

Searching for the right word: Hybrid visual and memory search for words

Sage E. P. Boettcher · Jeremy M. Wolfe

Published online: 19 March 2015
© The Psychonomic Society, Inc. 2015

Abstract In “hybrid search” (Wolfe *Psychological Science*, 23(7), 698–703, 2012), observers search through visual space for any of multiple targets held in memory. With photorealistic objects as the stimuli, response times (RTs) increase linearly with the visual set size and logarithmically with the memory set size, even when over 100 items are committed to memory. It is well-established that pictures of objects are particularly easy to memorize (Brady, Konkle, Alvarez, & Oliva *Proceedings of the National Academy of Sciences*, 105, 14325–14329, 2008). Would hybrid-search performance be similar if the targets were words or phrases, in which word order can be important, so that the processes of memorization might be different? In Experiment 1, observers memorized 2, 4, 8, or 16 words in four different blocks. After passing a memory test, confirming their memorization of the list, the observers searched for these words in visual displays containing two to 16 words. Replicating Wolfe (*Psychological Science*, 23(7), 698–703, 2012), the RTs increased linearly with the visual set size and logarithmically with the length of the word list. The word lists of Experiment 1 were random. In Experiment 2, words were drawn from phrases that observers reported knowing by heart (e.g., “London Bridge is falling down”). Observers were asked to provide four phrases, ranging in length from two words to no less than 20 words (range 21–86). All words longer than two characters from the phrase, constituted the target list. Distractor words were matched for length and frequency. Even with these strongly ordered lists, the results again replicated the curvilinear function of memory set size seen in hybrid search. One might expect to find serial position effects, perhaps reducing the RTs for the first (primacy) and/or

the last (recency) members of a list (Atkinson & Shiffrin, 1968; Murdock *Journal of Experimental Psychology*, 64, 482–488, 1962). Surprisingly, we showed no reliable effects of word order. Thus, in “London Bridge is falling down,” “London” and “down” were found no faster than “falling.”

Keywords Visual search · Long-term memory · Word recognition

Imagine searching through a list of names, looking for your own name. Assuming that the list is not alphabetical, the nature of this search is quite well understood. With one specific target in mind (your name), most visual search tasks will produce response times (RTs) that are a linear function of the number of items—in this case, the length of the list of names. This pattern is seen for many types of simple stimuli (Treisman & Gelade, 1980; Wolfe, 1994, 2007), and the pattern remains the same for visual search of words (Fisk & Schneider, 1983). That is, when observers are searching for a target word amongst varying numbers of distractor words, search times will increase as a function of the visual set size.

Suppose, instead, that you have focused your attention on one name on the list and you are trying to remember whether it is the name of one of 20 students in a class. How we perform this search through memory is less clear. Under some conditions, memory search patterns are similar to those found in the visual domain. Sternberg (1966, 1969) showed that the time that it takes to decide whether a single item is being held in memory is a linear function of the number of items in memory. Again, this result was replicated using various stimulus types, including words (Cavanagh, 1972; Fisk & Schneider, 1983; Juola & Atkinson, 1971; Sternberg, 1969). These studies had all used variable-mapping paradigms (Schneider & Shiffrin, 1977) in which the targets and distractors changed

S. E. P. Boettcher (✉) · J. M. Wolfe
Brigham & Women’s Hospital, Boston, MA, USA
e-mail: sage.boettcher@gmail.com

J. M. Wolfe
Harvard Medical School, Boston, MA, USA

on every trial in an unpredictable manner (e.g., a target on one trial could appear as a distractor on the next). In consistent-mapping tasks, in contrast, a particular stimulus will always appear as either a target or a distractor over a block of trials. In general, consistent-mapping tasks produce much more efficient memory search slopes, in some cases completely eradicating set size effects (Fisk & Schneider, 1983; Schneider & Shiffrin, 1977). Other studies have shown that consistent mapping causes decelerating, curvilinear functions of set size, rather than the linear functions of variable mapping (Donkin & Nosofosky, 2012; Kristofferson, 1972; McElree & Doshier, 1989; Monsell, 1978; Ratcliff, 1978). Simpson (1972) found that the mean RT was a linear function of the log of the memory set size in a consistent-mapping memory recognition task with extended practice for the observers. None of the work mentioned above, however, had used memory set sizes beyond eight items. Investigating searches through large memory sets, Juola, Fischler, Wood, and Atkinson (1971) and Atkinson and Juola (1973, 1974) found linear searches through sets ranging from ten to 60 words.

Finally, imagine that you are searching the entire list for the names of any one of those 20 students. This involves you bringing each of the students' names into some form of working memory, in order to compare it with each word or name in the list. This combination of visual and memory search is known as "hybrid search" (Schneider & Shiffrin, 1977). By combining multiple memory searches into a single trial—and therefore, a single RT—hybrid search can magnify small distinctions between set sizes that may otherwise be lost in a traditional recognition test. Moreover, the hybrid-search paradigm allows us to separate out the effects of visual set size from those of memory set size and to look at their interaction. In earlier work with visual objects, Wolfe (2012) found that the RT in hybrid search was a linear function of visual set size, as in other visual search tasks. However, Wolfe (2012) also found that RTs increased with the log of the memory set size. As we noted, Wolfe used photorealistic objects as the stimuli. Stimuli of this sort have been used to replicate the logarithmic search through memory with a search through time rather than space (Drew & Wolfe, 2013). Moreover, the full hybrid-search pattern applies to search through categories of objects, as well as through sets of specific objects (Cunningham & Wolfe, 2014). Leite and Ratcliff (2010) introduced a potentially useful version of a diffusion model to explain the logarithmically increasing RTs (diagrammed in Fig. 8 below). In the context of hybrid search, a diffuser is assigned to each member of the memory set. This diffuser accumulates evidence for the presence of its particular memory item as the trial progresses. If and when any of the diffusers cross a decision boundary, a response is given. Noise in the diffusion process might cause an incorrect diffuser to cross the decision bound, generating a false alarm error. More items in memory mean more diffusion processes, and thus a greater chance of such an error. Raising

the decision boundary, and therefore requiring more information before committing to a decision, can reduce errors. Higher decision boundaries take longer to reach, however, increasing the RT. If observers attempt to hold error rates constant, decision boundaries and RTs must rise with memory set size. Leite and Ratcliff (2010) showed that the resulting RT \times Set Size function will be logarithmic.

