of instruction (e.g. the split attention effect, redundancy effect, and the worked-example effect). Later, other researchers like Richard Mayer began to attribute learning effects to cognitive load.[10] Mayer and his associates soon developed a Cognitive Theory of Multimedia Learning.[11] [12] [13]

In the past decade, cognitive load theory has begun to be internationally accepted[14] and begun to revolutionize how practitioners of instructional design view instruction. Recently, human performance experts have even taken notice of cognitive load theory, and have begun to promote this theory base as the science of instruction, with instructional designers as the practitioners of this field.[15] Finally Clark, Nguyen and Sweller[16] published a textbook[17] describing how Instructional Designers can promote efficient learning using evidence-based guidelines of cognitive load theory.

Instructional Designers use various instructional strategies to reduce cognitive load. For example, they think that the onscreen text should not be more than 150 words or the text should be presented in small meaningful chunks. The designers also use auditory and visual methods to communicate information to the learner.

**Learning design**

The concept of learning design arrived in the literature of technology for education in the late nineties and early 2000s[18] with the idea that "designers and instructors need to choose for themselves the best mixture of behaviourist and constructivist learning experiences for their online courses"[19]. But the concept of learning design is probably as old as the concept of teaching. Learning design might be defined as "the description of the teaching-learning process that takes place in a unit of learning (e.g. a course, a lesson or any other designed learning event)"[20].

As summarized by Britain[21], learning design may be associated with:

- The concept of learning design
- The implementation of the concept made by learning design specifications like PALO, IMS Learning Design[22], LDL, SLD 2.0, etc...
- The technical realisations around the implementation of the concept like TELOS, RELOAD LD-Author, etc...
**Instructional design models**

**ADDIE process**
Perhaps the most common model used for creating instructional materials is the ADDIE Process. This acronym stands for the 5 phases contained in the model:

- **Analyze** – analyze learner characteristics, task to be learned, etc.
- **Design** – develop learning objectives, choose an instructional approach
- **Develop** – create instructional or training materials
- **Implement** – deliver or distribute the instructional materials
- **Evaluate** – make sure the materials achieved the desired goals

Most of the current instructional design models are variations of the ADDIE process.[21]

**Rapid prototyping**
A sometimes utilized adaptation to the ADDIE model is in a practice known as rapid prototyping.

Proponents suggest that through an iterative process the verification of the design documents saves time and money by catching problems while they are still easy to fix. This approach is not novel to the design of instruction, but appears in many design-related domains including software design, architecture, transportation planning, product development, message design, user experience design, etc.[23] [24] [25] In fact, some proponents of design prototyping assert that a sophisticated understanding of a problem is incomplete without creating and evaluating some type of prototype, regardless of the analysis rigor that may have been applied up front.[26] In other words, up-front analysis is rarely sufficient to allow one to confidently select an instructional model. For this reason many traditional methods of instructional design are beginning to be seen as incomplete, naive, and even counter-productive.[27]

However, some consider rapid prototyping to be a somewhat simplistic type of model. As this argument goes, at the heart of Instructional Design is the analysis phase. After you thoroughly conduct the analysis—you can then choose a model based on your findings. That is the area where most people get snagged—they simply do not do a thorough-enough analysis. (Part of Article By Chris Bressi on LinkedIn)

**Dick and Carey**
Another well-known instructional design model is *The Dick and Carey Systems Approach Model.*[28] The model was originally published in 1978 by Walter Dick and Lou Carey in their book entitled *The Systematic Design of Instruction*.[29]

Dick and Carey made a significant contribution to the instructional design field by championing a systems view of instruction as opposed to viewing instruction as a sum of isolated parts. The model addresses instruction as an entire system, focusing on the interrelationship between context, content, learning and instruction. According to Dick and Carey, "Components such as the instructor, learners, materials, instructional activities, delivery system, and learning
and performance environments interact with each other and work together to bring about the desired student learning outcomes. The components of the Systems Approach Model, also known as the Dick and Carey Model, are as follows:

- Identify Instructional Goal(s)
- Conduct Instructional Analysis
- Analyze Learners and Contexts
- Write Performance Objectives
- Develop Assessment Instruments
- Develop Instructional Strategy
- Develop and Select Instructional Materials
- Design and Conduct Formative Evaluation of Instruction
- Revise Instruction
- Design and Conduct Summative Evaluation

With this model, components are executed iteratively and in parallel rather than linearly.

