



Conceptual short term memory in perception and thought

Mary C. Potter *

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Edited by:

Arnon Cahen, Ben Gurion University
in the Negev, Israel
Michela C. Tacca, Heinrich-Heine
University Düsseldorf, Germany

Reviewed by:

Mark Nieuwenstein, University of
Groningen, Netherlands
Candice Coker Morey,
Rijksuniversiteit Groningen,
Netherlands

***Correspondence:**

Mary C. Potter, Department of Brain
and Cognitive Sciences, 46-4125,
Massachusetts Institute of
Technology, 77 Massachusetts
Avenue, Cambridge, MA 02139, USA.
e-mail: mpotter@mit.edu

Conceptual short term memory (CSTM) is a theoretical construct that provides one answer to the question of how perceptual and conceptual processes are related. CSTM is a mental buffer and processor in which current perceptual stimuli and their associated concepts from long term memory (LTM) are represented briefly, allowing meaningful patterns or structures to be identified (Potter, 1993, 1999, 2009). CSTM is different from and complementary to other proposed forms of working memory: it is engaged extremely rapidly, has a large but ill-defined capacity, is largely unconscious, and is the basis for the unreflective understanding that is characteristic of everyday experience. The key idea behind CSTM is that most cognitive processing occurs without review or rehearsal of material in standard working memory and with little or no conscious reasoning. When one perceives a meaningful stimulus such as a word, picture, or object, it is rapidly identified at a conceptual level and in turn activates associated information from LTM. New links among concurrently active concepts are formed in CSTM, shaped by parsing mechanisms of language or grouping principles in scene perception and by higher-level knowledge and current goals. The resulting structure represents the gist of a picture or the meaning of a sentence, and it is this structure that we are conscious of and that can be maintained in standard working memory and consolidated into LTM. Momentarily activated information that is not incorporated into such structures either never becomes conscious or is rapidly forgotten. This whole cycle – identification of perceptual stimuli, memory recruitment, structuring, consolidation in LTM, and forgetting of non-structured material – may occur in less than 1 s when viewing a pictured scene or reading a sentence. The evidence for such a process is reviewed and its implications for the relation of perception and cognition are discussed.

Keywords: CSTM, RSVP, scene perception, reading, short term memory, attentional blink

INTRODUCTION: CONCEPTUAL SHORT TERM MEMORY

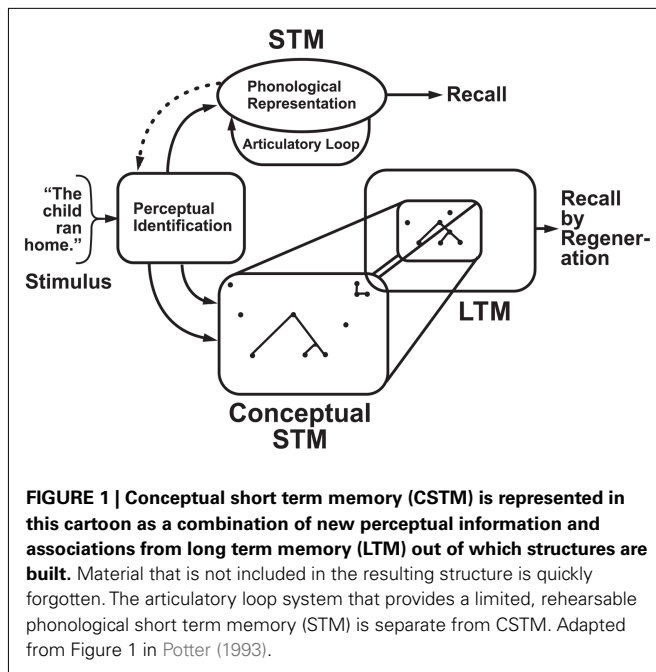
Conceptual short term memory (CSTM) is a construct based on the observation that most cognitive processing occurs without review or rehearsal of material in standard working memory and with little or no conscious reasoning. CSTM proposes that when one perceives a meaningful stimulus such as a word, picture, or object, it is rapidly identified and in turn activates associated information from long term memory (LTM). New links among concurrently active concepts are formed in CSTM, shaped by parsing mechanisms of language or grouping principles in scene perception, and by higher-level knowledge and current goals. The resulting structure is conscious and represents one's understanding of the gist of a picture or the meaning of a sentence. This structured representation is consolidated into LTM if time permits. Momentarily activated information that is not incorporated into such structures either never becomes conscious or is rapidly forgotten. **Figure 1** shows a cartoon of CSTM in relation to LTM and one component of conventional STM.

CSTM IN RELATION TO OTHER MEMORY SYSTEMS AND OTHER MODELS

Conceptual short term memory is a processing and memory system that differs from other forms of short term memory. In vision, *iconic memory* (Sperling, 1960) maintains a detailed visual

representation for up to about 300 ms, but it is eliminated by new visual stimulation. Meaning plays little or no role. *Visual short term memory* (VSTM) holds a limited amount of visual information (about four items' worth) and is somewhat resistant to interference from new stimulation as long as the information is attended to (Coltheart, 1983; Phillips, 1983; Luck and Vogel, 1997; Potter and Jiang, 2009). Although VSTM is more abstract than perception in that the viewer does not mistake it for concurrent perception, it maintains information about many characteristics of visual perception, including spatial layout, shape, color, and size. In audition, the *phonological loop* (Baddeley, 1986) holds a limited amount (about 2 s worth) of recently heard or internally generated auditory information, and this sequence can be maintained as long as the items are rehearsed (see **Figure 1**).

Conceptual short term memory differs from these other memory systems in one or more ways: in CSTM, new stimuli are rapidly categorized at a meaningful level, associated material in LTM is quickly activated, this information is rapidly structured, and information that is not structured or otherwise consolidated is quickly forgotten (or never reaches awareness). In contrast, standard *working memory*, such as Baddeley's articulatory/phonological loop and visuospatial sketchpad together with a central executive (Baddeley, 1986, 2007), focuses on memory systems that support cognitive processes that take place over several seconds or minutes. A



memory system such as the phonological loop is unsuited for conceptual processing that takes place within a second of the onset of a stream of stimuli: it takes too long to be set up and it does not represent semantic and conceptual information directly. Instead, Baddeley's working memory directly represents articulatory and phonological information or visuospatial properties: these representations must be reinterpreted conceptually before further meaning-based processing can occur.

More recently, Baddeley (2000) proposed an additional system, the *episodic buffer*, that represents conceptual information and may be used in language processing. The episodic buffer is "a temporary store of limited capacity. . . capable of combining a range of different storage dimensions, allowing it to collate information from perception, from the visuo-spatial and verbal subsystems and LTM. . . representing them as multidimensional chunks or episodes. . ." (Baddeley and Hitch, 2010). Baddeley notes that this idea is similar to CSTM as it was described in 1993 (Potter, 1993).

Although Baddeley's multi-system model of working memory has become the dominant model of short term memory, it neglects the evidence that stimuli in almost any cognitive task rapidly activate a large amount of potentially pertinent information, followed by rapid selection and then decay or deactivation of the rest. That can happen an order of magnitude faster than the setting up of a standard, rehearsable STM representation, permitting the seemingly effortless processing of experience that is typical of cognition. Of course, not all cognitive processing is effortless: our ability to engage in slower, more effortful reasoning, recollection, and planning may well draw on conventional short term memory representations.

