Distinct Neural Networks Underpin Distractor Sensitivity and Task Accuracy in a Complex Working Memory Task

Malayka Mottarella

A thesis
Submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington
2022

Committee:
Chantel Prat
Scott Murray

Program Authorized to Offer Degree:
Psychology

University of Washington

Abstract

Distinct Neural Networks Underpin Distractor Sensitivity and Task Accuracy in a Complex Working Memory Task

Malayka Mottarella

Chair of the Supervisory Committee:
Chantel Prat
Department of Psychology

It is unclear what role distractor sensitivity plays in performance on complex working memory tasks. Although several contemporary studies suggest that individual differences in working memory performance are driven by the ability to successfully resist distractors, the relation between these two processes is primarily observed when both measures are computed from performance on the same visuospatial task. The limited existing research relating distractor sensitivity to external complex span measures has provided mixed results. In the present study we use a higher-level semantic filtering neuroimaging paradigm with a visuospatial component to explore the relation between distractor sensitivity and performance on complex working memory tasks. Results indicated that individual differences in distractor sensitivity are underpinned by a distributed network of domain-general cognitive control – including the putamen and anterior cingulate cortex – and language processing areas such as the left inferior frontal gyrus, which were non-overlapping with the more limited regions associated with complex span performance. Behavioral results showed further evidence that distractor sensitivity and performance on complex working memory tasks were not related to one another. These findings provide novel contributions to our understanding of what drives individual differences in performance on complex working memory tasks. When viewed in light of the previous findings, we propose that both “distractor sensitivity” and “working memory capacity” are multifaceted constructs whose neurocognitive underpinnings vary as a function of task demands.

Distinct Neural Networks Underpin Distractor Sensitivity and Task Accuracy in a Complex Working Memory Task

Working memory is a core constraint of human information processing that is often described as a "bottleneck" for complex cognition. While many more specific definitions have been proposed, one theme that is shared across theoretical perspectives is that working memory capacity varies between individuals (*e.g.,* Baddeley, 2012; Cowan, 2014; Engle & Kane, 2004; Just & Carpenter, 1992). These individual differences, assessed by performance on a variety of working memory tasks, have been shown to predict variability in a wide array of contexts ranging from specific cognitive skills such as reading ability (*e.g.,* Daneman & Carpenter, 1980; Ferreira & Clifton, 1986; Just & Carpenter, 1992; Prat & Just, 2011) to more general fluid information processing capabilities such as performance on fluid intelligence tests (*e.g.,* Engle et al., 1999). Individual differences in working memory capacity also relate to non-cognitive processes such as emotion regulation (*e.g.,* Schmeichel et al., 2008) and predict real-world outcomes such as academic success (*e.g.,* Bull et al., 2008; Siegler et al., 2012). In short, whatever the driver or drivers of individual differences in working memory capacity may be, their implications are far reaching. Despite this, the underlying mechanisms that give rise to individual differences in performance on working memory tasks remain debated.

A major challenge in assessing what drives individual differences in performance is that working memory tasks rely on multiple sub-component processes such as attentional filtering, target storage, and goal maintenance. The processes work in tandem, often via overlapping timescales, to produce successful task performance. Thus, it can be difficult to determine the relative contributions of specific processes to task performance. For instance, someone who is able to maintain more information in working memory may have a larger capacity or they may be better able to filter out distracting information and store *only* information relevant to the task at hand. Neuroimaging paradigms provide a unique opportunity to identify the relative contributions of these processes to working memory performance by leveraging an understanding of how they are differentially deployed across time and space in the brain. In the present study, we adopt such an approach to systematically investigate the neural mechanisms underlying individual differences in working memory task performance.

Theories on the nature of individual differences in working memory

Research on the nature of individual differences in working memory has focused on two separable types of mechanisms: storage (*i.e.*, the temporary maintenance of information) and information manipulation (*i.e.*, the updating, reordering, or filtering of temporarily stored information). Capacity-based theories argue that individual differences in performance on working memory tasks stem from intrinsic differences in the amount of information an individual can hold in an active state at any given time (*e.g.*, Just & Carpenter, 1992). One conceptualization of capacity-based theories defines capacity as a series of “slots” in which information is temporarily stored (*e.g.*, Luck & Vogel, 2013). Accordingly, individual differences in capacity arise because some individuals have more slots available for storage than others. A more fluid conceptualization of capacity is that both storage and information processing tax the same pool of finite cognitive resources, such that storage ability is constrained as processing demands increase (Just & Carpenter, 1992). From this view, capacity can be defined as the “size” of this pool of resources such that individual differences in working memory result from some individuals having a larger resource pool to draw on than others. This theory is supported by a series of studies demonstrating the high-working-memory capacity individuals, who are presumed to have a larger pool of cognitive resources, are better able to perform competing or highly demanding cognitive tasks (*e.g.,* Daneman & Merikle, 1996; Just & Carpenter, 1992).

More contemporary theories have suggested that successful working memory performance is driven by *distractor sensitivity*, or the ability to discriminate between task-relevant and task-irrelevant information (*e.g.,* Engle, 2002; Engle & Kane, 2004; Vogel et al., 2005). From this view, individual differences in working memory capacity are not a reflection of inherent differences in the *amount* of resources an individual has, but rather a reflection of differences in the *strategic allocation* of resources towards relevant information and away from distracting information (*e.g.,* Awh et al., 2006; Fukuda & Vogel, 2009). Prior work using event-related potentials (ERP) and functional magnetic resonance imaging (fMRI) supports this theory by suggesting that higher-capacity individuals are less likely to store distractor items in working memory than lower-capacity individuals (*e.g.,* McNab & Klingberg, 2008; Vogel et al., 2005).

One complication with integrating the findings from these different viewpoints is that studies advocating for attentional filtering as the driver of individual differences in working memory performance tend to use visuospatial change detection tasks (*e.g.,* McNab & Klingberg, 2008; Vogel et al. 2005). In these paradigms, participants are shown a visual array of objects, commonly colored rectangles, and then after a brief retention interval are asked to judge whether the visual features of a probe are consistent with what they saw previously. Notably, different measures obtained from this same task are commonly used to index both distractor sensitivity ability and working memory capacity (*e.g.,* McNab & Klingberg, 2008; Vogel et al., 2005). This is potentially problematic as task-dependent characteristics may underlie the observed relation between working memory performance and distractor sensitivity (see Mall et al., 2014; Oberauer, 2019; and Robinson et al., 2018 for discussions).

Relating distractor sensitivity to performance on complex span tasks

Two recent studies have examined the relation between distractor sensitivity and performance on external measures of working memory capacity, namely complex span tasks (Mall et al., 2014; Robinson et al., 2018). Complex span tasks are one of the most predominant ways to measure working memory. In these tasks, participants alternate or “juggle” between processing information online (*e.g.*, solving a math equation) and storing to-be-remembered items in working memory (*e.g.*, letters). In one study, Mall and colleagues (2014) examined the relation between individual differences in distractor sensitivity and two measures of working memory capacity: one derived from performance on external complex span tasks and one derived from the same change detection task as the distractor sensitivity measure. Results indicated that when working memory was operationalized using a measure from the same change detection task, distractor sensitivity was positively related to working memory performance. However, when working memory was operationalized using complex span performance, there was not a significant relation between distractor sensitivity and working memory performance. These results are consistent with the idea that the relation observed in prior work between distractor sensitivity and working memory performance may be driven by task-characteristics rather than by a more general mechanism underlying distractor sensitivity.

In a related study, Robinson and colleagues (2018) conducted a series of three experiments further examining the relationship between distractor sensitivity, measured using a change detection task, and working memory performance, measured using a factor score derived from three complex span tasks. In two of their three experiments, they found significant relationships between complex span performance and distractor sensitivity. However, to observe a significant effect, the change detection task needed to meet several criteria: 1) load in the change detection task needed to be sufficiently high, ensuring capacity limits were taxed, and 2) target and distractor item designations needed to change across trials, requiring participants to flexibly update and maintain a new task goal for each trial. Even when these criteria were met, the magnitude of the relation between complex span performance and distractor sensitivity was modest in magnitude.

The observation that distractor sensitivity *only* related to complex span performance when the task goal changed across trials raises an intriguing possibility which – to the best of our knowledge – has not been systematically explored. Specifically, the shared variance in distractor sensitivity and working memory performance may be underpinned by individual differences in *proactive control,* the ability to prepare for a cognitively demanding task in advance based on a preparatory cue or instruction. Logically, the ability to successfully ignore distractors must depend on the ability to maintain and utilize information about the task goal to determine whether incoming information is relevant to the task at hand. Critically, Braver and colleagues (2012) have argued that such proactive control draws upon working memory capacity. Consistent with this notion, they showed that high-capacity individuals are more likely to engage in proactive control strategies than are lower-capacity individuals (*e.g.*, Braver, 2012).