**Instructional Development Learning System (IDLs)**

Another instructional design model is the Instructional Development Learning System (IDLs). The model was originally published in 1970 by Peter J. Esseff, PhD and Mary Sullivan Esseff, PhD in their book entitled IDLS—Pro Trainer 1: How to Design, Develop, and Validate Instructional Materials.

Peter (1968) & Mary (1972) Esseff both received their doctorates in Educational Technology from the Catholic University of America under the mentorship of Dr. Gabriel Ofiesh, a Founding Father of the Military Model mentioned above. Esseff and Esseff contributed synthesized existing theories to develop their approach to systematic design, "Instructional Development Learning System" (IDLs).

The components of the IDLS Model are:

- Design a Task Analysis
- Develop Criterion Tests and Performance Measures
- Develop Interactive Instructional Materials
- Validate the Interactive Instructional Materials

**Other models**

Some other useful models of instructional design include: the Smith/Ragan Model, the Morrison/Ross/Kemp Model and the OAR model, as well as, Wiggins theory of backward design.

Learning theories also play an important role in the design of instructional materials. Theories such as behaviorism, constructivism, social learning and cognitivism help shape and define the outcome of instructional materials.

**Influential researchers and theorists**

*Alphabetic by last name*

- Bloom, Benjamin – Taxonomies of the cognitive, affective, and psychomotor domains – 1955
- Bonk, Curtis – Blended learning – 2000s
- Bransford, John D. – How People Learn: Bridging Research and Practice – 1999
- Bruner, Jerome – Constructivism
- Carr-Chellman, Alison – Instructional Design for Teachers ID4T -2010
- Carey, L. – "The Systematic Design of Instruction"
• Clark, Ruth – Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load / Guided Instruction / Cognitive Load Theory
• Dick, W. – "The Systematic Design of Instruction"
• Gagné, Robert M. – Nine Events of Instruction [33] (Gagné and Merrill Video Seminar)
• Heinich, Robert – Instructional Media and the new technologies of instruction 3rd ed. – Educational Technology – 1989
• Jonassen, David – problem-solving strategies – 1990s
• Langdon, Danny G - The Instructional Designs Library: 40 Instructional Designs, Educational Tech. Publications
• Mager, Robert F. – ABCD model for instructional objectives – 1962
• Merrill, M. David - Component Display Theory / Knowledge Objects
• Papert, Seymour – Constructionism, LOGO – 1970s
• Piaget, Jean – Cognitive development – 1960s
• Piskurich, George – Rapid Instructional Design – 2006
• Simonson, Michael – Instructional Systems and Design via Distance Education – 1980s
• Schank, Roger – Constructivist simulations – 1990s
• Sweller, John - Cognitive load, Worked-example effect, Split-attention effect
• Reigeluth, Charles – Elaboration Theory, "Green Books" I, II, and III - 1999-2010
• Skinner, B.F. – Radical Behaviorism, Programed Instruction
• Vygotsky, Leon – Sociocultural and Constructivist simulations – 1990s
• Wiley, David – Learning Objects, Open Learning – 2000s

References
[29] http://books.google.com/books?hl=en&lr=&id=a0gdmLZwI8IC&oi=fnd&pg=PA71&dq=The+Systematic+Design+of+Instruction
[31] (http://esf-protrainer.com/Materials.html)

**External links**

- Instructional Design (http://www.instructionaldesign.org/) - An overview of Instructional Design
- Edutech wiki: Instructional design model (http://edutechwiki.unige.ch/en/Instructional_design_model)
Multimedia learning

Multimedia learning is the common name used to describe the cognitive theory of multimedia learning\textsuperscript{[1]}\textsuperscript{[2]}\textsuperscript{[3]} This theory encompasses several principles of learning with multimedia.

The Modality principle

When information is in fact better remembered when accompanied by a visual image,\textsuperscript{[4]} Baddeley and Hitch proposed a theory of working memory in 1974 which has two largely independent subcomponents that tend to work in parallel - one visual and one verbal/acoustic.\textsuperscript{[5]} This allows us to simultaneously process information coming from our eyes and ears. Thus a learner is not necessarily overwhelmed or overloaded by multimodal instruction, and it can in fact be beneficial.