In the present article, we ask whether the pattern generalizes beyond objects to words. With small set sizes (1–4) in combined memory and visual search through alphanumeric symbols, linear RT \times Memory Set Size functions have been reported (Briggs & Blaha, 1969; Burrows & Murdock, 1969; Nickerson, 1966; Schneider & Shiffrin, 1977), but the distinction between linear and logarithmic functions is not easy to detect with small set sizes.

Experiment 1 of the present article replicated Wolfe's (2012) findings, using words rather than photorealistic objects as the stimuli. Observers were asked to remember between two and 16 words and then to search for those amongst visual set sizes of two to 16 items. As in Wolfe (2012), RTs increased linearly with the visual set size, and log-linearly with the memory set size.

Words provide an opportunity to answer other questions about hybrid search that cannot be readily addressed with object stimuli. For example, up until this point the memorization process for all of the related studies has been the same: Observers have been asked to memorize new and arbitrary sets of items at the start of a block of trials. With words, it is possible to ask observers to search through previously memorized familiar texts. These are highly ordered, structured, meaningful sets of stimuli that have long been engrained in the observer's memory. Again, we can ask whether the basic hybrid search results are found, and again, there are good reasons to think that these stimuli might produce a different pattern of results.

For example, it is possible that an advantage might exist for targets at the beginning or end of an ordered list, mirroring the well-established primacy and recency effects (Atkinson & Shiffrin, 1968; Glanzer & Cunitz, 1966; Murdock, 1962). This U-shaped function has been found for lists stored in long-term memory up to several weeks (Neath & Brown, 2006). Moreover, serial position has been shown to be an important predictor of recalled targets in coherent passages, as well (Deese & Kaufman, 1957; Freebody & Anderson, 1981; Rubin, 1977, 1978). Kelley, Neath, and Surprenant (2013) found both primacy and recency effects in observers' memory of cartoon theme song lyrics, the seven Harry Potter books, and two different sets of movies, all of which were thought to be recalled from semantic memory stores. Although much of the work on serial position effects has focused on recall tasks, both early and late members of a list have shown RT benefits in recognition tasks, as well (McElree & Doshier, 1989; Monsell, 1978).

In addition to serial position, the rated importance of a particular section in a passage has also proved to be important for recall. Freebody and Anderson (1981) showed that the higher rating of the semantic importance of a particular subsection of a passage, the more likely it was to be recalled. Therefore, in an extension of this finding, we might expect to find RT benefits for words that were rated as being more semantically important in a given passage.

To test these possibilities, we asked observers to provide four phrases of varying lengths that they felt confident were firmly ensconced in their memory. Those words became the memory set in a hybrid search for any member of one of the target phrases amongst distractor words. Again, we replicated the pattern of a linear increase in RTs with visual set size and a logarithmic increase with memory set size. Interestingly, we found no reliable effect of serial position and almost no effect of the importance of the word in the phrase.

Experiment 1: Arbitrary memory sets

Observers

In Experiment 1, ten observers, 18 to 48 years of age, were tested (mean age: 29.2; six males, four females). The observers gave informed consent and were compensated \$10/h. All observers had at least 20/25 vision with correction, passed the Ishihara Color Blindness Test, and were fluent speakers of English.

Method

In Experiment 1, observers searched displays of 2, 4, 8, and 16 words for targets drawn from memory sets of 2, 4, 8, or 16 items (see Fig. 1). Observers were seated so that their eyes were approximately 57.4 cm from a 20-in. CRT monitor with an 85-Hz refresh rate. At a 57.4-cm viewing distance, 1 cm is equivalent to a visual angle of 1°. All experiments were written in MATLAB using the Psychophysics Toolbox (Brainard, 1997). Words were displayed in font size 40. At full height, the letters were 1.3° tall. Words were presented on a gray background ($LUM = 51.8 \text{ cd/ms}^2$); they were chosen at random from the Touchstone Applied Science Associates (TASA) database (Zeno, 1995), with the only constraint being that they had to be at least three letters long. The distractors were chosen to match the targets in word length and frequency.

In each of four blocks, observers memorized 2, 4, 8, or 16 words. During the memorization task, target words were presented one at a time centrally on the screen for 3 s at a time. Next, the observers were required to pass two recognition tests with 100% accuracy in order to proceed to the search portion

of the block. For this learning portion of each block, observers saw a set of words, one at a time, and labeled them as “old” (i.e., part of their memory set) or “new” (distractors). Distractors made up 50% of the recognition test; therefore, in total observers saw twice as many words as the memory set size. If observers failed the test, they reviewed the target words again for 3 s each and then attempted the memory test again. Word order was randomized during all portions of the memorization block, and the distractors were always novel.

After completing the memory portion of the block, observers moved on to a series of 330 search trials: 30 practice trials and 300 experimental trials. During the search task, observers saw displays with 2, 4, 8, or 16 printed words and were instructed to localize any one of their targets with a mouse click as quickly and accurately as possible. One random member from the target list was always present among an array of distractor words. The spatial locations of all of the words in the display were randomly chosen, with the only constraints being that words could not overlap with one another and that the entire word had to fit on the display. After clicking on the target, observers received “correct”/“incorrect” feedback before moving on to the next trial. Participants completed four blocks with memory set size pseudorandomized. From start to finish, the experiment lasted about 1.5 h.

Results and discussion

The results from the memory test showed that observers easily memorized the words (average accuracy of 97% over all possible tests). If observers were perfect in their memorization of their target words, they would have to complete exactly two recognition tests before moving on to the search trials. On average, observers complete 3.3 recognition tests in Experiment 1. This strong grasp on the memory sets is further demonstrated by the low error rates (2%), which were not modulated by either memory set size or visual set size. Figure 2a shows the mean RTs as a function of visual set size. As is typical in visual search experiments (e.g., Treisman, 1988), and as was seen in the basic hybrid-search results, RTs increased linearly with visual set size. Larger memory set sizes produced steeper slopes of the RT \times Visual Set Size functions [$F(1.993, 17.94) = 11.16, p = .0007$; using Greenhouse–Geisser correction], showing that when more target words are in memory, the cost of adding a distractor to the visual display is higher.

