



Published in final edited form as:

*J Exp Psychol Gen.* 2014 October ; 143(5): 1806–1836. doi:10.1037/a0036814.

## Central and Peripheral Components of Working Memory Storage

Nelson Cowan, J. Scott Saults, and Christopher L. Blume

Department of Psychological Sciences, University of Missouri

### Abstract

This study re-examines the issue of how much of working memory storage is central, or shared across sensory modalities and verbal and nonverbal codes, and how much is peripheral, or specific to a modality or code. In addition to the exploration of many parameters in 9 new dual-task experiments and re-analysis of some prior evidence, the innovations of the present work compared to previous studies of memory for two stimulus sets include (1) use of a principled set of formulas to estimate the number of items in working memory, and (2) a model to dissociate central components, which are allocated to very different stimulus sets depending on the instructions, from peripheral components, which are used for only one kind of material. We consistently find that the central contribution is smaller than was suggested by Saults and Cowan (2007), and that the peripheral contribution is often much larger when the task does not require the binding of features within an object. Previous capacity estimates are consistent with the sum of central plus peripheral components observed here. We consider the implications of the data as constraints on theories of working memory storage and maintenance.

---

A key issue in cognitive psychology is the nature of limitations in working memory, the small amount of information temporarily held and used in various cognitive tasks. An important question is to what extent working memory storage depends on a mental faculty that is general across domains as opposed to domain-specific (e.g., Kane et al., 2004). A long tradition of dual-task methods lends itself to the examination of this question, and has been used for many years (e.g., Allen, Baddeley, & Hitch, 2006; Baddeley, 1986; Baddeley & Hitch, 1974; Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002; Cowan & Morey, 2007; Fougne & Marois, 2011; Morey & Cowan, 2004, 2005; Morey, Cowan, Morey, & Rouder, 2011; Saults & Cowan, 2007; Stevanovski & Jolicoeur, 2007). With this method, it has been fairly well established that there is some dual-task interference between very different tasks. To some degree, for example, the need to retain visual items interferes with concurrent storage of verbal items, and vice versa. Yet, this methodology has never lived up to its promise because, we contend, a thorough analysis of the problem has not been available.

We examine the consequences of what we believe to be an improved methodology to ask how much of the working memory storage capacity is *central*, or capable of being allocated to different materials to be remembered according to task demands. This concept is separate from *peripheral* storage, which has different varieties, each of which can only be allocated

to a single type of stimulus regardless of task instructions. In our case, storage capacity could be allocated to colored objects (sometimes differing also in shape) and/or words (spoken or, in one experiment, written). Assuming that there is a limited amount of central storage to be allocated freely, the number of colored objects remembered should be reduced when the participant must also remember words at the same time, and vice versa.

A dual-task design, with memory for one or two different stimulus sets of different types required on a given trial, is required in order to estimate central and peripheral components of working memory. The reason is that memory for a single stimulus set theoretically can be based on multiple storage mechanisms, only some of which are shareable resources. With a single type of stimulus, multiple storage mechanisms cannot be separated. Here we apply a model in which memory for items in each modality is assumed to come from the sum of the contributions of central and peripheral components. The central component can be estimated as the portion of memory for stimuli of a certain type (e.g., colors) that have to be shared with stimuli of a second type (e.g., words) if that second type also is to be remembered. The peripheral component can be estimated as the portion of memory for stimuli of a certain type that does not have to be shared. Before the analysis of results into central and peripheral components can be accomplished, the number of items of each type retained in working memory must be estimated, as in past work (Cowan, 2001; Cowan, Blume, & Saults, 2013), separately for single- and dual-task conditions (cf. Saults & Cowan, 2007). Then certain subtractions between conditions can be carried out to estimate the central and peripheral components, as we explain below. Before explaining that, though, we discuss in more detail the theoretical implications of central and peripheral storage.

## THEORETICAL IMPLICATIONS OF CENTRAL AND PERIPHERAL STORAGE

The notion of central storage is assumed to involve categorical as opposed to sensory information. Categorical information refers to information that allows the assignment of each item to one of a finite number of categories. Considering the stimuli in our experiments, it is intended that categorical information could distinguish the different items. Information about a color is categorical in our experiments because it is enough to remember the conventional color group of each stimulus (red, blue, green, etc.) to perform correctly. We did not require non-categorical color information as we did not test multiple shades of the same basic color. The shapes we used included only one of each familiar category (star, square, circle, and so on) and therefore, again, could be retained as categorical information. The same was true of spoken digits or words; there are a finite number of categories and a very small number of different words were used over and over in any one experiment. Finally, although spoken voices must at first be distinguished on the basis of fine-grained acoustic information, the same few, quite discriminable voices were presented over and over in any one experiment, to allow categories reflecting the voices to be built up in memory quickly. Thus, we assess whether the amount of categorical information in a central part of working memory may be limited across modalities or types of stimuli, with interference even across different modalities.

Unlike categorical information, sensory information appears to be represented separately in each modality (vision and hearing in our case) with little interference between modalities

(e.g., Cowan, 1988). However, sensory information is quickly lost (e.g., Darwin, Turvey, & Crowder, 1972; Sperling, 1960), so categorical information is needed for recall in working memory tasks. We ensure that this is the case in most of our experimental conditions by including a pattern mask after a reasonable perceptual period, to overwrite any residual sensory information (cf. Saults & Cowan, 2007).

It is also possible to have peripheral information that is categorical rather than sensory in nature. A possible example is phonological information, which can be derived from either spoken or written words (e.g., Conrad, 1964). According to one traditional theory of working memory (Baddeley, 1986; Cocchini et al., 2002), all working memory storage could be peripheral. According to this theory, phonemes might be retained in one mechanism termed the phonological store, whereas visual or spatial features (line orientations, colors, spatial arrangements of objects, visual patterns, and so on) might be retained in another mechanism termed the visuospatial sketchpad. Provided that sufficient attention is made available to encode items of both types, the number of phonemes stored in working memory would not depend on the number of visuospatial elements stored concurrently, nor vice versa. Both would therefore be considered peripheral types of storage.

If the number of verbal and visuospatial units trade off, so that storage of one type limits concurrent storage of the other type, then according to our terminology there is a central storage mechanism that must be shared between the different types of information. Theories differ as to the amount of central storage versus peripheral storage to expect. One way in which central storage could work is that there could be a limit on the total number of items that can be remembered (cf. Cowan, 2001; Cowan, Rouder, Blume, & Saults, 2012). Each item in this approach is called a chunk of information (Miller, 1956), i.e., a group of elements that are strongly associated with one another and together form a member of a conceptual category. For example, a string of phonemes making up a known word together comprise a single chunk, as do an array of lines or curves making up a known geometrical shape such as a square, circle, or other familiar shape. According to this theoretical view, there would be a numerical limit of the central component of working memory to several chunks at once, and these chunks could be of any type (e.g., spoken words, shapes) or combination of types.

Peripheral stores could supplement central storage. To measure central storage, therefore, a method must be developed to separate it from peripheral storage, and soon we will explain our method to do so.

According to some other theoretical views, central storage should exist but it should be more severely limited. Attention could be focused on different memory representations in a rotating fashion, at a certain rate, to refresh them before a decay process makes them inaccessible to working memory (cf. Barrouillet, Portrat, & Camos, 2011). The working memory information could be stored as activated chunks in long-term memory, which is presumably not capacity-limited except that items soon become inactive because of decay, unless attention is used to refresh them (Camos, Mora, & Oberauer, 2011; Cowan, 1988, 1992). In that case, there would be a tradeoff between different types of storage; if attention cannot refresh all items at once, then some items will decay when they are not being

refreshed. Yet, the sharing of the refreshing process between modalities might be enough to produce only a relatively small amount of tradeoff between visual and verbal memory, and therefore only a small estimate of central storage.

In another popular view, we cannot make predictions because we believe that the amount of central storage to be obtained has been left unspecified. Specifically, Baddeley (2000, 2001) described an episodic buffer component that could hold several chunks and could supplement the phonological and visuospatial stores from his earlier model. This episodic buffer clearly is said to hold various types of information, including the binding between information from other stores, and semantic information. It is unclear, however, whether this buffer is used only when the other stores are inappropriate to hold the information. For example, if words and visual objects were held in separate buffers, could the episodic buffer hold a version of the items concurrently, perhaps in a more semantic form? Because we do not know, we will not suggest predictions of the total amount of central and peripheral storage based on this view.

The present experiments are designed to assess how much information is central and how much is peripheral, despite not knowing the exact nature of the central or peripheral memory representations. This assessment helps greatly to constrain models of working memory. Below, we discuss our analytic technique and its boundary conditions or limits, the historical context that motivated it, and the application of the technique to the present set of experiments.

## A NEW ANALYTIC TECHNIQUE AND ITS LIMITS

Our work provides a new analytic technique to assess central versus peripheral storage, as described above. The steps in doing so involve (1) the use of a discrete slot model to estimate the number of items in working memory appropriate to the test procedure; (2) examination of memory for two different stimulus sets, separately and also when memory for both sets is required at once; and (3) use of a new method to distinguish central and peripheral components of working memory, based on these data. To anticipate the results, relatively plentiful information about item features (in this study, color, shape, voice, and word) tends to be held in peripheral storage, whereas feature binding information does not seem to benefit from peripheral storage as much. For both item and binding information, a small but typically above-zero amount of information is held centrally.

### Limits of the Method

Before describing the analytic technique, we hasten to mention two important constraints in our investigation. First, we do not examine the conflict between two sets of stimuli with the same code (two nonverbal object sets or two verbal sets). Undoubtedly, more interference between sets at the time of memory maintenance would be obtained in studies with two stimulus sets of the same kind than with two different kinds (see for example Cowan and Morey, 2007; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012). It is of course also possible for visual stimuli to be encoded verbally (e.g., Conrad, 1964) or for verbal stimuli to be encoded visually (e.g., Logie, Della Sala, Wynn, & Baddeley, 2000), which would allow for more interference during maintenance, but any such recoding should be

minimized by the presentation of verbal and visual stimuli on the same trial. Still, we used articulatory suppression to avoid verbal rehearsal of either stimulus set. This technique should also prevent the verbal coding of visual materials (cf. Baddeley, Lewis, & Vallar, 1984). Our results should not be taken to indicate the most interference that is possible between two sets, only the amount of interference between two rather different sets at the time of memory maintenance.

A second constraint is roughly the converse of the first. We cannot rule out the possibility that if one stimulus set required the retention of object identities whereas another stimulus set required the retention of something different, such as the precise locations of a set of objects, there would be no interference between them at all. This, in fact, has been proposed in one recent investigation (Marois, 2013). Other investigations have similarly emphasized a separation between item capacity and precision as separate resources (e.g., Machizawa, Goh, & Driver, 2012; Zhang & Luck, 2008). However, one must also consider the possibility that a spatial arrangement of locations can be combined to form a spatial configuration, a kind of item amalgamation or chunking that would in effect reduce the number of items to be held in memory independently (Jiang, Chun, & Olson, 2004; Miller, 1956). Nevertheless, any conclusions we draw are specific to situations in which both sets to be remembered contain categorizable items or identities, not locations.

### Methods of Analysis for Several Slightly Different Procedures

Previous dual-task memory studies have examined the tradeoff between memory for one stimulus set and memory for the other using various metrics (e.g., Cocchini et al., 2002; Fougne & Marois, 2011; Morey et al., 2011). These metrics, however, have not been based on the estimated number of items in memory, which must take into account the effects of guessing in a recognition task. The importance of such estimates is that they allow a principled description of the allocation of storage to different kinds of items. We now have well-considered estimates of the number of items in storage for various versions of tasks in which a single set of items is followed by a single-item probe to be judged the same as one of the items in the set or different from all of them (Cowan, 2001; Cowan, Blume, & Saults, 2013; Rouder et al., 2008; Rouder, Morey, Morey, & Cowan, 2011).

In the tasks we will use, there is a set of  $N$  items to be remembered (with no duplicates in the set), followed by a probe to be judged the same as one item or different from all of them. The assumption is that  $k$  of them actually are remembered and we wish to estimate an individual's  $k$  value. Response to the probe can be based on a comparison of that probe to a single list or array item in working memory, if the probe appears in a way that makes clear which item may be identical to the probe. If, on the other hand, the task is such that the probe is always presented without an identifying cue, it must be compared to all items in working memory. Finally, if an item matching the probe is not found in working memory and some items were forgotten, the response can only be based on a guess. Based on this fundamental logic, Appendix A reviews the formulae used for slightly different procedures in this study, to estimate the number of items held in working memory. We will discuss these methods to estimate  $k$  before going on to explain how they are used to estimate central and peripheral components of working memory.

### Probe in the Location of a Specific Stimulus, with an Old or New Feature

In each case, the formula for items in working memory shown and explained in Appendix A represents a logical analysis of the task demands. In one visual task version we use, an array is followed by a single-item probe appearing at the location of an array item; the probe is either identical to that array item (e.g., both of them blue) or differs from it by one feature (e.g., the probe red unlike any item in the array). This task version is also extended here to verbal lists in some experiments, by presenting a probe word within a series of identical nonverbal sounds to indicate the serial position of the probed item.

For this kind of task in which a specific stimulus is probed, if the participant has the probed item in working memory, he or she presumably will correctly judge whether it is the same as or different than the probe. If the item at the probed location in the array is not in working memory, either because it was not presented or because it was presented but not remembered, the participant will not have any relevant information and must guess.

This is the model proposed by Cowan (2001) and, as show in the Appendix, the formula is simply  $k=N(H-F)$ , where  $H$  is the proportion of hits, i.e., changes successfully detected, and  $F$  is the proportion of false alarms, i.e., incorrect reports of a change. This kind of model makes predictions that have been strikingly confirmed for visual arrays (Rouder et al., 2008). It correctly predicts that the receiver operating characteristic (ROC) curve should be linear, whereas signal detection theory incorrectly predicts a curvilinear ROC curve, and Cowan's model also correctly predicts exactly how the effects of criterion bias should be combined with  $k$  for different values of  $N$ . Donkin, Nosofsky, Gold, and Shiffrin (2013) further show that models with discrete slots provide the preferred account of reaction times.

The essential assumption of the model is that the knowledge of a particular item is all-or-none; there can be no partial knowledge of an item in the model. This need be assumed only in so far as the tested feature is concerned. For example, if one always tests for color and never for shape, the model's assumptions could be satisfied sufficiently to make it appropriate as a measurement model for the colors, regardless of whether knowledge of the shapes is as good as knowledge of the colors. Luck and Vogel (1997) suggested that knowledge of an item in working memory carries with it knowledge of all of its features, but that assumption has been brought into question (Cowan, Blume, & Saults, 2013; Oberauer & Eichenberger, 2013; Xu, 2002) and we no longer hold it.

The assumption of discrete knowledge of the items is an oversimplification. Zhang and Luck (2008) devised a technique in which the items were drawn from a continuum: colors from any point on a color wheel or figures rotated to any orientation. The response was to reproduce the stimulus value as a location on a response wheel, so that the precision of the memory representation could be measured. Zhang and Luck found that participants did not spread their attention evenly over all items, with less precise representations as the number of array items increased; attention was spread only to about three items. In fact, participants are unable to spread attention to more objects in working memory at once even when the incentives favor it (Zhang & Luck, 2011). The evidence suggests that when there is only one item, precision is greatest, whereas precision is diminished with two items and further with

three items, reaching a stable plateau after that (for further evidence see Anderson, Vogel, & Awh, 2011).

The simpler model of Cowan (2001) and the other models used in this article do not include a precision factor. Our assumption based on the evidence is that the precision factor is likely to be important only when the stimuli are chosen from a continuum or differ by small amounts. We chose stimuli that were designed to be categorically different, with features including colors with different basic names (red, green, etc.), simple shapes (star, cross, etc.), simple familiar words, and voices that sounded very different from one another. If there is a partial loss of precision in the representation based on decay or interference, we simply assume that this degradation does not matter until it passes some threshold.

Not everyone in the field is convinced that the models with a limited number of discrete slots are correct. Bays and Husain (2008) argued that attention can be spread thinly across all items (but see the response by Cowan & Rouder, 2009). Van den Berg, Shin, Chou, George, and Ma (2012) proposed a model with a variable precision of items (as did Fougnie, Suchow, & Alvarez, 2012) and Sims, Jacobs, and Knill (2012) proposed a model in which the number of items in working memory can vary from trial to trial. Although there may well be some truth to these approaches that include precision, they do not seem to serve as an argument against using a measurement model with a simpler formula that just counts how many items are in working memory. An item might be considered out of working memory because there is no information about it, or because the working memory representation is too poor to be used to guide a response to the probe. Given that our aim is to determine how much of working memory is central between two different stimulus sets or peripheral to just one of them, it seems at worst a helpful and convenient simplification to use discrete slot models for this purpose, given the categorically distinct stimuli that we use in these experiments.

### **Probe in a Central Location, to be Matched to All List or Array Items**

**New-feature probe**—In another task version that we use, a probe is centrally presented and is identical to one array item or has a new feature not found in the list or array. In the verbal list version, a single probe is presented and could match any list item, or none of them. Then if there is a matching item from the array or list in working memory, the participant presumably will know that there was a match. If there is no matching item in working memory, however, and there are some array items not represented in working memory, the participant has no way to know whether the forgotten items do or do not include the probe, and therefore the participant must guess. This model was presented by Cowan, Blume, and Saults (2013). Its formula is  $k=N(H-F)/H$ , with terms as defined above; that is, it is the formula for in-location probes divided by the proportion of changes detected or hit rate. This formula is similar, but not identical, to a model for a whole-array probe provided by Pashler (1988), as Appendix A shows.

**Feature-binding probe**—Finally, we also used a task in which a central probe was either identical to one list item, or else consisted of a recombination of the features of two items (the shape of one item combined with the color of another, or the voice of one item

combined with the word identity of another). In this task, if there is a match, the participant can get the answer by having the matching item in working memory. If there is no match, then the participant can get the answer by having in working memory either of two items: for example, either the item with a shape matching the visual probe or the item with the color matching the visual probe. If neither of these items is in working memory, the participant must guess. This model, too, was presented by Cowan, Blume, and Saults (2013). Its form is complex, as indicated in Appendix A.

### Central and Peripheral Components Based on Single and Dual Tasks Together

The present work extends the discrete-slots analysis of items in working memory (Appendix A) to tasks with two sets of items potentially in competition for memory capacity. Figure 1 is a schematic diagram illustrating how this application of the capacity-estimation models to the dual-task situation occurs when there is one set of verbal items and one set of nonverbal objects to be remembered. The number of items in working memory is partitioned into four subsets: the number of verbal items that are retained no matter what the task instructions, or  $P_{verb}$ ; the number of nonverbal objects that are retained no matter what the task instructions, or  $P_{obj}$ ; the number of slots that are filled with verbal items when both sets are to be remembered, but with objects when only those are to be remembered, or  $C_{verb}$ ; and the number of slots that are filled with nonverbal objects when both sets are to be remembered, but with verbal items when only they are to be remembered, or  $C_{obj}$ . In this nomenclature,  $P$  stands for peripheral and  $C$  stands for central. Moreover, conceptually we care about the total number of slots that can be allocated freely to either verbal or nonverbal information, or  $C$ , such that  $C = C_{verb} + C_{obj}$ . These quantities can be calculated if one has estimates of the number of objects and verbal items retained both when the task is to retain only one set, and when the task is to retain both objects and verbal items at the same time. For example, the difference between the number of verbal items remembered alone versus the smaller number of verbal item remembered in a dual task yields an estimate of the number of verbal items that were sacrificed for the sake of object memory in the dual task,  $C_{obj}$ . When one adds this to the converse, namely the number of objects sacrificed for the sake of verbal information in a dual task,  $C_{verb}$ , then the sum  $C$  is obtained and is termed the central portion of working memory.  $C$  is thus the portion that can be reallocated depending on task demands. It also follows that, in a single-stimulus-set condition, memory results from the peripheral portion plus the central portion: for verbal items,  $C + P_{verb}$  and, for non-verbal objects,  $C + P_{obj}$ . These expressions allow the calculation of the peripheral portions of working memory,  $P_{verb}$  and  $P_{obj}$ . This basic description of the method is explained in a more mathematical form in Appendix B.

## CONTEXT FOR THE PRESENT ANALYSIS

We will discuss the relevant literature broadly speaking, and then the immediate precursor to our study. In an early popular model of working memory (e.g., Baddeley, 1986), information was said to be stored in code-specific buffers, the phonological loop and the visuo-spatial sketchpad. Cowan (1988) proposed that information was stored as the activated portion of long-term memory, which could exhibit feature-specific interference that would make storage seem domain-specific like Baddeley's buffers. Additionally, though, Cowan



suggested that there was some storage that took place in the focus of attention, regardless of the domain of coding. The peripheral storage in the present work might map onto the activated portion of long-term memory and the central storage might map onto the focus of attention. Note that this mapping depends, though, on the focus of attention being voluntarily directed. The mapping theoretically could break down if, for some reason, there is domain-specific information in the focus of attention that cannot be traded off between domains. For example, it is possible that some stimuli linger for a while in the focus of attention even after they become irrelevant to the task (Oberauer, 2002). To the extent that the allocation of attention is beyond the participant's control, any such information uncontrollably in focus contributes to the peripheral components of our model, not to the central component.