Relation to other cognitive models

Many models of cognition include some form of processing that relies on persistent activation or memory buffers other than standard working memory, tailored to the particular task

being modeled. CSTM may be regarded as a generalized capacity for rapid abstraction, pattern recognition, and inference that is embodied in a more specific form in models such as ACT-R (e.g., Budson and Anderson, 2004), the construction-integration model of discourse comprehension (Kintsch, 1988), the theory of long term working memory (Ericsson and Kintsch, 1995); and models of reading comprehension (e.g., Just and Carpenter, 1992; see Potter et al., 1980; Verhoeven and Perfetti, 2008).

EVIDENCE FOR CSTM

RAPID SERIAL VISUAL PROCESSING AS A METHOD FOR STUDYING CSTM

The working of CSTM is most readily revealed when processing time is limited. A simple way to limit processing time for visual stimuli is to use a single visual mask, immediately after a brief presentation. The idea is to replace or interfere with continued processing of the stimulus by presenting a new image that will occupy the same visual processors. Backward masks, if they share many of the same features (contours, colors, and the like) as the target stimulus, do produce interference and may prevent perception or continued processing of the target. Much has been learned about how we perceive, using backward masking. However, for complex stimuli such as pictures or written words, conceptual processing may continue despite the mask; to interfere with understanding or memory for the target, the mask itself must engage conceptual processing that will interfere with that of the target. An effective way to create such a limitation is to present a rapid sequence of visual stimuli, termed rapid serial visual processing (RSVP) by Forster (1970). By using RSVP in which all the stimuli (pictures or words) are meaningful and need to be attended, one can obtain a better measure of the actual processing time required for an individual stimulus or for the sequence as a whole (Potter, 1976; Intraub, 1984; Loftus and Ginn, 1984; Loftus et al., 1988). This method was used in many of the studies cited in the present review.

SUMMARY OF THE EVIDENCE

The evidence is summarized here before presenting some of it in more detail. Three interrelated phenomena give evidence for CSTM:

- (1) There is rapid access to conceptual (semantic) information about a stimulus and its associations. Conceptual information about a word or a picture is available within 100–300 ms, as shown by experiments using semantic priming (Neely, 1991), including masked priming (Forster and Davis, 1984) and so-called fast priming (Sereno and Rayner, 1992); eye tracking when reading (Rayner, 1983, 1992) or looking at pictures (Loftus, 1983); measurement of event-related potentials during reading (Kutas and Hillyard, 1980; Luck et al., 1996); and target detection in RSVP with letters and digits (Sperling et al., 1971; Chun and Potter, 1995), with pictures (Potter, 1976; Intraub, 1981; Potter et al., 2010), or with words (Lawrence, 1971b; Potter et al., 2002; Davenport and Potter, 2005). To detect a target such as an animal name in a stream of words, the target must first be identified (e.g., as the word *tiger*) and then matched to the target category, an animal name (e.g., Meng and Potter, 2011). Conceptually defined targets can

be detected in a stream of non-targets presented at rates of 8–10 items/s or higher, showing that categorical information about a written word or picture is activated and then selected extremely rapidly. These and other experimental procedures show that semantic or conceptual characteristics of a stimulus have an effect on performance as early as 100 ms after its onset. This time course is too rapid to allow participation by slower cognitive processes, such as intentional encoding, deliberation, or serial comparison in working memory.

- (2) New structures can be discovered or built out of the activated conceptual information, influenced by the observer's task or goal. Activated conceptual information can be used to discover or build a structured representation of the information, or (in a search task) to select certain stimuli at the expense of others. A major source of evidence for these claims comes from studies using RSVP sentences, compared with scrambled sentences or lists of unrelated words. Studies by Forster (1970), Potter (1984, 1993), Potter et al. (1980), and Potter et al. (1986) show that it is possible to process the structure in a sentence and hence to recall it subsequently, when reading at a rate such as 12 words/s. In contrast, when short lists of unrelated words are presented at that rate, only two or three words can be recalled (see also Lawrence, 1971a). For sentences, not only the syntactic structure, but also the meaning and plausibility of the sentence is recovered as the sentence is processed (Potter et al., 1986). Because almost all sentences one normally encounters (and all the sentences in these experiments) include new combinations of ideas, structure-building is not simply a matter of locating a previously encountered pattern in LTM: it involves the instantiation of a new relationship among existing concepts. The same is true when viewing a new pictured scene: not only must critical objects and the setting be identified, but also the relations among them: the gist of the picture. Structure-building presumably takes advantage of as much old structure as possible, using any preexisting associations and chunks of information to bind elements (such as individual words in a list) together.
- (3) There is rapid forgetting of information that is not structured or that is not selected for further processing. Conceptual information is activated rapidly, but the initial activation is highly unstable and will be deactivated and forgotten within a few hundred milliseconds if it is not incorporated into a structure. As a structure is built – for example, as a sentence is being parsed and interpreted – the resulting interpretation can be held in memory and ultimately stabilized or consolidated in working or LTM as a unit, whereas only a small part of an unstructured sequence such as a string of unrelated words can be consolidated in the same time period.

UNDERSTANDING PICTURES AND SCENES

In studies in which unrelated photographs are presented in RSVP, viewers can readily detect a picture when given a brief descriptive title such as *wedding* or *two men talking*, at rates of presentation up to about 10 pictures/s, even though they have never seen that picture before and an infinite number of different pictures could fit the description (Potter, 1976; Intraub, 1981; see Figure 2 and Potter (2009) Demo 1 for a demonstration). As Figure 2 shows,