The finding that items presented both visually and verbally are better remembered gave rise to dual-coding theory, first proposed by Paivio and later applied to multimedia by Richard Mayer and his associates. Mayer has shown learners are better able to transfer their learning given multimodal instruction. Mayer explains the modality effect from an information processing/cognitive load perspective.

In a series of studies Mayer and his colleagues tested Paivio's dual-coding theory, with multimedia. They repeatedly found that students learning given multimedia with animation and narration consistently did better on transfer questions than those who learn from animation and text-based materials. That is, they were significantly better when it came to applying what they had learned after receiving multimedia rather than mono-media (visual only) instruction. These results were then later confirmed by other groups of researchers.

Initially the instructional content of these multimedia learning studies was limited to logical scientific processes that centered on cause-and-effect systems like automobile braking systems, how a bicycle pump works, or cloud formation. But eventually it was found that the modality effect could be extended to other domains, which were not necessarily cause-and-effect based systems.

Information then can and should be encoded as both as visually and auditory (narration). If verbal information is encoded auditorily it reduces the cognitive load of the learner and they are better able to handle that incoming information. Mayer has since called this the "Modality effect," or the Modality Principle. This was one of the many principles of his "Cognitive Theory of Multimedia Learning"\textsuperscript{[1]}.

The Redundancy principle

According to this principle: "Students learn better from animation and narration than from animation, narration, and on-screen text."\textsuperscript{[3]}

Thus it's better to eliminate redundant material. This is because learners do not learn as well when they both hear and see the same verbal message during a presentation. This is a special case of the split attention effect of Sweller and Chandler.
Other Principles

- Spatial Contiguity Principle - "Students learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen."\[3\]
- Temporal Contiguity Principle - "Students learn better when corresponding words and pictures are presented simultaneously rather than successively."\[3\]
- Coherence Principle - "Students learn better when extraneous material is excluded rather than included."\[3\]
- Individual Differences Principle - "Design effects are stronger for low-knowledge learners than for high knowledge learners, and for high-spatial learners rather than for low-spatial learners."\[3\]

Challenges to the Application of Principles

Not all research has found that the principles of multimedia learning apply generally outside of laboratory conditions. For example, Muller, Lee, and Sharma found that the coherence principle did not transfer to an authentic learning environment. In their study, adding approximately 50% additional extraneous but interesting material did not result in any significant difference in learner performance.\[6\]

See also

- Cognitive load
- Split attention effect
- Worked-example effect
- Dual-coding theory

References


Split attention effect

The **split-attention effect** is a learning effect inherent within some poorly designed instructional materials. It is apparent when the same modality (e.g. visual and visual) is used for various types of information within the same display. To learn from these materials learners must split their attention between these materials to understand and use the materials provided.

A visual example of split attention

Consider the graphic below from Tarmizi and Sweller (1988). They used these graphics to compare the learning that takes place given split attention conditions. Each is a possibility of how one might arrange graphical material within a lesson. Ward and Sweller (1990) advise Instructional designers to be careful when they direct a learner's attention. Sweller and his associates found that learners had difficulty following some worked examples if they included diagrams that were separated from their formulas. In several studies and over a variety of experiments, they found that learners using integrated diagrams were better able to process that information, and significantly improved their performance relative to their peers (Ward & Sweller, 1990; Chandler & Sweller, 1991; Chandler & Sweller, 1992).

The split-attention effect is not limited to geometry, Chandler and Sweller (1991) found that this effect extends to a variety of other disciplines and is simply a limitation of human information processing. This overload is the result of high visual cognitive load and due to poor instructional design. By simply integrating formulas with diagrams, learners found it easier to integrate and process both forms of visual information and in turn they performed significantly better (Chandler & Sweller, 1991; Chandler & Sweller, 1992).

The left example produces split attention, however the right example enhances learning because it guides the learner's attention through the worked example. The split-attention effect is an important form of extraneous cognitive load that instructional designers should avoid.
Visual split-attention

Chandler and Sweller (1992) found through empirical study that the integration of text and diagrams reduces cognitive load and facilitates learning. They found that this effect is evident, when learners are required to split their attention between different sources of information (e.g., text and diagrams).