Other models also have proposed some type of central store in addition to some type of peripheral storage (e.g., Atkinson & Shiffrin, 1968; Baddeley, 2000; Baddeley & Hitch, 1974; Broadbent, 1958; Oberauer, 2002). The balance between peripheral and central storage, however, remains unclear. Cowan (2001) suggested that when mnemonic strategies like rehearsal and grouping are prevented and sensory memory is eliminated, adults are restricted to central storage and can remember only about 3-5 items, not about seven as Miller (1956) famously proposed. Oberauer et al. (2012) suggested that this capacity limitation is not the focus of attention, and that only one item at a time is in the focus of attention.

An important reason to raise the possibility of central storage is that peripheral storage cannot, in principle, include the links between very different types of material such as spatial and verbal materials. This reasoning led Baddeley (2000) to add to his model a central component called the episodic buffer, with a function that includes storage of the binding between the different attributes of a stimulus.

One can distinguish between two questions about any particular store: how many objects it can include, and how many attribute bindings it can include. The answer to these two questions is the same, according to Luck and Vogel (1997), based on their seminal task in which an array of objects is followed by a probe to be judged the same as the array or differing from it in one feature. Specifically, they found that objects with multiple features (e.g., color, orientation, length, and presence or absence of a gap within a bar) are retained in an all-or-none fashion, such that items that are remembered include all of their features. This result is, however, seemingly at odds with some subsequent work. Wheeler and Treisman (2002) found that the proportion correct for detecting feature changes was, under some circumstances, higher than for the binding between features. Vul and Rich (2010) found that features are remembered independently across objects, with the probability of remembering the binding between two features no higher than one would expect based on the probability that the two features were by chance retained for the same object (see also Bays, Wu, & Husain, 2011; Fougne & Alvarez, 2011). Cowan, Blume, and Saults (2013) replicated that effect with the finding that people retained about 3 array items in various conditions, but that if retention of two attributes (color and shape) were both required, some of the objects were retained with only color or only shape. (This model was confirmed and extended to more features by Oberauer & Eichenberger, 2013). In some of our experiments we used items

with two attributes (colored shapes, and digits spoken in different voices) in order to compare the means of storage of the attributes, and of the binding of attributes. With our new analysis method to be presented, we aim to examine the retention of both attributes and their binding in both peripheral and central stores.

### **Immediate Precursor to the Present Study: Saults & Cowan (2007)**

One study (Saults & Cowan, 2007) did result in estimates of the number of items in working memory in a dual task. Because the present study was designed to rectify issues remaining after that study, it will be described in some detail. In five experiments, spoken digits were presented from four loudspeakers concurrently in different voices to avoid rehearsal, while any array of several colored squares also was presented. Either the visual or the spoken array was then repeated as a probe, and was identical to the first array of that modality or differed in the identity of one item. The task was to indicate whether the studied array and the probe array were the same or different. In some trial blocks, the spoken digits were to be ignored; in others, the visual items were to be ignored; and in still others, both modalities were to be attended. In some trials, a difference consisted of a new feature (e.g., a red item where a green item had been, when no other studied item on that trial was red) whereas, in other trials, a difference consisted in a changed but familiar item (e.g., a red item where a green item had been, when another array item was also red).

In some of the experiments of Saults and Cowan (2007), after an amount of time that was shown to be enough to encode the stimuli (cf. Jolicoeur & Dell'Acqua, 1998; Vogel, Woodman, & Luck, 2006), a pattern mask was presented to eliminate any residual sensory memory. Although some evidence suggests that sensory memory ends in a matter of several hundred milliseconds (e.g., Efron, 1970a, 1970b, 1970c; Massaro, 1975; Sperling, 1960), there is other evidence of modality-specific memory that might be considered sensory lasting several seconds, both in audition (e.g., Cowan, Lichty, & Grove, 1990; Darwin et al., 1972) and in vision (e.g., Phillips, 1974; Sligte, Scholte, & Lamme, 2008). Massaro referred to a brief, literal sensory memory and a longer, synthesized sensory memory; Cowan (1988) similarly distinguished between two phases of sensory memory in every modality, a brief literal afterimage followed by a second, more processed sensory memory. The mask is included because of the longer form of sensory memory (or Massaro's synthesized sensory memory).

In some experiments of Saults and Cowan (2007), stimuli in the two modalities were presented one after the other in order to eliminate any interference at the time of encoding. In most experiments, digit-location associations and color-location associations could be used to carry out the task; in the last experiment in the series, though, the only possible information that could be used was digit-voice associations and color-location associations (because the digits were spatially shuffled between study and test).

The results of all of the experiments of Saults and Cowan (2007) showed a tradeoff between modalities. In the experiments with masks to eliminate sensory memory, people could remember about 4 visual items or 2 acoustic items in unimodal attention conditions, and about 4 items total in the bimodal condition (about 3 visual and 1 acoustic).

## THE PRESENT EXPERIMENTS

The present work follows up on this previous work of Saults and Cowan (2007) with the intent of improving the work in several ways. (1) First, we presented only a single item as a probe in order to limit the number of decisions that the participant must make (Luck & Vogel, 1997). (2) Second, we more consistently incorporated what we considered to be the best practices of Saults and Cowan (use of a mask, and sequential rather than concurrent presentation of the two stimulus sets to be remembered). (3) Third, inasmuch as Saults and Cowan found that only about two items could be remembered from concurrent arrays of sounds, we changed the verbal stimulus sets to sequences rather than arrays, and incorporated articulatory suppression (repetition of the word “the” during stimulus presentation and retention intervals) to prevent rehearsal. The use of verbal sequences avoids the perceptual bottleneck of encoding concurrent sounds into working memory.. (4) Fourth, we separated the procedures so that on a given trial, a participant only had to remember a single kind of feature of each object, or only had to remember the binding between features of each object. We did this to simplify the task so that the existing formulas of performance (Appendix A) would apply.

The other main change has to do with the manner in which the results were evaluated. Saults and Cowan (2007) assumed that, because unimodal visual capacity and bimodal capacity both were about 4 items, this must measure the capacity of the focus of attention. We now believe that logic to be faulty. Using the method of analysis of bimodal results described above and in Appendix B, we estimate that the peripheral storage of visual information included 2 to 3 items, the peripheral storage of acoustic information from spatial arrays of sounds included at most about 1 item, and the central storage component included only 1 to 2 items. Thus, the analysis method made a big difference even in the understanding of the results of Saults and Cowan.

### Overview of the Present Series of Experiments

Nine experiments were conducted to test the generality of the estimates obtained from Saults and Cowan (2007) according to our re-analysis of their data. The experiments can be divided into four sets based on the nature of the probe and conditions. Conditions in the first three sets of experiments are described in Tables 1-3, respectively. First, in Experiments 1a and 1b (see Table 1), the stimuli were the simplest we could devise but, as a result, the modalities were not closely comparable. Specifically, the acoustic probe (test) stimulus was a single digit (with no indication of serial position in the list) whereas the visual probe stimulus was a single item presented in the same spatial location as the corresponding array item. Different formulas were used to estimate items in working memory in the two modalities. In these experiments, effects of rehearsal were also investigated by the use of suppression versus tapping, but no difference was observed.

In Experiments 2a-2c (see Table 2), a serial position cue was added to the verbal probe so that the task became more directly comparable in the two modalities. Moreover, given that we observed a smaller central component than Saults and Cowan (2007), we considered whether aspects of the sequential presentation of verbal items (compared to their concurrent presentation as acoustic arrays in Saults & Cowan) were responsible. Specifically,

participants might combine temporally adjacent items to form larger chunks (Miller, 1956). To discourage this and thus make retention more dependent on attention, we introduced semantic-category and voice changes within lists. We tried lists of words from different semantic categories, thinking that words from the same semantic category might be chunked together (e.g., McElree, 1998; Miller, 1956), which could reduce the load on central processing. We also used lists with voice changes in some conditions of Experiment 2b, because Goldinger, Pisoni, and Logan (1991) found that under certain circumstances, it was harder to retain lists of words spoken in different voices (i.e., speaker variability). This finding may occur because the words in a variable-voice list are not perceived as falling within a single perceptual stream (Bregman, 1990; Macken, Tremblay, Houghton, Nicholls, & Jones, 2003). Despite these attempts, the same basic pattern of results was obtained in all conditions. Finally, we considered that the peripheral storage upon which participants relied might be a form of auditory sensory memory that somehow survived the mask. To rule this out, the acoustic stimuli were replaced by printed lists of verbal items in Experiment 2c. This, however, failed to reduce peripheral storage or increase central storage; the pattern of results remained basically the same.

Experiments 3a-3b and 4 (see Table 3) were conducted in order to apply the central / peripheral storage distinction to the case of attribute binding information. This purpose required some changes in design, the key one being the presentation of items with two physical attributes each (object color and shape; spoken word identity and voice). Also, the probe position cues were omitted; the spoken digit probe was presented without a serial position cue and the visual probe (on other trials) was presented at the center of the screen, not in the spatial location of one array item. This ensured that position information could not be used to reconstruct the binding between attributes.

Finally, Experiment 5 was conducted to assess the generality of the findings of the previous experiments. It included both item and binding trial blocks in one experiment. It also examined the features underlying binding in more detail: not only memory for spoken words but also, separately, for the voices in which they were spoken, and not only memory for object colors but also, separately, for the shapes in which they were presented.

These experiments included 26 different combinations of experimental variables yielding estimates of central, verbal, and visual nonverbal working memory parameters. To preview, these are the two most notable findings: First, the estimate of items in the central resource, *C*, is consistently small, in contrast to the conclusions of Saults and Cowan (2007). Second, estimates of peripheral storage are larger for item-change tasks than for binding-change tasks. These results call for a theoretical modification of a view in which individuals hold several items concurrently in the focus of attention continually from the time of the presentation of these items until the time of test (Cowan, 1988, 1999, 2001, 2005). We will argue that a multi-item focus does exist, but is not loaded continually during the retention interval with all the items in storage. Recent support for the concept of working memory items both inside and outside of the focus of attention comes from recent neuroimaging work showing stimulus-specific activation patterns for items to be considered for a current task (i.e., in the focus of attention), but not for all items to be remembered (Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012).

Fougnie and Marois (2011) also carried out a set of experiments similar to some of ours. They did not, however, modify the capacity formula in the manner we suggest, and they did not apply the formulas to separate central and peripheral components of storage. Moreover, they used articulatory suppression to prevent rehearsal in only one experiment. They observed more dual-task tradeoff when feature binding was required than when only feature information was required by the task. Their experiments in which feature binding was not required, however (their 3a and 3b) involved tonal stimuli, whereas their studies of memory for feature binding involved verbal acoustic stimuli. Using a wider variety of verbal stimuli, we find that the difference between attribute and binding memory rests not in the central component, but primarily in peripheral components.

## EXPERIMENT SERIES 1: EFFECTS OF ARTICULATORY SUPPRESSION

### Experiment 1a

In this experiment, we made use of the stimulus arrangement that seemed most simple and natural (Figure 2; Figure 3, top panel). Arrays of colored squares were combined with lists of spoken digits and the conditions included attention to visual stimuli only, acoustic stimuli only, or both modalities. Masks were presented in order to eliminate any residual sensory memory (Saults & Cowan, 2007). Then a probe was presented in an attended modality, a single spoken digit or a colored square in the spatial location of one of the array items. The item was always either identical to an array or list item (in the visual case, the replaced item) or different from all of the studied items. During the trial, either articulation was required, to suppress any covert verbal rehearsal, or repeated finger-tapping was required, as a secondary-task control.

### Method

**Participants:** The participants were 24 college students (14 female) who received course credit for their participation. Two female participants were excluded because of equipment problems.

**Apparatus and Stimuli:** Auditory stimuli were the spoken digits, 1-9, digitally recorded using an adult male voice. Recorded digits were temporally compressed to a maximum duration of 250 ms but presented at a pace of 2 digits per second. Temporal compression, from 65% to 95% of each word's original duration, was accomplished without altering pitch using the software Praat (Boersma & Weenink, 2009). The auditory mask combined all digits with onsets aligned. Digits and mask were presented to each ear with an intensity of 65-75 dB(A) using Telephonics TDH-39 headphones. Visual stimuli were presented on a 17-inch cathode ray tube monitor (1024 by 768 pixels). Each visual study array was presented for 500 ms. It consisted of 5 squares whose colors were sampled without replacement from 10 colors (black, white, red, blue, green, yellow, orange, cyan, magenta, and dark-blue-green). Squares were randomly positioned on the screen as in (Cowan et al., 2005): at a viewing distance of 50 cm, the array fell within 9.8 horizontal and 7.3 vertical degrees of visual angle. Squares of  $0.75^{\circ}$  on a side were randomly placed within this region except that the minimum center-to-center separation between squares was  $2.0^{\circ}$  and no square was located within  $2.0^{\circ}$  of the center of the array. Patterned masks consisted of

identical multicolored squares in the same locations as the items in the studied array, also presented for 500 ms.

**Procedure:** Each participant was tested individually in a quiet room. An experimenter was present throughout each session. Visual, auditory, and combined bimodal memory capacities were tested using a general procedure similar to Experiment 4 of Saults and Cowan (2007) with a few important differences. Most notably, in the current study, we presented sequences of 5 spoken digits before or after arrays of 5 colored squares and the probe was a single spoken digit or colored square the same or different from any stimulus in the study items. One of two secondary tasks, manual tapping or articulatory suppression, were performed during encoding and retention of the study items.

All conditions were blocked and their order counterbalanced between participants. Each session was divided into two subsessions, one with each secondary task, tapping or articulatory suppression; half had tapping first. Each of these subsessions was sub-divided into three blocks of memory load conditions: *remember visual*, *remember auditory*, and *remember both*. Each participant did these three memory load blocks in the same order within each secondary task subsession; the six possible orders of memory load were distributed equally across participants who did each order of the secondary tasks.

Each secondary task subsession began with general instructions about the memory task and the secondary task, as well as practice doing the secondary task at the proper rate. For the articulatory task, participants practiced whispering “the” in time with a click from the computer presented twice per second for 30 seconds. For the tapping task, participants practiced tapping the control key in time with a click presented twice per second for 30 seconds. During the memory task, they were instructed to begin the secondary task (tap or whisper “the”) when they saw the fixation cross, and stop when they saw or heard the probe.

Next, the participant began the first block of memory trials. An illustration of a trial is shown in Figure 2, which also summarizes the main differences between trials (inset). All stimuli were presented on a screen with a uniform medium gray background. Each trial began with fixation cross presented in the center of the screen for 1000 ms, followed by a 500-ms blank screen. This was followed by an array of 5 different colored squares lasting 500 ms, and a sequence of 5 different digits presented 2 per second over headphones. The two modalities were separated by an interstimulus interval (ISI) of 500 ms; one modality was completed before the other one started (Figure 2) and each modality occurred first on half of the trials. A blank screen appeared for 500 ms after the study items. Then five multicolored-mask squares appeared in the same locations as the squares in the study array. At the same time, an auditory mask, consisting of the combined digits, was presented via headphones. Thus, even though the target memoranda were presented consecutively by modality (all auditory before all visual or vice versa), the mask was presented with both modalities at once. Although this means that the target-mask interval varied between modalities used on a trial, the timing was counterbalanced across trials and the arrangement was considered important in order to avoid separate distracting switches of attention from one mask modality to the other.

Immediately after the 500-ms mask, a probe was presented. The auditory probe was a spoken digit identical to one in the study list or different from any in the study list. When an auditory probe occurred, a “?” appeared in the center of the screen. The visual probe was a square in the same location and the same color as a study square or in the same location as a square but a different color than any square in the study array. The participant was to press the “S” key if the probe was the same as a study item and the “D” key if the probe was different from all of the study items. The “?” or probe square remained on the screen until a response was recorded. Then participants saw “Correct” or “Incorrect” as feedback in the center of screen until they pressed “Enter” to begin the next trial.

For each subsession (tapping or articulatory suppression) there were three blocks of memory trials. Each visual-load block consisted of 8 practice and 40 experimental trials, in which participants were instructed to remember only the colored squares and were always tested with a probe square. Analogously, each auditory-load block consisted of 8 practice and 40 experimental trials, in which participants were instructed to remember only the spoken digits and were always tested with a probe digit. Finally, each bimodal-load block consisted of 16 practice and 80 experimental trials, in which participants were instructed to remember both the colored squares and the spoken digits. In these blocks there were an equal numbers of auditory- and visual-probe trials randomly intermixed. In every block of trials, half of the trials for each probe modality and presentation order (visual-first or auditory first) were change (different) trials. In experimental blocks with auditory probes, the probe digit in no-change (same) trials occurred in each serial position of the study list equally often.

**Results and Discussion**—The proportions correct are shown in Table 1 for each condition separately. In Figure 4, unimodal and bimodal proportions correct for verbal and object sets, collapsed across other conditions, is represented along with the next seven experiments. Clearly there was diminished performance in the bimodal conditions relative to the unimodal conditions, although the loss was modest for verbal items in Experiment 1a, and more notable for nonverbal objects, in support of the claim of an asymmetry in dual-task tradeoffs (Morey & Mall, 2012; Morey, Morey, van der Reijden, & Holweg, 2013; Vergauwe, Barrouillet, & Camos, 2010).

The verbal and object results in this experiment had to be analyzed with different formulas (Appendix A) because, in the verbal case, the probe could match any list item; whereas, in the object case, knowledge of a specific item was probed. The central and peripheral storage parameters are shown both in Table 1 and in Figure 5. The central parameter was less than 1 item in both the tapping and suppression conditions, and the auditory and visual peripheral storage parameters were both several items in both conditions. Articulation versus tapping made no statistical difference in *t* tests for the  $C$ ,  $P_{verb}$ , and  $P_{obj}$  parameters, but all six means were significantly above zero by *t* test (as shown in Table 1).

The absence of an effect of articulatory suppression suggests that, with suppression or without it, there was no effective use of verbal rehearsal to recode and rehearse the visual stimuli in phonological or verbal form. This result mirrors the finding of previous research (notably Morey & Cowan, 2004) showing no effect of articulatory suppression on visual array memory.

**Consistency with past capacity estimates:** Previous work has put the mean capacity estimate at 3-5 items and, for simple visual items, typically toward the lower end of that range (e.g., Cowan, 2001; Cowan et al., 2012; Luck & Vogel, 1997; Rouder et al., 2008). It should be emphasized that to obtain a comparison with the present results, one should not use only the central component but the sum of the central and peripheral components. For example, for the articulatory suppression condition of the present experiment, the verbal capacity (based on means in Table 1) is  $0.68(\text{central})+3.75(\text{peripheral})=4.43(\text{total})$ , and for objects it is  $0.68(\text{central})+2.23(\text{peripheral})=3.91(\text{total})$ . These numbers are quite consistent with past work. The finding of limited tradeoffs between modalities is also consistent with past work when an acoustic or verbal series is paired with a visual array (Cocchini et al., 2002; Fougine & Marois, 2011). These observations hold across the series of experiments taken as a whole.

## Experiment 1b

Perhaps the small size of the central resource parameter in Experiment 1a occurred because the acoustic stimuli all were presented in the same male voice, which elicit an auditory sensory means of storage for that modality despite the presence of a mask. Therefore, in the present experiment, we presented auditory stimuli in different voices and in alternating ears to make it more difficult to use a coherent auditory sensory stream (Figure 3, second panel).

### Method

**Participants:** The participants were 25 college students (17 female) who received course credit for their participation.

**Apparatus and stimuli:** This experiment replicated Experiment 1a in every way except for the auditory stimuli, which were the digits 1-9 spoken in five voices: those of a male adult, two distinctive female adults (one with a deeper voice than the other), a male child, and a female child. The temporal compression algorithm described for Experiment 1a was again applied so that none of the recordings was longer than 250 ms. Then, 20 different auditory masks were created by combining different digits spoken by all 5 voices in each channel (left and right) with aligned onsets. Auditory study items on each trial consisted of 5 digits randomly sampled without replacement, each digit spoken by a different voice. On each trial, study digits were presented alternately to the left and right ear. Which ear received the first item alternated from trial to trial. A 2-channel auditory mask was randomly selected on each trial, and presented at the same time as the visual mask.