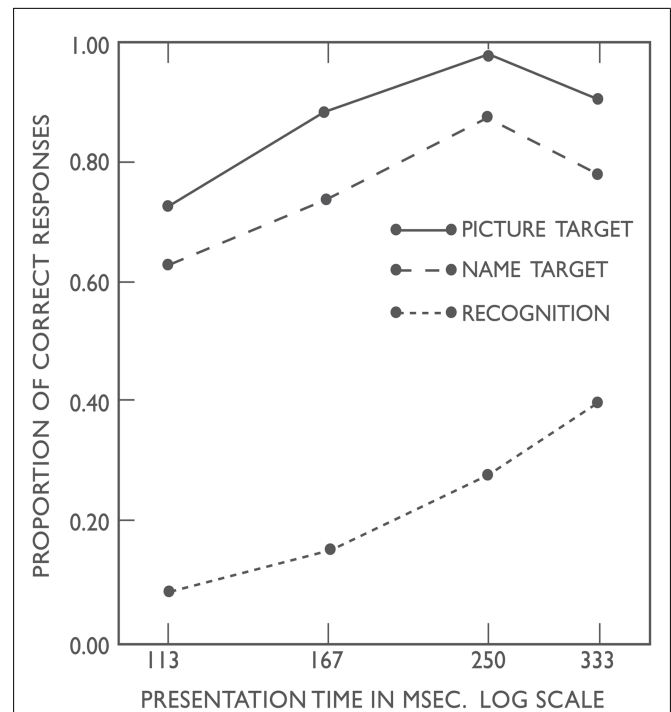
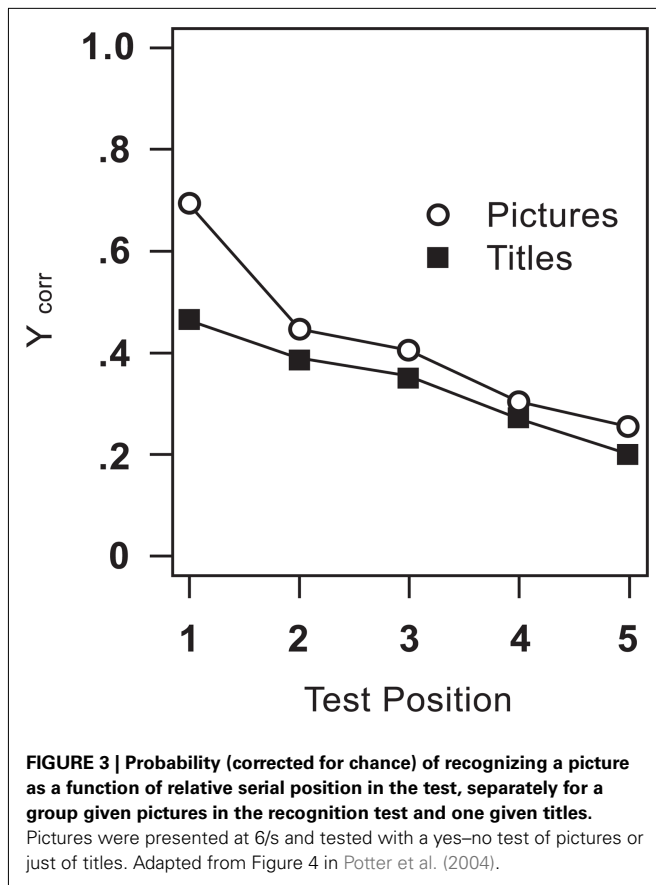


FIGURE 2 | Detection of a target picture in an RSVP sequence of 16 pictures, given a name or picture of the target, as a function of the presentation time per picture. Also shown is later recognition performance in a group that simply viewed the sequence, and then was tested for recognition. Results are corrected for guessing. Adapted from Figure 1 in Potter (1976).

detection is almost as accurate when given only a name, as when shown the target picture itself.

More recent research shows that viewers can detect named pictures in RSVP sequences at above-chance levels at still higher rates, even for durations as short as 13 ms (Potter, Wyble, and McCourt, in preparation). Evidently viewers can extract the conceptual gist of a picture rapidly, retrieving relevant conceptual information about objects and their background from LTM. Having spotted the target picture, viewers can continue to attend to it and consolidate it into working memory – for example, after the sequence they can describe the picnic scene they were looking for, or recognize it in a forced choice task. Yet, when they are not looking for a particular target, viewers forget most pictures presented at that rate almost immediately, as shown in Figure 2 (Potter and Levy, 1969; see also Intraub, 1980). The rate must be slowed to about 2 pictures/s for viewers to recognize as many as half the pictures as familiar when tested minutes after the sequence.

If a test picture is presented immediately after the sequence, however, viewers are usually able to recognize it, even if the pictures have been presented at a rate such as 6/s (Potter et al., 2004). That is, for a second they will remember most of the presented pictures, but memory drops off rapidly over the first few seconds thereafter, as memory is tested (Figure 3). Picture memory includes understanding of the gist of a picture, not just specific visual features, as shown by the ability of viewers to call a picture to mind when given a descriptive title as a recognition cue. Just as a viewer can



detect a picture that matches a target description given in advance, such as “baby reaching for a butterfly,” so can viewers recognize that they saw a picture that matches a name if they are given the name shortly *after* viewing the sequence. Recognition memory is not quite as good when tested by a title instead of showing the picture itself, however, and in both cases performance falls off rapidly (Figure 3). Thus, gist can be extracted rapidly, but is quickly forgotten if the presentation was brief and was followed by other stimuli.

Further evidence for the extraction of gist when viewing a rapid sequence of pictures is that viewers are more likely to falsely recognize a new picture with the same gist than a new picture with an unrelated gist. For example, if they saw a picture of a camel, they are more likely to falsely say yes to a very different picture of a camel than to a totally new picture (Figure 4), indicating that at some level they knew they had seen a camel (Potter et al., 2004). Other studies with single pictures presented briefly and masked by a following stimulus have shown that objects in the foreground are more easily recognized if they are consistent with the background (and the background is more readily recognized if it is consistent with the foreground), showing that relationships within a single picture are computed during initial recognition (Davenport and Potter, 2004; Davenport, 2007).

Conclusions: Pictures

Consistent with the CSTM hypothesis, the evidence shows that meaningful pictures can be understood extremely rapidly, permitting the detection of targets specified by name (i.e., by meaning)



FIGURE 4 | When participants has viewed the picture on the left and judged whether they had seen the picture on the right, they made more false yeses than when the new picture was unrelated in meaning to any pictures they had viewed. Adapted from Figure 7 in Potter et al. (2004).

and at least momentary understanding of the gist of a picture, although there is rapid forgetting if the picture is not selected for further processing.

UNDERSTANDING RSVP SENTENCES

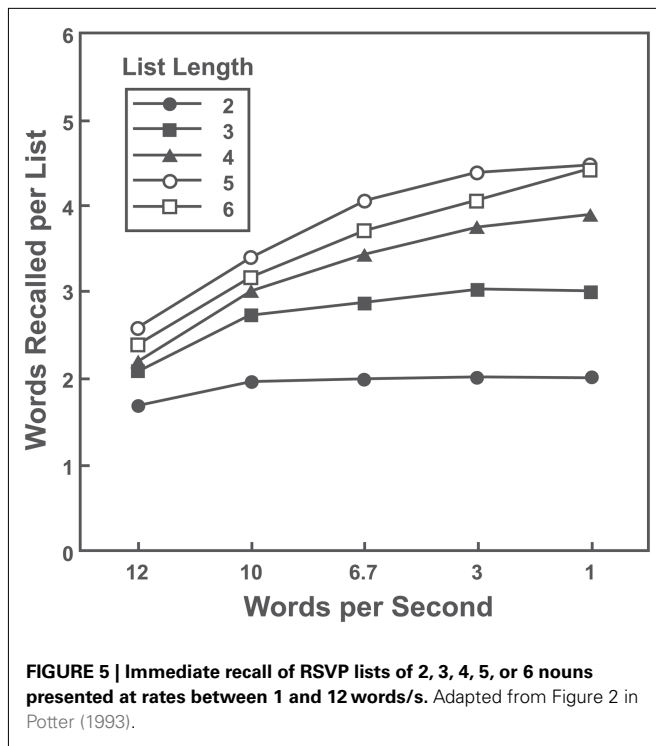
Our ability to read rapidly and continuously, with comprehension and substantial memory for the meaning of what we have read, is a strong indication that we are able to retrieve rapidly a lot of information not only about word meanings but also general knowledge about the world and specific episodic memories. Good readers can read about 300 words/min, or 5 words/s. Eye tracking studies have shown that the length of time that the eyes rest on a given word varies with the frequency of that word, its predictability at that point in the text, and other factors that indicate that the word’s meaning and its fit to the context is retrieved fast enough to affect whether the eyes linger on that word or move on (for a review, see Staub and Rayner, 2007).

