Working Memory Benchmarks: A Missed Opportunity

Comment on Oberauer et al. (2018).

André Vandierendonck

Ghent University, Belgium

Running Head: Working Memory Benchmarks

Author Note

André Vandierendonck, Department of Experimental Psychology, Ghent University, Belgium.

Correspondence about the paper can be directed to: André Vandierendonck, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Gent, Belgium. E-mail: Andre.Vandierendonck@UGent.be.

Abstract

This commentary addresses a number of problems with the benchmarks proposed for evaluating theories of short-term and working memory [Oberauer et al., this issue, Benchmarks for Models of Short Term and Working Memory. *Psychological Bulletin*]. First, it is shown that the proposed benchmarks intentionally exclude findings regarding the core of the working memory construct, and also miss some important findings from other subdomains. For these reasons, the benchmarks cannot be considered as a valid representation of the findings on short-term and working memory. Second, it is shown that although theory-neutrality of the benchmarks was aimed for, this goal was not achieved because theory-neutrality in the formulation of the benchmarks does not guarantee inclusion of all theory-dependent findings. For these reasons, the benchmarks miss their purpose and are defined in such a way as to encourage a future theory development that studies working memory in isolation from other cognitive activities and thus misses the opportunity to stimulate a better integrative understanding of working memory in the broader context of cognition.

Keywords: Working Memory, Short-Term Memory, Action, Attention, Cognitive Control

Observing that the amount of published findings in the area of short-term and working memory is steadily growing, and that no single theory can explain an important portion of these findings, a consortium of short-term and working memory researchers has decided to define benchmarks to bring order in this vast array of findings so as to provide a basis for evaluating new models and theories with respect to their contribution to the field (Oberauer et al., this issue). Bringing order in the vast array of short-term and working memory findings is an important, valuable, and most welcome endeavor. However, the upgrading of a set of findings to the status of privileged findings that should be accounted for by any new theory or model of working memory requires an agreement on the part of the scientific community. There are at least two issues that deserve the attention of any member of this community, namely the question whether the proposed benchmarks validly represent short-term and working memory and whether using a set of benchmarks in the way proposed by the consortium is a methodologically and rationally sound way to act. Both issues are considered in more detail in the present commentary.

Representativeness of the Benchmarks

The consortium—-they call themselves the Benchmark Team—-adopts the concept of working memory as it was introduced by Miller, Galanter and Pribram (1960). These authors used the term working memory to refer to a temporary quick-access memory used for a smooth continuation of plan execution after interruption. In expanding on this concept, the Benchmark Team observes that “Today it is generally accepted in cognitive psychology that working memory (WM) plays a central role in all deliberative cognition, from language comprehension and mental arithmetic to reasoning and planning.” (p. 3). Yet, in the section on the scope and criteria for the benchmarks, the Benchmark Team excludes the findings regarding interactions of WM and “deliberative cognition” (i.e., language comprehension, mental arithmetic and reasoning) from the benchmarks, with the argumentation that “explaining these findings relies at least as much on a model of the domain of application (e.g., a model of syntactic parsing, or of deductive reasoning) as on a model of WM" (p. 10). In the same paragraph they also exclude findings regarding the relations of WM to "executive function", a term they use to refer to cognitive control processes as studied in Stroop interference, task switching and verbal fluency, "because research on executive functions has become a field of its own" (p. 10).

Relations between WM and Deliberative Cognition

The Benchmark Team's choice to steer clear from findings on the interaction of WM with deliberative cognition is ill inspired: by excluding these areas of research, findings regarding the core business of working memory research are excluded from the benchmarks for models of working memory. The reason for this deliberate choice, namely that an explanation of such findings requires a theory of working memory as well as a theory of the task domain, is far from convincing. The problem with this argument is that if these findings are not considered relevant by WM researchers because another domain is included, and the researchers of that other domain argue that these findings are not relevant because WM is involved, then these findings end up in no man's land and are not worthy of any scientific explanation. One of the big advantages of a construct like working memory is that it allows to study links between several sets of cognitive tasks, so that investigations of these links may help to integrate the domain of cognitive psychology. In what follows, the three task domains mentioned by the Benchmark Team are considered in some more detail so as to find out to what extent their argument holds.

Mental Arithmetic. Finding the answer to an arithmetic problem basically involves retrieval and combination of information from long-term memory (LTM). More specifically, for a given problem, one of a number of available procedures, also known as strategies, is retrieved from LTM and used to guide the collection and combination of arithmetic facts that are also retrieved from LTM (e.g., Ashcraft, 1992; Ashcraft & Battaglia, 1978; Campbell, 1987, 1995; LeFevre, Bisanz, et al., 1996; LeFevre, Sadesky, & Bisanz, 1996; Siegler & Shipley, 1995). For example, to find the sum of 3 and 6 (3 + 6 = ?), most people will simply retrieve the arithmetic fact “3 + 6 = 9” from LTM and produce the correct answer. However, other persons will prefer to start with the largest of the two numbers (6 in this case) and count the number of “add 1” steps as specified in the smaller number (+1, +1, +1) to come to the same answer (minimum-addend strategy, Groen & Parkman, 1972). For more complex problems, such as the sum of 23 and 36 (23 + 36 = ?), many people will use their knowledge about tens and units present in LTM to decompose the sum into the partial sums 2 + 3 (tens) and 3 + 6 (units), for which the results are retrieved from LTM (5 and 9), and then combined (5 tens + 9 units) to produce the correct answer (59). In general, people prefer strategies in which they have confidence; as a result persons with a lower ability to solve arithmetic problems will more frequently use counting and simple transformation strategies whereas persons of higher ability will more frequently use complex transformation and retrieval strategies. Irrespective of the strategies used, an important question is whether WM is used to select the appropriate strategies from LTM, and how WM is involved in the application of these procedures to obtain the solution (for reviews, see DeStefano & LeFevre, 2004; LeFevre, DeStefano, Coleman, & Shanahan, 2005). Using the choice-no choice method (Siegler & Lemaire, 1997), strategy efficiency can be studied separately from strategy selection. Combined with different types of working memory load, it has been shown that strategy selection does not change under WM load, but strategy efficiency is poorer under a phonological and an executive (or domain-general) load (Imbo & Vandierendonck, 2007a, 2007b).