Split attention is important evidence of the cognitive load theory (that the working memory load of instructional materials is important in the design of instructional materials). Chandler and Sweller (1992) found that students viewing integrated instruction spent less time processing the materials but still outperformed students in the split attention condition.

Auditory split attention

Moreno and Mayer (2000) found evidence for auditory split attention when they tested learners with both ambient environmental sounds and music as they learned from instructional materials. Animation is processed in a visual channel but must be converted to the auditory channel. The extraneous cognitive load imposed by music or environmental sounds were not conducive to learning.

These researchers studied learners that learned within an environment that had additional (extraneous) environmental sounds or music and found that learners performed significantly poorer on retention and transfer tests (Moreno, 2001 [1]).

References


References

Worked-example effect

The **worked-example effect** is a learning effect predicted by cognitive load theory (Sweller, 1988). According to Sweller: “The worked example effect is the best known and most widely studied of the cognitive load effects” (Sweller, 2006, p. 165).

**What is a worked example?**

"A **worked example** is a step-by-step demonstration of how to perform a task or how to solve a problem" (Clark, Nguyen, Sweller, 2006, p. 190). Studying worked examples is an effective instructional strategy to teach complex problem-solving skills (van Merriënboer, 1997). This is because example-based instruction provides expert mental models, to explain the steps of a solution for novices.

**Evidence for the worked-example effect**

Sweller and Cooper were not the first to use this form of instruction, but certainly they were the first to describe it from a cognitive load perspective (Sweller & Cooper, 1985; Cooper & Sweller, 1987; Sweller, 1988).

While studying problem-solving tactics, Sweller and Cooper used worked examples as a substitute for conventional problem-solving for those learning algebra. They found learners that studied worked examples, performed significantly better than learners who actively solved problems (Sweller & Cooper, 1985; Cooper & Sweller, 1987). Sweller and Cooper (1985) had developed worked examples as a means of limiting problem solving search.
Developing effective worked examples

Ward and Sweller (1990) suggested that under some conditions "worked examples are no more effective, and possibly less effective, than solving problems" (p. 1). Thus it is important that worked examples be structured effectively, so that extraneous cognitive load does not impact learners. Chandler and Sweller (1992) suggested an important way to structure worked examples. They found that the integration of text and diagrams (within worked examples) reduces extraneous cognitive load. They referred to this single modality, attention learning effect as the split-attention effect (Chandler and Sweller, 1992). Tabbers, Martens, & Van Merrienboer (2000) proposed one may prevent split-attention by presenting text as audio.

Not all worked examples are print-based as those in the Tarmizi and Sweller study. Lewis (2005) proposed animated demonstrations are a form of worked example. Animated demonstrations are useful because this multimedia presentation combines the worked example, and modality effects within a single instructional strategy.

For Which Learners?

As it turns out, worked examples are not appropriate for all learners. Learners with prior knowledge of the subject find this form of instruction redundant, and may suffer the consequences of this redundancy. This has been described as the expertise reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003). It is suggested that worked examples be faded over time to be replaced with problems for practice (Renkl, Atkinson & Maier, 2000). Thus it is important to consider the learner as well as the media while developing worked examples, else learners may not perform as expected.

References

- Sweller, J. (2006). The worked example effect and human cognition. Learning and Instruction, 16(2) 165–169.

References

Expertise reversal effect

The expertise reversal effect is related to cognitive load theories of learning. It states that guidance provided to experts can hinder their ability to learn: "learners would have to relate and reconcile the related components of available long-term memory base and externally provided guidance. Such integration processes may impose an additional working memory load and reduce resources available for learning new knowledge."[1] The expertise reversal effect is a form of redundancy effect whereby added information places an additional load on working memory without providing any useful schema to direct executive function. The expertise reversal effect differs from the redundancy effect in that in the expertise reversal effect, "external information becomes redundant relative to a particular learner's internal knowledge structures," whereas in the redundancy effect, "only different external sources of information" are sources of cognitive load.[1]

Notes
Image Sources, Licenses and Contributors