Figure 2b shows RTs as a function of memory set size. Note that these are the same data points as in Fig. 2a, simply replotted. The functions appear to be curvilinear. Wolfe (2012) argued that RT was a linear function of $\log_2(\text{memory set size})$. One way to compare linear and \log_2 accounts of the data would be to use the three smaller memory set sizes to predict the data for set size 16. This is shown in Fig. 2b, with linear

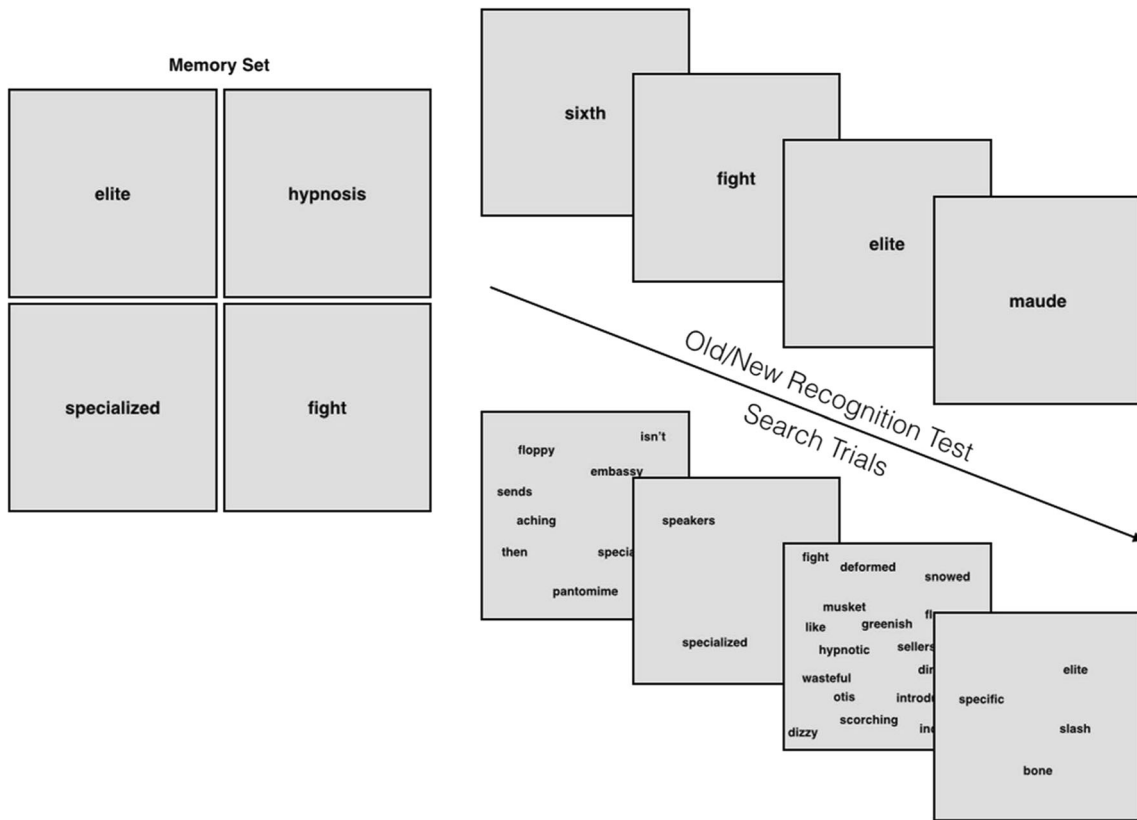


Fig. 1 Observers were first asked to memorize between two and 16 words. These words were presented centrally one at a time for 3 s each. After memorization, observers had to complete an old–new recognition

test at 100% accuracy before moving on to the search trials. All words used in this figure are actual words pulled from the experiment

predictions shown as Os and \log_2 predictions shown as Xs. The data and the \log_2 predictions are quite close (differences: visual set size 2, 31.7 ms; set size 4, 77.8 ms; set size 8, 87.5 ms; set size 16, 131.3 ms). The linear predictions overestimate the actual data, especially for the larger visual set sizes (differences: visual set size 2, 85.5 ms; set size 4, 43.13 ms; set size 8, 481 ms; set size 16, 691 ms). This illustrates one of the virtues of the hybrid-search paradigm in

comparison to a standard recognition task: Because observers need to search memory for each attended word, the larger the visual set size, the larger the number of memory searches that will contribute to the RT. This acts to magnify the differences between the predictions of linear and logarithmic processes. That is why the linear prediction is off by hundreds of milliseconds at the largest set size. The absolute errors (the differences between the predicted RT and actual RT) are

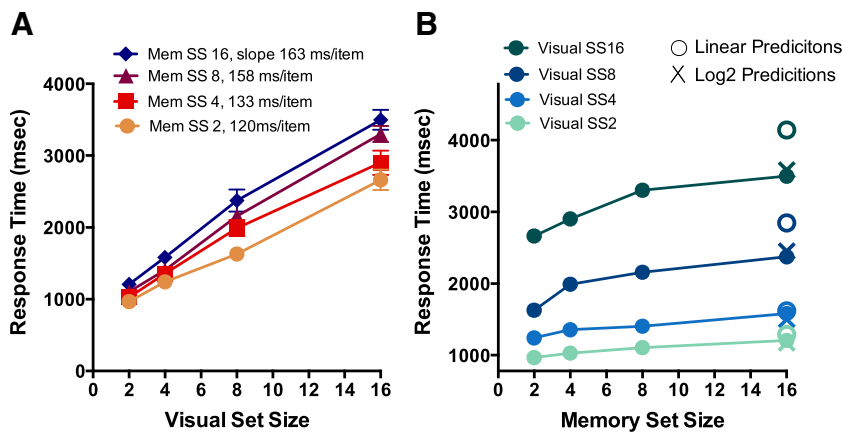


Fig. 2 (a) We can see that response times increase roughly linearly by visual set size and curvilinearly (b) by memory set size. In panel B, we can see predictions for memory set size 16 based on the results for

memory set sizes 2–8; the logarithmic model does a noticeably better job of predicting the actual data. Note that these graphs represent the same data simply replotted

significantly smaller for the logarithmic prediction at visual set sizes 2 and 16 [$t(9) > 2.2$, $p < .03$] and are marginally smaller at set size 8 [$t(9) = 1.718$, $p = .0599$].