This represents only a few highlights of the research on the interaction of WM and mental arithmetic. The questions addressed in such research show that there really is no need to come up with a special theory of the intersection of WM and mental arithmetic, as the issues all relate to memory and memory usage. If a theory of working memory cannot explain—-or in the Benchmark Team’s view, is not expected to explain—-such basic findings, then we are far away from a cognition-based technology.

Not only is a joint theory of WM and mental arithmetic not needed, such a theory would not help much to advance either of the involved disciplines as the interaction of WM and mental arithmetic is only a subset of research on mental arithmetic, which in its turn is a subset of research on mathematical cognition that includes areas such a number representation, numerical skill acquisition, mathematical learning disabilities, calculation, and math anxiety to name a few (cf. Handbook of Mathematical Cognition, Campbell, 2005). In short, it seems preposterous as well as useless to require the development of a joint theory of WM and mental arithmetic, but that does not mean that the relevant findings should not be accounted for by any theory of working memory. Moreover, recognition of one or more benchmarks relevant to this intersection would also help to advance research on the involvement of WM in mental arithmetic and by extension, on other issues under the broader header of mathematical cognition. Also note that the Benchmark Team is not always consistent in the choices that are made: although they explicitly exclude benchmarks related to the interaction of WM and mental arithmetic, they accept a number of benchmarks based on complex span tasks. Two of these complex span tasks, the counting span task (Camos & Barrouillet, 2004; Case, 1985), and the operation span task (Engle, Cantor, & Carullo, 1992) are based on the interaction of WM with mental arithmetic.

Deductive reasoning. Like mental arithmetic, deductive reasoning is an ability that heavily relies on retrieving information from LTM. By combining this information, inferences are made as to whether a particular state of affairs can or cannot be true given the information in the premisses. This seems at least to be the case in those particular reasoning formats that can be easily handled by using mental models (Johnson-Laird, 1983, 2001; Johnson-Laird, Byrne, & Schaeken, 1992) and in reasoning using heuristics (Tversky & Kahneman, 1973, 1974). For example, in three-term syllogisms, such as for example “Robert is taller than Sally” and “Ted is smaller than Sally”, it has been shown that people work with spatial images that represent the given information (Huttenlocher, 1968): the persons are ordered by size from left to right: Robert, Sally, Ted, so that the answer can be retrieved directly from this imagination (namely, Robert is taller than Ted). Not all reasoning is that simple of course, but in all cases of deductive reasoning, the information given is represented in WM and relevant additional information is retrieved from LTM and combined with the information in WM (Johnson-Laird, 1983; Rips, 1994). In short, just as was the case for mental arithmetic there is no need to develop a separate theory that combines WM with deductive reasoning, especially because the number of publications that studied the role of WM in deductive reasoning is rather scarce (some of the earlier work is reviewed by Johnson-Laird, 1999) and concerns only a small subset of research on deductive reasoning.

Language Comprehension. Comprehension of a text or even of a single sentence involves a range of processes, such as word recognition, retrieval of phonological, lexical, semantic and syntactic word features, elaboration of the syntactical structure, assignment of a thematic role to the noun phrases, etc. Whether and to what extent working memory is involved in one or more of these subprocesses, is still a matter of debate among psycholinguists. Some researchers (e.g., Caplan & Waters, 1999) claim that the interpretive processes (parsing and syntactical analysis) run off more or less automatically without involvement of (verbal) working memory, whereas the post-interpretive processes (usage of the sentence meaning in other tasks) may require verbal memory support. Other researchers, in contrast, defend the position that working memory plays an important role in sentence comprehension (e.g., Fedorenko, Gibson, & Rohde, 2006; Frazier, 1987; Gibson, 1998; Just & Carpenter, 1992). At which stages of processing, working memory is involved in sentence comprehension may be a matter of debate, an important number of findings show that comprehension depends on domain-general working memory processes (for some reviews, see Fedorenko & Thompson-Schill, 2014; Jäger, Engelmann, & Vasishth, 2017; Ye & Zhou, 2008, 2009). However, that WM and language comprehension do interact is beyond doubt and is supported by the extensive usage of complex span tasks such as the word span (Daneman & Carpenter, 1980) and the listening span (Siegel, 1994) tasks. Some studies have shown that recall is more impaired when during the retention interval syntactically more difficult sentences have to be processed (e.g., Loncke, Desmet, Vandierendonck, & Hartsuiker, 2011), and thus can be added to benchmark 5.2.

The number of publications that address the interaction of working memory and language comprehension is large, and the topics addressed as well as the methodologies used vary widely. For these reasons, a review focusing on how WM and language processing affect each other would certainly be most welcome. Given the state of this subfield, it is not straightforward to summarize the findings in terms of benchmarks as collected by the Benchmark Team, and this seems to be a good reason not to include this part of research in the set of benchmarks, but it is certainly not the case that this would only be possible in the context of a theory of the interaction of WM and sentence comprehension, as claimed by the team.

Summary. This short and schematic overview of the role of WM in mental arithmetic, language comprehension and deductive reasoning shows that benchmarks are difficult to delineate at present for the role of WM in language comprehension because there is a wealth of publications using a broad range of techniques; a specific review would no doubt help to bring order in these findings. Obviously, such a reviewing activity falls outside the scope of the project undertaken by the Benchmark Team. However, for the domains of mental arithmetic and deductive reasoning, it would have been possible to include them within the benchmarking project. In all three cases, the reason cited by the Benchmark Team for excluding benchmarks on working memory interactions with these three domains does not seem to be relevant.