Taken together, the findings of these two studies present a highly constrained and conditional picture of the conditions under which individual differences in distractor sensitivity might drive performance on complex working memory tasks (Mall at al., 2014; Robinson et al., 2018). We propose an alternative characterization of this relation, namely that the shared variance in distractor sensitivity and working memory performance may relate to individual differences in the ability to maintain and use task goals to proactively prepare for task filtering. The present study builds upon this existing literature by relating individual differences in brain and behavioral responses during a novel, cued complex working memory task to performance on traditional working memory measures.

A neuroimaging approach to investigating individual differences in working memory

To a certain extent, neuroimaging research has been leveraged to help uncover the mechanisms that underlie performance on working memory tasks. While the neural networks recruited do seem to vary as a function of task demands and stimuli modality, some common patterns of activation have also emerged. Most notably, activation in prefrontal regions has repeatedly been shown to increase along with working memory load irrespective of individual task-demand characteristics leading to the proposal that the prefrontal cortex acts as a domain-general executive control region (*e.g.,* Braver et al., 1997; Osaka et al., 2003; Osaka et al., 2004; Rypma et al., 1999).

However, as mentioned previously, a wide range of tasks are used to measure working memory. This variation leads not only to behavioral differences but also to divergent conclusions regarding the neural networks supporting working memory performance. For example, experiments that task participants with remembering the identity of objects more strongly recruit bilateral portions of the inferior frontal gyrus and primary visual cortex whereas experiments that task participants with remembering the location of objects show stronger recruitment of the posterior superior frontal gyrus, superior parietal lobule, precuneus, and right inferior parietal cortex (see Rottschy et al., 2012 for a meta-analysis). Likewise, tasks that require maintenance of verbal information recruit a distinct neural network including portions of the left ventral prefrontal cortices, left inferior frontal and temporal gyri, and left temporoparietal regions (*e.g.,* Emch et al., 2019; Nee et al., 2013; Rottschy et al., 2012) compared to those requiring the maintenance of visuospatial information which recruit regions including the left superior motor area, bilateral portions of the dorsolateral primary motor cortex and portions of the right dorsolateral prefrontal and parietal cortices (*e.g.*, Erikkson et al., 2015; Nee et al., 2013; Rottschy et al., 2012). Thus, group-level neuroimaging work demonstrates that a highly diverse and distributed set of brain regions can be recruited during working memory tasks due to task-specific demands.

A complementary possibility is that the diversity of regions recruited during a given working memory task may be in part a reflection of individuals using different cognitive processes to produce the same behavioral outputs. Despite a long-standing and robust literature surrounding how the brain responds during working memory tasks at the group-level, relatively few studies have investigated how individual differences in neural recruitment relate to variation in performance on working memory tasks. Considering the multidimensional nature of working memory, and the fact that successful performance on complex working memory tasks requires the recruitment of multiple subcomponent processes in tandem, it is likely that even when behavioral performance looks similar across participants, the underlying mechanisms driving these effects may differ across individuals. For example, several studies have demonstrated that individual differences in working memory capacity lead to differential recruitment of neural regions supporting cognitive control. In two studies Osaka and colleagues (2003; 2004) examined neural activation during listening and reading span tasks and related neural activation to individual differences in performance on these tasks obtained outside the scanner. Results for both tasks demonstrated that while the anterior cingulate cortex and left inferior frontal gyrus were recruited to some degree at the group level, these areas were differentially activated for high- compared to low-span subjects. Additionally, when compared to low-capacity individuals, high-capacity individuals have been shown to differentially recruit striatal circuitry in the basal ganglia when faced with increasingly complex sentence comprehension tasks (*e.g.,* Prat et al., 2007; Prat & Just, 2011). Together these results demonstrate that individual differences in working memory abilities lead to differential recruitment of neural regions implicated in cognitive control processes.

Consistent with the hypothesis that individual differences in distractor sensitivity and performance on complex span tasks may be jointly underpinned by proactive control mechanisms, McNab & Klingberg (2008) found that individual differences in the neural networks recruited during cued task preparation related to subsequent performance on a change detection task. Specifically, McNab & Klingberg (2008) found that differential recruitment of the striatum during the instruction phase of the task, when participants were presented with an informative cue instructing them as to what information would be task-relevant, was associated with the degree to which participants stored task-irrelevant representations in working memory. These findings are consistent with theories arguing that successful working memory performance is driven by individual differences in distractor sensitivity (*e.g.,* Engle, 2002; Engle & Kane, 2004; Vogel et al., 2005) as well as with theoretical understandings of the striatum as a “gating” mechanism that selectively routes task-relevant signals to the prefrontal cortex (*e.g.,* O’Reilly & Frank, 2006; Stocco et al., 2010). The proactive control framework proposed by Braver and colleagues (2008; 2012) is also relevant to consider here, as, similarly to the behavioral results of Robinson and colleagues (2018), the relation between distractor sensitivity and working memory capacity seems to be underpinned by one’s ability maintain and utilize the task goal.

The few existing studies investigating the neurocognitive bases of individual differences in working memory have been limited in several key ways. First, individual differences studies using fMRI are typically plagued by small sample sizes which can be underpowered to detect robust individual differences (see Yarkoni & Braver, 2010 for a discussion). While the small sample sizes typically seen in fMRI studies may pose less of an issue for group-level investigations, larger samples are required to robustly detect individual differences in neural activity. Second, working memory capacity is typically only measured with one task used both inside and outside of the scanner. This is problematic as it makes it difficult to determine whether the networks activated in the scanner are task-specific or generalize across diverse working memory measures. This issue is compounded in studies where the same task is used to assess both working memory performance and distractor sensitivity (*e.g.*, McNab & Klingberg, 2008). Finally, because most previously used measures of distractor sensitivity are obtained from visual change detection tasks it is challenging to determine if a task-specific (*i.e.,* the change-detection task), a domain-specific (*i.e.,* visuospatial processing), or a domain-general (*i.e.*, attention) characteristic are driving the observed relation between attentional filtering and working memory (Oberauer, 2019).

To address these limitations, in the present study we use a larger than average sample size (for neuroimaging research) and employ a variety of measures of working memory capacity. First, distractor sensitivity is assessed using a cued semantic filtering task. The Category Filtering paradigm we developed requires participants to maintain and manipulate serially presented visuospatial and verbal information based on semantic categories. This task operates at a level of processing more akin to that of complex span task. Thus, using this measure should provide clarity on if the failure to associate distractor sensitivity with complex span tasks is related to differences in task- or domain-specificity. Second, expanding on previously used methods (Mall et al., 2014; Robinson et al., 2018), we relate individual differences in distractor sensitivity to working memory measures derived from both the same task (*i.e.*, total accuracy on the Category Filtering task) as well as external working memory measures (a composite complex span score). This approach allows us to examine the extent to which any relation between distractor sensitivity and working memory capacity observed may be task-specific or may generalize across working-memory measures.

Together these variables will help address the question of whether performance on working memory tasks is underpinned by individual differences in distractor sensitivity. If individual differences in distractor sensitivity drive performance on complex working memory measures, we would expect similar networks in the brain to support both processes. To our knowledge, the present study is the first to systematically investigate the relation between neural indices of semantic filtering and behavioral measures of working memory and distractor sensitivity.

The present study

The present study examines both the behavioral relation between distractor sensitivity and performance on complex working memory tasks and investigates whether these processes are supported by distinct or overlapping neural networks. Considering the existing literature, we propose three potential hypotheses about the nature of the relation between working memory and distractor sensitivity.

First, we examine the predominant hypothesis that performance on working memory tasks is driven by individual differences in distractor sensitivity (*e.g.*, Engle, 2002; Engle & Kane, 2004; Vogel et al., 2005). This hypothesis would be supported by our behavioral measure of distractor sensitivity correlating both with internal (*i.e.*, task accuracy on the Category Filtering task) and external (*i.e.*, complex span composite) measures of working memory task performance. Moreover, a similar network of neural regions should be related to performance on all three behavioral measures.

Second, we investigate the possibility that the observed relation between distractor sensitivity and performance on complex working memory measures is largely a product of most research using the same task to measure both constructs (*e.g.*, Mall et al., 2014). Considering the wide range of tasks used to index working memory it is feasible that these tasks may rely to differing degrees on the ability to ignore distractor items. If this is the case, distractor sensitivity should selectively relate to task accuracy on the Category Filtering task but not to complex span composite score, indicating that task-specific characteristics are driving the observed relation. The neuroimaging results should also reflect this pattern such that the regions that relate to individual differences in distractor sensitivity and Category Filtering task accuracy should share more similarities than those that relate to complex span composite scores.