Experiment 1 replicated the work of Wolfe (2012) in the lexical domain, showing that the original result is not restricted to specific objects. Furthermore, by using the hybrid-search paradigm, we can perhaps shed some light on memory search in a consistent-mapping task through words. Since larger visual set sizes require multiple memory searches, deviations from a linear model become more evident, and this perhaps explains the difference of the present results from earlier work with smaller set sizes.

In the present experiment, observers had no trouble encoding the random words up to the maximum of 16. In Experiment 1, the words in memory had no meaning or grammatical structure. Moreover, they were not presented and tested in any fixed order in the first part of the study. Accordingly, we could not assess the effects of word order, including any analysis of serial position. Under normal circumstances, however, word order is important (at least in English). In Experiment 2, we asked whether the basic hybrid-search result would change if we switched the target sets from arbitrary lists of words, learned for the task, to structured lists of words, derived from well-learned text held in our observers' long-term memory.

Experiment 2: Familiar phrases as memory sets

Method

In Experiment 2, 12 observers 19 to 48 years of age (mean age: 27; four males, eight females) were initially asked to select and enter into the computer four phrases that they knew very well. For three of the phrases, they were instructed to think of passages that, as closely as possible, contained 2, 8, and 16 words. For the fourth and final target set, we asked for the longest phrase that the observer had fully committed to memory. These largest set sizes ranged in length from 19 to 86 words, with an average of 33.75 words. The participants were told that words less than three letters long and repeated words would not count toward their target list. Observers were also given a list of well-known phrases as suggestions (e.g., “Twinkle, Twinkle Little Star,” The US Pledge of Allegiance, etc.); they were instructed not to simply pick “popular” phrases, but rather to choose phrases that they were sure they knew well.

Observers were given a test of their memory for each phrase prior to the search portion of the experiment. Since these target sets were user-inputted, we expected the memory test not to cause many problems. However, in order to keep the two experiments as close as possible, as in Experiment 1, observers were shown the target and distractor words and were asked to make an “old” (part of the phrase) or “new”

(distractor) response to each word. Half of the words were target words. We lowered the threshold from Experiment 1; so that the experiment was not delayed due to motor errors, observers had to score 90% or higher twice in a row in order to pass this memory test. In other respects, the methods for Experiment 2 were the same as those for Experiment 1.

Results and discussion

Unsurprisingly, observers had little difficulty with the memory portion of the experiment, scoring an average of 98.2%. Figure 3 shows the data for all observers, with the slope of the $RT \times$ Visual Set Size function plotted as function of the memory set size. This slope represents the cost of each additional visual item as a function of how many items are held in memory. The functions would be linear on the log axis of Fig. 3 if the results were to replicate the logarithmic relationship of RT to memory set size seen in Wolfe (2012) and in Experiment 1.

In fact, because of the variety of memory set sizes and because of the variability of the individual observers' data, it is difficult to see the relationship in the individual observer data. Accordingly, in order to pool the data for purposes of analysis, the memory set sizes were binned into four distinct groups: short, medium, long, and extra long. All short memory sets were two words long. The medium bin consisted of phrases of 4–11 words (average: 7.2 words), the long bin of phrases of 12–19 words (average: 15.3 words), and the extra-long group of phrases of more than 19 words (average: 35.0 words). Using these criteria, all but one observer was tested on one phrase per memory group. (The longest phrase for that observer was merely “long,” not “extra long.”).

The results are shown in Fig. 4.

Even though it is a bit problematic to average across multiple set sizes, it is clear that the basic pattern of results in Experiment 2 mirrors that from Experiment 1; a linear effect

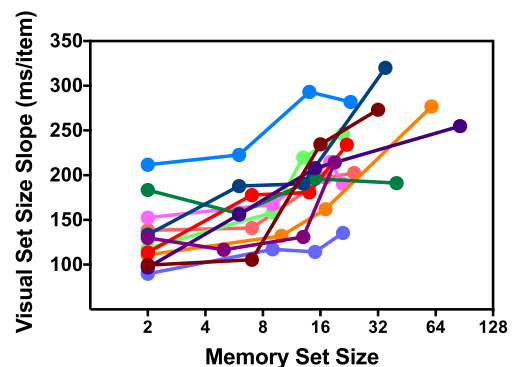


Fig. 3 Slopes of the visual set size function as a function of memory set size. Each line represents one observer. The slope represents the cost of each additional visual item as a function of how many items are held in memory. Note that the x-axis is logarithmically scaled

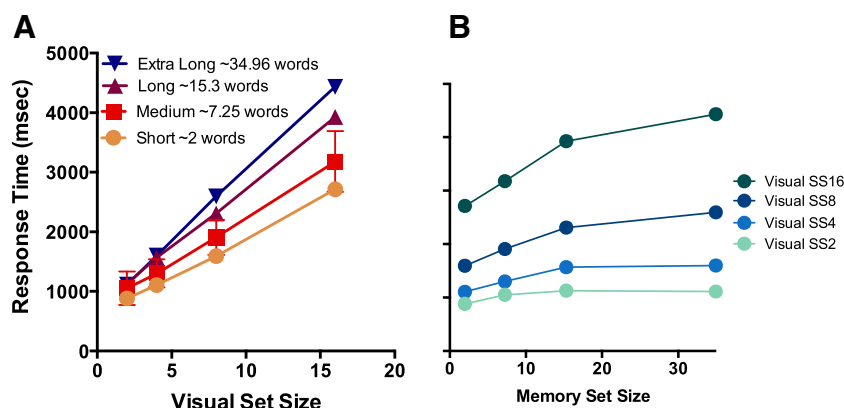


Fig. 4 Response times as a function of visual set size (a) and binned memory set size (b). Error bars, where visible, are $\pm 1SEM$