Working Memory and Executive Function

As already mentioned, the Benchmark Team also decided to exclude findings related to what they call “executive function”, which is not a very accurate label for findings on Stroop interference, task switching, and verbal fluency. Indeed, this is a rather odd collection of phenomena. Each of these areas listed will be considered in more detail in what follows, but first the issue of executive function is addressed, because the team uses this term as the overall label for these areas. Remarkably, the Benchmark Team does not provide any reference to research on executive functions (although they refer to research on cognitive control), not even to the seminal paper by Miyake et al. (2000). The latter paper defines three executive functions, namely set shifting, inhibition and updating as latent variables that represent inter-individual differences in executive functioning. Interestingly, the latent variable of *updating* is a composite score based on performance on three different memory updating tasks which play a prominent role in some of the benchmarks proposed by the Benchmark Team. The function of *set-shifting* refers to an updating of working memory specifically with respect to the current task set or mind set. This is exactly the kind of memory action Miller et al. (1960) referred to when they introduced the concept of working memory. Finally, *inhibition* is needed when particular memories are to be ignored or are no longer relevant and is present in the form of proactive interference also known as proactive inhibition in the benchmarks. What is the rationale of excluding findings with respect to executive functions, when these functions refer to processes performed on the contents of working memory, and when the tasks that are used to estimate at least one of these functions is one of the memory tasks on which the benchmarks are based? If not contradictory, this is at least confusing.

Stroop interference. It is interesting that the Benchmark Team refers to interference observed in all variations of Stroop tasks (Stroop, 1935), but it would be even better if they had also referred to a variation of other attention tasks such as the flanker task, the Simon task, negative priming, attentional blink and others. The evidence obtained in a range of studies of these tasks is consistent (for a review, see Vandierendonck, 2014), and shows that interference or conflict management effects in these tasks are larger in persons with lower working memory capacity (e.g., Kane & Engle, 2003; Morey et al., 2012) and in conditions with a higher WM load. When looking at the methodologies used (complex span tasks, Brown-Peterson tasks, tasks with a concurrent memory load), labeling these studies as investigations of executive function suggests the Benchmark Team failed to have a closer look at this part of the literature.

Task switching. For many researchers the role of working memory in task switching is evident. However, in their review of the task-switching literature, Vandierendonck, Liefooghe and Verbruggen (2010) showed that task switching relies on verbal and phonological WM representations to maintain and update the current plan, but that task switch costs and working memory load or working memory capacity do not seem to interact, except for the study by Liefooghe, Barrouillet, Vandierendonck and Camos (2008), which showed that recall is impaired when more task switches occur in the retention interval. The latter study is cited by the Benchmark Team as support for benchmark 5.2.4. Apart from this study, only the findings about the role of working memory as a support for controlling maintenance and updating of the current plan seem to be relevant. Shouldn't a theory of working memory be able to account for such a straightforward finding?

Yet, there is another issue in this context that deserves our attention. It concerns the findings reported by Logan (2004). Over a series of experiments, participants were given a list of task names such as hi-lo, odd-even, digit-word, for later recall and/or execution. After reading the list of names, the memory span was assessed as the number of names recalled in the correct order, and the task span was assessed as the number of tasks performed correctly in the correct order in response to stimuli (such as 1, four, seven, 3, … to which the remembered task was to be applied). Length of the sequence of task names was varied and order of testing counterbalanced. For each participant the span was estimated as the list length at which the complete sequence was correct on 50% of the trials. Estimated memory span and task span did not differ, and recall performance did not differ as a function of the number (or proportion) of task switches in the list. Assuming that planning and execution of the tasks taxes the same

WM resources as those required to remember the task names, a trade-off between maintenance of the task names and execution of the tasks was expected. No such trade-offs were observed. Although, this is a negative finding, in combination with the observation that verbal working memory supports task switching, this lack of trade-off between maintenance and performance suggests that as far as working memory is involved, and it seems to be, different sets of resources are required for the two components. In other words, a single-resource WM theory will fail to account for such observations. In my view, that deserves to be a benchmark and is a good reason for not excluding this small set of studies from the benchmark enterprise.

Verbal fluency. Finally, within this set of excluded “executive function” phenomena, it is quite surprising to see a mention of verbal fluency. Verbal fluency tasks are indeed sometimes used in a neuropsychological context to assess executive function capacities of a patient (e.g., Phillips, 1997). However, verbal fluency refers to a cognitive ability to retrieve a large number of words satisfying a given criterion in a short period of time; for instance, nouns starting with the letter N, or mammals. Such tasks clearly require fast but controlled retrieval: as many different elements as possible have to be retrieved from categorical/semantic LTM and some checking is required to avoid repetitions or perseverations of already recited elements and to adapt the retrieval plan or the retrieval strategies accordingly. Without performing an exhaustive search of the literature, several relevant publications are available in support of a relationship between WM and verbal fluency, showing that high WM capacity persons are more efficient in their use of retrieval strategies and consequently achieve better fluency performance than low WM individuals (Rosen & Engle, 1997; Unsworth, 2017; Unsworth, Brewer, & Spillers, 2013; Unsworth, Spillers, & Brewer, 2011) and that perseveration is increased under a working memory load (Azuma, 2004).

Summary. A closer look at the findings regarding Stroop interference (extended here to other attention tasks), task switching and verbal fluency shows that there is absolutely no good reason to exclude these areas of research from the working memory benchmarks. In fact, in all three areas sufficiently high quality studies of the role of working memory in these task settings are available and should be included in the list of benchmarks.