Finally, we consider that distractor sensitivity and performance on complex working memory tasks are jointly underpinned by a third factor – the ability to proactively activate and maintain a task goal. Prior work suggests that the relation between distractor sensitivity and working memory may be underpinned by goal maintenance ability (*e.g.*, McNab & Klingberg, 2008; Robinson et al., 2018). Neuroimaging work has shown that individual differences in frontostriatal activation at the filtering cue are associated with greater working memory capacity (indexed using an internal measure) and less storage of distractor items in load-sensitive parietal regions (McNab & Klingberg, 2008). Additionally, recent work has suggested that the behavioral relationship between distractor sensitivity and complex span performance is only observed when informative cues are present in the filtering task (Robinson et al., 2018). This reasoning is also in line with work by Braver and colleagues suggesting that utilizing proactive cues is capacity demanding and thus related to individual differences in working memory capacity (*e.g.*, Braver, 2012). If the relation between performance on complex working memory tasks and distractor sensitivity is modulated by goal maintenance ability, individual differences in the behavioral measures of interest should relate to differential preparation for filter cues compared to no filter cues. These differences should be most robust in regions important for goal maintenance (e.g., medial prefrontal cortex), anticipating conflict (e.g., anterior cingulate cortex) and selecting/filtering based on task-relevant features (e.g., striatum, dorsolateral prefrontal cortex), and visual and lexical-semantic areas important for reading and comprehending semantic categories.

Methods

Participants

Sixty-nine right-handed adults aged 18-34 (43 female, *M* = 20.75 years of age) were recruited for participation in this study. This sample was partially overlapping (*N* = 40) with prior work from our research group which focused on how performance on the Category Filtering task related to individual differences in reading skill (Mottarella et al., 2021). All participants were monolingual English speakers with no history of neurological or neuropsychiatric disorders. Language classification was assessed using the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). Monolinguals were classified as subjects with no prior exposure to a second language before age 10 and low proficiency in any language learned past age 10. Participants with below chance accuracy on the Category Filtering task (*N* = 1), and scanner or motion artifacts (*N* = 3 and *N* = 1, respectively) were excluded from all further analyses leaving a final sample size of sixty-four participants (39 female, *M* = 20.92

years of age). All participants provided informed consent in accordance with the University of Washington's Institutional Review Board and were paid for their participation.

Measures

Category Filtering task. In this paradigm, participants are tasked with remembering the order and spatial location of serially presented words (see Figure 1). All words belonged to one of ten taxonomic semantic categories (see the Appendix) and had an average Hyperspace Analogue to Language (HAL) frequency rating of 13374.73 (Balota et al., 2007).[1] At the start of each trial, participants were presented with a 2000ms cue interesting them to either remember all presented words (*i.e.*, WORDS; No-Filter trials) or only words in a particular category (*e.g.*, FRUIT; Filter trials). The cue was followed by a jittered interstimulus interval (ISI) varying from 1500ms - 7500ms.

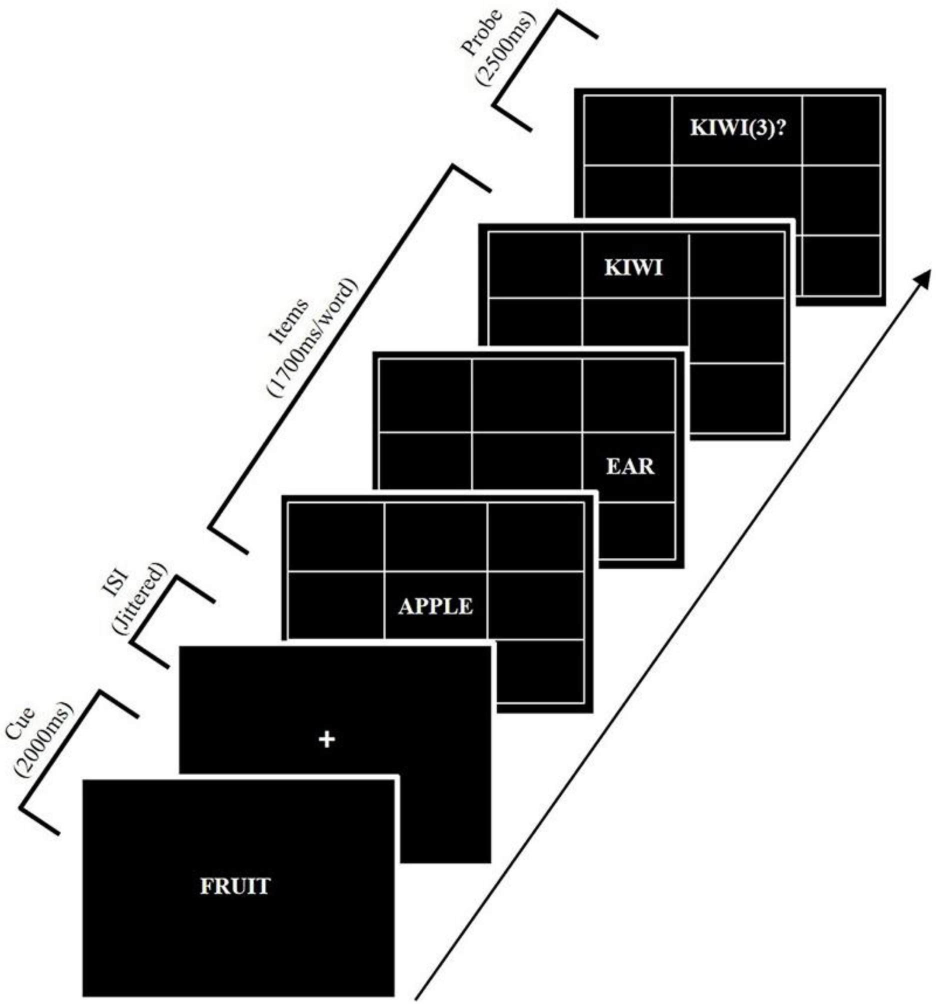
To-be-remembered words were then presented serially on a 3 x 3 grid in to-be-remembered locations for 1500ms each with a 200ms ISI between each word. Following the presentation of all words in a trial, participants judged whether or not a probe, consisting of one of the previously presented words in a grid location and a number indicating a proposed order in which that word occurred in the sequence, was correct. For the trial to be considered correct, the probe needed to be in *both* the same location on the grid and in the same order as one of the words in the preceding sequence. If only one of these criteria were met the trial was considered incorrect. Incorrect probes varied such that half were incorrect due to grid location and half were incorrect due to order. Participants had up to 2500ms to respond to the probe, if they did not respond within this time window their response was counted as incorrect and they advanced to the next trial. The number of presented words (*i.e.*, three or five words) and filtering demands

---

[1] The HAL frequency rating for "T-Shirt" was unavailable so the rating for "Shirt" was used in this average.

(*i.e.*, filter or no-filter cues) varied orthogonally across trials, resulting in four unique task conditions: Filter 3 (targets = 2, distractors = 1), No-Filter 3 (targets = 3, distractors = 0), Filter 5 (targets = 3, distractors = 2), and No-Filter 5 (targets = 5, distractors = 0). The task consisted of forty unique trials (*i.e.*, 10 per condition) presented in one of two pseudorandom orders.

Figure 1.



Note. Schematic of the Category Filtering task depicting an order-incorrect Filter 3 trial. In this trial “KIWI” was the second, not third word presented that fits into the cued category (“FRIUT”), making the order of the probe incorrect. Reproduced from Mottarella, Yamasaki, & Prat (2021).

Complex span composite. A complex span composite score was computed by entering absolute scores on three abbreviated versions of complex span tasks into a principal component analysis. The Operation Span and Symmetry Span tasks employed are described in Foster et al. (2015) and the Reading Span task employed is described in Oswald et al. (2015). Factor loadings on the complex span factor score were 57.65% for Operation Span, 25.53% for Reading Span, and 16.82% for Symmetry Span.

3-back task. In this version of the *n*-back task (*e.g.*, Prat et al., 2016), participants viewed a serially presented stream of letters and were tasked with determining if the current letter was the same or different as the letter presented three items previously. Participants responded that the current letter was the "same" or "different" by pressing one of the two designated buttons on a keyboard. For example, in the letter sequence "A B N A C B," the fourth letter ("A") would be considered a target because it matches the first letter ("A"). The fifth letter in the sequence ("C") would be considered a non-target because it does not match the second letter ("B"). The sixth letter ("B") would be considered a "distractor" or lure because it matches a letter presented four or two items previously. Total accuracy on all trials and accuracy only on distractor trials were both computed, arcsin transformed to account for non-normality in the distributions and used indices of working memory updating ability.

Response-Interference task (Simon task). The Simon task assesses an individual's ability to override a prepotent motor response. In this version of the task *(e.g.,* Prat et al., 2016) participants are instructed to press a key on either the right or left side of a keyboard to signal that they recognize one of two geometric shapes (*e.g.,* circle = right button; square = left button. For 75% of trials the shape appears on ipsilateral to the appropriate motor response (*i.e.,* congruent trials; square on the left side of the screen), on the remaining 25% of trials the shape appears contralateral to the appropriate motor response (*i.e.,* incongruent trials; square on the right side of the screen). Performance is computed by calculating the difference in response times for incongruent - congruent correct trials, where a smaller difference in response times is indicative of better response inhibition.