of visual set size, and a log-linear effect of memory set size. Rather than using the average data from low set sizes to predict the results for the largest set size, as in Fig. 2b, in Fig. 5a the three lower set sizes are used to predict the largest set size result for each observer at each visual set size. It is important to note that for Fig. 5, we are using observers' actual memory set sizes rather than the binned averages from Fig. 4. In the online (color) figure, solid purple symbols show linear predictions, and outlined green symbols show \log_2 predictions. If the prediction matched the data, the data point would lie on the diagonal of the graph plotting predicted against the real RT. The graph omits 28 data points greater than 4,000 ms in order to make the shorter RTs visible. However, the regression lines are based on all data from all participants. It is clear that the linear model predicts RTs that are longer than the actual RTs, whereas the \log_2 model predicts RTs that are, on average, closer to the measured RTs for the largest set sizes, though the data are noisy. Figure 5b replots these data as the differences between the predicted and actual RTs. It is evident that the error is much larger for the linear than for the logarithmic predictions [$F(1, 11) = 9.751, p = .0097$]; this is especially true for the larger visual set sizes.

One may suspect that the use of highly ordered memorized phrases like “Mary had a little lamb” could produce serial position effects. However, this proved not to be the case. Figure 6 shows the average RTs as a function of position relative to the beginning (left column) and relative to the end (right column) of the phrase for the medium and long phrases. Short, two-item phrases are uninteresting for these purposes. To test for primacy and recency effects, we compared the RTs of either the first five words (primacy) or the last five words (recency) using a one-way analysis of variance in the medium, long, and extra-long phrases.¹ We found no significant difference in RTs

¹ This analysis was repeated for varying groupings, including memory sets of between six and nine words [primacy, $F(4, 35) = 0.1249, p = .97$; recency, $F(4, 36) = 1.384, p = .259$], as well as memory sets of between 18 and 24 words [primacy, $F(4, 23) = 0.893, p = .48$; recency, $F(4, 19) = 2.631, p = .07$]. No reliable effects of primacy or recency were found.

between the first word and the next four words [medium, $F(4, 43) = 0.099, p = .98$; long, $F(4, 45) = 0.97, p = .43$], showing no evidence for a primacy effect, nor was there a difference between the last word and the previous four words [medium, $F(4, 31) = 1.16, p = .345$; long, $F(4, 36) = 0.82, p = .51$]. The results for extra-long phrases were noisy and are not plotted, but they also did not show effects [primacy, $F(4, 33) = 1.6, p = .19$; recency, $F(4, 22) = 0.74, p = .57$]. A two-way analysis of variance with Phrase Length and Serial Position as factors showed no overall effect of serial position [primacy, $F(4, 120) = 2.33, p = .06$; recency, $F(4, 124) = 0.80, p = .50$].

Although serial order did not seem have a significant effect on response times, it might be that some word or words in a phrase would be privileged in activated long-term memory (the supposed store for memory set items in hybrid search; Cowan, 1995; Cunningham & Wolfe, 2014; Drew & Wolfe, 2013), in a way that would affect RTs. For instance, one might expect words that are more semantically salient to be accessed more quickly (e.g., “America” might be found faster than “under” in the US Pledge of Allegiance). We asked eight new observers to rate each word of each phrase used in Experiment 2 on a semantic salience scale from 1 to 5. The observers were asked how strongly a particular word “contributes to the overall meaning of the phrase,” where 1 indicated *no contribution* and 5 indicated *a strong contribution*. They were instructed not to consider words that contributed to the grammar or structure of a phrase, but rather to focus on words that were necessary to understand the feeling of the phrase. With these ratings in mind, we looked for a negative correlation between the higher ratings (i.e., more semantically relevant words) and the RTs.

Figure 7 shows that the relationship between how important a particular target word was to the meaning of the phrase and the RTs was weak (medium, $r = -.002, p = .8819$; long, $r = .007, p = .6348$; extra long, $r = -.067, p < .0001$). There was no correlation for the shorter phrases, though a small but significant relationship does emerge for the extra-long set sizes. This may indicate that the semantically relevant words only

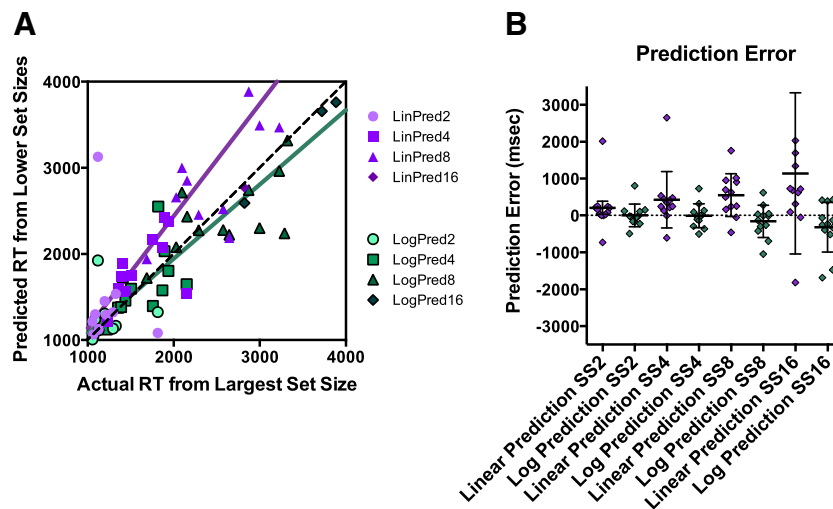


Fig. 5 (a) Actual response times (RTs) for the largest memory set sizes in Experiment 2 versus the RTs predicted by linear regression from the smaller memory set sizes. One set of points shows linear predictions, and the other shows logarithmic predictions. The log₂ predictions lie closer to the zero-error line (dashed black). The regression lines from

the linear and log prediction are color-matched to the relevant points. (b) Error data (predicted RT – actual RT) are plotted for each subject, with the linear and log predictions color-matched to those in panel A. Clearly, the log predictions are more accurate. SS2–SS16, set sizes 2–16

gain an advantage once there is a substantial memory load. The weakness of the effect and the absence of serial position effects indicate that, for the purposes of hybrid search, there is no clear advantage for any member of the target set on the basis of either serial position or semantic relevance.