Instead of measuring eye movements when reading, one can use RSVP to control the time available for processing each word in a sentence. The first to try this was Forster, who presented short sentences at 16 words/s, three times faster than a typical good reader would read spontaneously. He varied the syntactic complexity of the sentences and showed that recall was less accurate for more complex sentences, implying that sentence syntax was processed as sentences were read (Forster, 1970; Holmes and Forster, 1972).

As shown in studies reviewed above, searching for a specific target in an RSVP sequence can be easy even at high rates of presentation, whereas simply remembering all the items in the presentation can be difficult, if they are unrelated. A claim which is central to the CSTM hypothesis, however, is that associative and other structural relations among items can be computed rapidly, assisting in their retention. This section reviews some of the evidence for this claim.

Differences between lists and sentences

A person’s memory span is defined as the number of items, such as unrelated words or random digits, that one can repeat back accurately, after hearing or seeing them presented at a rate of about 1/s. For unrelated words a typical memory span is five or six words. The memory span drops, however, as the rate of presentation increases. In one experiment (Potter, 1982) lists of 2, 3, 4, 5, or 6 nouns were presented at rates between 1 and 12 word/s, for immediate recall. The results are shown in Figure 5.



With five words, recall accuracy declined from a mean of 4.5 at the 1 s rate to 2.6 words at the rate of 12 words/s. This was evidently not because participants could not recognize each of the words at that rate, because a list of just two words (followed by a mask) was recalled almost perfectly at 12/s: instead, some additional process is necessary to stabilize the words in short term memory. Note that this finding is similar to that for memory of rapidly presented pictures, in that one can detect a picture at a rate that is much higher than the rate required for later memory. In another study I found that the presentation of two related words on a five-word RSVP list (separated by another word) resulted in improved recall for both words, suggesting that the words were both activated to a level at which an association could be retrieved. This hinted at the sort of process that might stabilize or structure information in CSTM.

In contrast to lists, 14-word sentences presented at rates up to at least 12 words/s can be recalled quite accurately (see Potter, 1984; Potter et al., 1986). The findings with sentences versus lists or scrambled sentences strongly support the CSTM assumption that each word can be identified and understood with an 83- to 100-ms exposure, even when it is part of a continuing series of words. (See Potter, 2009, Demo 2, for a demonstration.) The results also support the second assumption that representations of the words remain activated long enough to allow them to be bound into whatever syntactic and conceptual structures are being built on the fly. When, as with a list of unrelated words, there is no ready structure to hand, all but two or three of the words are lost.

How are RSVP sentences recalled? The regeneration hypothesis

Before addressing the question of how rapidly presented sentences are retained, one should address the prior question of why

sentences heard or read at *normal* rates are easy to repeat immediately, even when they are two or three times as long as one's memory span (the length of list that can be repeated accurately). The difference in capacity between lists and sentences is thought to be due to some form of chunking, although it has also been assumed that sentences can be stored in some verbatim form temporarily (see the review by Von Eckardt and Potter, 1985). Before continuing, try reading the sentence below once, cover it, and then read the five words on the next line, look away, and write down the sentence from memory. (We will come back to this exercise shortly.)

The knight rode around the palace looking for a place to enter.

Anchor forest castle oven stocking [look away and write down the sentence].

Instead of assuming that people remember sentences well because they hold them in some verbatim form, we (Potter and Lombardi, 1990) proposed a different hypothesis: immediate recall of a sentence (like longer-term recall) is based on a conceptual or propositional representation of the sentence. The recaller regenerates the sentence using normal speech-production processes to express the propositional structure (what the sentence means). That is, having understood the conceptual proposition in a sentence, one can simply express that idea in words, as one might express a new thought. We proposed that recently activated words were likely to be selected to express the structure. In consequence, the sentence is normally recalled verbatim.

To test this hypothesis we (Potter and Lombardi, 1990) presented distractor words (like the five words below the sentence you read above) in a secondary task immediately before or after the to-be-recalled sentence, and on some trials one of these distractor words was a good substitute for a word in the sentence (such as "castle" for "palace"). As we predicted, that word was frequently intruded in recall, as long as the sentence meaning as a whole was consistent with the substitution. (Did you substitute "castle" for "palace"?) In the experiments, half the participants had lure words like "castle" on the word list, and half did not, allowing us to show that people are more likely to make the substitution when that word has appeared recently. Thus, recall is guided by a conceptual representation, not by a special verbatim representation such as a phonological representation.

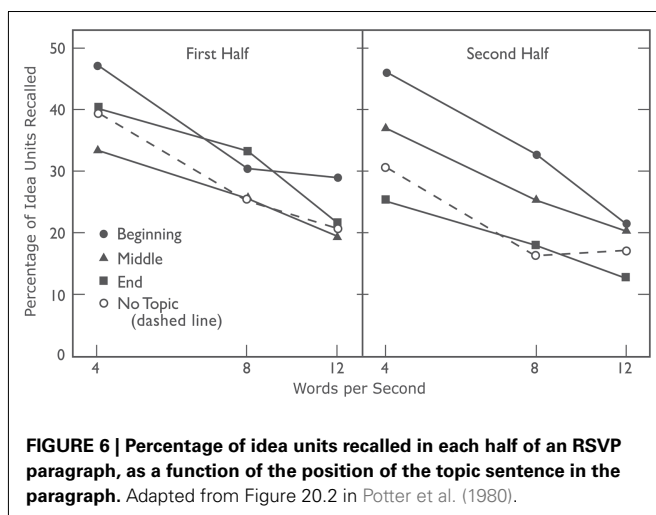
Further studies (Lombardi and Potter, 1992; Potter and Lombardi, 1998) indicated that syntactic priming from having processed the sentence plays a role in the syntactic accuracy of sentence recall. Syntactic priming (e.g., Bock, 1986) is a temporary facilitation in the production of a recently processed syntactic structure, as distinguished from direct memory for the syntactic structure of the prime sentence.

The Potter–Lombardi hypothesis that sentences are comprehended and then regenerated rather than "recalled verbatim" is consistent with the CSTM claim that propositional structures are built rapidly, as a sentence is read or heard. In one of the Potter and Lombardi (1990) experiments the sentences were presented at a rate of 10 words/s, rather than the moderate 5 words/s of our other experiments: the intrusion results were essentially the same, showing that the relevant conceptual processing had occurred at the higher rate, also.

Reading RSVP paragraphs: More evidence for immediate use of structure

A single RSVP sentence is apparently easy to comprehend and recall when presented as fast as 12 words/s, so that recall of a single sentence at that rate is close to ceiling. Does that mean that longer-term retention of the sentence will be as good as if it had been presented more slowly? To answer that question Potter et al. (1980) presented RSVP paragraphs of about 100 words at three rates, 4, 8, and 12 words/s, with the equivalent of a two-word pause between sentences (the net rate averaged 3.3, 6.7, or 10 words/s). (See Demo 3 in Potter, 2009, for a demonstration.) Participants wrote down the paragraph as accurately as possible, immediately after presentation. To evaluate both single-word perception and use of discourse-level information, paragraphs were used that appeared to be ambiguous and poorly integrated unless the reader knew the topic (see Dooling and Lachman, 1971; Bransford and Johnson, 1972). Only one sentence mentioned the one-word topic (e.g., “laundry”), and this sentence appeared either at the beginning, the middle, or the end of the paragraph, or was omitted.