Raven's advanced progressive matrices (RAPM). The RAPM is a non-verbal measure commonly used to index fluid reasoning. Participants are tasked with choosing which option, out of eight alternatives, best completes a pattern matrix. The measure used herein was abbreviated based on the procedure described in Prat et al. (2016) to include a total of 18 items that increased in difficulty as participants moved through the task. Participants were given a total of 20 minutes to complete the test items. Performance was computed by calculating an accuracy score [100 * (number of items correct / 18 possible points)].

Behavioral data analyses

Distractor sensitivity was assessed by computing a difference in response times for Filter 5 (targets = 3, distractors = 2) - No-Filter 3 trials (targets = 3, distractors = 0). Because the number of targets is consistent in both conditions, this measure can index the added cost in response time when distractor items are present compared to when distractor items are not present. Distractor sensitivity was only computed for response times as the variance of the accuracy data was truncated due to high performance across task conditions. Correlations were computed between distractor sensitivity and the other behavioral measures of interest to determine if: 1) distractor sensitivity relates to Category Filtering task performance, and if 2) distractor sensitivity relates to an external measure of complex working memory task performance (*i.e.,* complex span composite score).

Procedure

The data reported herein were collected as part of a larger study investigating individual differences in complex skill learning (Prat et al., 2020). The Category Filtering task was broken into two eight-minute blocks and collected during a scan session which also included a five-minute structural (MP-RAGE) and seven-minute resting-state scan. Prior to completing the Category Filtering task, participants completed a practice version of the task both outside the scanner, on a lab-provided computer, and inside the scanner, during their structural scan. Behavioral measures were collected in a separate experimental session which occurred within one week of the scan session.

fMRI acquisition. MRI data were collected using a 3.0 T Phillips Ingenia scanner at the Bio-Molecular Imaging Center operated by the University of Washington. The data were aligned to the anterior-posterior commissure and collected using a gradient echo-planar pulse sequence with a TR = 2000ms, TE = 25ms, a 79° flip angle, and a field of view = 240mm. Each volume consisted of 38 oblique-axial slices, each of which was 3mm thick, with a 0mm gap between slices. The acquisition matrix was 80 x 80 voxels with an in-plane resolution of 3 x 3 voxels. For some participants this did not constitute full coverage, in such cases, no data was collected from the cerebellum or the regions surrounding the vertex.

fMRI data processing

Image preprocessing. Preprocessing was performed in SPM12 (Wellcome Trust Center of Neuroimaging, Cambridge, UK). All volumes were corrected for slice timing acquisition, realigned to the first image within each run, normalized to the Montreal Neurological Institute (MNI) template, resampled to $2mm^3$ voxels, and smoothed using an 8mm Gaussian kernel. Signal-to-noise ratio maps were visually inspected, and participants with abnormal scanner artifacts were removed (*N* = 3). Artifact correction was applied using Artifact Detection Tools

(ART), where volumes with more than 1mm of movement or a global mean signal intensity greater than three standard deviations above the mean were flagged as outliers. Participants for whom outliers made up 10% or more of the data were excluded from further analyses ($N = 1$). Only fMRI data from correct trials were used in subsequent analyses.

General linear modeling (GLM) analysis. Statistical analyses were performed on the data using the general linear model as implemented in SPM12 (Wellcome Trust Center of Neuroimaging, Cambridge, UK). First-level models were computed for each participant based on six regressors: two indexing activity at the cue (*i.e.,* Filter Cue and No-Filter Cue) and four indexing activity during the task (*i.e.,* Filter 3, No-Filter 3, Filter 5, and No-Filter 5).

Second-level group analyses were then conducted using voxel-level non-parametric permutation testing as implemented in SnPM13 (http://warwick.ac.uk/snpm). Non-parametric approaches benefit from requiring minimal assumptions and are more robust than standard parametric approaches (Nichols & Holmes, 2001). Second-level SnPM analyses were conducted on first-level models for both *Filtering Cue Activity* (Filter Cue - No-Filter Cue) and *Filtering Task Activity* [Filter Task (averaged across Filter 3 and Filter 5) - No-Filter Task (averaged across No-Filter 3 and No-Filter 5)] contrasts based on 10,000 permutations and locally pooled variance estimates with 4mm smoothing. The resulting pseudo-*t* maps were thresholded at both family-wise error (FWE) corrected $p < 0.05$ and uncorrected $p < 0.001$ with an extent threshold of 20 voxels.

Individual differences in the relation between neural activity and behavioral task performance were also assessed using permutation testing on a voxel-wise basis as implemented in SnPM. Three separate regression models in which distractor sensitivity, task accuracy, and complex span composite were used as predictor variables, and Filtering Cue Activity and

Filtering Task Activity were used as the dependent variables were computed. The resulting models were based on 10,000 permutations and locally pooled variance estimates with 4mm smoothing. The resulting pseudo-*t* maps are reported at uncorrected $p < 0.001$ with an extent threshold of 20 voxels.

Results

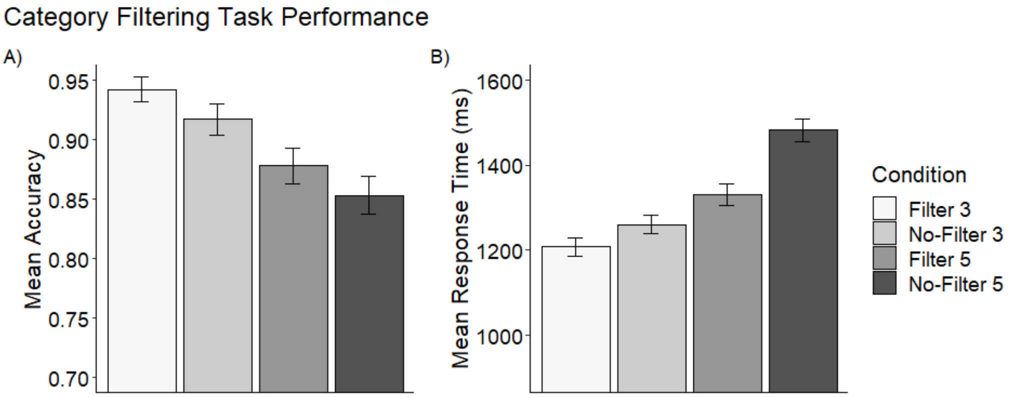
Behavioral results.

Category Filtering task performance. Total accuracy on the Category Filtering task was high across subjects (*M* = 0.90, *SD* = 0.08, *range* = 0.68-1.00). Group-level accuracy and response time data are depicted in Figure 2. As expected, accuracy was higher and response times were faster for conditions with fewer target items. The response times and accuracies seem to be influenced by both filtering demands and the number of to-be-remembered items. For example, performance in the Filter 5 condition (targets = 3, distractors = 2) was faster and more accurate than in the No-Filter 5 condition (targets = 5, distractors = 0), suggesting a benefit conferred by the reduction in target items, but was also slower and less accurate than performance in the No-Filter 3 condition (targets = 3, distractors = 0), suggesting a decrement to performance as a result of filtering demands.

To quantify the demands of the task on performance a 2 (number of items) x 2 (filtering demands) repeated measures ANOVA was calculated to assess the effects of number of items (three vs five) and filtering (filter vs no-filter) on Category Filtering task response times. Results indicated a significant main effect of the number of items, $F(1, 63) = 271.45$, $p < 0.001$. A post-hoc paired-samples *t*-test further elucidated that participants responded faster on three-item trials (*M* = 1234.38, *SD* = 167.02) than on five-item trials (*M* = 1406.67, *SD* = 191.97); $t(63) = 16.48$, $p < 0.001$, indicating that as the number of items presented increased, behavioral response times to the probe slowed. Additionally the ANOVA revealed a main effect of filtering $F(1, 63) = 74.94$, $p < 0.001$, such that participants responded faster for filter ($M = 1269.27$, $SD = 1371.78$) compared to no-filter trials ($M = 1371.78$, $SD = 183.48$); $t(63) = 8.66$, $p < 0.001$. There was also a significant number of items by filtering interaction $F(1, 63) = 21.24$, $p < 0.001$, such that the effect of filtering on response times was greater for five-item trials ($M = 152.42$, $SD = 150.19$) than three-item trials ($M = 52.59$, $SD = 102.02$), $t(63) = 4.61$, $p < 0.001$. This result suggests that filtering becomes more imperative at greater set sizes and is likely a product of the two filtering conditions differing in the number of distractor items that must be filtered out. Specifically, the Filter 5 condition required participants to filter out two distractors whereas the Filter 3 condition only required one distractor to be filtered out.