Admittedly, the conditions of the present study were not ideal for finding serial position effects. In a replication of Sternberg (1966), Donkin and Nosofsky (2012) showed that serial position curves tend to flatten during recognition tests with enough time for rehearsal. Furthermore, since the phrases used in Experiment 2 were so well-known, they may have been stored as single units of information, and could therefore be retrieved as such. This would suggest that arbitrary words presented in a specific order (e.g., in a list) are more likely to

show primacy and recency effects. However, the goal of this experiment was to investigate the efficiency of memory search in well-known phrases. The lack of serial position effects is in fact consistent with an efficient logarithmic search through the items in the memory, with little or no effect of either serial position or the significance of the words.

General discussion

The apparently logarithmic increases in memory set size found in both Experiments 1 and 2, accompanied with the failure to find primacy or recency effects in Experiment 2, suggest that hybrid search is not performed by a serial search

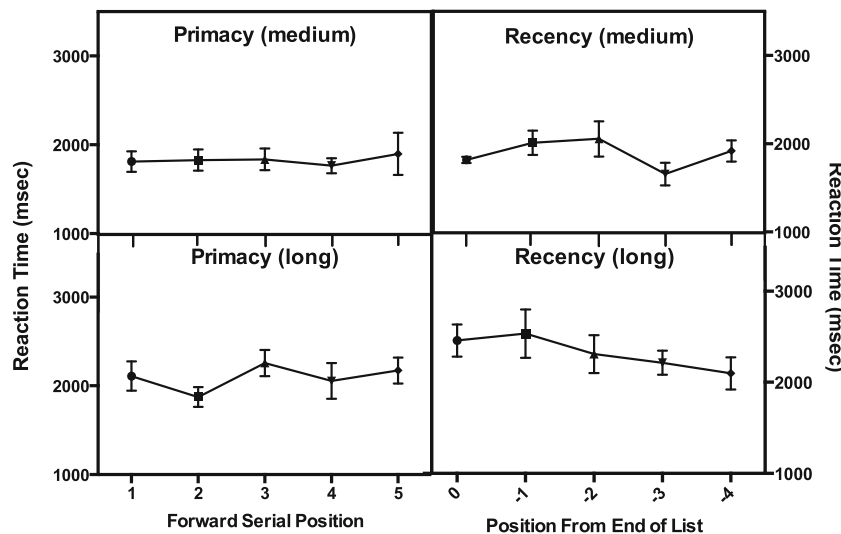


Fig. 6 No evidence for serial position effects is apparent in hybrid search for words drawn from well-known phrases. Neither the first nor the last word in a phrase is reported more quickly than neighboring words

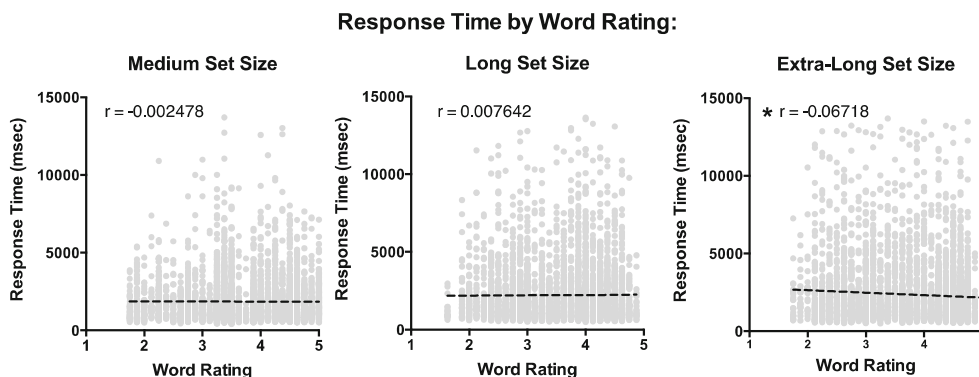


Fig. 7 Individual target words were rated on a scale of 1 to 5 on how strongly they contributed to the overall meanings of the phrases. No strong correlations are apparent between a word’s rating and the response time; however, with the extra-long memory set, this correlation is significant

through the memory set. The added attribute of large visual set sizes, which was not found in the earlier memory search literature, allowed us to more easily distinguish logarithmic from linear increases in RT as a function of memory set size. It is also clear that Wolfe’s (2012) findings are not restricted to photorealistic objects, which are easier to memorize (Gehring, Toggia, & Kimble, 1976; Brady, Konkle, Alvarez, & Oliva, 2008) and to find (Paivio & Begg, 1974) than words.

The memory searches required to complete a hybrid-search task may be viewed as multiple-alternative decisions akin to those that produce Hick’s law behavior in motor RT tasks. In those tasks, the time required to chose between N response keys rises according to the equation $k[\log_2(N)]$ (Hick, 1952; Hyman, 1953). Similarly, the time required to decide whether a word in a visual display is one of N words in a list or phrase

also rises according to $k[\log_2(N)]$. As we discussed in the introduction, Leite and Ratcliff (2010) have provided one potential model for such a pattern of RTs. Figure 8 shows a cartooned version of that model in the context of hybrid search. Recognition of each of the potential memory items can be modeled as a diffusion process or accumulator, with a target-present response being generated when one of those accumulators reaches a decision bound (Fig. 8a and b). With even small amounts of noise in the system, it is intuitively clear that the chance of reaching a boundary by mistake, and thus generating a false alarm, will increase as the number of items in the memory set grows (Fig. 8c). Raising the bound to reduce these errors increases RTs (Fig. 8d). The bound needed to hold errors roughly constant produces RTs that increase linearly as a function of \log_2 the number of diffusors.