Miscellaneous Oversights

Thus far the focus was on selectivity in the set of benchmarks based on a deliberate decision of the Benchmark Team to exclude some categories of findings. However, the benchmarks also suffer from a selectivity that is not announced. In the present section, I will address some findings that I missed in the paper; this concerns findings that may help to broaden the scope of some of the benchmark or findings that do not seem to fit at all in any of the proposed benchmarks. This overview is not intended to be exhaustive; no comprehensive search of the literature was made to find out what is missing. On the contrary, it rather concerns a few blind spots that struck me and that aroused my concern.

Already in the first benchmark elaborated by the team, namely the size effects, it is so strange that apart from findings based on the change detection task, only findings with verbal content are addressed. Yet, a lot of research has been performed on limitations in visuospatial processing and representation. In particular with respect to Benchmark 1.3, an important number of studies using either the Corsi blocks task (Berch, Krikorian, & Huha, 1998; Corsi, 1972), the Dots task (Jones, Farrand, Stuart, & Morris, 1995), and the Visual Patterns Test (Wilson, Scott, & Power, 1987) have shown that the number of spatial positions that can be remembered in correct order is limited (Ball, Pearson, & Smith, 2013; Brown, 2016; Cornoldi & Mammarella, 2008, 2011; Davis, Rane, & Hiscock, 2013; Hamilton, 2011; Logie, 1986, 1989, 1995, 2011; Logie & Pearson, 1997; Logie, Zucco, & Baddeley, 1990; Pickering, Gathercole, Hall, & Lloyd, 2001; Tremblay, Saint-Aubin, & Jalbert, 2006; Vandierendonck, Kemps, Fastame, & Szmalec, 2004; Vecchi & Cornoldi, 1999; Weicker, Hudl, & Thone-Otto, 2017; Zimmer & Liesefeld, 2011).

Also Benchmark 2 (2.1 - 2.4) on retention interval and presentation duration is completely concerned with verbal materials, as if there are no findings available about retention intervals filled with visuospatial materials (Brooks, 1968; Jones et al., 1995; Logie, 1995; Logie et al., 1990; Vergauwe, Barrouillet, & Camos, 2009) or about interactions of forgetting with different types of distractor materials in the spatial domain (e.g., Baddeley, Grant, Wight, & Thomson, 1975; Farmer, Berman, & Fletcher, 1986; Pearson & Sahraie, 2003; Smyth & Scholey, 1994).

Similarly, the changing state effect (Benchmark 6.2) is only mentioned for the verbal domain, although several studies have examined the effect of changing states also in the visuospatial domain by means of matrix tapping (e.g., Jones et al., 1995; Vandierendonck et al., 2004), eye movements (Pearson & Sahraie, 2003) or other movements (Smyth & Scholey, 1992), irrelevant pictures (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999; Logie, 1986; Logie & Marchetti, 1991; Quinn & McConnell, 1996), and dynamic visual noise (Darling, Della Sala, & Logic, 2009; Dent, 2010; McConnell & Quinn, 2000; Quinn & McConnell, 1996, 1999; Rossi-Arnaud, Cortese, & Cestari, 2004; Zimmer, Speiser, & Seidler, 2003), although effect boundaries have also been reported (Andrade, Kemps, Werniers, May, & Szmalec, 2002; Avons & Sestieri, 2005).

Benchmark 7 is another interesting case where the Benchmark Team seems to be interested only in the verbal domain. It concerns the syllable-based word length effect. Although syllables do not occur in the visuospatial domain, it may be considered that syllable-length expresses some form of complexity of the presented materials. In a search for parallels between the verbal and the visuospatial domain, effects of complexity have also been studied in the visuospatial domain; but again this does not seem to be known to the Benchmark Team. In particular, using Corsi blocks tasks and the Dots task, effects of path complexity and good continuation (Kemps, 1999, 2000, 2001; McConnell & Quinn, 2004), of path length and path crossings (Guerard, Tremblay, & Saint-Aubin, 2009; Parmentier, 2011; Parmentier, Elford, & Maybery, 2005), of symmetry (Pieroni, Rossi-Arnaud, & Baddeley, 2011; Rossi-Arnaud, Pieroni, & Baddeley, 2006; Rossi-Arnaud, Pieroni, Spataro, & Baddeley, 2012), and of spatial clustering (De Lillo, 2004; De Lillo, Kirby, & Poole, 2016; De Lillo & Lesk, 2010) have been demonstrated. By and large, these effects are similar to the syllable-based word length effect and subsuming the complexity effects in both domains jointly, a benchmark could be defined that is qualified with an A rating because it then no longer holds that the effect is only present in one single domain.

Utility of Benchmarks

The second main issue addressed in this commentary concerns the question whether it is methodologically and rationally sound to use the proposed set of benchmarks for judgments about the adequacy of new theories of short-term and working memory. Indeed any theory, not only a new one, should preferably account for all of the relevant observations available, and if it does account for all these findings, it should give a better account than its competitors. Moreover, if possible, new theories are also expected to predict new phenomena. Hence, which one of two WM theories should be preferred depends on comparisons based on several criteria, such as how well each theory accounts for the benchmarks, how well each theory accounts for other known findings, and to what extent it predicts new phenomena. One additional difficulty in this process concerns the degree of theory-independence of findings and benchmarks. In what follows, each of these concerns is discussed in more detail.

Theory-independence of the Findings

Empirical findings are sometimes closely linked to the theory that generated the prediction of these findings. The Benchmark Team is clearly aware of this limitation as appears from the following quote, “To serve its purpose, the set of benchmarks needs to be as unbiased and theory-neutral as possible” (Oberauer et al., this issue, p. 5). They further claim to “aim for a description of each benchmark that is not biased in favor or against one contemporary theoretical view” (p. 5). As will become clear from a few examples, this commitment is not sufficient to guarantee theory-neutrality. One obvious example concerns a finding which is present in all introductory textbooks, but which does not figure in the benchmarks, namely the finding that short-term memory can be largely improved by using mnemonics, such as the peg-word mnemonic. Although this observation is not strictly linked to any memory theory, it was probably ignored due to a too strict focus on published findings supported by theories. It would seem that the most obvious finding which is not at all linked to any theory was ignored in the effort towards theory-neutrality.