An ANOVA was not calculated for Category Filtering task accuracy due to non-normality in the distributions for all conditions even after applying an arcsine transformation [Filter 3: $W(63) = 0.73$, $p < 0.001$; No-Filter 3: $W(63) = 0.78$, $p < 0.001$; Filter 5: $W(63) = 0.85$, $p < 0.001$; No-Filter 5: $W(63) = 0.87$, $p < 0.001$]. The skew of the accuracy data is likely a product of allowing several task practices.

Figure 2.



Note. Category Filtering task A) mean accuracy, and B) response times by condition. Error bars depict standard error.

Relating Category Filtering task performance to established cognitive measures. To ground our understanding of what performance on the Category Filtering task measures, we first correlated Category Filtering total accuracy with performance on a battery of external cognitive measures designed to assess working memory, fluid intelligence, and inhibitory control. To do so, bivariate correlations were computed between Category Filtering total accuracy, complex span composite score, 3-back measures, Raven's APM, and Simon performance (see Table 1). Total accuracy on the Category Filtering task was significantly correlated with the complex span factor score and remained significant after a False Discovery Rate (FDR) correction for multiple comparisons was applied ($r = 0.42$, $p = 0.001$, $pfdr = 0.005$). Category Filtering total accuracy was also significantly correlated with performance on Raven's Advanced Progressive Matrices, though the relation was marginal after correcting for multiple comparisons ($r = 0.27$, $p = 0.03$, $pfdr = 0.08$). Neither total or distractor accuracy on the 3-back task nor Simon task performance were significantly correlated with Category Filtering total accuracy ($ps > 0.10$). Taken together these results suggest that despite the truncated range of Category Filtering accuracy scores, there is sufficient variability in the Category Filtering total accuracy measure to predict individual differences in traditional complex span measures and tests of fluid reasoning ability.

To examine the hypothesis that individual differences in performance on complex working memory tasks are driven by differences in distractor sensitivity, correlations were then computed between distractor sensitivity, Category Filtering total accuracy, and our external cognitive measures of interest (see Table 1). Contrary to the idea that individual differences in working memory are underpinned by the ability to resist distractors, distractor sensitivity did not significantly correlate with Category Filtering total accuracy, complex span composite score, or any of the other cognitive measures of interest ($ps > 0.40$).

Table 1. Bivariate correlations and descriptive statistics of behavioral variables.

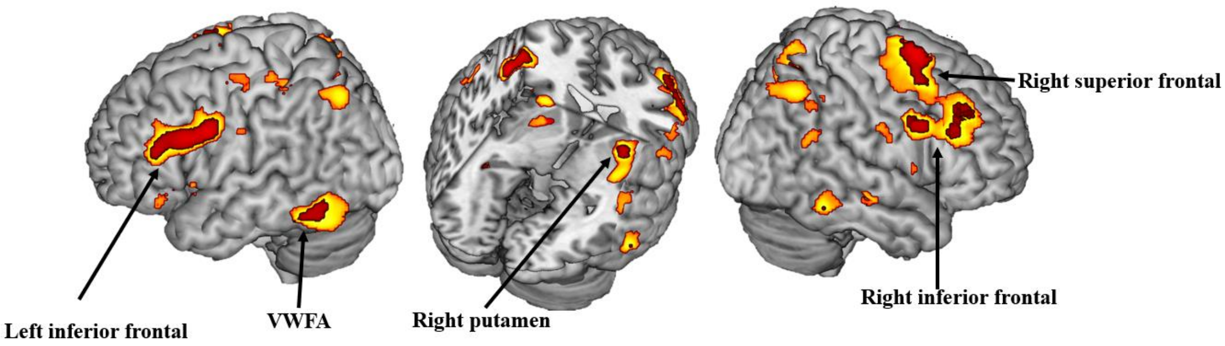
Measures	*N*	*M*	*SD*	1	2	3	4	5	6
1. Category Filtering total accuracy	64	0.90	0.08	-					
2. Distractor sensitivity	64	69.79	102.72	-0.073	-				
3. Complex span composite	62	0.09	0.95	__0.42**__	0.14	-			
4. 3-back total accuracy	63	0.84	0.11	0.20	0.11	__0.27*__	-		
5. 3-back distractor accuracy	63	0.76	0.16	0.18	0.014	__0.28*__	__0.78**__	__-__	
6. Simon	64	53.77	47.05	-0.12	0.095	-0.14	0.081	0.074	-
7. RAPM	64	0.69	0.14	__0.27*__	0.019	__0.36**__	__0.55**__	__0.41**__	__0.057__

Note. $p < 0.10$, *$p < 0.05$ uncorrected, **$p < 0.05$ fdr corrected

Group-level fMRI results

Filtering Task Activity. Group-level results revealed a differential increase in neural activity for Filter > No-Filter trials across a distributed network of lexico-semantic and domain-general cognitive control regions. Specifically, Filter > No-Filter trials led to greater recruitment of portions of the bilateral inferior frontal regions, visual word form area, precuneus, right premotor, and right putamen (see Figure 3 and Table 2A). Virtually all clusters of activation survived FWE correction (see Figure 3 and Table 2A). These results suggest that filtering trials recruit additional executive control and word processing resources in order to determine if a given stimulus is a target and should be maintained in working memory or is a distractor and should be ignored as task-irrelevant.

Figure 3.



Group-level *Filtering Task Activity* at uncorrected $p < 0.001$ and FWE corrected $p < 0.05$ thresholds

Note. Group-level GLM activiation for the Filtering Task Activity contrast (Filter Task > No-Filter Task). Orange denotes $p < 0.001$ uncorrected with an extent threshold = 20 voxels, red denotes $p < 0.05$ FWE corrected.

Filtering Cue Activity. Group-level activation at the cue revealed a network of bilateral visual areas in the occipital lobe and left-lateralized memory and lexical-decision areas including portions of the left inferior frontal gyrus, left inferior parietal lobule, left fusiform gyrus, and left hippocampus such that activation in these regions was greater for informative Filter Cues (*e.g.*, FRUIT) than for non-informative No-Filter Cues (*i.e.*, WORDS) (See Table 2B). Only the bilateral occipital and left fusiform clusters survived FWE correction.

Table 2. Group-level grey matter activation for whole-brain GLM analysis.

Peak cortical region	Brodmann's area	Cluster size	Pseudo t value	Peak MNI coordinates		
				x	*y*	*z*
(A) Filtering Task Activity						
Left inferior frontal	44	1466 (522)	5.72	-48	6	30
Left fusiform	20	1909 (100)	5.72	-36	-14	-28
Left fusiform*	20	(68)	5.47	-42	-32	-18
Left inferior temporal*	37	(102)	4.76	-50	-56	-14
Left insula	47	222	4.36	-30	24	-4
Right superior frontal	6	2925 (250)	6.07	28	8	68
Right middle frontal*	46	(48)	4.74	42	34	38
Right inferior frontal*	44	(96)	5.10	44	8	30
Right putamen	48	723 (52)	5.04	26	10	-6
Right fusiform	37	200 (25)	4.98	32	-34	-24
Right precuneus	7	1845 (362)	6.27	6	-64	48
(B) Filtering Cue Activity						
Left orbital frontal	47	49	4.12	-40	32	-14
Left inferior frontal	48	106	3.87	-44	20	26
Left SMA	6	115	3.84	-8	12	54
Left hippocampus	20	53	3.88	-30	-20	-16
Left fusiform	37	253 (11)	4.83	-36	-42	-24
Left inferior parietal	7	139	3.72	-36	-64	48

Left occipital	18	596 (59)	4.80	-16	-92	-10
Right occipital	18	387 (138)	6.29	18	-94	-2

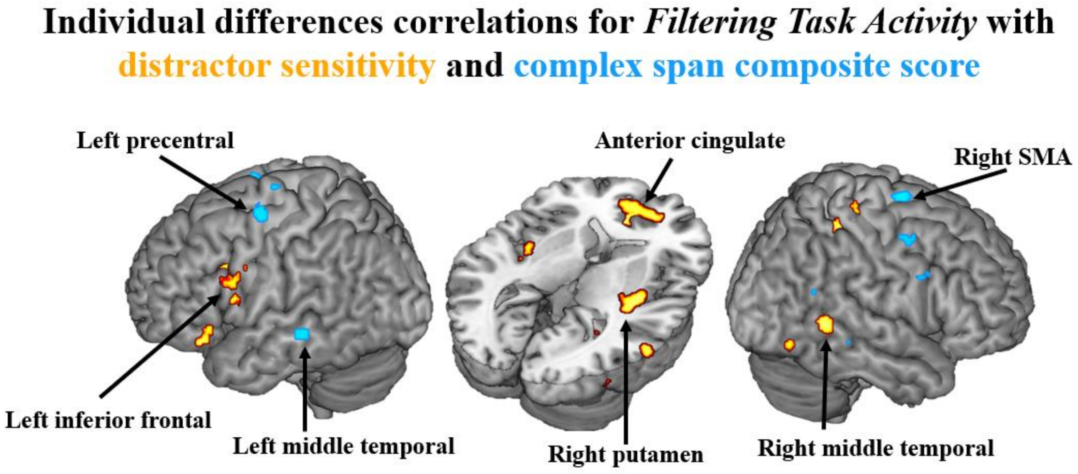
Note. For clusters that survive FWE correction, the cluster size following FWE is provided in parentheses. Areas denoted with a * became separable clusters when FWE correction was applied; these clusters are nested under the original cluster they were a part of.