**Results and Discussion—**As shown in Table 1 and Figures 4 (overall proportion correct) and 5 (parameter values), the results were very similar to Experiment 1a. Most importantly, it was again found that there was a dual-task tradeoff but that most of the storage was peripheral (modality- or code-specific), and only a smaller amount was central in nature. The fact that the results were similar across experiments goes against the notion that an auditory sensory stream in Experiment 1a was an important contributing factor to the results.



In this experiment, unlike the first experiment, the articulation condition produced lower estimates of  $P_{verb}$ ,  $t(24)=3.57$ ,  $p=.002$ , and lower estimates of  $P_{obj}$ ,  $t(24)=3.33$ ,  $p=.003$ , compared to the tapping condition. This finding, in combination with the absence of such an effect in Experiment 1a, suggests that the presentation of changing voices may cause distraction. In the tapping condition, rehearsal may help to overcome or compensate for that distraction.

Of the six parameters shown in Table 1, five were significantly above zero (Table 1), the exception being the  $C$  parameter in the tapping condition. Perhaps in that condition, verbal rehearsal tended to replace central storage as a mnemonic strategy.

As in Experiment 1a and Morey and Cowan (2004), the absence of an effect of articulatory suppression suggests that, with suppression or without it, there was no effective use of verbal rehearsal to recode and rehearse the visual stimuli in phonological or verbal form. (This would not prevent phonological encoding of the spoken verbal stimuli, and phonological storage could be the basis of peripheral verbal storage; see for example Baddeley et al., 1984). Nevertheless, we continue to use suppression in all of the experiments to make sure that rehearsal is discouraged in all test circumstances.

## EXPERIMENT SERIES 2: EFFECTS OF SEMANTIC CATEGORY MIXTURE

### Experiment 2a

Next we explored whether semantic differences between the acoustic items would increase central storage (Figure 3, third panel). Mixed-category lists might help to prevent a potential mnemonic strategy that worried us, the combination of list items to form semantically coherent, multi-word chunks (McElree, 1998; Miller, 1956), which would lower the load on the central store.

In this experiment, we also modified the stimuli so that the same formula could be used to assess the verbal and object stimulus sets. Specifically, acoustic markers were used to indicate the serial position of the list to which the probe word was to be compared. This is conceptually comparable to putting a visual probe object at a specific item location to indicate to which array item the probe was to be compared.

### Method

**Participants:** The participants were 24 college students (13 female) who received course credit for their participation.

**Apparatus and stimuli:** The visual stimuli were reduce to 4 per array, and the same was true of spoken lists, to accommodate the greater amount of information per spoken list in this experiment. Acoustic stimuli were spoken in a male voice and were items from four categories: directions (*up, down, in, out, on, off, left, right*), letters (*B, F, H, J, L, Q, R, Y*), digits (*1-9*), and body parts (*head, foot, knee, wrist, mouth, nose, ear, toe*). Each list consisted of either four words drawn from the same category, or one word drawn from each category for a total of four (as illustrated in the third panel of Figure 3). The word “go”, in the same voice, was recorded and used as an auditory mask. The durations of all auditory

stimuli were less than 500 ms. They were presented to participants with intensities between 65 and 75 dB(A) over Telephonics TDH-39 headphones.

**Procedure:** All trials in this experiment included articulatory suppression to prevent rehearsal. Each session began with instructions and a practice example of each kind of trial. The instructions included a step-by-step demonstration of exactly what a participant would see and hear during each trial. Then there was a real-time practice example of each kind of trial so that the participant had a chance to see and hear each kind of cue and stimulus. More extended practices followed this instruction phase. First the participant got instruction and practice doing articulatory suppression. As in the previous experiments, the participants practiced for 30 s whispering “the” twice a second in time to a click. Then they did practice blocks of each memory condition while also doing articulatory suppression. They first did 8 auditory practice trials, each preceded by instruction to attend to and remember only the auditory stimuli and ending with an auditory probe. Next they did 8 visual practice trials, each preceded by instructions to attend to and remember only the visual stimuli and ending with a visual probe. Finally they did a practice block of bimodal memory trials, each preceded by instruction to attend to and remember both auditory and visual stimuli. Half of these trials ended with an auditory probe and half ended with a visual probe. Each block of practice trials included examples of the same combinations of stimuli the participant would see and hear in the experimental trials. For example, half of the trials in each practice block presented auditory study stimuli first, and half presented visual study stimuli first. In half of the practice trials, the four auditory study stimuli in a list were from the same category, and the remaining practice trials were mixed lists with one word from each category. After participants practiced a block of each kind of memory trial, with articulatory suppression, the experimenter left the booth but continued to monitor participants via a camera and microphone while they completed 128 experimental trials, including all kinds of memory trials intermixed in a different random order for each participant.

The stimuli and their relative timings were similar for all trials, with the exception of the pretrial memory cue, the order of study modality (auditory-first or visual-first) and the modality of the probe stimuli. Each trial began with the display of one of the following memory cues in the center of the screen that told the participant “Attend to and remember ONLY AUDITORY STIMULI”, “Attend to and remember ONLY VISUAL STIMULI”, or “Attend to and remember BOTH VISUAL and AUDITORY STIMULI”, and “Press enter to continue.” When participants pressed the enter key they saw a fixation cross in the center of the screen for 1000 ms, followed by a 500 ms blank screen. Then the first study stimuli occurred. If it was an auditory-first trial, the screen remained blank while the participant heard 4 words with onset-to-onset times of 500 ms. Next, the visual array of 4 colored squares appeared 1500 ms after the onset of the last auditory stimulus. The visual array was displayed for 500 ms, followed by a 500 ms blank screen and then by the bimodal mask consisting of the word “Go” and multicolored squares in the same locations as the colored squares. The mask was displayed for 500 ms, followed by a 500 ms blank screen and then the probe stimuli. If it was a visual-first trial, the 500 ms blank screen was followed by a visual study display for 500 ms and, 1500 ms after the onset of the visual array, the participant heard 4 words with onset-to-onset times of 500 ms. The bimodal mask occurred

1000 ms after the onset of the last auditory stimulus. The 500 ms mask was followed by 500 ms blank screen and then the probe stimuli. If the probe modality was also the modality that had been presented first, then the onset of the probe occurred 1000 ms after the onset of the mask. If the probe modality was the modality that had been presented last, then there was an additional delay before the probe, depending on its modality, so that the retention delay always was a constant 5000 ms from the onset of the study stimuli to the onset of the probe.

After the mask and the additional delay for the 5000 ms retention interval, the probe occurred. The visual probe was a single square in the same location and the same color as a study square or a different color than any in the study array. The probe square remained on the screen for 2 s or until the participant responded. The auditory test consisted of sequence of four sounds with the same timing as the study list. The auditory probe was presented in a serial position that corresponded to the relative location of the probed study item. Other stimuli in the test list consisted of a brief buzz. Each buzz served as a placeholder for a non-probed word and indicated the serial position of the probed item (e.g., when probing the third item for the word *foot*, the test would proceed as *buzz – buzz – foot – buzz*). The probe stimulus was identical to the auditory stimulus presented in the same serial position of the study list or it was a word from the same category that was not in the test list. Immediately after the test sequence a “?” appeared in the center of the screen and remained on the screen for 2 s or until the participant responded. Immediately after each response, accuracy feedback (correct or incorrect) was displayed for 1 s.

The experimental block consisted of 128 trials with all types randomly intermixed; half were bimodal memory trials and half were unimodal memory trials. Of the 64 bimodal (remember-both) trials, 32 had auditory probes and 32 had visual probes. Of the 64 unimodal trials, 32 were remember-auditory trials and 32 were remember-visual trials. Each trial-type included 8 change and 8 no-change trials for each modality order, visual-first and auditory-first. An important characteristic of auditory lists was whether all words were from the same category (same-list) or each word was from a different category (mixed-list). There were equal numbers of same- and mixed-list trials for each memory condition, even when auditory stimuli were irrelevant. Also, the words in same-list trials were composed of each of the four categories equally often for each trial type. The serial position of the probed item was another important variable balanced across the different kinds of auditory probe trials. Of the 32 auditory probe trials for each memory load condition, each serial position was probed once for each combination of probe order (visual- or auditory-first), category (same- or mixed-list), and correct response (same or different).

**Results and Discussion**—Table 2 and Figures 4 and 5 show that mixed semantic information and changed method did not alter the results in a notable way from previous experiments; the central information included only 0.67 item on average, again with several items in peripheral storage in each modality. When a single category per list was used, the average central storage was increased to 1.15 items, but this difference did not reach significance by *t* test. Fundamentally, the results are unchanged. In *t* tests for each parameter, none showed a difference between single-category and mixed-category lists. All six parameter values were significantly above zero (Table 2). Thus, the results cannot be explained in terms of semantic grouping in uniform lists.

## Experiment 2b

In this experiment we attempted to give as much distinctiveness as possible to each acoustic item, by varying not only the semantic category within mixed lists, but also the voice and ear of presentation in those lists (Figure 3, fourth panel). Perhaps that would increase the central storage component and prevent any chunking that might occur. As mentioned, voice variability does make recall more difficult in many situations (Goldinger et al., 1991).

### Method

**Participants:** The participants were 24 college students (7 female) who received course credit for their participation.

**Apparatus and stimuli:** The equipment and setting were all same as the previous experiment. The only difference in the stimuli was that the spoken stimuli were synthesized in four different voices (whereas they were naturally spoken in previous experiments), so that each stimulus could be presented in a different voice in each mixed-category list. Speech synthesis was done using the built-in speech synthesis capabilities of the Macintosh operating systems 10.6.7 (Snow Leopard). Synthesized words were saved for playback as wave files with 16-bit resolution and 44.1 khz sample rate. To obtain four high-quality and distinctive voices, we selected two built-in voices (Alex and Victoria) and added two from a third-party source, the Infovox iVox voices Ryan and Heather from AssistiveWare (<http://www.assistiveware.com/>). The duration and intensity of stimuli, reproduced over the same headphones, were similar to the previous experiment. Each stimulus was presented with equal intensity to the two ears. For each trial presenting different-category auditory stimuli, each study stimulus in the four-word list was a word from a different category presented in a different voice, randomly selected from the four available voices. Also, the presentation side alternated between right and left ear for each successive study stimulus. The side of the first stimulus was randomly selected for each trial. Likewise, at the time of test, the “buzz” and probe stimuli alternated, side-to-side, in the same way as the study stimuli.

For each trial presenting same-semantic-category auditory stimuli, all speech stimuli were presented in the same voice, which was randomly selected from the four possible voices.

The auditory mask was the word “go”, spoken in the same voice(s) as the study stimuli. The auditory mask in trials with different-category stimuli consisted of the word “go” simultaneously presented to each ear, spoken by each of the two voices, superimposed, that had presented study stimuli to the same ear on that trial. The test stimulus was the same word or a different word than the study stimulus in the probed serial position, also spoken in the same category, voice, and location as the probed stimulus, in both conditions. Half the time the probe stimulus was the same word and half the time it was a different word from the same category as the target stimulus.

Counterbalancing and randomization of conditions, both within and between subjects, were the same as for the previous experiment. Likewise, the numbers of trials in each block and the assignments of trials to each condition within each block were the same, including assignments of same- and different-category conditions for the four auditory stimuli. The

only difference from Experiment 2a was the assignment in Experiment 2b of different voices and spatial channels to auditory stimuli in the mixed semantic category condition.

**Results and Discussion**—Table 2 and Figures 4 and 5 show that the results were similar to the previous experiment. The central component was quite small for both the uniform and mixed-category list conditions in this experiment, and the peripheral visual and auditory storage components again held several items each. Apparently, auditory chunking cannot explain why a larger central component is not found.

As in the previous experiment, *t* tests for each parameter revealed no significant effects of pure versus mixed lists. The  $P_{verb}$  and  $P_{obj}$  parameter values were significantly above zero for both kinds of list but, in this experiment, the *C* parameter was not significantly above zero for either kind of list (Table 2). The low *C* values suggest that for participants to use central storage to best advantage, there must be a certain amount of predictability about the nature of materials not only within a list, but also between lists within a session.

### Experiment 2c

To this point, the central storage component is quite small but it is not clear why. It could be that peripheral acoustic and visual stores can retain the information. Alternatively, the peripheral storage may reflect different verbal and nonverbal codes rather than sensory modalities. In this study, therefore, the acoustic digital lists were replaced with printed lists (Figure 3, fifth panel), in order to determine what would happen with two visual domains, verbal versus nonverbal.

#### Method

**Participants:** The participants were 24 college students (5 female) who received course credit. An additional male was omitted from the study for falling asleep.

**Apparatus and stimuli:** The experiment was an all-visual version of the previous two experiments. Virtually everything about this experiment was the same except that the words were all presented as text, on screen (instead of as spoken words over headphones), and some parameters and stimuli associated with visual presentation were necessarily different. The same words from four different categories were used as to-be-remembered stimuli. Study and probe words appeared in a white box in the center of the screen about 10 mm high by 24 mm wide (subtending a visual angle of about  $1.15 \times 2.75$  degrees at 50 cm). The text was displayed in a Courier New bold font. The text was all lower case, about half the height of the box, and vertically and horizontally centered. Each text item was presented for 250 ms, separated by a 250 ms blank screen. The array of colored squares was presented before or after the word list, as in the previous experiments with auditory lists. The same restrictions regarding the minimum separation of the squares from each other and the center of the screen prevented any of the squares in the array from overlapping the area when the text boxes appeared. The mask consisted of multicolored squares in the locations of the target array squares, as before, combined with a text box, identical to the target boxes, in the center of the screen filled with “XXXXX”, presented for 500 ms. The text probe was presented within a sequence, analogous to the way the auditory probe had been presented in

previous bimodal experiments. That is, the text probe consisted of the presentation of four white boxes, each displayed for 250 ms and separated by 250 ms, all empty except for the probe stimulus. The text probe stimulus was either the same as the word that had appeared in the study list at that serial position or else it was different from all words in the study list, but still a word from the same category as the target word. A “?” appeared in the center of the screen 500 ms after the last word and remained there for 2000 ms or until a response was made. The visual array was probed with a single colored square, in the same location as the target item, just as it was in the previous bimodal experiments.

**Results and Discussion**—Table 2 and Figures 4 and 5 show that the results were similar to before. The central store was here quite similar in magnitude to some of the previous experiments: about 1 item in central storage and about 2 items in each peripheral storage domain. It is clear that the peripheral storage in question here is a matter of coding domain (verbal versus nonverbal), not a matter of sensory modality per se.

In this experiment, there was no significant effect of category mixture for any parameter, and all six parameter values were significantly above zero (Table 2).

### EXPERIMENT SERIES 3: MEMORY FOR COLOR-SHAPE BINDING

The following two experiments return to the use of bimodal (auditory and visual) stimuli and ask whether a larger central component of storage might be found by requiring the maintenance of binding information, namely the association between the colors and shapes of objects, and between digits and the voices in which they are spoken. This possibility was suggested by Fournie and Marois (2011) who reported finding a larger central storage component when bound feature representations were required. However, their analysis was not oriented toward estimating the number of items in working memory with the intent of quantifying the contribution of central and peripheral components of performance as in the present work (Figure 1 and Equations 6-8). In our initial experiment on binding (3a), we omitted the mask (Figure 3, sixth panel) in order to examine binding in the simplest possible situation. The results included a central component similar to the previous studies, but the peripheral components were much smaller than before. In case this latter finding could have been caused by contributions from sensory memory, masks were added in Experiment 3b, but it produced similar results regardless of the type of mask, ruling out that possibility.

We changed several other aspects of Experiment 3a and 3b compared to Experiments 1-2, looking for a larger central component for binding information. First, we presented the visual array objects at fixed locations to eliminate location as a basis of reconstructing the binding between attributes at the time of test. Second, we presented the spoken digits from four loudspeakers so that they would be array-like, similar to Saults and Cowan (2007); but with the digits temporally overlapping rather than totally concurrently, so as to allow better acoustic encoding than Saults and Cowan obtained with concurrent stimuli. Experiment 4 was then run as a control experiment that included the exact same stimulus arrangement as Experiment 3b, but with probes that changed to a new color rather than to a new color-shape binding.

None of the following experiments include any cue as to which list item or array item is being probed; in both modalities, the probe must be compared to all items in that modality (see above or Appendix A for the formula for items in working memory). We used this method to test binding memory because it emphasizes the association between the two features intrinsic to the item (color-shape or word-voice) and removes location of the probe as an indirect cue to that binding.

### Experiment 3a

#### Method

**Participants:** The 22 participants (11 female) received course credit for participation.

**Apparatus and stimuli:** Each trial was preceded by a 2000-ms “Get Ready” signal, followed by a 2000-ms fixation cross and then the first target stimulus array or sequence. Visual stimuli consisted of an array of four colored shapes presented simultaneously in a diamond-shaped array. Each array included a square, a triangle, a circle, and a five-pointed star, arranged at each of four locations (Figure 3). The shapes were 12 mm × 12 mm at their widest and tallest points and were presented with the centers of the top and bottom shapes 33 mm apart, and with the centers of the left and right shapes 33 mm apart as well. The relative locations of the four shapes were randomly shuffled among these four locations across trials. Each shape could be one of seven colors: *white, red, blue, green, yellow, violet, or black*. Each shape in an array was a different color. The target array was presented for 500 ms. Memory for the visual array was tested by a single probe shape, presented for 1000 ms at its center, with an onset 3000 ms after the onset of the target array. On no-change trials, the probe was the same color as that shape was in the target array. On change trials, the probe shape was the same color as one of the other shapes in the target array. Auditory stimuli consisted of a list of four digits randomly selected without replacement from the set 1-9, spoken in 4 different voices, a male or female adult or a male or female child voice. On each trial, each voice spoke a different digit from a different loudspeaker location, with two lateral and two central loudspeaker locations as described by Saults and Cowan (2007). Spoken digit onsets were separated by 250 ms and the actual durations of digits were 300 to 600 ms, so that they overlapped in time. The location and serial order of voices were randomly determined for each trial. The auditory memory probe consisted of one voice saying a digit, presented on the two front speakers so it seemed to originate in the center. The digit was always one of the digits in the target set. On no-change trials, the digit was spoken by the same voice that it had been in the target list. On change trials, the digit was spoken by a different voice that it had been in the target list. The onset of the probe digit occurred 3000 ms after the onset of the first auditory target stimulus.

The auditory list and visual array were offset so they did not overlap. When the visual stimuli occurred first, the onset of the visual target array occurred 750 ms before the onset of the first auditory stimulus. When the auditory stimuli occurred first, the onset of the visual target array occurred 750 ms after the onset of the last auditory target stimulus.

**Procedure:** There were three blocks of trials with different instruction conditions: attend to the visual items, attend to the sounds, or attend to both. These three trial blocks occurred in

all possible orders, counterbalanced across participants. The attend-visual and attend-auditory trial blocks began with 4 practice trials followed by 32 test trials, whereas the attend-both trial blocks began with 4 practice trials followed by 64 test trials. The attend-both trials were evenly divided among visual and auditory probes, and all trial blocks were evenly divided among change and no-change probes. The trial types were randomized with the restriction that there were no more than 4 consecutive trials with the same answer and, in the attend-both trial blocks, no more than 4 consecutive trials with the same probe modality. The target in each location was used as a probe equally often in each modality, condition, and presentation order.

**Results and Discussion**—The results and parameter values can be seen in Table 3 and in Figures 4 and 5. Based on the proportion correct in Figure 4, clearly there was a pronounced effect of sharing attention between modalities, especially in the case of objects. Similar to the previous experiments, the mean central parameter value was less than 1 item. Unlike the previous experiments, however, the peripheral parameters were much smaller than before, both less than a single item as opposed to several items (as illustrated in the bottom two panels of Figure 5). Nevertheless, all three parameters were significantly above zero (Table 3).

Whereas participants seem to store considerable stimulus feature information in peripheral types of memory, this peripheral storage apparently cannot store much information about the binding between features. This result is consistent with the view that attention must be used to maintain the binding between features in working memory (Wheeler & Treisman, 2002).

### Experiment 3b

In this experiment, we replicated the previous experiment, sometimes with a mask in the same modality as the following probe and sometimes with a mask in the opposite modality (Figure 3, seventh panel). A mask in the same modality could either degrade peripheral information that will be needed, or it could distract attention from the memoranda (Hawkins & Presson, 1986), whereas a mask in the opposite modality can only serve as a distraction.

Theoretically, degradation of peripheral information could result in more reliance on central memory, boosting its parameter value. This could happen if the total amount of the central resource is not fixed, but increases as participants exert more effort (Kahneman, 1973) when there is less available peripheral information. The total number of participants was increased from previous experiments because the number of potential comparisons between conditions also was increased.