Consistent with the prediction that knowing the topic would enhance recall of the other sentences (and that the topic word itself would be recalled because of its relevance to the paragraph), recall was improved for the part of the paragraph after the topic was presented (but not the part before), at all three rates of presentation. Even at the highest rate the discourse topic could be used to structure the paragraph (Figure 6). This suggests that the discourse topic, once it became evident, remained active as a source of structure as the rest of the paragraph was read. (The topic word was perceived and recalled by more than 80% of the subjects regardless of rate or condition.) At the same time, there was a marked main effect of rate: recall declined as rate of presentation increased, from 37% of the idea units at 4 words/s to 26% at 8 words/s to 20% at 12 words/s, averaging over all topic conditions. Clearly, even though there was internal evidence that discourse-level structuring was occurring at all rates of presentation, some process of consolidation was beginning to fail as rate increased.



Conclusions: words, sentences, and paragraphs.

Putting the paragraph results together with those for word lists and single sentences, we see that structuring can occur rapidly, and that more structure results in better memory (comparing lists to sentences, or comparing a string of seemingly unrelated sentences to sentences structured by having a topic). Nonetheless, rapid conceptual processing is not sufficient for accurate retention if there is no additional time for consolidation: the gist may survive, but details will be lost in immediate recall, just as they are in longer-term memory.

MECHANISMS OF STRUCTURING IN RSVP SENTENCE PROCESSING

Although I have repeatedly invoked the idea that there is rapid structuring of information that is represented in CSTM, I have had little to say about just how this structuring occurs. In the case of sentences, it is evident that parsing and conceptual interpretation must occur virtually word by word, because any substantial delay would outrun the persistence of unstructured material in CSTM. Here I will describe briefly three studies that have investigated the process of selecting an appropriate interpretation of a given word in an RSVP sentence, a key process in comprehension given the extent of lexical ambiguity in English and in most other languages.

The influence of sentence context on word and non-word perception

One study (Potter et al., 1993) took advantage of the propensity of readers to misperceive a briefly presented non-word as an orthographically similar word. Non-words such as *dack* that are one letter away from two other words (*deck*, *duck*) were presented in RSVP sentences biased-toward one or the other of these words or neutral between them, as in the following examples. Note that when we presented a real word in the biased sentences, it was always the mismatching word. The sentence was presented at 10 words/s. Participants recalled the sentence; they were told to report misspelled words or non-words if they saw them.

Neutral: “The visitors noticed the *deck/duck/dack* by the house”

Biased: “The child fed the *deck/dack* at the pond”

“The sailor washed the *duck/dack* of that vessel.”

Our main interest is what participants reported when shown a non-word. Non-words were reported as the biased-toward word (here, *duck*) on 40% of the trials, compared with only 12% with the neutral sentence and 3% with the biased-against sentence (the non-words were reported correctly as non-words on 23% of the neutral trials). Similar although smaller effects of context were shown when the biased-against word (rather than a non-word) appeared in the sentence. Thus, context can bias word and non-word perception even when reading at 10 words/s. More surprisingly, even selective context that did not appear until as much as three words (300 ms) after the critical word or non-word influenced perceptual report, suggesting that multiple word candidates (and their meanings) are activated as the non-word or word is perceived, and may remain available for selection for at least 300 ms after the word or non-word has been masked by succeeding words. This supposition that multiple possible words and their meanings are briefly activated during word perception accords with the Swinney (1979) hypothesis that multiple meanings of ambiguous

words are briefly activated: both results are consistent with the CSTM view. In the present study and in the case of ambiguous words the process of activation and selection appears to occur unconsciously for the most part, an issue considered in a later section.

Double-word selection

In another study (Potter et al., 1998) two orthographically distinct words were presented simultaneously (one above and one below a row of asterisks) in the course of an RSVP sentence, as illustrated below. Participants were instructed to select the one that fit into the sentence and include it when immediately recalling the sentence. We regarded this as an overt analog of lexical ambiguity resolution. The sentence was presented for 133 ms/word and the two-word (“double word”) display for 83 ms.

Maggie carried the kitten in a

pencil

basket

 to her house

Sentence context had a massive influence on selection: the appropriate word was included in recall in 70% of the sentences and the non-matching word in only 13%. This ability to pick the right word was evident both when the relevant context arrived before the double words and when it arrived later (up to 1 s later, in one experiment), showing that readers could activate and maintain two distinct lexical possibilities. Subjects were asked to report the “other” word (the mismatching word) after they recalled the sentence, but most of the time they were unable to do so, showing that the unselected word was usually forgotten rapidly. Again, this illustrates the existence of fast and powerful processors that can build syntactically and pragmatically appropriate structures from briefly activated material, leaving unselected material to be rapidly forgotten.

Lexical disambiguation

Miyake et al. (1994) carried out two experiments on self-paced reading of sentences with ambiguous words that were not disambiguated until two to seven words after the ambiguous word. They found that readers with low or middling reading spans were slowed down when the disambiguation was toward the subordinate meaning, especially with a delay of seven words. (High-span readers had no problems in any of the conditions.) In an unpublished experiment we presented subjects with a similar set of sentences that included an ambiguous word, using RSVP at 10 words/s; the task was to decide whether or not the sentence was plausible, after which we gave a recognition test of a subset of the words, including the ambiguous word. The sentence was implausible with one of the meanings, plausible with the other. Our hypothesis was that sentences that eventually turned out to require the subordinate meaning of an ambiguous word would sometimes be judged to be implausible, implying that only the dominant reading had been retrieved. Unambiguously implausible and plausible sentences were intermixed with the ambiguous sentences.

Subjects were more likely to judge a plausible sentence to have been implausible (1) when a subordinate meaning of the ambiguous word was required (27 versus 11% errors), (2) when the disambiguating information appeared after a greater delay (23

versus 16% errors), and especially (3) when there was both a subordinate meaning and late disambiguation (32% errors, versus 9% for the dominant/early condition). A mistaken judgment that the sentence was implausible suggests that on those trials only one meaning, the wrong one, was still available at the point of disambiguation. Interestingly, the ambiguous word itself was almost always correctly recognized on a recognition test of a subset of words from the sentence, even when the sentence had mistakenly been judged implausible. The results suggest that although multiple meanings of a word are indeed briefly activated (in CSTM), the less frequent meaning will sometimes fall below threshold within a second, when sentences are presented rapidly.

Conclusion: lexical interpretation and disambiguation

The results of all three studies show that context is used immediately to bias the perception or interpretation of a word, consistent with the CSTM claim that processing to the level of meaning occurs very rapidly in reading.

SELECTIVE SEARCH AND THE ATTENTIONAL BLINK

The attentional blink

In brief, the attentional blink (AB) is a phenomenon that occurs in RSVP search tasks in which two targets are presented among distractors. When the rate of presentation is high but still compatible with accurate report of a single target (e.g., a presentation rate of 10/s, when the task is to detect letters among digit distractors), two targets are also likely to be reported accurately except when the second target appears within 200–500 ms of the onset of the first target. This interval during which second target detection drops dramatically was termed an AB by Raymond et al. (1992).