Individual-level fMRI results

To assess whether individual differences in distractor sensitivity and working memory performance were underpinned by similar mechanisms, we correlated our behavioral measures of interest – namely distractor sensitivity, Category Filtering total accuracy, and complex span composite score – with both the Filtering Task Activity and Filtering Cue Activity contrasts. It is important to note that for the distractor sensitivity measure, a smaller value indicates better performance (*i.e.*, response time is less impacted by distractors) whereas for the Category Filtering total accuracy and complex span composite score a larger value is indicative of better performance.

Correlating Filtering Task Activity with behavioral outcomes. Correlations between Filtering Task Activity and behavioral measures of interest revealed a striking dissociation. Specifically, Filtering Task Activity and distractor sensitivity demonstrated a negative correlation in a distributed network of domain-general cognitive control regions including the anterior cingulate cortex and right putamen as well as lexical decision areas including portions of the left inferior frontal gyrus (see Figure 4 and Table 3A). The directionality of this result suggests that greater sensitivity to distractors, as indexed by a smaller difference in response times between Filter 5 and No-Filter 3 trials, is associated with more recruitment of cognitive control and lexical decision regions during Filter > No-Filter task trials. Conversely, correlational analyses computed between Filtering Task Activity and complex span composite revealed a negative correlation in a much smaller network of regions including the right supplementary motor area (SMA), left precentral gyrus, and left middle temporal gyrus, such that more recruitment of these regions during Filter > No-Filter trials was associated with lower span scores. Notably, as depicted in Figure 4, none of the regions that predicted complex span composite overlapped with those implicated in predicting distractor sensitivity. Correlational analyses between Filtering Task Activity and Category Filtering task accuracy revealed no significant clusters of activation, possibly due to a lack of variability in task performance.

Figure 4.



Individual differences correlations for *Filtering Task Activity* with distractor sensitivity and complex span composite score

Note. Individual-level correlations between the Filtering Task Activity contrast (Filter Task > No-Filter Task) and behavioral measures at $p < 0.001$ uncorrected with an extent threshold = 20 voxels. The negative correlation between distractor sensitvity and Filtering Task Activity is depicted in orange and the negative correlation between complex span composite score and Filtering Task Activity is depicted in blue.

Correlating Filtering Cue Activity with behavioral outcomes. Correlational analyses between activity at the cue and behavioral outcome measures revealed minimal significant clusters. Specifically, Filtering Cue Activity was positively correlated with distractor sensitivity in the right rolandic operculum such that greater activity in this area for Filter Cues > No-Filter Cues was associated with worse sensitivity to distractors (see Table 3B). Furthermore, Filtering Cue Activity was negatively correlated with task accuracy in the right SMA such that greater activity in this area for Filter Cues > No-Filter Cues was associated with worse task accuracy (see Table 3B). Correlational analyses between Filtering Cue Activity and complex span composite scores revealed no significant clusters of activation.

Table 3. Correlations between GLM activation in grey matter and behavioral measures

Peak cortical region	Behavioral Measure	Correlation	Brodmann's area	Cluster size	Pseudo t value	Peak MNI coordinates		
						x	*y*	*z*
(A) Filtering Task Activity								
Anterior cingulate	Distractor sensitivity	Negative	10	765	4.12	-4	44	-2
Middle cingulate	Distractor sensitivity	Negative	24	80	3.81	-6	0	32
Left inferior frontal	Distractor sensitivity	Negative	44	468	4.31	-32	2	30
Left orbital frontal	Distractor sensitivity	Negative	38	45	3.83	-46	30	-12
Right amygdala	Distractor sensitivity	Negative	34	52	4.07	28	2	-18
Right putamen	Distractor sensitivity	Negative	48	349	3.97	34	-16	-4
Right inferior parietal	Distractor sensitivity	Negative	40	154	3.92	32	-42	54
Right inferior temporal	Distractor sensitivity	Negative	37	131	3.85	44	-62	-10
Right middle temporal	Distractor sensitivity	Negative	21	58	3.82	66	-48	6
Left precentral	Complex span composite	Negative	6	31	3.94	-40	-6	62
Left middle temporal	Complex span composite	Negative	21	70	3.82	-68	-28	-2
Right SMA	Complex span composite	Negative	6	69	4.57	14	-2	72
(B) Filtering Cue Activity								

Right SMA	Category Filtering total	Negative	6	73	3.62	8	-10	64
Right rolandic operculum	Distractor sensitivity	Positive	48	103	4.21	52	4	4

Note. Both the correlation between the GLM Filtering Task Activity and Category Filtering total accuracy and the correlation between GLM Filtering Cue Activity and complex span composite score failed to produce any activation clusters > 20 voxels, uncorrected $p < 0.001$.

Discussion

The results of this experiment provide some critical clues for answering a central question in cognitive psychology: what do individual differences in complex span tasks reflect? With an emphasis on task characteristics, we explored the extent to which individual differences in distractor sensitivity, or the ability to ignore task-irrelevant information, drives performance on complex span tasks. Our findings, particularly when viewed in light of the existing literature, suggest that distractor sensitivity and performance on complex span tests are largely separable – at least when the tasks involve higher-level semantic decisions. In the sections below, we outline our behavioral and neuroimaging support for this claim, while highlighting the limitations in our design and those that have come before us.

Relating distractor sensitivity to an “external” complex working memory measure

Consistent with the findings of Mall and colleagues (2014) our behavioral measure of distractor sensitivity did not significantly correlate with the external complex span composite score. This result suggests that the ability to resist distractors in one complex task does not seem to drive performance on an external complex span task. Contrary to our prediction, the addition of a cue indicating which type of information was relevant for a given trial did not increase the relation between distractor sensitivity and performance on complex span tasks. These results contrast with the findings of Robinson and colleagues (2018), who found that individual differences in distractor sensitivity related to complex span performance when the filtering task's: 1) goal changed dynamically across trials, and 2) load was sufficiently taxing. In the present study the instruction of whether or not to filter incoming information was unpredictable across trials, satisfying the former requirement. However, it is possible that the Category Filtering task did not sufficiently tax load to the extent necessary to observe this effect, as total behavioral accuracy was quite high across participants (*M* = 0.90, *SD* = 0.08, *range* = 0.68-1.00). While the set sizes used in the Category Filtering task were comparable with those used in previous change detection tasks (*e.g.,* McNab & Klingberg, 2008), the conceptual nature of the stimuli and their serial presentation may have facilitated chunking or verbal rehearsal strategies, resulting in lower capacity demands.

The conclusion that distractor sensitivity and performance on complex working memory tasks are separable processes is further supported by our neuroimaging work, which showed that non-overlapping neural networks were correlated with individual differences in these processes. Specifically, individuals who were less susceptible to distractors, showed greater recruitment of domain general cognitive control regions, including the anterior cingulate cortex (ACC) and right putamen, as well as of classic language areas like the left inferior frontal gyrus (IFG) during filtering trials. Both the left IFGs and the right putamen were also more active for Filter > No-Filter trials at the group level. This suggests that the process of semantic filtering recruits these regions across participants (at least in-so far as we can draw conclusions from group-averages), and that *increased* recruitment in these areas in an individual is associated with decreased behavioral sensitivity to distractors. These results are consistent with prior work implicating the striatum, which includes the caudate and putamen, in gating access to working memory (*e.g.,* O'Reilly & Frank, 2006) or routing contextually relevant signals to prefrontal cortex (*e.g.,* Stocco et al., 2010). In contrast, the increased activation of the ACC for Filter > No-Filter trials associated with decreased distractor sensitivity was only observed in the individual differences analyses, not at the group-level. It is not uncommon for areas where involvement differs between individuals to “wash out” at the group-level (for a discussion see Prat & Just, 2011). However, it is interesting to note that the ability to successfully resist distractors seems to be jointly underpinned by quantitative (greater recruitment of select areas; caudate and left IFG) and qualitative (recruitment of distinct areas; ACC) differences in the activation of cognitive control areas.

Critically for the questions on which this investigation was centered, the neural network associated with individual differences in distractor sensitivity was completely non-overlapping with the regions related to individual differences in working memory performance. When activation for Filter > No-Filter trials was correlated with complex-span composite scores, greater activation of bilateral frontotemporal areas was related to higher complex-span scores. Together, the behavioral and neuroimaging results converge to suggest that when distinct measures are used to index distractor sensitivity and working memory performance these processes are largely separable.