Avoidance of bias is very difficult when some findings are inherently accounted for by some theories but not by others. For example, some theories distinguish working memory modules for representation of modality-specific information (e.g., Baddeley & Hitch, 1974) while others explicitly do not (e.g., Cowan, 1999). Recognition of findings that specifically support differences in storage and processing between verbal and visuospatial modalities can be seen as a bias towards the former theories; however, keeping silent about such differences to avoid such bias, evidently implies a bias towards the latter type of theories. Similar problems arise with theories that assume that executive control processes are part of the operation of working memory (e.g., Engle, 2002) while other theories postulate executive control to be part of the processing system outside working memory (e.g., Oberauer, 2009). Again a decision to include executive control effects in the benchmarks is biased towards the former type of theory, whereas excluding these is a bias towards the latter type of theory. Working memory theories differ from each other in many assumptions (see e.g., Miyake & Shah, 1999, who focused on differences between theories), implying that there are many occasions to introduce bias by including or not including particular findings. Besides, from the biases which were already documented in the first part of the present commentary, it is clear that the Benchmark Team is faced with a task which is impossible to solve, because any decision they take against or in favor of particular assumption suggests a bias, unless they would have really succeeded in including such findings with a formulation that is acceptable for adepts of both types of theory.

In short, although the Benchmark Team may have honestly tried to compose a set of theory-neutral benchmarks, it seems that they did not completely succeed in doing so, because decisions about including or excluding particular findings are not theory-neutral. It could be argued, though, that the team performed a survey among a large number of researchers active in the field, which should normally result in more theory-neutrality. Unfortunately, these researchers usually also favor particular theories, and without any effort towards making these beliefs explicit, there is no way to check on potential conflicts of interest or potential biases. Hence, notwithstanding the efforts on the side of the Benchmark Team, it cannot be excluded that bias may be present in the formulation of some of the benchmarks. However, the locus of potential biases is most obvious in the absence of particular findings from the benchmark set.

Validity of the Benchmark Set

Related to the issue of bias, the question must also be considered whether the set of benchmarks completely and exhaustively covers the domain of study. In practical terms, are all the known findings regarding the subject at hand, in the present case, short-term and working memory, represented in the benchmarks? This is a concern of construct validity that can be approximated by an assessment of the content validity. Clearly, as discussed in the first part of the present commentary, some areas of short-term and working memory content have been deliberately left out and some other areas were missed by the Benchmark Team. As a result the answer to the question of construct/content validity is clear: the benchmarks do not yield a complete coverage of the contents. Even if one is willing to accept that completeness is not feasible, the benchmarks are not representative of the domain, because there is selectivity is in the exclusions.

This shortcoming in the validity of the set of benchmarks implies that a correct evaluation of a theory is difficult to achieve. First, consider the comparison of two theories: one that accounts for some part of the benchmarks in the set, and another one that accounts for fewer benchmarks but that additionally accounts for some findings that have been excluded from the benchmarks. Having knowledge only about the benchmarks in the set (or only accepting the benchmarks in the proposed set), the former theory will be judged as the better one. Whether that conclusion is correct depends on the importance of the benchmarks accounted for by the second theory. The situation could even be worse, as in a comparison of a theory that accounts for some benchmarks and another one that accounts for findings that were excluded from the benchmark set. By all means, the Benchmark Team will conclude that the former theory is better because it accounts for some of the benchmarks, while the latter theory does not account for any single benchmark. Again, a conclusion is reached that is not trustworthy. Without agreement about the construct validity of the benchmarks, assessment of the comparative value of theories cannot be trusted.

One important side-effect of excluding almost all findings regarding the interaction of WM with other cognitive tasks, such as mental arithmetic, language comprehension and deductive reasoning, but also attention, cognitive control, task switching and verbal fluency is that the benchmarks favor theories that are designed to account for short-term memory and working memory findings that can occur in isolation of any other cognitive activity. In other words, new formal models with a limited scope can easily account for an important subset of the benchmarks, and it may be expected, therefore, that these benchmarks will invite more of that type of modeling. The question is whether encouragement of this route of theoretical development is what we really need to advance our knowledge and understanding of short-term and working memory. Considering that the notion of working memory was introduced to enable studies of the link between memory and action in other cognitive tasks, following that particular route of theory development may lead us away from a better understanding of working memory. Instead of avoiding theoretical advances regarding the links of WM with other cognitive domains, an encouragement of developing unified theories of cognition seems to be the future for a better understanding of these phenomena. The proposed set of benchmarks thus rather seems to miss the opportunity for taking the route to better and more encompassing theories.

Using Benchmarks to Evaluate Theories

The Benchmark Team has collected these benchmarks as a tool to support comparative evaluation of theories. However, irrespective of whether one or more of the benchmarks is theory-biased and irrespective of whether the set of benchmarks is exhaustive or representative of the field, using the benchmarks to evaluate theories is not a simple and straightforward exercise. First, not all benchmarks are equally important. The Benchmark Team addressed this issue by assigning priority ratings (A, B, or C) to each benchmark, based on the results of the survey and an estimate of the scope of the findings represented in the benchmarks. However, to some extent importance of a benchmark also depends on theoretical considerations as some benchmarks are more important for one theory than for another. Second, benchmarks are not static. New findings will result in changes in the set of benchmarks and as already suggested in the first part of this commentary and in the evaluation of their validity, the set of benchmarks cannot be considered to be complete or exhaustive at present, so that even today there is already a need for including more findings in the set of benchmarks.