### Method

**Participants:** There were 46 participants (28 female), who received course credit for their participation.

**Apparatus and stimuli:** These were identical to Experiment 3a except for the inclusion of a mask on some trials. The visual mask consisted of four multicolored shapes presented for 200 ms in the same location as the target stimuli, and the auditory mask was a random



combination of the digits spoken by all four voices, simultaneous over the four speakers and compressed to a duration of 220 ms. The duration from the onset of the target stimuli in the same modality as the probe to the onset of the probe itself was always 3250 ms, and the mask was presented before the probe such that there was always a 750-ms period between mask onset and probe onset.

As in Experiment 3a, there were three trial blocks presented in counterbalanced order: attend-visual, attend-auditory, and attend-both trial blocks. At the beginning of each attend-visual or attend-auditory trial block there were 6 practice trials, which included one change trial and one no-change trial for each masking condition: no mask, the mask in the same modality as the test probe, and the mask in the opposite modality. These were followed by 48 trials. In the attend-both trial blocks there were 12 practice trials followed by 96 test trials. The trials were evenly divided between the three masking conditions and between change and no-change trials, and the restrictions on trial randomization described in the previous experiment applied here as well.

**Results and Discussion**—As shown in Table 3 and in Figures 4 and 5, the results were comparable to the previous experiment. (In Figure 4, the two binding experiments are represented with dashed lines.) There was very little effect of the mask. Once more the central parameter was less than 1, similar to the other experiments, and once more (as in Experiment 3a) the peripheral parameters were also less than 1, in stark contrast to the experiments in which the task did not require maintenance of feature bindings. Masking did not significantly change any parameter, and all parameter values were significantly above zero (Table 3).

## EXPERIMENT 4: FEATURE-MEMORY CONTROL EXPERIMENT

The experiments requiring feature binding (Experiments 3a and 3b) produced much lower peripheral storage parameters than all of the previous experiments, which required only feature information. Those binding experiments, however, differed in several methodological details from Experiments 1-2. This next experiment examined memory for features instead of bindings but otherwise used the same method as Experiment 3b; the design of this experiment was identical to Experiment 3b, except that the change trials all involved changes to a new feature value that was not in the studied list or array (Figure 3, last panel). If the results of this experiment are like Experiments 1-2 and not like Experiments 3a-3b, that will confirm that it was the need for binding information in Experiments 3a-3b that was the critical difference between them and Experiments 1-2, and not some other methodological detail. Such a result would confirm the much smaller estimates of peripheral storage for binding memory than for feature memory.

### Method

**Participants**—The 30 participants (15 female) received course credit for participation.

**Apparatus, stimuli, and procedure**—There were 12 practice trials at the beginning of every trial block. All else was identical to Experiment 3b, with the key exception that the probe on digit-change trials was a digit that had not been included in the studied list (in a

voice that had been presented), and the probe on object-change trials had a color that had not been included in the studied array (but in a shape that had been presented). On no-change trials, the probe was an item from the spoken list or from the visual array.

**Results and Discussion:** Table 3 and Figures 4 and 5 show the results, which were similar to the other, previous experiments in which only feature maintenance was required, not the binding between features of an object. Specifically, the central component of maintenance is about 1 item, whereas the peripheral components are again several items, in stark contrast to the much smaller estimates of peripheral components in Experiments 3a and 3b. Once more, masking did not change any parameter and the parameter values were all significantly above zero (Table 3).

**Cross-experiment comparisons:** In order to test the effect of the type of probe in comparable experiments, we combined the data from Experiments 3b and 4 in a separate ANOVA for each parameter. Even with the larger total sample size in this combined analysis, there was no overall effect of the masking condition on parameter values and, for the central parameter, there was no overall effect of the type of probe. Concerning effects of experiment, the test for feature-binding changes in Experiment 3b produced a much smaller peripheral verbal parameter  $P_{verb}$  than did the test for new-feature changes in Experiment 4 (Mean=0.69, SEM=0.13 vs. Mean=2.96, SEM=0.17 for these experiments, respectively;  $F(1,74)=112.32$ ,  $\eta_p^2=.60$ ,  $p=.000$ ). The same was true for the peripheral parameter for visual objects,  $P_{obj}$  (Mean=0.88, SEM=0.15 vs. Mean=2.60, SEM=0.18 for Experiments 3a vs. 4, respectively;  $F(1,74)=55.29$ ,  $\eta_p^2=.43$ ,  $p=.000$ ). This pattern of results, with reduced peripheral storage of feature bindings compared to features themselves, but no such reduction in central storage, was supported in a general ANOVA including all data, which yielded an interaction of the parameter with the experiment,  $F(2,146)=17.31$ ,  $\eta_p^2=.19$ ,  $p=.000$ . Newman-Keuls tests showed a difference between item and binding experiments for  $P_{verb}$  and  $P_{obj}$ ,  $p=.000$  in both cases, but not for  $C$ ,  $p=.83$ . This pattern seems consistent with the proposal that there is not much retention of attribute binding outside of attention (e.g., Wheeler & Treisman, 2002).

This difference between experiments cannot be accounted for by the different formulas that were applied in order to estimate items in working memory for item-change versus binding-change experiments (Appendix A). To make sure, the results of Experiment 3b were re-analyzed (incorrectly, in our view) using the same formula that was applied to Experiment 4 as explained in Appendix A, a modification of Pashler's (1988) formula for a centrally-presented, single-item probe. This application resulted in central parameters for the same-modality mask, no-mask, and different-modality mask situations of 1.98, 2.04, and 2.01 units, respectively, somewhat larger than before. However, the peripheral auditory scores were still low (0.44, 1.00, and 1.01 units in the three masking situations, respectively), as were the peripheral visual scores (0.66, 0.70, and 0.46 units, respectively). Clearly, the use of a different formula cannot explain the difference between experiments.

The present finding reveals proportionally greater dependence on a central component of working memory for the retention of binding information than for the retention of feature information (Figure 5). This aspect of the results may appear, at first glance, to contradict

previous studies showing that the effect of a distracting secondary task on binding information is no greater than its effect on feature information (Allen et al., 2006; Allen, Hitch, Mate, & Baddeley, 2012; Cowan, Blume, & Saults, 2013; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006). The contradiction is, however, more apparent than real. The effect of distraction may be interference with central but not peripheral information. If so, the detrimental effect on performance should be comparable for features and bindings because central information is comparable in magnitude for these types of information. Even though the central, attention-dependent component of working memory is a larger proportion of memory for binding information than for attribute information, the central component includes about the same number of items no matter whether it is attribute or binding information that is being considered (see Figure 5). The latter property could make the effects of distraction similar in magnitude for feature and binding information. (It also must be considered, though, that in the present study we presented visual and acoustic stimuli successively in either order rather than concurrently, eliminating any division of attention between modalities during encoding, unlike many of the studies with distracting tasks.)

Last, it is also worth mentioning that the results shown in Figures 4 and 5 are consistent with an asymmetry between the visual and verbal modalities. Figure 4 plots the results across all conditions and shows that for both attributes and binding, there is an asymmetry. Specifically, it appears that dividing attention between verbal and visual objects is more injurious to visual information than it is to verbal information. The attention effect for the means across experiments (solid points in Fig. 4) was, for visual attribute and binding conditions, 0.55 and 0.46 items; whereas it was smaller for verbal attribute and binding conditions, 0.20 and 0.15 items. Other evidence of this asymmetry has been reported previously (Morey & Mall, 2012; Vergauwe et al., 2010). Verbal rehearsal requires relatively little attention (Guttentag, 1984; Naveh-Benjamin & Jonides, 1984), but here the asymmetry is seen to some extent even in the presence of articulatory suppression. Verbal encoding may result in a form of memory that is resilient without the devotion of attention for maintenance, with rich support from lexical long-term memory (Cowan et al., 2012); whereas the maintenance of nonverbal visual representations in working memory could require more attention-based processing, such as refreshing (Camos, Mora, & Oberauer, 2011; Raye, Johnson, Mitchell, Greene, & Johnson, 2007) or maintenance in the focus of attention (Oberauer, 2002; Saults & Cowan, 2007).

## **EXPERIMENT 5: FURTHER ASSESSMENT OF FEATURE AND BINDING CAPACITIES**

One limitation of the previous experiments is that new-feature and feature-binding changes occurred in different experiments. Another limitation is that not all features were tested. In Experiments 3b and 4 taken together, we tested memory for spoken items and their binding with voices, and we tested memory for colors and their binding with shapes. We did not, however, test memory for voices or shapes per se in any of the experiments. Instead, there were four voices and four shapes, all of which appeared within the stimulus sets on every trial. It is possible that the poor peripheral storage for binding within a modality, compared to feature storage, could have occurred because we omitted testing of the more difficult

features contributing to the binding. Perhaps binding performance is no worse than performance on the worst feature.

In the present experiment, we addressed both of these limitations. We included memory for colors, shapes, and their binding, and also for spoken letters, the voices in which they were spoken, and their binding. The use of letters rather than digits allowed us to construct a situation in which it was natural to have seven possible values of each feature from which to draw on each trial. Seven of each feature type was adequate, and yet a small enough number that we could find seven distinct, not-easily-confused values of each (i.e., seven colors, shapes, voices, and letters). The sample size was comparable to Experiment 3b given the large number of potential comparisons between conditions.

## Method

**Participants**—The 44 participants (34 female) received course credit for their participation.

**Apparatus, stimuli, and procedure**—Articulatory suppression was used on each trial. The arrangement of stimuli on a trial included arrays of 4 colored shapes and lists of 4 spoken letters, with timing as in Experiments 3b and 4. The mask was always bimodal; specifically, the pattern of multicolored squares was accompanied by a garbled voice composed of digits from all voices on every trial. When the probe was visual, it was presented at the center of the screen, and the timing of the visual or acoustic probe was as in Experiments 3b and 4.

In order to test all features, we generated seven values of each feature from which to draw the four visual and four acoustic stimuli. There were seven possible colors as in Experiments 3 and 4, but also seven possible shapes (cross, triangle, square, circle, star, convex lens, concave lens). The spoken stimuli were drawn from the set of letters *BFJLQRY*, and they were presented in seven possible voices (computer-generated speech voices, chosen to be as clear and distinct from one another as possible, were Macintosh OSX 10.6.8 text-to-speech voices Victoria, Princess, Agnes, Alex, Junior, Bruce, and Ralph). No color, shape, letter, or voice was repeated more than once within the stimuli on a trial, except in the case of the probe insofar as it matched a prior stimulus on the trial.

Trials were divided into nine trial blocks of the following types: (1) attend visual, with new-color changes on some trials; (2) attend visual, with new-shape changes on some trials; (3) attend auditory, with new-letter changes on some trials; (4) attend auditory, with new-voice changes on some trials; (5) attend both modalities, with new-color changes on some trials and new-digit changes on other trials; (6) attend both modalities, with new-voice changes on some trials and new-shape changes on other trials; (7) attend visual, with color-shape binding changes on some trials; (8) attend auditory, with digit-voice binding changes on some trials; and (9) attend both modalities, with color-shape binding changes on some trials and digit-voice binding changes on other trials. Each participant received either all six of the feature-change trial blocks first, or all three of the binding-change trial blocks first, with the order of trial blocks random within the feature-change and within the binding-change sets of

trial blocks. This trial ordering scheme was adopted so that the instructions would have to change from new features to feature binding or vice versa only once per participant.

Each trial block began with 4 practice trials illustrating the possible trial types. This was followed by 24 test trials, for blocks in which only one modality was to be attended, or 48 trials, in blocks in which both modalities were to be attended. In the latter case, half of the trials were visual-probe trials and half were auditory-probe trials. In each trial block, half of the trials of each type were change trials and half were no-change trials in which the probe was identical to one item in the set to be remembered. In all, there were thus 36 practice trials and 288 test trials.

## Results and Discussion

The proportions correct are summarized in Figure 6 and the parameters are shown in Figure 7. Overall, the pattern was consistent with what was found in the previous experiments. Performance on voices and shapes, not tested previously, was somewhat poorer than on the other features (especially shapes). Nevertheless, it seems clear from the data that the sole limit to binding detection could not be item detection. In particular, peripheral storage levels for letter-voice binding and color-shape binding were almost identical to one another, even though there was much more peripheral information for features in the auditory case.

An ANOVA on the  $C$  parameter showed no difference between trial blocks in which the test objects were letter and color, voice and shape, or within-modality binding (letter-voice and color-shape),  $F(2,88) < 1$ ,  $\eta_p^2 = .02$ . We believe, however, that this null effect should not be over-interpreted given that a trend appears in which  $C$  is at a lower level. ANOVA procedures assess algebraic differences between conditions, not performance ratios and, given the overall low level of  $C$ , the trend could be real without reaching a level sufficient to produce a significant effect. Note, though, that inasmuch as the  $C$  contribution here is numerically small in all conditions, the possible further reduction of  $C$  in the binding condition makes relatively little difference for the overall capacity.

For  $P_{verb}$ , there was a significant effect,  $F(2,88) = 21.44$ ,  $\eta_p^2 = .33$ ,  $p < .001$ , and Newman-Keuls tests showed that it occurred because storage of letter and voice were both above the storage of the binding between them,  $p < .001$  in both cases. For  $P_{obj}$ , there was a significant effect,  $F(2,88) = 10.58$ ,  $\eta_p^2 = .19$ , which, according to Newman-Keuls tests, occurred because color storage was higher than either shape or color-shape binding storage,  $p < .001$  in both cases.

An exception to the superiority of feature information compared to binding information throughout the study was memory for the shape feature (Figures 6 and 7), which did produce low  $P_{obj}$  values, more similar to feature binding than to the other features. We can speculate on the basis of this inferiority of shape information, as follows. Acoustic features contributing to both voice and word identification might be retained as an integrated sequence or pattern across time. Similarly, the color feature might be stored as a configuration of colors across space. These can be viewed as a type of chunking (Miller, 1956) or combination of item information to form a multi-item unit. In contrast, it might be difficult to remember a multi-item configuration of shapes across space. The reason would

be that shapes are themselves defined on the basis of spatial features: the cross is spatially extensive at the vertical and horizontal axes, the circle is equally extensive in all directions, and so on. So the shape feature, and the binding between features, would have to be retained on an individual item basis, with no chance of chunking unlike the color, voice, and spoken word features.

In sum, the levels of central storage were consistent with the values obtained in the previous experiments; they were in the range of a single item only, much smaller than Saults and Cowan (2007) surmised. Further, on the basis of the auditory results, it is clear that the limiting factor for peripheral information about binding is not always the quality of feature information in memory; both letters and voices were recalled relatively well in terms of peripheral information reflected in the  $P_{verb}$  parameter, yet the binding between them was much poorer in that regard. We cannot say with certainty that peripheral information for binding storage is always below the value for feature storage, given the relatively poor memory for shapes, which was almost as low as color-shape binding. However, it is noteworthy that the level of binding information in the peripheral parameter was nevertheless comparable for verbal (auditory) and object (visual) information, suggesting that binding information may have its own small limit.

## GENERAL DISCUSSION

In dual-task methodology, it has long been realized that some aspects of performance are susceptible to tradeoffs between tasks, whereas others are not so susceptible (e.g., Norman & Bobrow, 1975). Yet, until now, to our knowledge, no study has converted this wisdom to an analysis of working memory tradeoffs separating domain-specific resources from cross-domain resources. In doing so, we have changed our view from that espoused by Saults and Cowan (2007), who proposed that under confined circumstances (when rehearsal was prevented by presenting acoustic items concurrently in arrays rather than sequentially in lists, and sensory memory was eliminated with a same-modality mask), acoustic verbal and visual objects were remembered using only a common resource. That view would be less successful in interpreting evidence from procedures in which verbal lists were used (e.g., Cocchini et al., 2002; Fournie and Marois, 2011; the present experiments).

Our new analysis technique (illustrated in Figure 1) provides a more principled approach. We consistently find that the central component of working memory storage that is shifted from one task to another depending on task difficulty is about 1 item, or sometimes less. In contrast, the peripheral components are often several items when stimulus features can be used to carry out the task. When feature binding is needed, the peripheral component we observed was consistently 1 item or less (see Figures 4, 5, 6, and 7).

There is a key conundrum to be pondered in trying to understand these results. How is it possible for participants to encode items from a list or array into working memory only up to a limit of typically 3-4 items, and yet be capable of holding most of these items in working memory even while encoding and maintaining a comparable number of items from a second list or array, with only about a 1-item conflict between the sets?

In the remainder of our discussion we first demonstrate the consistency of our results with past results. Next, we examine potential inconsistencies of our results with past results, and we suggest resolutions for all such apparent inconsistencies. Last, we consider alternative theoretical accounts of our results, favoring a two-process account that includes interference and attention components. As we indicate, the present data do not allow us to determine definitively the way in which attention processes result in dual-task tradeoffs, but we present several alternatives and demarcate a path for future research to resolve the issue.

### Consistency with Past Results

**Feature information**—The feature-memory results are fully consistent with past estimates of capacity in working-memory recognition tasks, which have focused on features rather than binding and have placed capacity at 3-5 items (Cowan, 2001; Cowan et al., 2012) and toward the lower end of that range for arrays of simple visual items (Luck & Vogel, 1997; Rouder et al., 2008). As discussed in Experiment 1a, the overall capacity estimate is obtained from the present data by summing the peripheral and central components, which yields estimates mostly in the 3-5-item range. The small central component is consistent with past results combining an acoustic list with a visual array (Cocchini et al., 2002; Fougne & Marois, 2011) and similar results were obtained even when we substituted a printed verbal list for the acoustic list, but we have explored more factors and are the first to use an item-capacity-based metric to assess such dual-task results.

**Binding information**—The color-shape binding results are comparable to past studies. One study (Allen et al., 2006, Experiment 5) used a procedure very similar to the visual portion of our Experiment 5. Both experiments used sets of four colored shapes per trial, with articulatory suppression, followed by a probe in the central location. Allen et al. found proportions correct (based on their Table 5, averaged across hits and correct rejections) for color memory of .82; for shape memory, .75; and for color-shape combinations, .69. Although these proportions are somewhat higher than we found, the amount of separation between conditions in proportion correct is quite similar to what we found (Figure 6, right-hand panel, data for one stimulus set: for color, 0.79, for shape, .67, and for their binding, .62).

Alvarez and Thompson (2009) used a different kind of task with a complex moving display, but examined capacity for color-location binding. They presented two arrays and used estimates of the number of bindings in working memory with formulas constructed for a 4-alternative forced choice (about which color in the more recent array was in a probed location or which location had a probed color) and for a 2-alternative forced choice (about whether there was a switch in the color-location binding between the earlier array and the more recent array). Formulas were presented to estimate the number of items in working memory in these situations. The two-alternative forced choice procedure is more comparable to ours and yielded a comparable result, with slightly more than one item in working memory. That is nearly the same as one obtains in the binding condition of any of our experiments, by adding the central and peripheral values for object binding to obtain an estimate of items in working memory.

In Alvarez and Thompson (2009), also, much higher performance levels were found with the 4-alternative procedures, and these higher levels were attributed to the requirement that participants remember the second array but not the first. It was suggested that in the same-different task, the second, probe array contained new bindings that could overwrite the ones in the first array. This finding raises the interesting possibility that in our study, as well, the relatively poor peripheral memory for binding occurs because the probe interferes with memory by including a new binding that is confused with the ones in the set to be remembered.

**Dual-task tradeoffs for features versus binding**—Fougnie and Marois (2011) found that the tradeoff between modalities was much more severe for binding than for features. Their measure was designed to examine the proportion of capacity that is lost in a dual-task situation. Given that we obtain smaller peripheral components with binding than with features, but comparable central information, our data are consistent with this kind of analysis. They, by the way, had only one experiment in which articulatory suppression was used and, in that experiment, item feature capacity was examined but there was nevertheless no dual-task cost. In that experiment, though, the acoustic stimuli were tones, and single-task capacity was only at about one item, so one could argue that there was not much room for a tradeoff. Morey and Bieler (2013) used stimuli similar to ours (regular shapes presented in categorical colors) but in some conditions with a tone categorization task rather than our verbal list memory task. Their results were fairly similar to ours, with higher performance levels for features than for feature bindings but with effects of divided attention for both.

### Potential Discrepancies with Other Evidence

Given the complexity of the working memory literature, it should come as no surprise that there are findings apparently in conflict with our own findings. Below, we suggest ways to reconcile various results in the literature with our own results, with attention to several theoretical issues.

**Inferiority of Binding Memory**—A few studies report results that seem to conflict with our finding of a general superiority of feature memory over binding memory, but methodological differences can explain the discrepancy. Delvenne and Bruyer (2004, Experiment 2A) examined memory for arrays of 2 or 4 textured shapes using a subsequent central probe, similar to the visual stimuli in our Experiments 3a, 3b, and 4. They found that the proportion correct for the binding between texture and shape was comparable to the memory for the textures and shapes. In contrast, we found that color, voice, and word information were consistently above binding information. There are two points to be made here. First, instead of colors and regular shapes, they used irregular shapes and subtle textures that might be more poorly represented as features. The second point is that the guessing basis is different for item and binding information, a point that we take into account with our formulas. As explained above and in Appendix A, a change in binding in this procedure logically can be detected on the basis of knowledge of either of the original two objects contributing features to the recombined probe object, whereas knowledge of only one object indicates whether the color was in the array. Consequently, for example, the



performance level of about 0.65 shown for 4-item trials (Delvenne & Bruyer, Figure 7) corresponds to about 0.8 objects in the binding condition, versus about 1.85 objects in the feature condition.