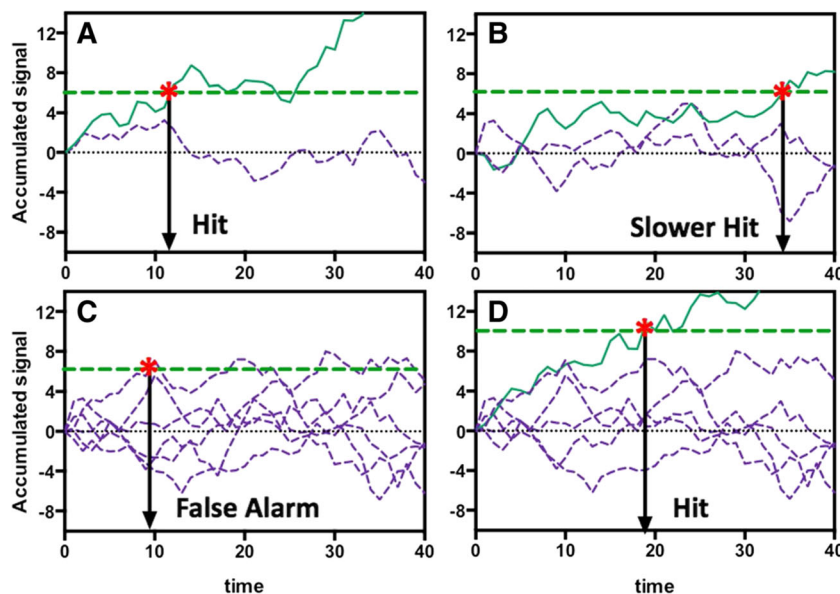


Fig. 8 Examples of a multiple-accumulator model of hybrid search. (a) With a memory set size of 2, two accumulators collect information. The accumulator corresponding to the target (top line) reaches threshold and generates a hit. The other item accumulates noise. (b) Memory set size 3. Here, by chance, the target takes longer to reach threshold, generating a

slower hit. (c) Memory set size 6. Here, an accumulator reaches threshold by chance, generating a false alarm error. The chance of these errors grows with the number of accumulators, so (d) at higher set sizes, the threshold is raised (see the text for more details)

Alternatively, Nosofsky, Cox, Cao, and Shiffrin (2014) proposed a model that also accounts for logarithmic RTs in memory search. As in Leite and Ratcliff (2010), items are racing to a decision boundary. However, in the Nosofsky version, these memory templates are not affected by the size of the memory set, but rather are dictated by their “memory strength,” or the time that they were presented last.

The present results show that this pattern of results holds for words, including words in ordered phrases, in a manner that is qualitatively similar to the results seen with specific objects. Hybrid search for words in the present experiments

is substantially slower than was hybrid search for objects in Wolfe (2012). For instance, in Experiment 1, 3,500 ms were required to find the target when the memory set size and visual set size were both 16. In the comparable experiment of Wolfe (2012), the comparable RT was 2,700 ms. Not too much should be made about the differences between experiments. The most obvious source of difference is that word reading probably requires fixation on each item, but object recognition does not. What is important is that the pattern of log memory search and linear visual search is a general phenomenon.

Appendix

Table 1 A list of all user-inputted phrases from Experiment2. Duplicate words and words less than 3 letters were not counted as part of the memory set. The order of the phrases has been randomized such that the row does not necessarily indicate a single observer

SHORT	MEDIUM	LONG	X-LONG
happy birthday	let me live in a house by the side of the road and be a friend to man	the muscle and bone they encase my heart but never touch my soul I'll save that for the water and its shore	I pledge allegiance to the flag of the united states of America and to the republic for which it stands one nation under god indivisible with liberty and justice for allyou
happy birthday	sea shells on the sea shore Shelly sells sea shells on the sea shore	twinkle twinkle little star how I wonder what you are up above the world so high like a diamond in the sky twinkle twinkle little star how I wonder what you are	never understood why we did this the audience knows the truth the world is simple miserable solid all the way through but if you can fool them even for a second then you can make them wonder and then you got to see something very special you really don't know it was the look on their faces
thank you	baa baa black sheep have you any wool	jack and jill went up the hill to fetch a pail of water jack fell down and broke his crown and jill came tumbling after	I pledge allegiance to the flag of the united states of America and to the republic for which it stands one nation under god indivisible with liberty and justice for all
good luck	I'll get you my pretty and your little dog too	yankee doodle went to town riding on a pony stuck a feather in his hat and called it macaroni	I pledge allegiance to the flag of the united states of america and to the republic for which it stands one nation under god indivisible with liberty and justice for all
happy birthday	ask not what your country can do for you but what you can do for your country	jack and jill went up the hill to fetch a pail of water jack fell down and broke his crown	alcoholic anonymous is a fellowship of men and women who share their experience strength and hope with each other so they can solve their common problem and help other to recovery from alcoholism there are no dues or fee for membership the only requirement is the desire to stop drinking we are self supporting declining any outside contribution AA doesn't align itself with any outside enterprise hence the name never be drawn into public controversy our primary purpose is to stay sober and help other alcoholics to achieve sobriety
good evening	Mary had a little lamb little lamb Mary had a little lamb and it's fleece was white as snow	humpty dumpty sat on the wall humpty dumpty had a great fall all the king's horse and all the king's men couldn't put humpty together again	I had a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character
happy birthday	baa baa black sheep have you any wool	twas the night before christmas when all through the house not a creature was stirring not even a mouse	I went out to the hazel wood because a fire was in my head and cut and peeled a hazel wand and hooked a berry to a thread and when white stars were on the wing and moth like stars were flickering out I dropped the berry in the stream and caught a little silver trout when I had laid it on the floor I stooped to blow the fire aflame and

Table 1 (continued)

SHORT	MEDIUM	LONG	X-LONG
			something rustled on the floor and someone called me by my name it had become a glimmering girl with apple blossoms in her hair who called me by my name and ran and faded through the brightening air though I am old with wandering through hollow lands and hilly lands I will find out where she has gone and kiss her lips and take her hands and walk among long dappled grass and pluck till time and times are done the silver apples of the moon the golden apples of the sun
great scott	sorry but your princess is in another castle	yankee doodle went to town riding on a pony stuck a feather in his hat and called it macaroni	our father who art in heaven hallowed be thy name thy kingdom come thy will be done on earth at it is in heaven give us this day our daily bread and forgive us our trespasses as we forgive those who trespass against us
good bye	I don't want to set the world on fire	X on your hand but today it's a beer why do you bother just get out of here	amazing grace how sweet the sound that saved a wretch like me I once was lost but now I'm found, was blind but now I see it was grace that taught my heart to fear and grace my fears relieved how precious did that great appear the hour I first believed
rolling stone	and the cow jumped over the moon	I will lift up mine eyes to the hills from whence cometh my help my help cometh from the lord the lord who made heaven and earth	my man inf left a tec and a nine at my crib turned himself in had to do a bid of one to three he'd be home at end of ninety-three I'm ready this paper g you with me
yes indeed	once upon a time you dressed so fine	my mama always said life is like a box of chocolate you never know what you're going to get	I'm singing in the rain just singing in the rain what a glorious feeling I'm happy again I laugh at the clouds so dark up above the sun's in my heart and I'm ready for love let the stormy clouds chase everyone from the place come on with the rain I've a smile on my face I walk down the lane with a happy refrain just singing singing in the rain
that's great	she sells all art really is quite useless	people used to call say beware doll you're bound to fall you thought they were all kidding you	I'm not sad anymore I'm just tired of this place the weight of the world would be okay if it would pick a shoulder to lean on so I could stand up straight *