Comparison of two (or more) theories that account for overlapping but incomplete subsets of the benchmarks, is bound to be quite a difficult undertaking. On what basis can a decision be made that one theory's account is better than that of another theory. It clearly does not suffice to count the number of benchmarks each theory covers, because some benchmarks are more important than others. It also makes a difference, if the evaluation is made today or after the next update of the set of benchmarks. In fact, the existence of an explicitly agreed upon set of benchmarks does not make theory evaluation less complex than without benchmarks. Even more so, when one takes into account that theories may also differ in their scope: some theories will be addressing for example verbal short-term memory whereas other theories may be focusing on the interrelations of short-term memory with other domains of cognition. Evaluation becomes even more tricky when it is taken into account that the benchmarks are not completely theory-neutral and that they also are not completely representative for the scope of short-term and working memory theories.

Conclusion

The delineation of the concept of working memory in the introductory paragraph of the benchmark paper (Oberauer et al., this issue) leaves no doubt that working memory pervades almost all cognitive processes and is completely intertwined with cognition. In other words, working memory is not just a kind of memory that operates in isolation. Benchmarks based on studies of working memory in isolation, therefore, cannot do justice to the core ideas that are the basis of the working memory concept. Such memory-in-isolation benchmarks are only useful to constrain theories about short-term memory and about isolated aspects of working memory and are therefore not valid as a representation of the field of short-term and working memory findings. As working memory is so closely tied up with cognition, what is really needed to advance the field is a theory that accounts for working memory in relation to cognition; in other words, a unified theory of cognition. Clearly, the team missed an important opportunity to encourage such more integrative theory development.

References

Andrade, J., Kemps, E., Werniers, Y., May, J., & Szmalec, A. (2002). Insensitivity of visual short-term memory to irrelevant visual information. *Quarterly Journal of Experimental Psychology, 55A*, 753-774. doi: 10.1080/02724980143000541

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75-106.

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 527-538.

Avons, S. E., & Sestieri, C. (2005). Dynamic visual noise: No interference with visual short-term memory or the construction of visual images. *European Journal of Cognitive Psychology, 17*(3), 405-424. doi: 10.1080/09541440440000104

Azuma, T. (2004). Working memory and perseveration in verbal fluency. *Neuropsychology, 18*(1), 69-77. doi: 10.1037/0894-4105.18.1.69

Baddeley, A. D., Grant, S., Wight, E., & Thomson, N. (1975). Imagery and visual working memory. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 205-217). London: Academic Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Ball, K., Pearson, D. G., & Smith, D. T. (2013). Oculomotor involvement in spatial working memory is task-specific. *Cognition, 129*(2), 439-446. doi: 10.1016/j.cognition.2013.08.006

Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition, 38*(3), 317-338. doi: 10.1006/brcg.1998.1039

Brooks, L. R. (1968). Spatial and verbal components in the act of recall. *Canadian Journal of Psychology, 22*, 349-368.

Brown, L. A. (2016). Spatial-Sequential Working Memory in Younger and Older Adults: Age Predicts Backward Recall Performance within Both Age Groups. *Frontiers in Psychology, 7*. doi: 10.3389/fpsyg.2016.01514

Camos, V., & Barrouillet, P. (2004). Adult counting is resource demanding. *British Journal of Psychology, 95*, 19-30.

Campbell, J. I. D. (1987). Network interference and mental multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 109-123.

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*, 121-164.

Campbell, J. I. D. (2005). Handbook of Mathematical Cognition. New York: Psychology Press.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-126.

Case, R. (1985). Intellectual development: Birth to adulthood. New York: Academic Press.

Cornoldi, C., & Mammarella, I. C. (2008). A comparison of backward and forward spatial spans. *Quarterly Journal of Experimental Psychology, 61*(5), 674-682. doi: 10.1080/17470210701774200

Cornoldi, C., & Mammarella, I. C. (2011). The organization of visuospatial working memory: Evidence from the study of developmental disorders. In A. Vandierendonck & A. Szmalec (Eds.), *Spatial Working Memory* (pp. 102-121). Hove: Psychology Press.

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain.* (Doctoral dissertation), McGill University, Montreal, Canada.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory. Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466. doi: 10.1016/S0022-5371(80)90312-6

Darling, S., Della Sala, S., & Logic, R. H. (2009). Dissociation between appearance and location within visuo-spatial working memory. *Quarterly Journal of Experimental Psychology, 62*(3), 417-425. doi: 10.1080/17470210802321984

Davis, L. C., Rane, S., & Hiscock, M. (2013). Serial recall of visuospatial and verbal information with and without material-specific interference: Implications for contemporary models of working memory. *Memory, 21*(7), 778-797. doi: 10.1080/09658211.2012.756037

De Lillo, C. (2004). Imposing structure on a Corsi-type task: Evidence for hierarchical organisation based on spatial proximity in serial-spatial memory. *Brain and Cognition, 55*(3), 415-426. doi: 10.1016/j.bandc.2004.02.071

De Lillo, C., Kirby, M., & Poole, D. (2016). Spatio-Temporal Structure, Path Characteristics, and Perceptual Grouping in Immediate Serial Spatial Recall. *Frontiers in Psychology, 7*. doi: 10.3389/fpsyg.2016.01686

De Lillo, C., & Lesk, V. E. (2010). Spatial clustering and hierarchical coding in immediate serial recall. *European Journal of Cognitive Psychology, 22*(2), 216-246. doi: 10.1080/09541440902757918

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo-spatial memory. *Neuropsychologia, 37*(10), 1189-1199. doi: 10.1016/S0028-3932(98)00159-6

Dent, K. (2010). Dynamic Visual Noise Affects Visual Short-Term Memory for Surface Color, but not Spatial Location. *Experimental Psychology, 57*(1), 17-26. doi: 10.1027/1618-3169/a000003

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353-386. doi: 10.1080/09541440244000328

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23. doi: 10.1111/1467-8721.00160

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972-992. doi: 10.1037//0278-7393.18.5.972

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *Quarterly Journal of Experimental Psychology, 38A*, 675-688. doi: 10.1080/14640748608401620

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language, 54*(4), 541-553. doi: 10.1016/j.jml.2005.12.006

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences, 18*(3), 120-126. doi: 10.1016/j.tics.2013.12.006

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII* (pp. 559-586). Hillsdale, NJ: Erlbaum.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1-76.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329-343.