Cowan et al. (2006) used a procedure in which arrays of colored squares were presented, such that each color could occur more than once in an array (as in Luck & Vogel, 1997). The probe was a repetition of the entire array, but with a single encircled probe item that may have changed from the studied array. In Experiment 1, a change in the probe could consist of either a new color or a color that was in the array, but not at the location shown. Half of the trials instead were no-change trials. In that experiment, performance on color-location binding trials was considerably below performance on new-color trials, as in our findings. In Experiment 2 of Cowan et al., the new-color and color-location binding trials occurred in separate trial blocks. In that situation, performance on new-color and color-location binding were comparable. Moreover, the difference in formulas does not apply because multiples of the same color could occur in the display in this experiment, so that in both item and binding trial blocks, one had to know information about the tested location in order to be sure to answer correctly. There were, however, important limitations of this method, as follows.

In this second experiment of Cowan et al. (2006), special precautions had to be taken to ensure that the probe display did not give away the results. Unfortunately, these precautions could have altered the relative performance in color-change versus color-location binding change blocks. In color-change blocks, changes could not occur for some of the items in the array. The authors wanted to make sure that the correct result could not be obtained on the basis of the probe display alone. Consequently, for no-change trials in the color-change blocks, the probed item had a unique color not found elsewhere in the array, inasmuch as that was necessarily always the case in change trials. Similarly, for no-change trials in the binding-change blocks, the color of the probed location was always a duplicate of another location, inasmuch as that was necessarily the case in a change trial. Unfortunately, these constraints meant that, in order to detect the changes, participants only had to note which colors were unique in color-change blocks and which colors were non-unique in binding-change blocks. For example, in an item-change block, if the array contained a red, a green, and two blue squares, one would only have to note that red and green were the unique colors. If a blue square turned orange, then one would simply note that orange was not one of the unique colors in the array; one would not have to memorize the color of the two blue squares to know this. Similarly, if the green square turned orange, one would only have to note that the green color was now missing. Perhaps participants learned this statistical information about the stimuli while carrying out the experiment. This may be problematic for comparing item and binding performance because no metric is available and, in hindsight, it underscores the superiority of single-item probe methods used in the present study.

**Verbal-Visual Asymmetry?**—There also are studies that seem at odds with our finding of approximately equal peripheral information for verbal codes and one visual code (color), as shown in Figures 5 and 7. Some studies have shown that in dual-task situations, there is less interference with verbal memory from the visual task than there is with visual memory

from the verbal task (Morey & Mall, 2012; Morey et al., 2013; Vergauwe et al., 2010). That asymmetry in interference would be realized here as a larger amount of peripheral information in the verbal task than in the visual task. There are a number of reasons why an asymmetry could occur. First, some comparisons involve verbal information that is intrinsically more meaningful than the visual information, which could lead to the need to consider long-term storage in working memory tasks (cf. Cowan et al., 2012; Unsworth & Engle, 2007). However, Morey and Mall reproduced the asymmetry in a task of order memory for visual items and pronounceable nonwords, so this cannot be the full explanation. Second, it is well known that verbal rehearsal is a general technique that can be of assistance for visual as well as verbal materials. Nevertheless, prevention of rehearsal does not appear to be very important in the retention of arrays of visual objects; prevention of rehearsal has had little effect on this kind of visual memory (e.g., Morey & Cowan, 2004). There has been an asymmetry between modalities in studies in which rehearsal was not prevented, such as the one by Vergauwe et al., but preventing rehearsal has not eliminated the asymmetry (Morey et al., 2013). Third, there is an effect of sensory memory that is often greater in the acoustic modality. Morey and Mall found that their asymmetry disappeared in the presence of a mask to eliminate sensory memory. This does not necessarily imply that sensory memory lasts longer in the auditory modality; it may be that there is a type of sensory memory that lasts several seconds in all modalities but that auditory sensory memory is richer for temporal features, whereas visual sensory memory is richer for spatial features (Cowan, 1988; Massaro, 1975; Penney, 1989). Morey et al. (2013) found an asymmetry that persisted despite masking.

Recall that we used the same small set of items over and over, required articulatory suppression, and presented a post-perceptual mask. An asymmetry is nevertheless visible in terms of proportion correct in our studies, in that the decline from single tasks to dual tasks is somewhat smaller for verbal items than for nonverbal objects (Figures 4 and 6). In our study, the asymmetry could come from the support of verbal long-term memory, with not as much long-term memory support for the visual objects. Nevertheless, neither the literature nor our experiments have yet determined just what factors can contribute to this asymmetry.

**Severely Limited Central Storage?**—Central storage may not always be limited to about a single item. Morey et al. (2011) carried out a study in which an array of colored squares was to be remembered at the same time as a series of tones. Instead of manipulating task difficulty, the payoffs for remembering visual versus acoustic information were varied. This study was evaluated with a Bayesian measure of items in working memory and revealed a tradeoff between modalities. Applying Equations 6-8 to their data (based on the figures) yields estimates of  $C=2.7$ ,  $P_{tones}=0.6$ , and  $P_{colors}=2.8$ . This larger estimate of the central component is compatible with the notion that about three items can be held at once in the focus of attention. Although participants may most comfortably limit the central storage to about one item, central storage may be capable of taking on a larger role when the peripheral, automatic storage of one modality is particularly poor, as in the study by Morey et al. Central storage could do this by more effort being expended (Kahneman, 1973) and/or the focus of attention zooming out or expanding to apprehend more items (Cowan et al., 2005; Oberauer, 2013; Verhaeghen et al., 2004). In this zooming-out process, just as a

spotlight on a stage might be broadened, some precision of any one representation in working memory might be sacrificed in order to accommodate more items total, up to a limit of about 3 items (Zhang & Luck, 2008).

Contrary to this finding but more similar to our own finding, Fougne and Marois (2011, Experiment 3A) also used colors and tones as the memoranda. Based on the means reported for that experiment,  $C=0.93$ ,  $P_{tones}=2.60$ , and  $P_{colors}=2.67$ , similar to what we have observed for colors and verbal stimuli. In that experiment, though, tones and colors were presented in concurrent sequences. Similar to how perceptual organization affects unimodal working memory (Woodman, Vecera, & Luck, 2003), it is theoretically possible that multimodal objects can be perceived, in this case concurrent tone-color pairings (i.e., grouping across modalities, much as a chirping bird might be perceived as a multimodal object). Clearly, there is room for additional follow-up research to ascertain the boundary conditions for finding larger versus smaller values of  $C$ .

Our findings of a limited central component must be reconciled with unimodal studies in which individuals are shown to retain about 3-4 items in visual working memory (Anderson et al., 2011; Cowan, 2001; Luck & Vogel, 1997) and in verbal working memory (Chen & Cowan, 2009a; Cowan et al., 2012; Jarrold, Tam, Baddeley, & Harvey, 2010). In a general review, Cowan (2001) favored a view that the central component based on the focus of attention included 3-5 items, a view that Saults and Cowan (2007) thought they had reinforced but that seems inconsistent with the present data and analysis. Given our present findings, this view is inadequate and we now believe that ordinary capacity must include the peripheral components. Next, we consider alternative theoretical accounts of our findings.

### Alternative Theoretical Accounts of the Present Findings

We can find several basic hypotheses to explain the present evidence. The main distinction we wish to draw is between two different simple accounts and a compound account. The simple accounts are based on storage of information in the focus of attention and interference between items in memory, respectively. We believe both of these accounts to be inadequate and present a compound account in which, in addition to interference, there is also a role of attention in working-memory maintenance. We then consider several variants of this type of account but leave them to be distinguished in future research.

**Storage in the Focus of Attention?**—The theory of Saults and Cowan (2007) was a one-process account in which items were held in the focus of attention. Note that the one-process account was only meant to apply in a situation in which both sensory memory and covert articulation were eliminated. In this situation, the theory was that items are held in the focus of attention and that the limit in performance occurs because there are too many items in the two sets so that they cannot all fit within the focus of attention. Attention to both sets results in fewer items recalled from either set than when a single set is to be recalled.

The present data conclusively rule out this hypothesis, provided that one accepts that articulatory suppression eliminates covert verbal rehearsal for lists of verbal items. The number of items recalled in each modality in a bimodal memory situation was not cut in half relative to a unimodal situation; far from it, the conflict amounted to only about one item,

the *C* parameter we have estimated. A peripheral form of storage seems to be used for each stimulus modality or type of code (in this study, verbal versus nonverbal storage).

**Interference During Memory Maintenance?**—The present results might alternatively be explained by interference alone, during the retention interval following the presentation of both stimulus sets, in a way that has nothing to do with limits of attention-as-storage. Consider for example the function of interference within the theory of Oberauer et al. (2012), ignoring for the time being their one-item focus of attention. In their theory, working memory capacity limits arise from interference between similar items held in a limited-access region of memory. The limited-access region (Oberauer, 2002) contains new bindings between items and context. Alternatively, Cowan (1999, 2001, 2005) considered these new bindings to be included in the activated portion of long-term memory, and according to that view the similarity-based interference could occur in this part of memory. In either version of this account, the verbal and nonverbal visual items could be considered dissimilar enough that they do not interfere with one another very much, whereas within-set items interfere with each other enough to produce a within-modality capacity limit of 3-5 items. The roughly 1-item central component could be explained by residual similarity between items in different domains (cf. Marois, 2013), though the nature of that similarity is unknown and we tried to minimize it.

The interference account of bimodal capacity, described above, cannot easily explain results of Cowan and Morey (2007). It shows evidence for a role of attention not only in the process of encoding information into working memory, but also in the process of maintaining information in working memory after it is encoded. They found that after two sets of items were encoded into working memory on a trial, a cue to continue to retain one or both of these sets, during the subsequent retention interval, affected performance. (For retro-cueing procedures see also Griffin & Nobre, 2003; Makovski & Jiang, 2007.) The need to continue to retain both sets impaired performance on the tested set compared to the need to continue to retain only that one set. These sets were both auditory-verbal, both visual-nonverbal, or one of each. Although there was more interference between sets when they were two of a kind (both verbal or both nonverbal), the similarity effect occurred during encoding into working memory and did not influence the magnitude of the effect of a cue (i.e., the detrimental effect of having to retain both sets as opposed to just one). This suggests that there is a maintenance process that does not depend on a similarity-based interference process, but does depend on attention.

To account for Cowan and Morey (2007) it would still be possible to have an interference-based account in which the interference only applies to items that have been “moved in” to working memory using attention, and not to items that have been “moved out” using attention. However, this idea is not typical of the literature on interference, in which items, once encoded, are subject to mutual interference if they are close to one another in their time of encoding and are otherwise similar (e.g., Brown, Neath, & Chater, 2007). This is also the case for information maintained in Cowan's (1988) activated portion of long-term memory; one presumably can use that information for recall and cannot use attention to remove information from it.

**Attention-based Mnemonic Processing Plus Interference?**—Having pointed to the shortcomings of two simple accounts, we advocate an alternative account in which both interference and selective attention play a role in determining the limits of working memory storage and the tradeoff between tasks.

**Feature information and interference:** Following Oberauer (2013), interference between items with similar features would provide the main explanation of why there is a capacity limit of 3-5 items within a modality, but why there is only a little interference between modalities. On another level, interference could explain why the capacity for bindings between features is typically lower than the capacity for the features themselves. The capacity limit could arise because there can be only a few items associated unambiguously with a present-trial cue (e.g., the previous trial's stimuli included red, green, and blue; the present trial's stimuli included red, orange, and brown). Thus, proactive interference effects between trials have been noted in the visual array procedure (Shipstead & Engle, 2013).

**Feature-binding information and interference:** The more severe limit for binding information could result from a further assignment of binding. To get a binding question right, one must retain two levels of binding, the trial contextual level and the within-trial binding level (e.g., the cross was green and the square was red on this trial; the cross was red and the square was blue on a previous trial). That added complexity could consume more space in working memory than retaining a single feature on a trial.

Recent evidence sheds light on the confusability of feature-binding information. Bae and Flombaum (2013) studied the precision of memory for arrays of one or two objects. When there were two objects, they were sometimes distinguished by features that were integral with the feature being tested (Garner, 1974). Integral features are perceptually dependent: size judgments depend on shape of the object, luminance judgments depend on the hue, and tone amplitude judgments depend on the frequency of the tone. This is different from more separable features (e.g., size and color, luminance and orientation, tone amplitude and location). In a trial within one experiment of Bae and Flombaum, for example, a participant might have to judge the size of a triangle that appeared in an array with another triangle, or the size of a triangle that appeared in an array with a circle. In another experiment, the judgment was luminance in the presence of same or different hues, and in another, it was the amplitude of tones presented in the same or different frequencies. In all of these situations, when there was a discriminating feature that was integral to the feature being tested, the precision was as good for two-item arrays as it was for one-item arrays. In contrast, in studies in which the discriminating feature was perceptually separable from the feature being judged, precision declined as the array size increased from one to two items (Anderson et al., 2011; Zhang & Luck, 2008). These findings suggest that the loss of precision comes from the confusion of binding among separable features. One cannot use a feature like color to keep separate the sizes of different objects in an array; the judgment of one size is influenced by the sizes of other objects in the same array (Brady & Tenenbaum, 2013). In the same way, we suggest that voice and digit are separable features, as are color and shape, leading to the inability to use a peripheral component of working memory to retain more than about one voice-digit or color-shape binding at a time.

**Role of attention:** . The need for attention in retaining already-encoded visual items has been reinforced by the finding of interference with maintenance of even a few visual items when a tone categorization task is to be carried out (Morey & Bieler, 2013; Stevanovski & Jolicoeur, 2007). Ricker, Cowan, and Morey (2010) found that working-memory maintenance of unfamiliar characters was impaired by the retrieval of verbal information from long-term memory, even when the retrieval was covert in that no response was to be made in the secondary retrieval task. More generally, tradeoffs have been found between working memory and other tasks that do not require working memory, but do require attention. Chen and Cowan (2009b) showed a tradeoff between a verbal list memory task and a very simplistic visual choice reaction time task. In serial recall, not only is storage affected by the rate of attention-demanding processing (Barrouillet et al., 2011), but the rate of processing is affected by the amount of information being stored (Vergauwe, Camos, & Barrouillet, in press). Note, however, that some other studies have found no effect of general attention in working memory (e.g., Woodman, Vogel, & Luck, 2001; Fougny & Marois, 2006; Hollingworth & Maxcey-Richard, 2013) so researchers disagree on the issue.

Several brain imaging studies have shown a tradeoff between visual and verbal working memory information (Chein, Moore, & Conway, 2011; Cowan et al., 2011; Majerus et al., 2010) or between working memory and attention (Majerus et al., 2012). Working memory for both verbal and visual information appears to depend on an area of the brain, the intraparietal sulcus (Cowan et al., 2011; Todd & Marois, 2004; Xu & Chun, 2006) that also is found to be heavily involved in non-mnemonic attention processes; a part of the intraparietal sulcus seems to act as a “hub” with functional connections to other brain areas that varies depending on what information is attended (e.g., J. Anderson, Ferguson, Lopez-Larson, & Yurgelun-Todd, 2010; Majerus et al., 2006). This brain area appears to help maintain multiple items at once for some seconds, in that blood oxygen-level dependent (BOLD) activity increases with the visual working memory load until capacity is reached, at which point it levels off (Xu & Chun, 2006; Todd & Marois, 2004). This same area (on the left side) is the only area that Cowan et al. found to increase activity with both verbal and nonverbal memory loads, so it seems that in some way, attention is used for the central portion of working memory storage.

There are several ways in which attention could be deployed in the present ensemble of working memory tasks. First, attention could be used to encode the list or array items in parallel. Alternatively, even for arrays, attention could be shifted rapidly from a single item to another or from one subset of items to another, until the items are encoded. The exact method of encoding does not matter for the present results because stimulus sets were presented one at a time.

What is more germane is the attention function that accounts for the tradeoff between tasks. One possibility is that the storage of information in peripheral (modality- or code-specific) storage faculties is incomplete; and that some of the information therefore must be stored in the focus of attention. This is the model promoted by Sauls and Cowan (2007). For the present data set, however, the central component is smaller than we thought, amounting to about one item, fewer than the 3 or 4 items that we expected based on Cowan (2001), so this explanation is less compelling than we previously thought. It could make sense,

nevertheless, according to theories in which there is a focus of attention that is limited to one item at a time (e.g., Garavan, 1998; McElree, 1998; Oberauer, 2002; Öztekin, Davachi, & McElree, 2010; Nee & Jonides, 2011). Oberauer et al. (2012) included a one-item focus of attention in addition to interference, and therefore could serve as one such two-process theory.

Another possibility is that all of the items in working memory are represented in peripheral storage but that this storage is short-lived unless the representations are reactivated, either through attention or through covert articulation (Barrouillet et al., 2011; Camos et al., 2011). Given that we used articulatory suppression to prevent covert articulation, it would presumably be attention that is used to reactivate, i.e. refresh, the representations. According to this theory, though, by retaining two sets of stimuli at once, the rate at which either set can be refreshed is in effect cut in half.

Refreshing one item at a time in a circulating fashion would seem to predict a more severe loss in memory than we observed. If only one item can be refreshed at a time and the result is just a one-item difference between conditions in what can be recalled from each modality, why circulate attention rather than just maintain the focus of attention on one item the entire time? Refreshing could still make sense, depending on the rate of loss or decay of information from working memory that is assumed, in the absence of refreshing. Suppose, for example, that there are three verbal and three visual items in working memory and that they decay at such a slow rate that during the retention interval, without refreshing only one item in each set would be lost. Further suppose that the refreshing process is not selective to the representation in danger of loss, but rotates between all objects. In that scenario, the likelihood that refreshing will save a particular item that would otherwise be lost could be 1.0 in a unimodal retention situation, versus 0.5 in a bimodal situation. Loss over time is still a matter of dispute in the verbal arena (e.g., Barrouillet, De Paepe, & Langerock, 2012; Barrouillet, Plancher, Guida, & Camos, 2013; Oberauer & Lewandowsky, 2008, 2013) but has been demonstrated for nonverbal items (Ricker & Cowan, 2010; Zhang & Luck, 2009).

An intriguing recent theory by Oberauer (2013; cf. Oberauer & Hein, 2012) takes into account evidence that the focus of attention is often restricted to a single item, but other times can expand to include more than one item (cf. Cowan, 2005). According to the theory, multiple items can be in the focus of attention at the same time if the task requires it or favors it, but this tends not to occur otherwise because there is interference between items that are in the focus of attention concurrently. What is said to be critical is the binding between items and context, with confusions between those bindings taking place.

Some evidence that the focus of attention was limited to a single item came from a study using memory search with a cued response deadline, and showing a faster retrieval dynamic (increase in accuracy as a function of the response cue latency) of the most recent item compared to other items (McElree, 1998). Additional evidence came from a procedure in which items in a set (e.g., numbers) have to be remembered while one item is updated (e.g., through addition or subtraction), the finding being that it takes less time to update the most recent item than to update a different item. Presumably, this repetition benefit occurs because the most recent item is already in the focus of attention. Next, there was evidence

that multiple items could be advantaged, but presumably only because they could be chunked together. McElree (1998) found advantaged recall of the last three items when they all came from the same semantic category. Oberauer and Bialkova (2009) found repetition benefits for two items that they thought could be chunked together. In retrospect, however, it seems unlikely that items in the same semantic category can be chunked together as McElree thought, given that the response probe could include a lure of the same category that was absent from the list. Modifications of the procedure of Oberauer resulted in repetition benefits for multiple items even when there was proof that these items were not chunked together (Gilchrist & Cowan, 2011; Oberauer & Bialkova, 2011). With practice, as well, the pattern of repetition benefits indicating a single-item focus of attention can change, expanding from an advantage for one item to a graded advantage for several at once (Verhaeghen, Cerella, & Basak, 2004). The availability of the focus of attention for multiple items at once has been supported by Beck, Hollingworth, and Luck (2012). The observed eye movements as participants searched an array either for instances of a single color, or for instances of two colors. When they searched for instances of a single color, eye movements landed on that color repeatedly, whereas when they searched for two colors, eye movements landed on either of the colors.