* “Extra long” phrase actually falls into the bin “Long”

References

- Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583–612). New York, NY: Academic Press.
- Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Learning, memory, and thinking, Vol. 1, pp. 243–293). San Francisco, CA: W. H. Freeman.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York, NY: Academic Press.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*, 14325–14329. doi:10.1073/pnas.0803390105
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. doi:10.1163/156856897X00357
- Briggs, G. E., & Blaha, J. (1969). Memory retrieval and central comparison times in information processing. *Journal of Experimental Psychology*, *79*, 395–402.
- Burrows, D., & Murdock, B. B., Jr. (1969). Effects of extended practice on high-speed scanning. *Journal of Experimental Psychology*, *82*, 231–237. doi:10.1037/h0028145
- Cavanagh, J. P. (1972). Relation between the immediate memory span and the memory search rate. *Psychological Review*, *79*, 525–530. doi:10.1037/h0033482
- Cowan, N. (1995). *Attention and memory: An integrated framework*. New York, NY: Oxford University Press.
- Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General*, *143*, 1585–1599. doi:10.1037/a0036313
- Deese, J., & Kaufman, R. A. (1957). Serial effects of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology*, *54*, 180–187.
- Donkin, C., & Nosofsky, R. M. (2012). The structure of short-term memory scanning: An investigation using response time distributions. *Psychonomic Bulletin & Review*, *19*, 363–394. doi:10.3758/s13423-012-0236-8

- Drew, T., & Wolfe, J. M. (2013). Hybrid search in the temporal domain: Evidence for rapid, serial logarithmic search through memory. *Attention, Perception, & Psychophysics*, *76*, 296–303. doi:10.3758/s13414-013-0606-y
- Fisk, A. D., & Schneider, W. (1983). Category and word search: Generalizing search principles to complex processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 177–195. doi:10.1037/0278-7393.9.2.177
- Freebody, P., & Anderson, R. C. (1981). *Serial position and rated importance in the recall of text*. Urbana, IL: University of Illinois.
- Gehring, R. E., Toglia, M. P., & Kimble, G. A. (1976). Recognition memory for words and pictures at short and long retention intervals. *Memory & Cognition*, *4*, 256–260. doi:10.3758/BF03213172
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Behavior*, *5*, 351–360. doi:10.1016/S0022-5371(66)80044-0
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11–26. doi:10.1080/17470215208416600
- Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *45*, 188–196.
- Juola, J. F., & Atkinson, R. C. (1971). Memory scanning for words versus categories. *Journal of Verbal Learning and Verbal Behavior*, *10*, 522–527. doi:10.1016/S0022-5371(71)80024-5
- Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. (1971). Recognition time for information stored in long-term memory. *Perception & Psychophysics*, *10*, 8–14. doi:10.3758/BF03205757
- Kelley, M. R., Neath, I., & Surprenant, A. M. (2013). Three more semantic serial position functions and a SIMPLE explanation. *Memory & Cognition*, *41*, 600–610. doi:10.3758/s13421-012-0286-1
- Kristofferson, M. W. (1972). When item recognition and visual search functions are similar. *Perception & Psychophysics*, *12*, 379–384.
- Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics*, *72*, 246–273. doi:10.3758/APP.72.1.246
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, *118*, 346–373. doi:10.1037/0096-3445.118.4.346
- Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, *10*, 465–501.
- Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482–488. doi:10.1037/h0045106
- Neath, I., & Brown, G. D. A. (2006). SIMPLE: Further applications of a local distinctiveness model of memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 201–243). San Diego, CA: Academic Press.
- Nickerson, R. S. (1966). Response times with a memory-dependent decision task. *Journal of Experimental Psychology*, *72*, 761–769. doi:10.1037/h0023788
- Nosofsky, R. M., Cox, G. E., Cao, R., & Shiffrin, R. M. (2014). An exemplar-familiarity model predicts short-term and long-term probe recognition across diverse forms of memory search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1524–1539. doi:10.1037/xlm0000015
- Paivio, A., & Begg, I. (1974). Pictures and words in visual search. *Memory & Cognition*, *2*, 515–521. doi:10.3758/BF03196914
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108. doi:10.1037/0033-295X.85.2.59
- Rubin, D. C. (1977). Very long-term memory for prose and verse. *Journal of Verbal Learning and Verbal Behavior*, *16*, 611–621. doi:10.1016/S0022-5371(77)80023-6
- Rubin, D. C. (1978). A unit analysis of prose memory. *Journal of Verbal Learning and Verbal Behavior*, *17*, 599–620. doi:10.1016/S0022-5371(78)90370-5
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66. doi:10.1037/0033-295X.84.1.1
- Simpson, P. J. (1972). High-speed memory scanning: Stability and generality. *Journal of Experimental Psychology*, *96*, 240–246.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654. doi:10.1126/science.153.3736.652
- Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientists*, *57*, 421–457.
- Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *40A*, 201–237. doi:10.1080/02724988843000104
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. doi:10.1016/0010-0285(80)90005-5
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238. doi:10.3758/BF03200774
- Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford University Press.
- Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, *23*, 698–703. doi:10.1177/0956797612443968
- Zeno, S. M. (1995). *The educator's word frequency guide*. New York, NY: Touchstone Applied Science Associates.