Relating distractor sensitivity to an “internal” complex working memory measure

In the present study, even when both measures of distractor sensitivity and working memory performance were derived from the Category Filtering task, these processes did not relate to one another. This finding contrasts prior work (*e.g.*, Mall et al., 2014; Robinson et al., 2018; Vogel et al., 2005) showing that indices of distractor sensitivity and working memory performance relate to one another when they are derived from the same task. One possible explanation for these conflicting results is that there is a fundamental difference between the visuospatial change detection tasks used in previous research, and the conceptual filtering task employed herein.

While both types of filtering tasks require maintaining the location and identity of relevant-cued-items in the face of distractors, it is important to note several differences between the tasks that may be driving this result. One consideration is that the Category Filtering task operates at a higher-level of information processing than typical change-detection tasks, which rely on simple visual features. The semantic nature of the Category Filtering task requires participants to process the visual features of the word (*e.g.,* “C”, “A”, “T”), before performing higher-level processing steps including extracting the meaning (*e.g.,* those letters together mean “CAT”), retrieving the cued category (*e.g.,* the cued category was FRUIT), determining if the meaning of the stimuli belongs in the cued category (*e.g.*, CAT is not a fruit), and if it does not, actively disengaging cognitive resources from the distractor (e.g., ignore the location of CAT and update the order of to-be-remembered items). Thus, successful performance on the Category Filtering task necessitates that all incoming stimuli first “get into” working memory, to extract the semantic meaning of the words, before distractors can be filtered out. Once items are deemed distractors, attention must be actively disengaged from those items to expunge them from working memory. This process is akin to the idea of working memory “updating,” a term used by Miyake and colleagues (2000; 2012) to refer to the rapid and dynamic addition and deletion of items from working memory. Conversely, in visuospatial change detection tasks attentional filtering occurs in early stages of perceptual processing *before* stimuli enter working memory. An additional consideration is that the serial presentation of the Category Filtering task requires participants to maintain targets across time and dynamically update ordinal information for each trial. This manipulation was added to replicate the processing demands required in complex span tasks more closely. However, it is a noteworthy conceptual deviation from visuospatial change detection tasks in which all stimuli are simultaneously presented – resulting in a shorter retention interval with no ordinal information to maintain.

Our neuroimaging results suggested that individual differences in Category Filtering task accuracy were not associated with differential activation for Filter > No-Filter trials, whereas differences in distractor sensitivity recruited the distributed network of domain general cognitive control and language-related areas discussed in the previous section. It is possible that the truncated range of Category Filtering total accuracy scores – stemming from high performance across participants – impeded our ability to both detect individual variability in the neural networks associated with task accuracy and to behaviorally correlate task performance with distractor sensitivity. However, our concerns about this are somewhat ameliorated by the fact that Category Filtering task accuracy *was* robustly correlated with our complex span composite score ($r = 0.42$, $p = 0.001$, $pfdr = 0.005$) as well as other standard behavioral cognitive control measures in the predicted directions (see Table 1). Additionally, an examination of the Filtering Cue Activity contrast showed that Category Filtering task accuracy *was* significantly correlated with differential activation in right SMA when preparing for Filter > No-Filter trials. This finding, in combination with the robust behavioral results, suggest that the task accuracy measure did have sufficient variability to detect individual differences in both behavioral and brain measures. Thus, the lack of activation differences associated with Filtering Task Activity contrast remains unexplained. One possibility, consistent with our wider body of results suggesting that complex span task performance is separable from distractor sensitivity, is that the processes differentially engaged by higher-capacity individuals were similarly recruited across all trials irrespective of filtering demands. We see this as an interesting area for future exploration.

Where does distractor sensitivity come from?

Together, both the behavioral and neuroimaging results converge to suggest that when a semantic filtering task is used, distractor sensitivity and performance on complex working memory tasks are separable processes underpinned by distinct neural networks. This pattern of results conflicts with the hypotheses derived from previous work suggesting that individual differences in distractor sensitivity drive performance on complex working memory tasks (*e.g.*, Robinson et al., 2018; Vogel et al., 2005), or that task-specific (*e.g.*, Mall et al., 2014) or domain-specific (*e.g.,* Oberauer, 2019) factors lead to a relation between measure of working memory performance and distractor sensitivity. Instead, the results of the present study suggest that – at least within the context of a semantic filtering task – the ability to resist distractors and performance on complex working memory tasks are largely separable.

Integrating our results with those using visuospatial change detection tasks, one interpretation is that distractor sensitivity may be underpinned by distinct cognitive components in the two types of tasks. While tasks that allow distractors to be filtered by low-level visual features may allow for early attention-based gating of information into working memory, semantic content requires at least some processing in working memory before a decision can be made about whether a particular piece of information is relevant for the task or is a distractor. In other words, filtering can occur much earlier for simple-visual than for complex-semantic information.

What is less clear is whether the difference between distractor sensitivity in different domains is solely a product of *when* filtering occurs or if the *nature* of the distractor sensitivity process itself is categorically distinct in different processing domains. From this perspective a similar debate emerges regarding distractor sensitivity as has been long argued in the working memory literature. Namely, is the ability to resist distractors domain-general regardless of the type of distractor presented? Prior work from our laboratory using the Category Filtering task has shown that neural responses when preparing for filtering trials is tied to individual differences in language skill (Mottarella et al., 2021). Thus, it appears that the ability to prepare for upcoming distractors may be driven by the ease of access to category information that dictates whether or not incoming stimuli are distractors. However, it is an open question as to how much of the variance in resisting distractors can be explained by domain-specific versus domain-general factors.

A final consideration is that distractor sensitivity itself may not be a unidimensional construct. This consideration is born out of the observations that: 1) the extent to which distractor sensitivity relies on the task goal, and 2) the amount of information processing necessary to decide whether something is task-relevant, changes across different tasks. This raises the question of whether distractor sensitivity is an appropriate level of description, or whether individual differences are better characterized by combined contributions of smaller executive function factors. For example, the unity-diversity model of executive function proposed by Miyake and colleagues (2000; 2012) has used latent variable analyses to robustly demonstrate that three sub-components of executive function – updating (*i.e.*, rapid addition and/or deletion of to-be remembered items), inhibition (*i.e.*, overriding a prepotent response), and shifting (*i.e.*, juggling between different tasks or goals) – can be behaviorally dissociated from one another despite sharing some common variance. Applying a similar latent variable methodological approach to multiple measures of distractor sensitivity may provide further clarity into the extent to which it is a unitary construct across task domains. We see this as an important avenue for future research.

Limitations

The present study provides novel insights into how the neural and behavioral indices of distractor sensitivity are largely separable from performance on complex working memory tasks, however, several key limitations are important to consider when interpreting these results. First, the serial presentation of stimuli in the Category Filtering task makes it challenging to compare the neural timecourses for conditions with three versus five items. For example, an attempt to directly compare brain activation in the No-Filter 3 and Filter 5 conditions would be confounded by the fact that targets in five-item conditions must be maintained longer than those in three-item conditions. Thus, in the present study we could not create the neural contrast to directly index distractor sensitivity which has been central to neuroimaging studies of visual distractor sensitivity (*e.g.,* McNab & Klingberg, 2008).

Second, the presentation of the cue in our study was very brief (*i.e.,* 2000ms) and may not have been well captured by the slow signaling of the BOLD response in light of our event-related design with variable jitter. It is possible that having a longer cue presentation, or a longer interval between the end of the previous trial and the cue presentation may have increased our sensitivity to the proactive control signal. Future studies either using an elongated cue presentation or a faster TR may be better able to capture the neural dynamics associated with preparing to filter within an fMRI paradigm. Alternatively, other neuroscience methodologies with better temporal resolution such as electroencephalography (EEG) or magnetoencephalography (MEG) may be able to better capture the rapid timecourse of neural responses associated with cue presentation.

Third, it is important to note that the neural networks associated with distractor sensitivity and complex span performance are correlations between neural activity on the Category Filtering task, specifically the difference between Filter > No-Filter trials, and behavioral measures obtained outside of the scanner. Thus, the present study cannot determine whether neural recruitment may look similar or distinct between these processes when completing the tasks in real-time.

Finally, we did not include a visuospatial change detection task in our behavioral task battery. While the primary goal of the present study was to examine the relation between distractor sensitivity and task performance within a higher-level semantic filtering task, our results raise the possibility that semantic filtering may be a fundamentally different process than visuospatial filtering. Future work should directly compare distractor sensitivity measures obtained from the Category Filtering task to those obtained from a visuospatial change detection task, to examine how much of the ability to resist distractors is domain general – perhaps by using a latent variable method as discussed previously.

Conclusion

From our perspective there are at least two possibilities as to why our findings may have differed from previous theories proposing that individual differences in distractor sensitivity drive performance on working memory tasks: 1) relations seen in prior work may be an artifact of using the same task to measure both distractor sensitivity and working memory, and 2) filtering in lower-level perceptual tasks may be fundamentally different than in the higher-level semantic task employed herein. Note that these explanations need not be mutually exclusive.