Finally, Cowan, Donnell, and Saults (2013) presented lists of 3, 6, or 9 items for an orienting task (indicating which word in each list is most interesting) and later tested memory for whether two words, from nearby serial positions, came from the same list or different lists. Participants could do this task with better-than-chance accuracy only when the lists were 3 items long. The theoretical suggestion was that only 3 items typically are held in the focus of attention at once, and associations are formed between items that are in the focus concurrently.

In the present study, according to Oberauer (2013) one could imagine that the focus of attention expands, of necessity, to take in items from the first list or array on a trial and encode them in memory, only up to the point that limits confusions; on average, about 3 items. After a representation of those items is set up in memory (in the activated portion of long-term memory according to Cowan, 1988, or in a currently-relevant subset according to Oberauer), the focus of attention is then free to take in the second list or array. The great difference between the nature of the two stimulus sets on a trial and the separate times of their presentation then presumably results in only a little cross-interference between the sets.

Given two sets of stimuli represented in activated memory, what could the focus of attention do in the retention interval to enhance performance? It might be used to bolster some part of the representation that is weak, whether through a circulating process of refreshing or through the constant holding of about one item in the focus throughout the retention interval. Alternatively, the focus might be used to keep pointers to the two stimulus sets. None of these hypotheses, however, lead specifically to the prediction of the magnitude of interference between stimulus sets, nor do any rule it out. Thus, although we have eliminated two different single-process theories in favor of a class of theories in which interference and attention both operate, we do not yet have enough information to determine exactly how attention is used in the task, such that a central component greater than zero, of about a single item, is obtained.



The discussion has applied to items that are simple enough so that all items are comparable in terms of the role of attention in storage. Alvarez and Cavanagh (2004) showed that working memory capacity was smaller for arrays of complex items (e.g., Chinese characters for non-Chinese readers, or cubes of different orientations) than for arrays of simple items (e.g., colors or known letters). Awh, Barton, and Vogel (2007) showed, however, that the complexity of items does not affect the number of items in working memory; the same number of items is held, but with the complex items maintained at a lower level of resolution. Thus, in a list of items of mixed item types (e.g., some colors, some cubes, and some Chinese characters), a participant might recall that there was a Chinese character in a certain array location without recalling exactly which Chinese character it was. The available attention might be divided among array items, limiting the resolution of each item, or attention might be concentrated on a single complex item, perhaps by encoding the item by parts in working memory, representing it as more than one chunk per character to allow greater resolution.

**Are there different kinds of selective attention?:** Throughout this discussion we have implied that there is a single kind of selective attention, with the sole attentional mechanism used for different purposes in encoding and maintenance periods. We do make that assertion, tentatively at least. Some evidence for doing so is the finding that similar mechanisms are used for visual search and visual working memory. Anderson et al. (2013) found that searching for an L among Ts placed in different orientations was a process that depended on the N2pc component of event-related potentials in a manner that depended on the array size. The pattern was quite similar to what has been observed in an array memory task (except when the stimuli in the visual search task allowed grouping similar distracting objects into a single pattern or chunk). Specifically, the pattern suggested that participants search a number of items, up to the working memory limit, all at once and therefore that the N2pc component reaches an asymptotic level when the search set size corresponds to the participant's working memory capacity.

Hollingworth and Maxcey-Richard (2013) found what they interpreted as the use of different attentional mechanisms for visual search and visual working memory. We, however, question this conclusion. They did find an effect of the search task on working memory, but they placed more weight on another finding: the absence of modulation of this effect on working memory by a post-array cue that specified one item to be remembered, thus reducing the memory load from 6 items to 1. It is not clear to us, however, just what the theories predict here. On one hand, the post-array cue could allow better performance on the cued item by focusing attention on it, which perhaps should make that memory more vulnerable to distraction. On the other hand, performance on that cued item required attention to at most one slot, which might free up more attention for the search task compared to the uncued condition. These factors could offset one another. The final piece of evidence Hollingworth and Maxcey-Richard deemed relevant was the absence of an effect of working memory on visual search. That null finding, however, could occur because the visual search was for a target that did not change throughout the session. Moreover, the search array may allow for some chunking in the visual search task, inasmuch as multiple objects in the search array were identical; that should decrease the need for attention during

search (cf. Anderson et al., 2013) and it might allow the switching of attention between tasks with no cost to the search task. Overall, we do not find compelling evidence of different kinds of attention that operate independently for encoding and maintenance.

### Concluding Remarks

In sum, we have shown that the notion of the focus of attention as a central holding area is too simplistic to account for a wide range of experimental results involving the combination of verbal and visual stimuli. We have delineated reasons why we still believe that attention is intricately involved in working memory maintenance and why a simple interference view of capacity limits may not be sufficient. If attention is involved in working memory maintenance, it also appears that some information is instead held in a more automatic form of storage, allowing a wider degree of flexibility for the use of attention. The findings are compatible with our favored notion, which is that the focus of attention zooms out to encompass several items at once in the encoding phase (Cowan et al., 2005) and then zooms in to maintain fewer, maybe even one item at a time (Barrouillet et al., 2011) more intensively during the retention interval.

It seems fair to suggest that the field is making progress in that various models are becoming more similar. For example, Baddeley's (2000) addition of the episodic buffer provides it with a central store to complement the peripheral buffer stores, making it closer to Cowan (1988). The present finding is that items are apparently not always held en masse in central storage, but are in many circumstances relegated to peripheral storage when possible except for about 1 item at a time engaged in focal attention. This finding brings our model closer to one in which storage limits have to do largely with feature-specific interference (Oberauer, 2013; Oberauer et al., 2012), though attentional capacity limits still play an important role (Cowan, Blume, & Saults, 2013), and consistently so in the encoding of information into working memory (Vogel et al., 2006). This evidence should be important in the near future for continuing efforts to understand the nature of working memory capacity limits.

Our results are relevant to the theoretical question of how much material can be held in conscious attention at once. Peripheral storage is, in effect, storage in one modality that cannot be traded in for more storage in the other modality, which appears to make it non-attentional in nature. But if attention is generally as constrained as in our study, comparisons of items to one another must occur either with attention in an expanded state (Cowan et al., 2005; Oberauer, 2013) or with the assistance of peripheral storage. Practical implications of our results also remain to be explored. Much of the focus of the field since Daneman and Carpenter (1980) has been on working memory storage in the presence of processing or manipulation of the materials, whereas we examine storage of one or two sets of materials with no other processing; different constraints might apply in this case. Specifically, storage of two stimulus sets may in some way cause more conflict than is found between storage of one set and processing of another. In a broader sense, working memory also pertains to how much can be stored in order to facilitate learning and comprehension. Cowan, Donnell, and Saults (2013) showed that associations between adjacent list items are formed better within short lists, and assumed that the associations are formed between items in the focus of attention concurrently; one could now ask if they form in peripheral rather than central

storage. Halford, Cowan, and Andrews (2007) argued that as children develop, increases in working memory storage capacity underlie the increase in the complexity of ideas that they can understand, but Halford et al. do not talk about materials presented in mixed modalities. There are recent debates in the educational field regarding the best way to present educational material (Cowan, 2013) and the present results reinforce the possibility that, to make best use of peripheral storage capabilities, it may be most effective to present sets of concurrent concepts to be integrated together partly in verbal form and partly in nonverbal, visual form.

## Acknowledgments

We thank Alexander P. Boone, D. Johnson, Jacob Nicholson, and Suzanne Redington for assistance. This work was completed with support from NICHD Grant R01-HD21338.

## Appendix A

### Models for Items in Working Memory

The full derivation of models of the number of items in working memory suitable for several different tasks was presented by Cowan, Blume, and Saults (2013) and is briefly summarized here. All of the models are designed to account for situations in which an array of objects is studied and the studied array is followed by a probe array or a probe item. The task is always to indicate whether there has been a change between the studied array and the probe array or probe item. In the latter case, the question is whether the probe item differs from the item in the same location within the studied array. According to all the models, an individual is assumed to hold  $k$  items in working memory on a particular trial and the model is designed to allow an estimate of  $k$  based on the proportion of hits, or correctly detected changes from a studied array, and false alarms, or incorrect responses indicating that there was a change when in fact there was none. If  $k$  is smaller than the number of array items, then  $k$  also estimates the participant's working memory capacity.

Model 1: Studied array followed by a probe array (Pashler, 1988)

The logic of Pashler (1988) was designed for a task in which each array was followed by a second array identical to the first or differing in the identity of one item in the array. According to this model, if there are  $N$  array items and there is a change in one item from the studied array to the probe array, the chances of answering correctly ( $h$ , probability of hits) include the chance that the changed item is in working memory plus the chance that it is not in working memory but the participant guesses correctly given a guessing rate  $g$  for that set size:

$$f = g.$$

There is presumably no information about a possible change when there is no change, so performance ( $f$ , false alarm rate) is based on guessing:

$$k = N(h - f) / (1 - f).$$

Combining these equations produces

$$f = (1 - k/N)g.$$

Notice that the  $g$  term drops out of the equation.

It may seem puzzling to learn that an individual has no information when there is no change. Clearly, if  $N$  is less than capacity, the participant will know all of the items and therefore will know there has been no change. In this situation,  $g=0$  so no mistakes are made, and that is how the issue is handled in the model; and  $g$  changes with set size.

#### Model 2: Studied Array Followed by a Central Probe

This model is new with Cowan, Blume, and Saults (2013) and is designed for a commonly-used situation in which the studied array is followed a single-item probe appearing at the center of the array. The probe is identical to one item that occurred at one location in the array or is identical to no array item. In the latter case, for this model the probe contains a feature or attribute not present in any other item in the studied array (e.g., a new color or shape).

The rational means to carry out the task is to search memory to find a match to the probe. If no match is found the participant does not know whether the item was missing from the array, or was in the array as an item that did not make it into working memory. In that case, the only course of action is to guess with a certain guessing rate. That is,

$$h = g.$$

that is, when there is no knowledge of the match there can be a guess that there was a new item, and

$$k = N(h - f) / h$$

Putting these equations together,

$$h = k/N + (1 - k/N)g$$

#### Model 3: From Cowan (2001)

This model is meant for the situation in which the probe is a single item presented at the same location as one of the array items, with the participants informed that if there is a change anywhere in the array, it is in the item in the probe's location. Given this limiting of the necessary information to a known location, there is information if the probe is the same

as the targeted array item and there is information if the probe differs from it, provided that the targeted item is in working memory. With no information the participant guesses. That is,

$$f = (1 - k/N) g.$$

and

$$k = N(h - f).$$

Putting these equations together,

$$h = c + (1 - c)g$$

#### Model 4: Two-Attribute Binding Model

This model is meant for a situation in which the probe is a single item that either is identical to one of the array items, or is the combination of attributes from two items (e.g., the shape of one array item combined with the color of another array item). There is an asymmetry in the information that is needed to detect the absence of change, for which one array item is relevant, namely the item matching the probe, versus the presence of a recombination change, for which either of two array items provides the necessary information, Cowan, Blume, and Sauls (2013) did not find a closed-form solution to the problem but provided a lookup table for arrays of four unique items, each with two task-relevant attributes (e.g., shape and color). When there is a recombination change, either of two items is relevant to detecting the change:

$$c = (k/N) [(k - 1) / (N - 1)] + 2(k/N) (1 - [(k - 1) / (N - 1)])$$

where  $c$  is the probability that the participant knows at least one of the two items contributing features to the recombination probe. The equation for  $c$  is

$$f = (1 - k/N) g.$$

where the first part of the sum reflects knowledge of both items and the second part reflects knowledge of one item or the other (hence the factor of 2).

When there is no recombination, only one item is relevant, the item identical to the probe, so

$$C = C_{verb} + C_{obj} \quad (1)$$

All these equations were used together to calculate a  $k$  corresponding to each combination of  $h$  and  $f$ , with the values comprising a table in Cowan, Blume, and Sauls (2013).

## Appendix B

### Calculation of Central and Peripheral Components of Working Memory from Dual Tasks

Figure 1 shows how retention of information from two sets in working memory can be divided into peripheral and central parts; the central parts are the parts that can be allocated to either set, in our situation an object set or a verbal set, depending on the task instructions. The peripheral parts are the parts that stay allocated to a single stimulus set regardless of instructions. Thus, if nonverbal objects are not to be retained on a particular trial, the area for peripheral storage of nonverbal objects cannot be used instead to retain verbal objects, and vice versa.

We will call the number of items held from a particular modality in peripheral form  $P_x$ , where  $x$  is the modality or domain (in our case, *verb* for verbal information or *obj* for nonverbal visual objects). The mental faculty that holds the information may not be capacity-limited per se, so the value of  $P_x$  could be specific to a certain array size.

By definition, central storage is a working memory resource that can be devoted to either modality or domain in a dual-task situation. (Given that this central information is at the discretion of the participant, we believe that it requires a form of attention.) It is quantified as  $C$  and is composed of two parts:

$$k_{U_{verb}} = P_{verb} + C_{verb} + C_{obj} = P_{verb} + C \quad (2)$$

where  $C_{verb}$  is the reduction in nonverbal objects retained when the participant is also responsible for verbal items and, conversely,  $C_{obj}$  is the reduction in verbal items retained when the participant is also responsible for visual objects. It is reasonable to add these components,  $C_{verb}$  and  $C_{obj}$ , inasmuch as both reductions occur on the same trials (bimodal trials). Any cost of dividing attention between modalities per se is absorbed in  $C$ .

The number of verbal stimuli stored in working memory is the total of peripheral and central resources allocated to digits:  $P_{verb} + C_{verb} + C_{obj}$  in the unimodal condition, and  $P_{verb} + C_{verb}$  in the bimodal condition. Similarly, the number of visual objects stored is  $P_{obj} + C_{obj} + C_{verb}$  in the unimodal condition, and  $P_{obj} + C_{obj}$  in the bimodal condition. In the case of unimodal attention ( $U$ ), each modality presumably makes use of all of the central attentional resources:

$$k_{U_{obj}} = P_{obj} + C_{verb} + C_{obj} = P_{obj} + C \quad (3)$$

and

$$k_{B_{verb}} = P_{verb} + C_{verb} \quad (4)$$

For each individual, the number of spoken digits in working memory with bimodal attention ( $k_{B_{verb}}$ ) is composed of its peripheral automatic component ( $P_{verb}$ ) plus its attention-

dependent central component ( $C_{verb}$ ) that remains allocated to that modality during the dual task:

$$k_{Bobj} = P_{obj} + C_{obj} \quad (5)$$

Similarly, with bimodal attention the number of visual objects encoded into working memory is

$$C = C_{verb} + C_{obj} = k_{Uverb} + k_{Uobj} - k_{Bverb} - k_{Bobj} \quad (6)$$

By combining Equations 2-5, the key central and peripheral storage components can be calculated:

$$P_{verb} = k_{Uverb} - C \quad (7)$$

$$P_{obj} = k_{Uobj} - C \quad (8)$$

## REFERENCES

- Allen RJ, Baddeley AD, Hitch GJ. Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*. 2006; 135:298–313. [PubMed: 16719655]
- Allen RJ, Hitch GJ, Mate J, Baddeley AD. Feature binding and attention in working memory: A resolution of previous contradictory findings. *Quarterly Journal of Experimental Psychology*. 2012; 65:2369–2383.
- Alvarez GA, Cavanagh P. The capacity of visual short term memory is set both by visual information load and by number of objects. *Psychological Science*. 2004; 15:106–111. [PubMed: 14738517]
- Alvarez GA, Thompson TW. Overwriting and rebinding: why feature-switch detection tasks underestimate the binding capacity of visual working memory. *Visual Cognition*. 2009; 17:141–159.
- Anderson DE, Vogel EK, Awh E. Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *The Journal Of Neuroscience*. 2011; 31:1128–1138. [PubMed: 21248137]
- Anderson DE, Vogel EK, Awh E. A common discrete resource for visual working memory and visual search. *Psychological Science*. 2013; 24:929–938. [PubMed: 23572280]
- Anderson JS, Ferguson MA, Lopez-Larson M, Yurgelun-Todd D. Topographic maps of multisensory attention. *Proceedings of the National Academy of Science of the United States of America*. 2010; 107:20110–20114.
- Atkinson, RC.; Shiffrin, RM. Human memory: A proposed system and its control processes.. In: Spence, KW.; Spence, JT., editors. *The psychology of learning and motivation: Advances in research and theory*. Vol. 2. Academic Press; New York: 1968. p. 89-195.
- Awh E, Barton B, Vogel EK. Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*. 2007; 18:622–628. [PubMed: 17614871]
- Baddeley, AD. *Oxford Psychology Series #11*. Clarendon Press; Oxford: 1986. *Working memory..*
- Baddeley A. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*. 2000; 4:417–423. [PubMed: 11058819]
- Baddeley A. The magic number and the episodic buffer. *Behavioral and Brain Sciences*. 2001; 24:117–118.
- Baddeley, AD.; Hitch, G. Working memory.. In: Bower, GH., editor. *The psychology of learning and motivation*. Vol. 8. Academic Press; New York: 1974. p. 47-89.

- Baddeley A, Lewis V, Vallar G. Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology*. 1984; 36A:233–252.
- Bae GY, Flombaum JI. Two items remembered as precisely as one: How integral features can improve visual working memory. *Psychological Science*. 2013; 24:2038–2047. [PubMed: 23938276]
- Barrouillet P, De Paepe A, Langerock N. Time causes forgetting from working memory. *Psychonomic Bulletin & Review*. 2012; 19:87–92. [PubMed: 22184034]
- Barrouillet P, Plancher G, Guida A, Camos V. Forgetting at short term: When do event-based interference and temporal factors have an effect? *Acta Psychologica*. 2013; 142:155–167. [PubMed: 23337080]
- Barrouillet P, Portrat S, Camos V. On the law relating processing to storage in working memory. *Psychological Review*. 2011; 118:175–192. [PubMed: 21480738]
- Bays PM, Husain M. Dynamic shifts of limited working memory resources in human vision. *Science*. 2008; 321:851–854. [PubMed: 18687968]
- Bays PM, Wu EY, Husain M. Storage and binding of object features in visual working memory. *Neuropsychologia*. 2011; 49:1622–1631. [PubMed: 21172364]
- Beck VM, Hollingworth A, Luck SJ. Simultaneous control of attention by multiple working memory representations. *Psychological Science*. 2012; 23:887–898. [PubMed: 22760886]
- Boersma, P.; Weenink, D. Praat: doing phonetics by computer [Computer program]. Version 5.1.02. 2009. retrieved March 2009 from <http://www.praat.org/>
- Brady TF, Tenenbaum JB. A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review*. 2013; 120:85–109. [PubMed: 23230888]
- Bregman, AS. Auditory scene analysis. MIT Press; Cambridge, MA: 1990.
- Broadbent, DE. Perception and communication. Pergamon Press; New York: 1958.
- Brown GDA, Neath I, Chater N. A temporal ratio model of memory. *Psychological Review*. 2007; 114:539–576. [PubMed: 17638496]
- Camos V, Mora G, Oberauer K. Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*. 2011; 39:231–244. [PubMed: 21264630]
- Chein JM, Moore AB, Conway ARA. Domain-general mechanisms of complex working memory span. *NeuroImage*. 2011; 54:550–559. [PubMed: 20691275]
- Chen Z, Cowan N. Core verbal working memory capacity: The limit in words retained without covert articulation. *Quarterly Journal of Experimental Psychology*. 2009a; 62:1420–1429.
- Chen Z, Cowan N. How verbal memory loads consume attention. *Memory & Cognition*. 2009b; 37:829–836. [PubMed: 19679862]
- Cocchini G, Logie RH, Della Sala S, MacPherson SE, Baddeley AD. Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*. 2002; 30:1086–1095. [PubMed: 12507373]
- Conrad R. Acoustic confusion in immediate memory. *British Journal of Psychology*. 1964; 55:75–84.
- Cowan N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*. 1988; 104:163–191. [PubMed: 3054993]
- Cowan N. Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*. 1992; 31:668–684.
- Cowan, N. An embedded-processes model of working memory.. In: Miyake, A.; Shah, P., editors. *Models of Working Memory: Mechanisms of active maintenance and executive control*. Cambridge University Press; Cambridge, U.K.: 1999. p. 62-101.
- Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*. 2001; 24:87–185. [PubMed: 11515286]
- Cowan, N. Working memory capacity. Psychology Press; Hove, East Sussex, UK: 2005.
- Cowan N. online before print). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*. 2013 DOI: 10.1007/s10648-013-9246-y.