Guerard, K., Tremblay, S., & Saint-Aubin, J. (2009). The processing of spatial information in short-term memory: Insights from eye tracking the path length effect. *Acta Psychologica, 132*(2), 136-144. doi: 10.1016/j.actpsy.2009.01.003

Hamilton, C. (2011). The nature of visuospatial representation within working memory. In A. Vandierendonck & A. Szmalec (Eds.), *Spatial working memory* (pp. 122-144). Hove: Psychology Press.

Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review, 75*, 550-560.

Imbo, I., & Vandierendonck, A. (2007a). Do multiplication and division strategies rely on executive and phonological working-memory resources? *Memory & Cognition, 35*, 1759-1784.

Imbo, I., & Vandierendonck, A. (2007b). The role of phonological and executive working memory resources in simple arithmetic strategies. *European Journal of Cognitive Psychology, 19*(6), 910-933. doi: 10.1080/09541440601051571

Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language, 94*, 316-339. doi: 10.1016/j.jml.2017.01.004

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.

Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual Review of Psychology, 50*, 109-135. doi: 10.1146/annurev.psych.50.1.109

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, 5*, 434-442.

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99*, 173-182.

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(4), 1008-1018. doi: 10.1037/0278-7393.21.4.1008

Just, M. A., & Carpenter, G. A. (1992). A capacity theory of comprehension: Individual-differences in working memory. *Psychological Review, 99*, 122-149. doi: 10.1037/0033-295X.99.1.122

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to

Stroop interference. *Journal of Experimental Psychology: General, 132*(1), 47-70. doi: 10.1037/0096-3445.132.1.47

Kemps, E. (1999). Effects of complexity on visuo-spatial working memory. *European Journal of Cognitive Psychology, 11*(3), 335-356. doi: 10.1080/713752320

Kemps, E. (2000). Structural complexity in visuo-spatial working memory. *Current Psychology Letters, 1(3)*, 59-70.

Kemps, E. (2001). Complexity effects in visuo-spatial working memory: Implications for the role of long-term memory. *Memory, 9*(1), 13-27. doi: 10.1080/09658210042000012

LeFevre, J. A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General, 125*(3), 284-306. doi: 10.1037/0096-3445.125.3.284

LeFevre, J. A., DeStefano, D., Coleman, B., & Shanahan, T. (2005). Mathematical cognition and working memory. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 361-377). New York: Psychology Press.

LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 216-230.

Liefooghe, B., Barrouillet, P., Vandierendonck, A., & Camos, V. (2008). Working memory costs of task switching. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*(3), 478-494. doi: 10.1037/0278-7393.34.3.478

Logan, G. D. (2004). Working memory, task switching, and executive control in the task span procedure. *Journal of Experimental Psychology: General, 133*(2), 218-236. doi: 10.1037/0096-3445.133.2.218

Logie, R. H. (1986). Visuo-spatial processes in working memory. *Quarterly Journal of Experimental Psychology, 38A*, 229-247. doi: 10.1080/14640748608401596

Logie, R. H. (1989). Characteristics of visual short-term memory. *European Journal of Cognitive Psychology, 1*, 275-284.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Logie, R. H. (2011). The visual and the spatial of a multicomponent working memory. In A. Vandierendonck & A. Szmalec (Eds.), *Spatial Working Memory* (pp. 19-45). Hove: Psychology Press.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 105-115). Amsterdam: Elsevier.

Logie, R. H., & Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: Evidence from developmental fractionation. *European Journal of Cognitive Psychology, 9*(3), 241-257. doi: 10.1080/713752559

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica, 75*(1), 55-74. doi: 10.1016/0001-6918(90)90066-o

Loncke, M., Desmet, T., Vandierendonck, A., & Hartsuiker, R. J. (2011). Executive control is shared between sentence processing and digit maintenance: Evidence from a strictly timed dual-task paradigm. *Journal of Cognitive Psychology, 23*(7), 886-911. doi: 10.1080/20445911.2011.586625

McConnell, J., & Quinn, J. G. (2000). Interference in visual working memory. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53(1), 53-67. doi: 10.1080/027249800390664

McConnell, J., & Quinn, J. G. (2004). Complexity factors in visuo-spatial working memory. *Memory, 12*(3), 338-350. doi: 10.1080/09658210344000035

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Henry Holt and Company.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Miyake, A., & Shah, P. (1999). Models of working memory. Mechanisms of active maintenance and executive control. Cambridge: Cambridge University Press.

Morey, C. C., Elliott, E. M., Wiggers, J., Eaves, S. D., Shelton, J. T., & Mall, J. T. (2012). Goal-neglect links Stroop interference with working memory capacity. *Acta Psychologica, 141*(2), 250-260. doi: 10.1016/j.actpsy.2012.05.013

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), Psychology of Learning and Motivation: Advances in Research and Theory, Vol 51 (Vol. 51, pp. 45-100). San Diego: Elsevier Academic Press Inc.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., . . . Ward, G. (this issue). Benchmarks for Models of Short Term and Working Memory. *Psychological Bulletin*.

Parmentier, F. B. R. (2011). Exploring the determinants of memory for spatial sequences. In A. Vandierendonck & A. Szmalec (Eds.), *Spatial Working Memory* (pp. 67-86). Hove: Psychology Press.