The AB is relevant to CSTM because it provides evidence for rapid access to categorical information about rapidly presented items and at the same time shows that selective processing of specified targets has a cost. When the task is to pick out targets from among distractors, the experimental findings suggest that there is a difference in time course between two stages of processing, a first stage that results in identification of a stimulus (CSTM) and a second stage required to select and consolidate that information in a reportable form (Chun and Potter, 1995).

Consider a task in which targets are any letter of the alphabet, presented in an RSVP sequence of digit distractors. Presumably a target letter must be identified as a specific letter in order to be classified as a target (see Sperling et al., 1971). At rates as high as 11 items/s the first letter target (T1) is detected quite accurately, consistent with evidence that a letter can be identified in less than 100 ms. This initial identification is termed Stage 1 processing, which constitutes activation of a conceptual but short-lasting representation, i.e., a CSTM representation.

But a second target letter (T2) that arrives soon after the first one is likely to be missed, suggesting that a selected target (T1) requires additional processing beyond identification: Stage 2 processing. Stage 2 processing is necessary to consolidate a selected item into some form of short term memory that is more stable than CSTM. However, Stage 2 processing is serial and limited in capacity. The items following the first target (T1) continue to be processed successfully in Stage 1 and remain for a short time in CSTM; the problem is that as long as Stage 2 is tied up with T1,

a second target may be identified but must wait, and thus may be lost from CSTM before Stage 2 is available. When this happens, T2 is missed, producing an AB. Although the duration of the AB varies, it is strong at 200–300 ms after the onset of the first target, diminishes thereafter, and is usually gone by 500 ms.

Consistent with the CSTM hypothesis, there is both behavioral and ERP evidence that stimuli that are not reported because of an AB are nonetheless momentarily comprehended, because they activate an ERP mismatch marker when they are inconsistent with prior context (Luck et al., 1996). Similarly, word targets that are related in meaning are more accurately detected even when the second word occurs within the time period that produces an AB (e.g., Chun et al., 1994; Maki et al., 1997; Vogel et al., 1998; Potter et al., 2005).

Attention only blinks for selection, not perception or memory

As shown in earlier sections, meaningful items in a continuous stream, such as the words of a sentence, are easy to see and remember, which makes the difficulty of reporting a second target (the AB) surprising. When there is an uninterrupted sequence of targets, as happens when a sentence is presented and recalled as a whole, there is no AB, whereas if the task is to report just the two words in a sentence that are marked by color or by case, as in the following example, there is an AB (Potter et al., 2008).

Our tabby cat CHASED a MOUSE all around the backyard

In a more recent study (Potter et al., 2011), participants did both tasks simultaneously: they reported the sentence and then marked the red (or uppercase) words. In another block they only reported the two target words. An AB for marking or reporting the second target word was observed in both blocks. Surprisingly, the target words were highly likely to be reported as part of the sentence even when the participant could not mark them correctly. What seemed to happen was that the feature (color or case) that defined the target was detected, but in the AB interval that feature was displaced to a different word: the AB interfered with the binding of the target feature to the correct word.

In subsequent experiments the targets (Arabic digits or digit words) were inserted between words of the sentence as additional items, and again participants either reported the digits and then the surrounding sentence, or just reported the digits. The following is an example:

Our 6666 tabby cat 2222 chased a mouse all around the backyard

When participants reported just the digits, they were overall more accurate, but showed an AB for the second digit string or number word. When they reported the digits and then the whole sentence, they did not show an AB for the digits. Evidently the continuous attention associated with processing the sentence included the inserted digits, allowing them to be selected afterward rather than during initial processing. We concluded that on-line, immediate selection generates an AB, whereas continuous processing with delayed selection does not (cf. Wyble et al., 2011).

Summary: CSTM and the attentional blink

Studies of the visual AB demonstrate a dissociation between an early stage of processing sufficient to identify letters or words

presented at a rate of about 10/s, and a subsequent stage of variable duration (up to about 400 ms) required to stabilize a selected item in reportable STM. The AB thus provides evidence for the central claims of CSTM.

FURTHER QUESTIONS ABOUT CSTM

HOW DOES STRUCTURING OCCUR IN CSTM?

Structuring in CSTM is not different in principle from individual steps in the slower processes of comprehension that happen as one gradually understands a difficult text or an initially confusing picture, or solves a chess problem over a period of seconds and minutes. But CSTM structuring occurs with a relative absence of awareness that alternatives have been weighed and that several possibilities have been considered and rejected, at least implicitly. As in slower and more conscious pattern recognition and problem solving, a viewer's task set or goal makes a major difference in what happens in CSTM, because one's intentions activate processing routines such as sentence parsing, target specifications in search tasks, and the like. Thus the goal partially determines what enters CSTM and how structuring takes place.

Conceptual short term memory is built on experience and learned skills. Seemingly immediate understanding is more likely for material that is familiar, as becomes evident when one learns a new language or a new skill such as chess. Our ability to understand a new pictured scene in a fraction of a second also depends on our lifetime of visual experience.

Compound cuing and latent semantic analysis

The presence of many activated items at any moment, in CSTM, allows for compound cuing – the convergence of two or more weak associations on an item. The visual system is built on converging information, with learning playing a major role, at least at higher levels in the visual system (e.g., DiCarlo et al., 2012), enabling familiar combinations of features to converge on an interpretation in a single forward pass. The power of converging cues, familiar to any crossword puzzle fan, is likely to be central to structure-building in CSTM. A radical proposal for the acquisition and representation of knowledge, latent semantic analysis (LSA; Landauer and Dumais, 1997), provides a suggestive model for how structure may be extracted from loosely related material. However, there is no syntactic parser in LSA and it is clear from RSVP research that we do parse rapidly presented sentences as we read; thus, the LSA approach is at best a partial model of processing in CSTM.

IS CSTM CONSCIOUS?

The question is difficult to answer, because we have no clear independent criterion for consciousness other than availability for report. And, by hypothesis, report requires some form of consolidation; therefore, only what persists in a structured form will be reportable. Thus, while the evidence we have reviewed demonstrates that there is conceptual processing of material that is subsequently forgotten, it does not tell us whether we were briefly conscious of that material, or whether the activation and selection occurred unconsciously.

It seems unlikely that multiple competing concepts (such as the multiple meanings of a word) that become active simultaneously could all be conscious in the ordinary sense, although

preliminary structures or interpretations that are quickly discarded might be conscious. For example, people do sometimes become aware of having momentarily considered an interpretation of a spoken word that turns out to be mistaken. As noted earlier, most words are ambiguous, yet we are only rarely conscious of multiple meanings (except when someone makes a pun). In viewing rapid pictures, people have a sense of recognizing all the pictures but forgetting most of them. But such experiences in which we are aware of momentary thoughts that were immediately lost seem to be the exception, rather than the rule. Thus, much of CSTM activation, selection, and structuring happens before one becomes aware. It is the structured result, typically, of which one is aware, which is why perception and cognition seem so effortless and accurate.