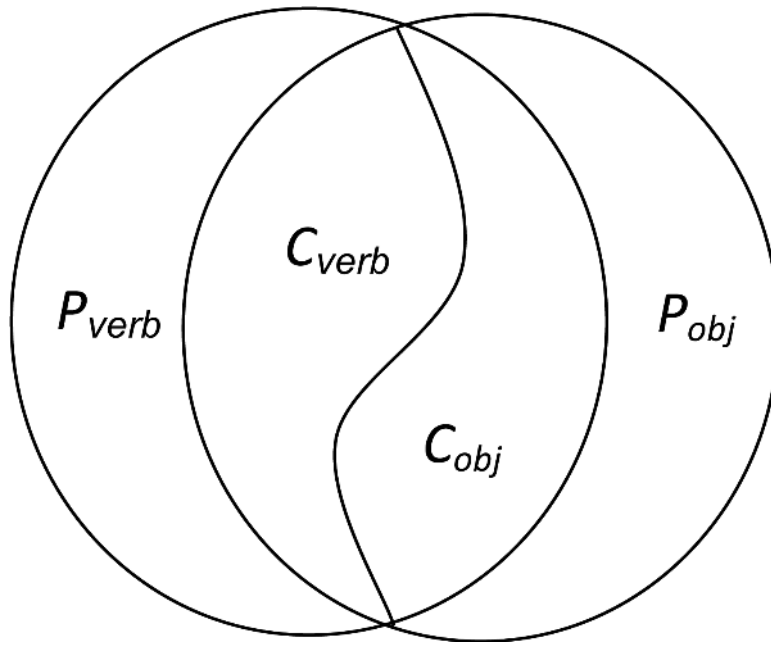
- Cowan N, Blume CL, Saults JS. Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2013; 39:731–747.
- Cowan N, Donnell K, Saults JS. A list-length constraint on incidental item-to-item associations. *Psychonomic Bulletin & Review*. 2013; 20:1253–1258. [PubMed: 23653420]
- Cowan N, Elliot EM, Saults JS, Morey CC, Mattox S, Hismjatullina A, Conway ARA. On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*. 2005; 51:42–100. [PubMed: 16039935]
- Cowan N, Li D, Moffitt A, Becker TM, Martin EA, Saults JS, Christ SE. A neural region of abstract working memory. *Journal of Cognitive Neuroscience*. 2011; 23:2852–2863. [PubMed: 21261453]
- Cowan N, Lichty W, Grove TR. Properties of memory for unattended spoken syllables. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1990; 16:258–269.
- Cowan N, Morey CC. How can dual-task working memory retention limits be investigated? *Psychological Science*. 2007; 18:686–688. [PubMed: 17680938]
- Cowan N, Naveh-Benjamin M, Kilb A, Saults JS. Life-Span development of visual working memory: When is feature binding difficult? *Developmental Psychology*. 2006; 42:1089–1102. [PubMed: 17087544]
- Cowan N, Rouder JN. Comment on “Dynamic shifts of limited working memory resources in human vision.”. *Science*. 2009; 323(5916):877. [PubMed: 19213899]
- Cowan N, Rouder JN, Blume CL, Saults JS. Models of verbal working memory capacity: What does it take to make them work? *Psychological Review*. 2012; 119:480–499. [PubMed: 22486726]
- Daneman M, Carpenter PA. Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*. 1980; 19:450–466.
- Darwin CJ, Turvey MT, Crowder RG. An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*. 1972; 3:255–267.
- Delvenne J-F, Bruyer R. Does visual short term memory store bound features? *Visual Cognition*. 2004; 11:1 – 27.
- Donkin C, Nosofsky RM, Gold JM, Shiffrin RM. Discrete slot models of visual working-memory response times. *Psychological Review*. 2013; 4:873–902. [PubMed: 24015956]
- Efron R. Effects of stimulus duration on perceptual onset and offset latencies. *Perception & Psychophysics*. 1970a; 8:231–234.
- Efron R. The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia*. 1970b; 8:37–55. [PubMed: 5522546]
- Efron R. The minimum duration of a perception. *Neuropsychologia*. 1970c; 8:57–63. [PubMed: 5522547]
- Fougnie D, Alvarez GA. Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal Of Vision*. 2011; 11:1–12.
- Fougnie D, Marois R. Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. *Psychological Science*. 2006; 17:526–534. [PubMed: 16771804]
- Fougnie D, Marois R. What limits working memory capacity? Evidence for modality-specific sources to the simultaneous storage of visual and auditory arrays. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2011; 37:1329–1341.
- Fougnie D, Suchow JW, Alvarez GA. Variability in the quality of working memory. *Nature Communications*. 2012; 3:1229.
- Garavan H. Serial attention within working memory. *Memory & Cognition*. 1998; 26:263–276. [PubMed: 9584434]
- Garner, WR. *The processing of information and structure*. Erlbaum; Potomac, MD: 1974.
- Gilchrist AL, Cowan N. Can the focus of attention accommodate multiple separate items? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2011; 37:1484–1502.
- Goldinger SD, Pisoni DB, Logan JS. On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1991; 17:152–162.

- Griffin IC, Nobre AC. Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*. 2003; 15:1176–1194. [PubMed: 14709235]
- Guttentag RE. The mental effort requirement of cumulative rehearsal: A developmental study. *Journal of Experimental Child Psychology*. 1984; 37:92–106.
- Halford GS, Cowan N, Andrews G. Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*. 2007; 11:236–242. [PubMed: 17475538]
- Hawkins, HL.; Presson, JC. Auditory information processing.. In: Boff, KR.; Kaufman, L.; Thomas, JP., editors. *Handbook of perception and human performance*. Vol. 2. Wiley; New York: 1986.
- Hollingworth A, Maxcey-Richard AM. Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance*. 2013; 39:1047–1058. [PubMed: 23067118]
- Jarrold C, Tam H, Baddeley AD, Harvey CE. The nature and position of processing determines why forgetting occurs in working memory tasks. *Psychonomic Bulletin & Review*. 2010; 17:772–777. [PubMed: 21169567]
- Jiang Y, Chun MM, Olson IR. Perceptual grouping in change detection. *Perception & Psychophysics*. 2004; 66:446–453. [PubMed: 15283069]
- Jolicoeur P, Dell'Acqua R. The demonstration of short-term consolidation. *Cognitive Psychology*. 1998; 36:138–202. [PubMed: 9721199]
- Kahneman, D. *Attention and effort*. Prentice Hall; Englewood Cliffs, NJ: 1973.
- Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, Engle RE. The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*. 2004; 133:189–217. [PubMed: 15149250]
- Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*. 2012; 24:61–79. [PubMed: 21955164]
- Logie RH, Della Sala S, Wynn V, Baddeley AD. Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology*. 2000; 53A:626–646. [PubMed: 10994221]
- Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature*. 1997; 390:279–281. [PubMed: 9384378]
- Machizawa MG, Goh CCW, Driver J. Human visual short-term memory precision can be varied at will when the number of retained items is low. *Psychological Science*. 2012; 23:554–559. [PubMed: 22527526]
- Macken WJ, Tremblay S, Houghton RJ, Nicholls AP, Jones DM. Does auditory streaming require attention? Evidence from attentional selectivity in short term memory. *Journal of Experimental Psychology: Human Perception and Performance*. 2003; 29:43–51. [PubMed: 12669746]
- Majerus SS, Attout L, D'Argembeau AD, Degueldre C, Fias W, Maquet P, et al. Attention supports verbal short-term memory via competition between dorsal and ventral attention networks. *Cerebral Cortex*. 2012; 22:1086–1097. [PubMed: 21765184]
- Majerus S, D'Argembeau A, Martinez Perez T, Belayachi S, Van der Linden M, Collette F, Salmon E, et al. The commonality of neural networks for verbal and visual short term memory. *Journal of Cognitive Neuroscience*. 2010; 22:2570–2593. [PubMed: 19925207]
- Majerus S, Poncelet M, Van der Linden M, Albouy G, Salmon E, Sterpenich V, Vandewalle G, Collette F, Maquet P. The left intraparietal sulcus and verbal short term memory: Focus of attention or serial order? *NeuroImage*. 2006; 32:880–891. [PubMed: 16702002]
- Makovski T, Jiang YV. Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review*. 2007; 14:1072–1078. [PubMed: 18229477]
- Marois, R. Invited presentation to *Attention & Performance XXV: Sensory Working Memory*. Saint-Hippolyte, Québec, Canada: Jul. 2013 The evolving substrates of working memory..
- Massaro, DW. *Experimental psychology and information processing*. Rand McNally; Chicago: 1975.
- McElree B. Attended and non attended states in working memory: Accessing categorized structures. *Journal of Memory and Language*. 1998; 38:225–252.

- Miller GA. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*. 1956; 63:81. [PubMed: 13310704]
- Morey CC, Bieler M. Visual short-term memory always requires attention. *Psychonomic Bulletin & Review*. 2013; 20:163–170. [PubMed: 23002024]
- Morey CC, Cowan N. When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*. 2004; 11:296–301. [PubMed: 15260196]
- Morey CC, Cowan N. When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2005; 31:703–713.
- Morey CC, Cowan N, Morey RD, Rouder JN. Flexible attention allocation to visual and auditory working memory tasks: Manipulating reward induces a tradeoff. *Attention, Perception, & Psychophysics*. 2011; 73:458–472.
- Morey CC, Mall JT. Cross-domain costs during concurrent verbal and spatial serial memory tasks are asymmetric. *Quarterly Journal of Experimental Psychology*. 2012; 65:1777–1797.
- Morey CC, Morey RD, van der Reijden M, Holweg M. Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*. 2013; 69:324–348.
- Naveh Benjamin M, Jonides J. Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1984; 10:369–385.
- Nee DE, Jonides J. Dissociable contributions of prefrontal cortex and the hippocampus to short-term memory: Evidence for a 3-state model of memory. *Neuroimage*. 2011; 54:1540–1548. [PubMed: 20832478]
- Norman DA, Bobrow DG. On data-limited and resource-limited processes. *Cognitive Psychology*. 1975; 7:44–64.
- Oberauer K. Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2002; 28:411–421.
- Oberauer K. The focus of attention in working memory—from metaphors to mechanisms. *Frontiers in Human Neuroscience*. 2013; 7:1–16. doi: 10.3389/fnhum.2013.00673. [PubMed: 23355817]
- Oberauer K, Bialkova S. Accessing information in working memory: Can the focus of attention grasp two elements at the same time? *Journal of Experimental Psychology: General*. 2009; 138:64–87. [PubMed: 19203170]
- Oberauer K, Bialkova S. Serial and parallel processes in working memory after practice. *Journal of Experimental Psychology: Human Perception and Performance*. 2011; 37:606–614. [PubMed: 21463089]
- Oberauer K, Eichenberger S. Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*. 2013; 41:1212–1227. [PubMed: 23716004]
- Oberauer K, Hein L. Attention to information in working memory. *Current Directions in Psychological Science*. 2012; 21:164–169.
- Oberauer K, Lewandowsky S. Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review*. 2008; 115:544–576. [PubMed: 18729591]
- Oberauer K, Lewandowsky S. Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General*. 2013; 142:380–411. [PubMed: 22866686]
- Oberauer K, Lewandowsky S, Farrell S, Jarrold C, Greaves M. Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*. 2012; 19:779–819. [PubMed: 22715024]
- Öztekin I, Davachi L, McElree B. Are representations in working memory distinct from representations in long term memory? Neural evidence in support of a single store. *Psychological Science*. 2010; 21:1123–1133. [PubMed: 20624934]
- Pashler H. Familiarity and visual change detection. *Perception & Psychophysics*. 1988; 44:369–378. [PubMed: 3226885]
- Penney CG. Modality effects and the structure of short term verbal memory. *Memory & Cognition*. 1989; 17:398–422. [PubMed: 2668697]

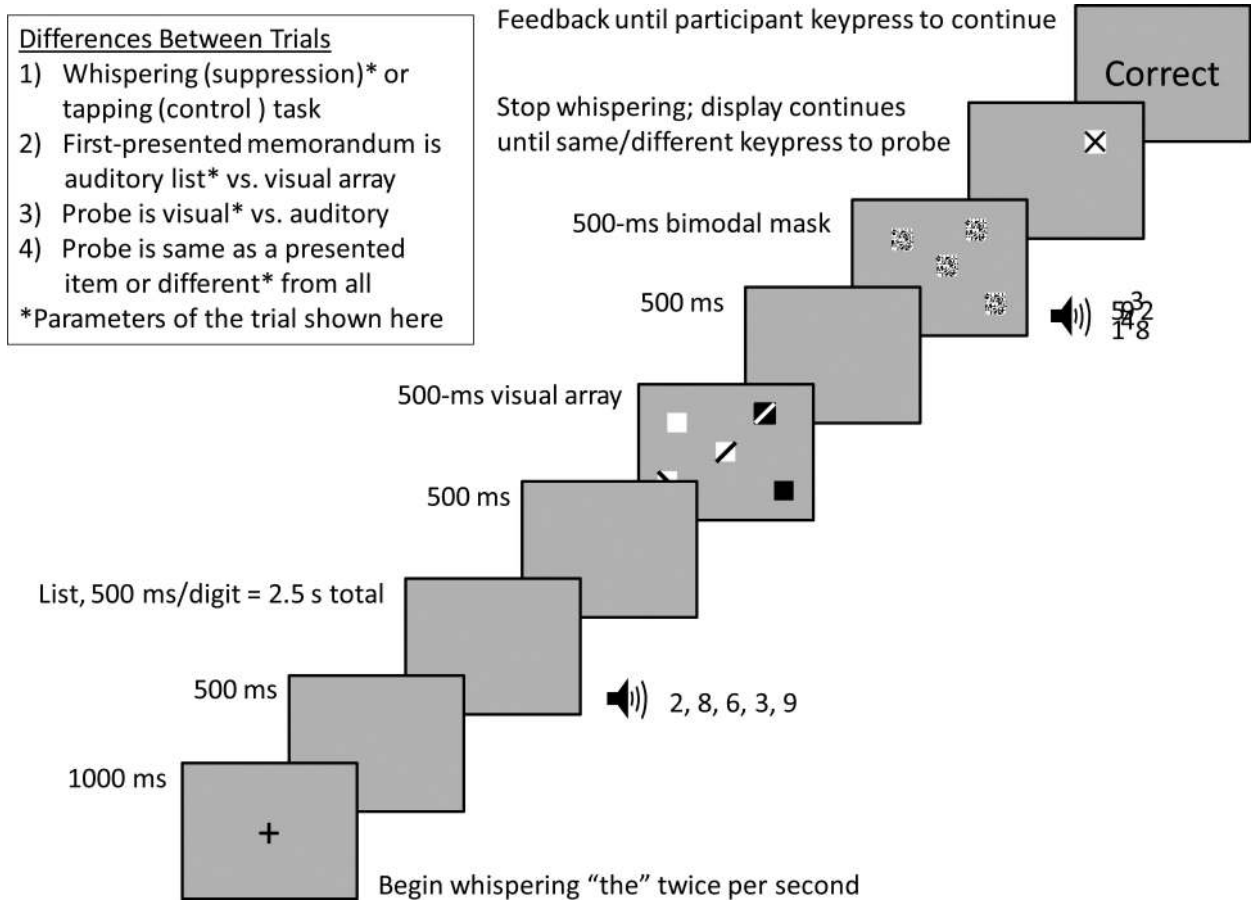
- Raye CL, Johnson MK, Mitchell KJ, Greene EJ, Johnson MR. Refreshing: A minimal executive function. *Cortex*. 2007; 43:135–145. [PubMed: 17334213]
- Phillips WA. On the distinction between sensory storage and short term visual memory. *Perception & Psychophysics*. 1974; 16:283–290.
- Ricker TJ, Cowan N. Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2010; 36:1355–1368.
- Ricker TJ, Cowan N, Morey CC. Visual working memory is disrupted by covert verbal retrieval. *Psychonomic Bulletin & Review*. 2010; 17:516–521. [PubMed: 20702871]
- Rouder JN, Morey RD, Cowan N, Zwilling CE, Morey CC, Pratte MS. An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences (PNAS)*. 2008; 105:5975–5979.
- Rouder JN, Morey RD, Morey CC, Cowan N. How to measure working-memory capacity in the change-detection paradigm. *Psychonomic Bulletin & Review*. 2011; 18:324–330. [PubMed: 21331668]
- Saults JS, Cowan N. A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology: General*. 2007; 136:663–684. [PubMed: 17999578]
- Shipstead Z, Engle RW. Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2013; 39:277–289.
- Sims CR, Jacobs RA, Knill DC. An ideal observer analysis of visual working memory. *Psychological Review*. 2012; 119:807–830. [PubMed: 22946744]
- Sligte IG, Scholte HS, Lamme VAF. Are there multiple visual short term memory stores? *PLOS One*. 2008; 3:1–9.
- Sperling G. The information available in brief visual presentations. *Psychological Monographs*. 1960; 74 (Whole No. 498).
- Stevanovski B, Jolicoeur P. Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition*. 2007; 15:532 – 563.
- Todd JJ, Marois R. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*. 2004; 428:751–754. [PubMed: 15085133]
- Unsworth N, Engle RW. The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*. 2007; 114:104–132. [PubMed: 17227183]
- van den Berg R, Shin H, Chou W-C, George R, Ma WJ. Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*. 2012; 109:8780–8785.
- Vergauwe E, Barrouillet P, Camos V. Do mental processes share a domain general resource? *Psychological Science*. 2010; 21:384–390. [PubMed: 20424075]
- Vergauwe E, Camos V, Barrouillet P. The impact of storage on processing: How is information maintained in working memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition*. in press.
- Verhaeghen P, Cerella J, Basak C. A Working-memory workout: How to expand the focus of serial attention from one to four items, in ten hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 30:1322–1337.
- Vogel EK, Woodman GF, Luck SJ. The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*. 2006; 32:1436–1451. [PubMed: 17154783]
- Vul E, Rich AN. Independent sampling of features enables conscious perception of bound objects. *Psychological Science*. 2010; 21:1168–1175. [PubMed: 20631320]
- Wheeler ME, Treisman AM. Binding in short term visual memory. *Journal of Experimental Psychology: General*. 2002; 131:48–64. [PubMed: 11900102]
- Woodman GF, Vogel EK, Luck SJ. Visual search remains efficient when visual working memory is full. *Psychological Science*. 2001; 12:219–224. [PubMed: 11437304]

- Woodman GF, Vecera SP, Luck SJ. Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*. 2003; 10:8087.
- Xu Y. Limitations in object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*. 2002; 28:458–468. [PubMed: 11999866]
- Xu Y, Chun MM. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*. 2006; 440:91–95. [PubMed: 16382240]
- Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. *Nature*. 2008; 453:23–35.
- Zhang W, Luck SJ. Sudden death and gradual decay in visual working memory. *Psychological Science*. 2009; 20:423–428. [PubMed: 19320861]
- Zhang W, Luck SJ. The number and quality of representations in working memory. *Psychological Science*. 2011; 22:1434–1441. [PubMed: 21987693]



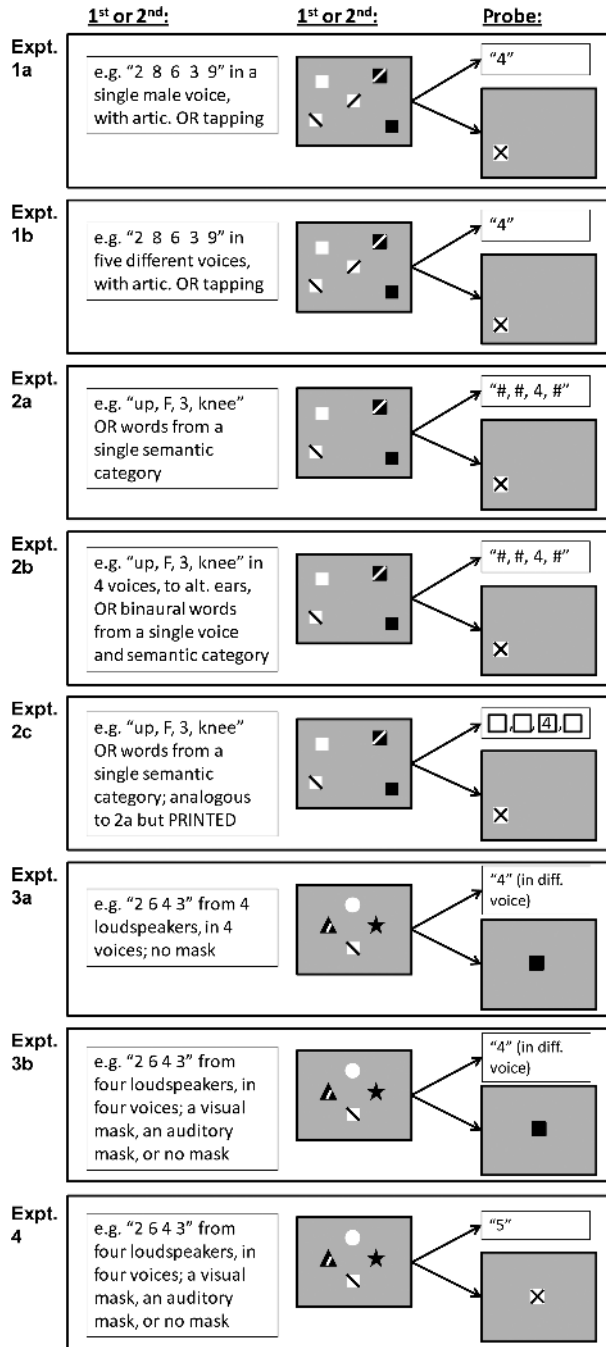
**Figure 1.**

A simple theoretical model of the distribution of working memory resources in the present tasks, including peripheral resources specific to memory of the verbal items ( $P_{verb}$ ), peripheral resources specific to memory of the visual objects ( $P_{obj}$ ), central resources devoted to verbal items during divided attention ( $C_{verb}$ ), and central resources devoted to visual objects during divided attention ( $C_{obj}$ ). In divided-attention situations, the number of items retained for modality  $x$  is  $P_x + C_x$ , where  $x=verb$  or  $x=obj$ . In a single-modality-attention situation, the number of items retained is  $P_x + C$ , where  $C = C_{verb} + C_{obj}$ .