Parmentier, F. B. R., Elford, G., & Maybery, M. (2005). Transitional information in spatial serial memory: Path characteristics affect recall performance. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*(3), 412-427. doi: 10.1037/0278-7393.31.3.412

Pearson, D. G., & Sahraie, A. (2003). Oculomotor control and the maintenance of spatially and temporally distributed events in visuo-spatial working memory. *Quarterly Journal of Experimental Psychology, 56A*, 1089-1111. doi: 10.1080/02724980343000044

Phillips, L. H. (1997). Do ``frontal tests'' measure executive function? Issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 191-213). Hove: Psychology Press.

Pickering, S. J., Gathercole, S. E., Hall, M., & Lloyd, S. A. (2001). Development of memory for pattern and path: Further evidence for the fractionation of visuo-spatial memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54*(2), 397-420. doi: 10.1080/02724980042000174

Pieroni, L., Rossi-Arnaud, C., & Baddeley, A. D. (2011). What can symmetry tell us about working memory? In A. Vandierendonck & A. Szmalec (Eds.), *Spatial Working Memory* (pp. 145-158). Hove: Psychology Press.

Quinn, J. G., & McConnell, J. (1996). Irrelevant pictures in visual working memory. *Quarterly Journal of Experimental Psychology, 49A*, 200-215. doi: 10.1080/027249896392865

Quinn, J. G., & McConnell, J. (1999). Manipulation of interference in the passive visual store. *European Journal of Cognitive Psychology, 11*(3), 373-389. doi: 10.1080/713752322

Rips, L. J. (1994). Deduction and its cognitive basis. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 150-178). New York: Academic Press.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126*(3), 211-227. doi: 10.1037/0096-3445.126.3.211

Rossi-Arnaud, C., Cortese, A., & Cestari, V. (2004). Memory span for movement configurations: The effects of concurrent verbal, motor and visual interference. *Cahiers De Psychologie Cognitive-Current Psychology of Cognition, 22*(3), 335-349.

Rossi-Arnaud, C., Pieroni, L., & Baddeley, A. (2006). Symmetry and binding in visuo-spatial working memory. *Neuroscience, 139*(1), 393-400. doi: 10.1016/j.neuroscience.2005.10.048

Rossi-Arnaud, C., Pieroni, L., Spataro, P., & Baddeley, A. (2012). Working memory and individual differences in the encoding of vertical, horizontal and diagonal symmetry. *Acta Psychologica, 141*(1), 122-132. doi: 10.1016/j.actpsy. 2012.06.007

Siegel, L. S. (1994). Working Memory and Reading: A Life-Span Perspective. *International Journal of Behavioral Development*, 109-124.

Siegler, R. S., & Lemaire, P. (1997). Older and Younger Adults' Strategy Choices in Multiplication: Testing Predictions of ASCM Using the Choice/No-Choice Method. *Journal of Experimental Psychology: General, 126*, 71-92.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In G. Halford & T. Simon (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Erlbaum.

Smyth, M. M., & Scholey, K. A. (1992). Determining spatial span: The role of movement time and articulation rate. *Quarterly Journal of Experimental Psychology, 45A*, 479-501.

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition, 22*, 1-13. doi: 10.3758/BF03202756

Stroop, J. R. (1935). Studies of inteference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662. doi: 10.1037/h0054651

Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review, 13*(3), 452-457. doi: 10.3758/bf03193869

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207-232. doi: 10.1016/0010-0285(73)90033-9

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131. doi: 10.1126/science.185.4157.1124

Unsworth, N. (2017). Examining the dynamics of strategic search from long-term memory. *Journal of Memory and Language, 93*, 135-153. doi: 10.1016/j.jml.2016.09.005

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: the role of controlled search. *Memory & Cognition, 41*(2), 242-254. doi: 10.3758/s13421-012-0261-x

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology, 64*(3), 447-466. doi: 10.1080/17470218.2010.505292

Vandierendonck, A. (2014). Symbiosis of executive and selective attention in working memory. *Frontiers in Human Neuroscience, 8*. doi: 10.3389/fnhum.2014.00588

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology, 95*, 57-79. doi: 10.1348/000712604322779460

Vandierendonck, A., Liefooghe, B., & Verbruggen, G. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin, 136*(4), 601-626. doi: 10.1037/a0019791

Vecchi, T., & Cornoldi, C. (1999). Passive storage and active manipulation in visuo-spatial working memory: Further evidence from the study of age differences. *European Journal of Cognitive Psychology, 11*(3), 391-406. doi: 10.1080/713752324

Vergauwe, E., Barrouillet, P., & Camos, V. (2009). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 1012-1028. doi: 10.1037/a0015859

Weicker, J., Hudl, N., & Thone-Otto, A. (2017). Underlying Cognitive Processes of the Corsi Block-Tapping Task. *Zeitschrift Fur Neuropsychologie, 28*(1), 45-54. doi: 10.1024/1016-264X/a000194

Wilson, J. T. L., Scott, J. H., & Power, K. G. (1987). Developmental differences in the span of visual memory for pattern. *British Journal of Developmental Psychology, 5*, 249-255. doi: 10.1111/j.2044-835X.1987.tb01060.x

Ye, Z., & Zhou, X. L. (2008). Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research, 1203*, 103-115. doi: 10.1016/j.brainres.2008.01.090

Ye, Z., & Zhou, X. L. (2009). Executive control in language processing. *Neuroscience and Biobehavioral Reviews, 33*(8), 1168-1177. doi: 10.1016/j.neubiorev. 2009.03.003

Zimmer, H. D., & Liesefeld, H. R. (2011). Spatial information in (visual) working memory. In A. Vandierendonck & A. Szmalec (Eds.), *Spatial Working Memory* (pp. 46-66). Hove: Psychology Press.

Zimmer, H. D., Speiser, H. R., & Seidler, B. (2003). Spatio-temporal working-memory and short-term object-location tasks use different memory mechanisms. *Acta Psychologica, 114*(1), 41-65. doi: 10.1016/s0001-6918(03)00049-0