THE LIMITS OF CSTM

Conceptual short term memory is required to explain the human ability to understand and act rapidly, accurately, and seemingly effortlessly in response to the presentation of richly structured sensory input, drawing on appropriate knowledge from LTM. Working memory as it is generally understood (e.g., Baddeley and Hitch, 2010) comes into play when a first pass in CSTM does not meet one's goal. Then, more conscious thought is required, drawing on working memory together with continued CSTM processing. Systematic reasoning, problem solving, recollection, and planning are slower and more effortful, however. They typically involve a series of steps, each of which sets up the context for the next step. CSTM may be involved in each step.

SUMMARY: RAPID CONCEPTUAL PROCESSING FOLLOWED BY RAPID FORGETTING

In each of the experimental domains discussed – comprehension and retention of RSVP word lists, sentences, and paragraphs; studies of word perception and selection; experiments on picture perception and memory; and the AB – there is evidence for activation of conceptual information about a stimulus early in processing (possibly before conscious awareness), followed by rapid forgetting unless conditions are favorable for retention. The two kinds of favorable conditions examined in these studies were selection for attention (e.g., the first target in the AB procedure, or selection of a target picture from among rapidly presented pictures) and the availability of associations or meaningful relations between momentarily active items (as in sentence and paragraph comprehension and in perception of the gist of a picture). The power of these two factors – selective attention that is defined by conceptual properties of the target, and the presence of potential conceptual structure – is felt early in processing, before conventional STM or working memory for the stimuli has been established, justifying the claim that CSTM is separate from those forms of working memory.

Outside of the laboratory, we usually have control over the rate of presentation: we normally read at a rate sufficient not only for momentary comprehension, but also for memory of at least the gist of what we are reading. Although we cannot control the rate of input from radio, TV, movies, or computer games, producers are adept at adjusting the rate to fit the conceptual and memory requirements of their audience. Similarly, in a conversation

the speaker adjusts his or her rate of speech to accommodate the listener's signals of comprehension. Rapid structuring can only occur if the material permits it and if the skills for discovering latent structure are highly practiced: object and word recognition, lexical retrieval, sentence parsing, causal inference, search for a target, and the like. These cognitive skills, built up over a lifetime, make comprehension seem trivially easy most of the time.

Conceptual short term memory is the working memory that supports these processes, lasting just long enough to allow multiple interpretation to be considered before one is selected and the remaining elements evaporate, in most cases without entering awareness. The labored thoughts and decisions we are aware of pondering are a tiny fraction of those we make effortlessly. Even these worked-over thoughts may advance by recycling the data through CSTM until the next step comes to mind. We may be aware of slowly shaping an idea or solving a problem, but not of precisely how each step occurs. More work will be needed to gain a full understanding of this largely preconscious stage of cognitive processing.

IMPLICATIONS FOR THE RELATION BETWEEN PERCEPTION AND COGNITION

In the present account, perception is continuous with cognition. Information passes from the sense organs to the brain, undergoing transformations at every stage, combining with information from other senses, evoking memories, leading to conscious experience and actions shaped by one's goals. These events are continually renewed or replaced by new experiences, thoughts, and actions. From the moment the pattern of a falling cup appears on the retina (to take one example) to the moment when one reaches to catch it, a series of events has passed through the visual system and to the motor system by way of the conceptual and goal-directed systems, at every stage influenced by prior experience already represented in these systems. It is to some extent a matter of convention that we break up such an event into perceptual, cognitive, goal-directed, and motor parts, when in reality these parts are not only continuous but also interact. In this example, perception of the tipping cup combines with conceptual knowledge to elicit the goal and action almost simultaneously, and the action anticipates the subsequent perceptual sequence. There is no clear separation between mental/brain activity originating from outside the observer ("perception") and that from the observer's internally generated thoughts and memories ("cognition"); any experience is likely to be a blend of these sources. Thinking with the eyes closed in a dark, silent room might seem to come close to a "pure" cognitive experience, but our propensity to augment our thoughts with visual, auditory, and other sense images suggests that the perceptual system is ubiquitous in cognition. Whether such self-generated perceptual images are the heart of cognition (as embodiment theories suggest) or only play a supporting role, is in dispute, however. I keep Helen Keller in mind as an apparent counterexample to the claim that cognition is entirely embodied in perception and action: Keller could neither see nor hear, but those who knew her as an adult had no doubt about her cognitive abilities and her command not only of language but also of world knowledge. Was Keller a kinesthetic/proprioceptive

zombie who simply simulated human understanding, or did she have an intact capacity for cognitive abstraction, once an access route for language was established? The latter seems more likely.

Authorization for the use of experimental animals or human subjects: The experimental studies described here were approved

by the Internal Review Board of Massachusetts Institute of Technology, and all participants signed consent forms.

ACKNOWLEDGMENTS

This work was supported by Grant MH47432 from the National Institute of Mental Health.