The present study adds important insights to the ongoing debate of whether the ability to successfully resist distractors drives individual differences in performance on working memory tasks. Specifically, to the best of our knowledge, our study is the first to investigate this question using a higher-level semantic filtering task and a combination of neuroimaging and a battery of behavioral measures. Both our behavioral and neuroimaging results suggest that within our higher-level semantic filtering task, distractor sensitivity and performance on complex working memory tasks are largely separable processes. This result has important implications for our theoretical understanding of the factors that underpin individual variation in complex working memory task performance and raises important questions about the degree to which distractor sensitivity reflects a unitary cognitive process.

References

Awh, E., Vogel, E. K., & Oh, S.H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage*, *5,* 49-62. https://doi.org/10.1006/nimg.1996.0247

Braver, T. S., Gray, J. R., & Burgess, G. C. (2008). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in Working Memory* (pp. 76–106). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195168648.003.0004

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–228. https://doi.org/10.1080/87565640801982312

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. https://doi.org/10.3758/BF03214546

Emch, M., von Bastian, C. C., & Koch, K. (2019). Neural correlates of verbal working memory: An fMRI meta-analysis. *Frontiers in Human Neuroscience, 13*. https://doi.org/10.3389/fnhum.2019.00180

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

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In *Psychology of Learning and Motivation* (vol. 44, pp. 145–199). Elsevier. https://doi.org/10.1016/S0079-7421(03)44005-X

Engle, R. W., Laughlin, J. E., Tuholski, S. W., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331. https://doi.org/10.1037/0096-3445.128.3.309

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron*, *88*(1), 33–46. https://doi.org/10.1016/j.neuron.2015.09.020

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*(3), 348–368. https://doi.org/10.1016/0749-596X(86)90006-9

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, *43*(2), 226–236. https://doi.org/10.3758/s13421-014-0461-7

Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *Journal of Neuroscience*, *29*(27), 8726–8733. https://doi.org/10.1523/JNEUROSCI.2145-09.2009

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149. https://doi.org/10.1037/0033-295X.99.1.122

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Mall, J. T., Morey, C. C., Wolff, M. J., & Lehnert, F. (2014). Visual selective attention is equally functional for individuals with low and high working memory capacity: Evidence from accuracy and eye movements. *Attention, Perception, & Psychophysics*, *76*(7), 1998–2014. https://doi.org/10.3758/s13414-013-0610-2

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*(1), 103–107. https://doi.org/10.1038/nn2024

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

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. https://doi.org/10.1006/cogp.1999.0734

Mottarella, M., Yamasaki, B. L., & Prat, C. S. (2021). Relating individual differences in reading skill to neural indices of proactive control and online filtering during a working memory task. *Discourse Processes*, *58*(5–6), 569–591. https://doi.org/10.1080/0163853X.2021.1926407

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex*, *23*(2), 264–282. https://doi.org/10.1093/cercor/bhs007

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058

Oberauer, K. (2019). Working memory and attention – A conceptual analysis and review. *Journal of Cognition*, *2*(1), 36. https://doi.org/10.5334/joc.58

O’Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*(2), 283–328. https://doi.org/10.1162/089976606775093909

Osaka, M., Osaka, N., Kondo, H., Morishita, M., Fukuyama, H., Aso, T., & Shibasaki, H. (2003). The neural basis of individual differences in working memory capacity: An fMRI study. *NeuroImage*, *18*(3), 789–797. https://doi.org/10.1016/S1053-8119(02)00032-0

Osaka, N., Osaka, M., Kondo, H., Morishita, M., Fukuyama, H., & Shibasaki, H. (2004). The neural basis of executive function in working memory: An fMRI study based on individual differences. *NeuroImage*, *21*(2), 623–631. https://doi.org/10.1016/j.neuroimage.2003.09.069

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, *47*(4), 1343–1355. https://doi.org/10.3758/s13428-014-0543-2

Prat, C. S., & Just, M. A. (2011). Exploring the neural dynamics underpinning individual differences in sentence comprehension. *Cerebral Cortex*, *21*(8), 1747–1760. https://doi.org/10.1093/cercor/bhq241

Prat, C. S., Keller, T. A., & Just, M. A. (2007). Individual differences in sentence comprehension: A functional magnetic resonance imaging investigation of syntactic and lexical processing demands. *Journal of Cognitive Neuroscience*, *19*(12), 1950–1963. https://doi.org/10.1162/jocn.2007.19.12.1950

Prat, C. S., Madhyastha, T. M., Mottarella, M. J., & Kuo, C.-H. (2020). Relating natural language aptitude to individual differences in learning programming languages. *Scientific Reports*, *10*(1), 3817. https://doi.org/10.1038/s41598-020-60661-8

Prat, C. S., Yamasaki, B. L., Kluender, R. A., & Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. *Brain and Language*, *157–158*, 44–50. https://doi.org/10.1016/j.bandl.2016.04.007

Robison, M. K., Miller, A. L., & Unsworth, N. (2018). Individual differences in working memory capacity and filtering. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(7), 1038–1053. https://doi.org/10.1037/xhp0000513

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., Fox, P. T., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, *60*(1), 830–846. https://doi.org/10.1016/j.neuroimage.2011.11.050

Rypma, B., Prabhakaran, V., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Load-dependent roles of frontal brain regions in the maintenance of working memory. *NeuroImage*, *9*(2), 216–226. https://doi.org/10.1006/nimg.1998.0404

Schmeichel, B. J., Volokhov, R. N., & Demaree, H. A. (2008). Working memory capacity and the self-regulation of emotional expression and experience. *Journal of Personality and Social Psychology*, *95*(6), 1526–1540. https://doi.org/10.1037/a0013345

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. https://doi.org/10.1177/0956797612440101

Stocco, A., Lebiere, C., & Anderson, J. R. (2010). Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological Review*, *117*(2), 541–574. https://doi.org/10.1037/a0019077

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. https://doi.org/10.1038/nature04171

Yarkoni, T., & Braver, T. S. (2010). Cognitive neuroscience approaches to individual differences in working memory and executive control: Conceptual and methodological issues. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook of Individual Differences in Cognition: Attention, Memory, and Executive Control* (pp. 87-107). Springer New York. https://doi.org/10.1007/978-1-4419-1210-7

Appendix

BIRDS	BODYPARTS	CLOTHING	FRUIT	FURNITURE
CHICKEN	ARM	BELT	APPLE	BED
CROW	EYE	BLOUSE	BANANA	BENCH
DOVE	EAR	BRA	BLACKBERRY	BOOKCASE
DUCK	ELBOW	DRESS	CHERRY	CABINET
EAGLE	FINGER	HAT	COCONUT	CHAIR
FLAMINGO	FOOT	JACKET	GRAPE	COUCH
HAWK	HAND	JEANS	KIWI	DESK
OWL	HEAD	PANTS	LEMON	DRESSER
PARROT	KNEE	SANDAL	LIME	HUTCH
PIDGEON	LEG	SCARF	ORANGE	LAMP
ROBIN	LIP	SHOE	PEACH	OTTOMAN
ROOSTER	MOUTH	SKIRT	PEAR	RUG
SEAGULL	NECK	SOCK	PINEAPPLE	SHELF
SPARROW	NOSE	SUIT	RASPBERRY	SOFA
SWAN	SHOULDER	SWEATER	STRAWBERRY	STOOL
STORK	TOE	T-SHIRT	WATERMELON	TABLE

INSECTS	INSTRUMENTS	MAMMALS	TOOLS	VEHICLES
ANT	ACCORDION	BEAR	AXE	BICYCLE
BEE	BANJO	CAMEL	CHISEL	BOAT
BEETLE	CLARINET	CAT	CLAMP	BUS
BUTTERFLY	CYMBAL	COW	DRILL	CAR
CATERPILLAR	DRUM	DEER	FILE	CANOE
CENTIPEDE	FLUTE	DOG	HAMMER	JEEP
COCKROACH	GUITAR	ELEPHANT	LEVEL	MOTORCYCLE
CRICKET	HARP	GOAT	MALLET	PLANE
DRAGONFLY	PIANO	HAMSTER	PLIERS	SCOOTER
FLEA	SAXOPHONE	LION	RULER	SHIP
GRASSHOPPER	TAMBOURINE	MONKEY	SANDER	TANK
FLY	TROMBONE	MOOSE	SAW	TRACTOR
MOSQUITO	TRUMPET	MOUSE	SCREW	TRAIN
MOTH	TUBA	PIG	SHEARS	TRUCK
SPIDER	VIOLIN	SQUIRREL	SHOVEL	VAN
WASP	XYLOPHONE	TIGER	WRENCH	YACHT

Note. Reproduced from Mottarella et al. (2021)