REFERENCES

- Baddeley, A. (1986). *Working Memory*. Oxford: Clarendon.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci. (Regul. Ed.)* 4, 417–423.
- Baddeley, A. (2007). *Working Memory, Thought, and Action*. USA: Oxford University Press.
- Baddeley, A., and Hitch, G. J. (2010). Working memory. *Scholarpedia* 5, 3015.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cogn. Psychol.* 18, 355–387.
- Bransford, J. D., and Johnson, M. K. (1972). Contextual prerequisites for understanding: some investigations of comprehension and recall. *J. Verbal Learn. Verbal Behav.* 11, 717–726.
- Budiu, R., and Anderson, J. R. (2004). Interpretation-based processing: a unified theory of semantic sentence processing. *Cogn. Sci.* 28, 1–44.
- Chun, M. M., Bromberg, H. S., and Potter, M. C. (1994). “Conceptual similarity between targets and distractors in the attentional blink,” *Poster presented at the 35th Annual Meeting of the Psychonomic Society*, St. Louis, MO.
- Chun, M. M., and Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 109–127.
- Coltheart, M. (1983). Iconic memory. *Philos. Trans. R. Soc. Lond. B* 302, 283–294.
- Davenport, J. L. (2007). Contextual effects in object and scene processing. *Mem. Cognit.* 35, 393–401.
- Davenport, J. L., and Potter, M. C. (2004). Scene consistency in object and background perception. *Psychol. Sci.* 15, 559–564.
- Davenport, J. L., and Potter, M. C. (2005). The locus of semantic priming in RSVP target search. *Mem. Cognit.* 33, 241–248.
- DiCarlo, J. J., Zoccolan, D., and Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron* 73, 415–434.
- Doolling, D. J., and Lachman, R. (1971). Effects of comprehension on retention of prose. *J. Exp. Psychol.* 88, 216–222.
- Ericsson, K. A., and Kintsch, W. (1995). Long-term working memory. *Psychol. Rev.* 102, 211–245.
- Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Percept. Psychophys.* 8, 215–221.
- Forster, K. I., and Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 680–698.
- Holmes, V. M., and Forster, K. I. (1972). Perceptual complexity and underlying sentence structure. *J. Verbal Learn. Verbal Behav.* 11, 148–156.
- Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 1–12.
- Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 604–610.
- Intraub, H. (1984). Conceptual masking: the effects of subsequent visual events on memory for pictures. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 115–125.
- Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychol. Rev.* 95, 163–183.
- Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.
- Landauer, T. K., and Dumais, S. T. (1997). A solution to Plato’s problem: the latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychol. Rev.* 104, 211–240.
- Lawrence, D. H. (1971a). Temporal numerosity estimates for word lists. *Percept. Psychophys.* 10, 75–78.
- Lawrence, D. H. (1971b). Two studies of visual search for word targets with controlled rates of presentation. *Percept. Psychophys.* 10, 85–89.
- Loftus, G. R. (1983). “Eye fixations on text and scenes,” in *Eye Movements in Reading: Perceptual and Language Processes*, ed. K. Rayner (New York: Academic Press), 359–376.
- Loftus, G. R., and Ginn, M. (1984). Perceptual and conceptual masking of pictures. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 435–441.
- Loftus, G. R., Hanna, A. M., and Lester, L. (1988). Conceptual masking: how one picture captures attention from another picture. *Cogn. Psychol.* 20, 237–282.
- Lombardi, L., and Potter, M. C. (1992). The regeneration of syntax in short term memory. *J. Mem. Lang.* 31, 713–733.
- Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281.
- Luck, S. J., Vogel, E. K., and Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature* 383, 616–618.
- Maki, W. S., Frigen, K., and Paulson, K. (1997). Associative priming by targets and distractors during rapid serial visual presentation: does word meaning survive the attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 23, 1014–1034.
- Meng, M., and Potter, M. C. (2011). An attentional blink for nontargets? *Atten. Percept. Psychophys.* 73, 440–446.
- Miyake, A., Just, M. A., and Carpenter, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neutral contexts. *J. Mem. Lang.* 33, 175–202.
- Neely, J. H. (1991). “Semantic priming effects in visual word recognition: a selective review of current findings and theories,” in *Basic Processes in Reading: Visual Word Recognition*, eds D. Besner and G. W. Humphreys (Hillsdale, NJ: Erlbaum), 264–336.
- Phillips, W. A. (1983). Short-term visual memory. *Philos. Trans. R. Soc. Lond. B* 302, 295–309.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *J. Exp. Psychol. Hum. Learn. Mem.* 2, 509–522.
- Potter, M. C. (1982). Very short-term memory: in one eye and out the other. *Paper presented at the 23rd Annual Meeting of the Psychonomic Society*, Minneapolis.
- Potter, M. C. (1984). “Rapid serial visual presentation (RSVP): a method for studying language processing,” in *New Methods in Reading Comprehension Research*, eds D. Kieras and M. Just (Hillsdale, NJ: Erlbaum), 91–118.
- Potter, M. C. (1993). Very short-term conceptual memory. *Mem. Cogn.* 21, 156–161.
- Potter, M. C. (1999). “Understanding sentences and scenes: the role of conceptual short term memory,” in *Fleeting Memories: Cognition of Brief Visual Stimuli*, ed. V. Coltheart (Cambridge, MA: MIT Press), 13–46.
- Potter, M. C. (2009). Conceptual short term memory. *Scholarpedia* 5, 3334.
- Potter, M. C., Dell’Acqua, R., Pesciarelli, F., Job, R., Peressotti, F., and O’Connor, D. H. (2005). Bidirectional semantic priming in the attentional blink. *Psychon. Bull. Rev.* 12, 460–465.
- Potter, M. C., and Jiang, Y. V. (2009). “Visual short-term memory,” in *Oxford Companion to Consciousness*, eds T. Bayne, A. Cleeremans, and P. Wilken (Oxford: Oxford University Press), 436–438.
- Potter, M. C., Kroll, J. F., and Harris, C. (1980). “Comprehension and memory in rapid sequential reading,” in *Attention and Performance VIII*, ed. R. Nickerson (Hillsdale, NJ: Erlbaum), 395–418.
- Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., and Sherman, J. (1986). Pictures in sentences: understanding without words. *J. Exp. Psychol. Gen.* 115, 281–294.
- Potter, M. C., and Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *J. Exp. Psychol.* 81, 10–15.
- Potter, M. C., and Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *J. Mem. Lang.* 29, 633–654.
- Potter, M. C., and Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *J. Mem. Lang.* 38, 265–282.
- Potter, M. C., Moryadas, A., Abrams, I., and Noel, A. (1993). Word perception and misperception in context. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 3–22.
- Potter, M. C., Nieuwenstein, M. R., and Strohminger, N. (2008). Whole report versus partial report in RSVP sentences. *J. Mem. Lang.* 58, 907–915.

- Potter, M. C., Staub, A., and O'Connor, D. H. (2002). The time course of competition for attention: attention is initially labile. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 1149–1162.
- Potter, M. C., Staub, A., and O'Connor, D. H. (2004). Pictorial and conceptual representation of glimpsed pictures. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 478–489.
- Potter, M. C., Stiefbold, D., and Moryadas, A. (1998). Word selection in reading sentences: preceding versus following contexts. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 68–100.
- Potter, M. C., Wyble, B., and Olejarczyk, J. (2011). Attention blinks for selection, not perception or memory: reading sentences and reporting targets. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1915–1923.
- Potter, M. C., Wyble, B., Pandav, R., and Olejarczyk, J. (2010). Picture detection in RSVP: features or identity? *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1486–1494.
- Raymond, J. E., Shapiro, K. L., and Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 18, 849–860.
- Rayner, K. (ed.). (1983). *Eye Movements in Reading: Perceptual, and Language Processes*. New York: Academic Press.
- Rayner, K. (ed.). (1992). *Eye Movements and Visual Cognition: Scene Perception and Reading*. New York: Springer-Verlag.
- Sereno, S. C., and Rayner, K. (1992). Fast priming during eye fixations in reading. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 173–184.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr. Gen. Appl.* 74, 1–29.
- Sperling, G., Budiansky, J., Spivak, J. G., and Johnson, M. C. (1971). Extremely rapid visual search: the maximum rate of scanning letters for the presence of a numeral. *Science* 174, 307–311.
- Staub, A., and Rayner, K. (2007). “Eye movements and on-line comprehension processes,” in *The Oxford Handbook of Psycholinguistics*, ed. G. Gaskell (Oxford: Oxford University Press), 327–342.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *J. Verbal Learn. Verbal Behav.* 18, 645–659.
- Verhoeven, L., and Perfetti, C. (2008). Advances in text comprehension: model, process and development. *Appl. Cogn. Psychol.* 22, 293–301.
- Vogel, E. K., Luck, S. J., and Shapiro, K. L. (1998). Electrophysiological evidence for a post-perceptual locus of suppression during the attentional blink. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1656–1674.
- Von Eckardt, B., and Potter, M. C. (1985). Clauses and the semantic representation of words. *Mem. Cognit.* 13, 371–376.
- Wyble, B., Potter, M., Bowman, H., and Nieuwenstein, M. (2011). Attentional episodes in visual perception, *J. Exp. Psychol. Gen.* 140, 488–505.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 January 2012; paper pending published: 24 January 2012; accepted: 29 March 2012; published online: 27 April 2012.

Citation: Potter MC (2012) Conceptual short term memory in perception and thought. *Front. Psychology* 3:113. doi: 10.3389/fpsyg.2012.00113

This article was submitted to *Frontiers in Perception Science*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Potter. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.