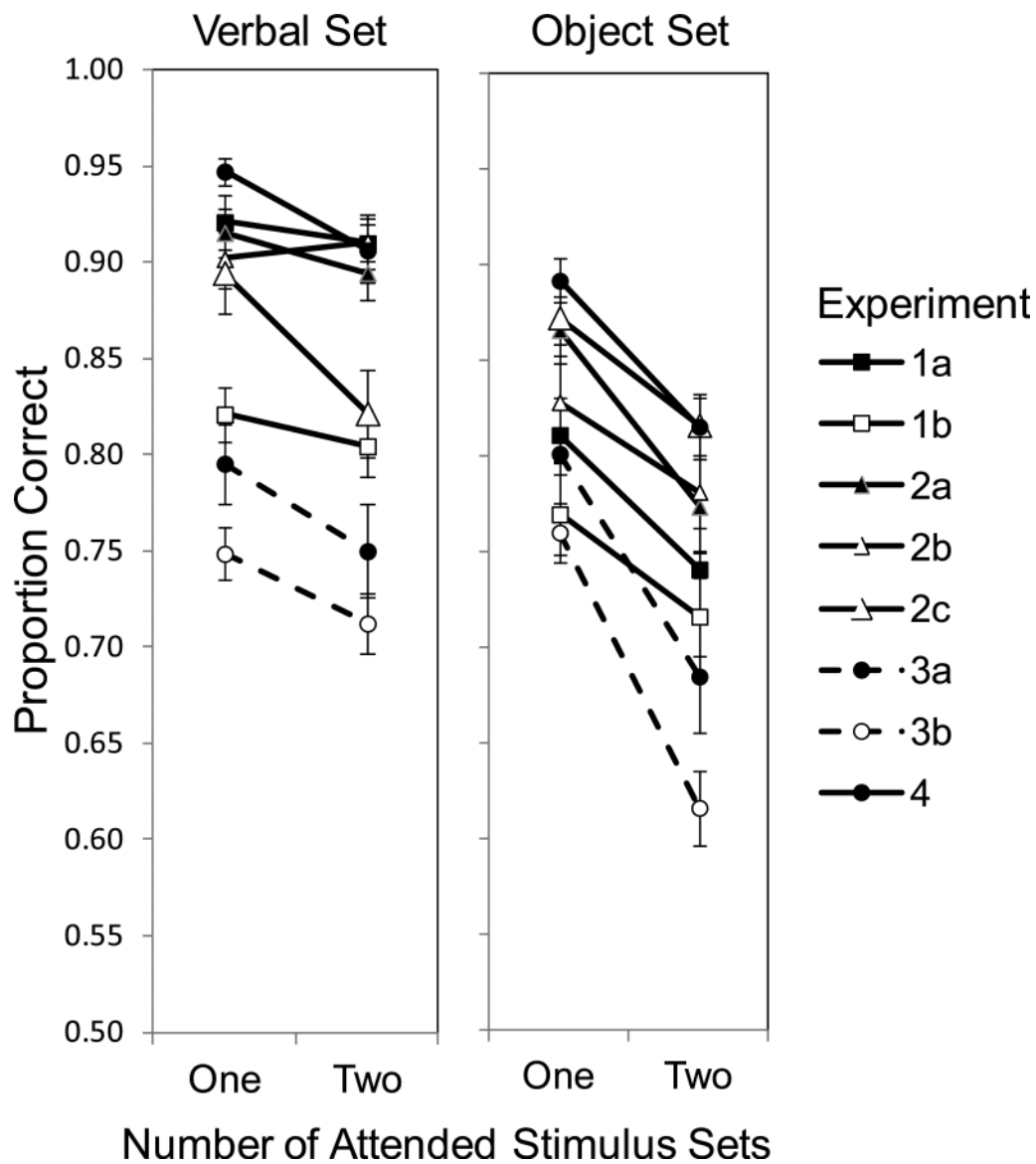
**Figure 2.**

A detailed illustration of a trial in Experiment 1a. The inset explains the different trial types in this experiment.

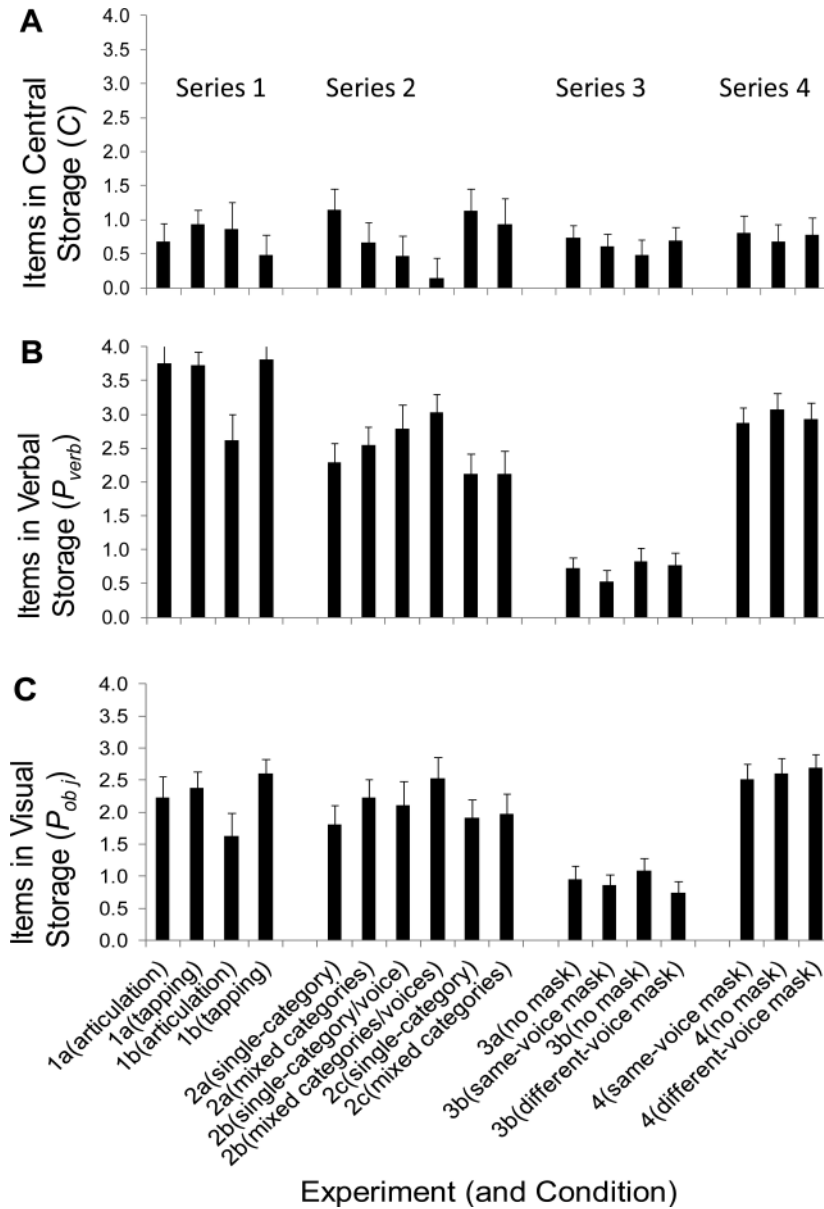


**Figure 3.** Illustration of two change trials, one for each stimulus type, in Experiments 1-4. On half the trials the nonverbal visual display preceded the verbal display (unlike what is shown in the figure). There was a bimodal mask before the probe, except as otherwise stated. In bimodal trial blocks, the probe could be either verbal or nonverbal. Objects' patterns within the arrays represent different colors.

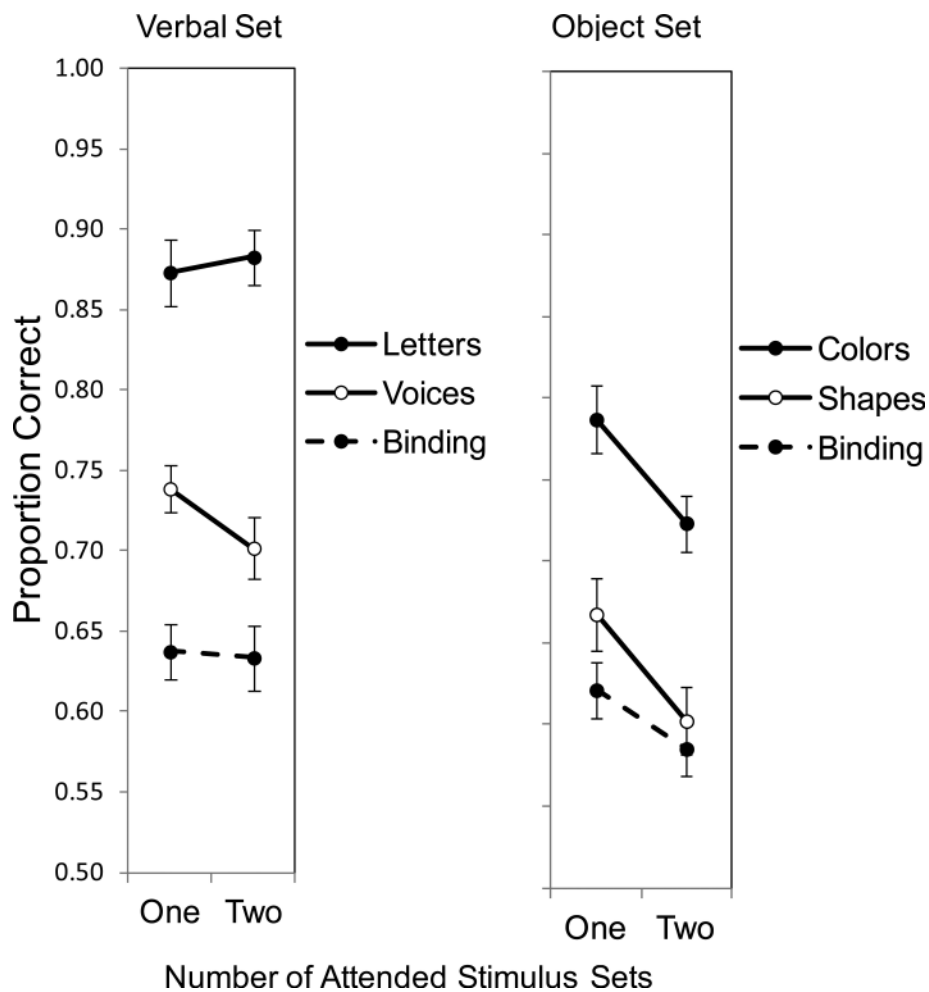




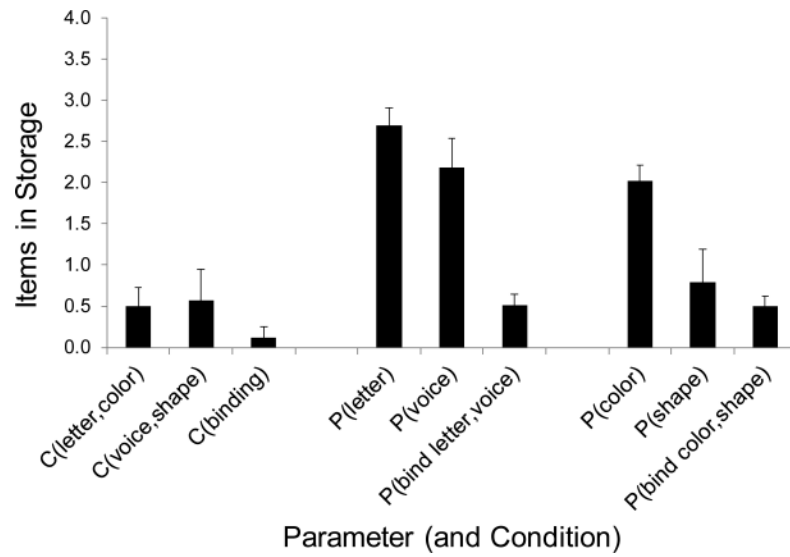
**Figure 4.** Mean proportion correct across conditions in each of the first eight experiments (graph parameter). The left panel shows the tests on verbal items and the right panel shows the tests on nonverbal objects. On the X axis is the number of stimulus sets attended: one (unimodal) versus two (bimodal) conditions. Error bars are standard errors.



**Figure 5.** Mean values of parameters of working memory for 17 conditions from the first 8 experiments. (A) Top panel: Parameter  $C$ , items in central storage; (B) middle panel, Parameter  $P_{verb}$ , items in peripheral verbal storage; (C) bottom panel, Parameter  $P_{obj}$ , objects in peripheral visual storage. Experiments in Series 3 (3a and 3b) are the only ones that required binding of visual features (colors with shapes) and verbal features (digits with spoken voices), and they produced much lower estimates of peripheral information storage than did the other experiments.



**Figure 6.** Proportion correct across conditions in the last, ninth experiment (Experiment 5). The left panel shows tests on acoustic verbal items and the right panel shows tests on nonverbal objects. On the X axis is the number of stimulus sets attended: one (unimodal) versus two (bimodal). The graph parameter is specific to each stimulus set. Error bars are standard errors.



**Figure 7.**

Parameter values obtained in Experiment 5. *C* refers to central storage and *P* refers to peripheral storage, specifically  $P_{verb}$  in the middle cluster of bars and  $P_{obj}$  in the right-hand cluster of bars.

Accuracy and working memory parameter values in experiments with a single acoustic probe (no position cue) or with a visual probe presented in a target location.

**Table 1**

Condition	Unimodal			Bimodal			Parameters				
	Digits( $h_cfa$ )	Objects( $h_cfa$ )	Objects( $h_cfa$ )	Digits( $h_cfa$ )	Objects( $h_cfa$ )	Objects( $h_cfa$ )	C	$P_{verb}$	$P_{obj}$		
Experiment 1a											
Articulatory Suppression											
M	0.87	0.09	0.82	0.24	0.87	0.11	0.79	0.32	<b>0.68</b>	<b>3.75</b>	<b>2.23</b>
SEM	0.03	0.02	0.03	0.03	0.02	0.02	0.04	0.03	0.27	0.29	0.34
Tapping											
M	0.97	0.06	0.89	0.23	0.96	0.08	0.82	0.33	<b>0.93</b>	<b>3.73</b>	<b>2.39</b>
SEM	0.01	0.01	0.02	0.02	0.01	0.01	0.03	0.04	0.21	0.20	0.24
Experiment 1b											
Articulatory Suppression											
M	0.77	0.21	0.81	0.30	0.76	0.24	0.75	0.37	<b>0.87</b>	<b>2.62</b>	<b>1.63</b>
SEM	0.04	0.02	0.03	0.04	0.03	0.03	0.03	0.03	0.39	0.39	0.36
Tapping											
M	0.87	0.15	0.83	0.26	0.83	0.14	0.83	0.34	<b>3.83</b>	<b>2.56</b>	
SEM	0.02	0.02	0.03	0.03	0.02	0.02	0.03	0.03	0.24	0.23	0.24

**Note.** Parameter values that are above zero by  $t$  test,  $p < .05$ , are shown in **bold**.

Accuracy and working memory parameter values in experiments with a single verbal probe with a serial position cued, or a visual probe presented in a target location.

Table 2

Condition	Unimodal			Bimodal			Parameters				
	Words( $h_vfa$ )	Objects( $h_vfa$ )	Objects( $h_vfa$ )	Words( $h_vfa$ )	Objects( $h_vfa$ )	Objects( $h_vfa$ )	C	$P_{verb}$	$P_{obj}$		
Experiment 2a											
Same Category Within a Spoken List											
M	0.93	0.07	0.91	0.17	0.86	0.11	0.87	0.31	<b>1.15</b>	<b>2.29</b>	<b>1.81</b>
SEM	0.02	0.02	0.02	0.04	0.03	0.03	0.02	0.05	0.30	0.29	0.29
Different Categories within a Spoken List											
M	0.93	0.13	0.92	0.20	0.92	0.09	0.81	0.28	<b>0.67</b>	<b>2.54</b>	<b>2.23</b>
SEM	0.02	0.03	0.02	0.04	0.02	0.02	0.04	0.04	0.29	0.28	0.29
Experiment 2b											
Same Voice and Category Within a Spoken List											
M	0.90	0.09	0.88	0.23	0.91	0.09	0.83	0.31	0.46	<b>2.79</b>	<b>2.11</b>
SEM	0.03	0.02	0.04	0.04	0.03	0.04	0.05	0.05	0.30	0.35	0.38
Different Voices and Categories Within a Spoken List											
M	0.90	0.11	0.90	0.23	0.91	0.09	0.88	0.27	0.14	<b>3.04</b>	<b>2.54</b>
SEM	0.03	0.04	0.04	0.04	0.02	0.02	0.03	0.06	0.29	0.26	0.32
Experiment 2c											
Same Category Within a Printed List											
M	0.93	0.12	0.94	0.18	0.80	0.17	0.89	0.23	<b>1.13</b>	<b>2.12</b>	<b>1.91</b>
SEM	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.32	0.29	0.29
Different Categories Within a Printed List											
M	0.90	0.14	0.92	0.19	0.81	0.16	0.88	0.27	<b>0.94</b>	<b>2.13</b>	<b>1.98</b>
SEM	0.03	0.04	0.02	0.04	0.05	0.03	0.03	0.04	0.37	0.33	0.32

Note. Parameter values that are above zero by  $t$  test,  $p < .05$ , are shown in **bold**.

Accuracy and working memory parameter values in Experiments 3a, 3b, and 4, with a single acoustic probe with no position cue or a visual probe presented at the center of the screen.

Table 3

Condition	Unimodal			Bimodal			Parameters				
	Digits( $h_sfa$ )	Objects( $h_sfa$ )	Objects( $h_cfa$ )	Digits( $h_sfa$ )	Objects( $h_sfa$ )	Objects( $h_cfa$ )	C	$P_{verb}$	$P_{obj}$		
Experiment 3a (Changes in Digit-Position or Color-Location Binding)											
No Mask											
M	0.72	0.13	0.85	0.24	0.67	0.17	0.74	0.37	<b>0.73</b>	<b>0.73</b>	<b>0.96</b>
SEM	0.03	0.02	0.03	0.04	0.04	0.03	0.05	0.19	0.15	0.15	0.20
Experiment 3b (Changes in Digit-Position or Color-Location Binding)											
Same-Modality Mask											
M	0.63	0.22	0.75	0.23	0.64	0.26	0.65	0.44	<b>0.61</b>	<b>0.52</b>	<b>0.87</b>
SEM	0.03	0.02	0.03	0.02	0.03	0.03	0.04	0.03	0.18	0.17	0.16
No Mask											
M	0.70	0.17	0.77	0.23	0.69	0.19	0.64	0.39	<b>0.48</b>	<b>0.82</b>	<b>1.09</b>
SEM	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.22	0.20	0.20
Different-Modality Mask											
M	0.73	0.17	0.75	0.24	0.63	0.22	0.64	0.39	<b>0.69</b>	<b>0.78</b>	<b>0.75</b>
SEM	0.03	0.02	0.03	0.03	0.03	0.03	0.04	0.03	0.20	0.17	0.17
Experiment 4 (Changes to a New Item)											
Same-Modality Mask											
M	0.96	0.08	0.94	0.15	0.91	0.15	0.88	0.23	<b>0.80</b>	<b>2.87</b>	<b>2.51</b>
SEM	0.01	0.02	0.02	0.03	0.02	0.03	0.03	0.03	0.25	0.23	0.24
No Mask											
M	0.98	0.06	0.95	0.17	0.95	0.12	0.85	0.26	<b>0.68</b>	<b>3.08</b>	<b>2.60</b>
SEM	0.01	0.01	0.01	0.04	0.01	0.03	0.03	0.03	0.25	0.24	0.25
Different-Modality Mask											
M	0.95	0.07	0.91	0.12	0.97	0.13	0.86	0.21	<b>0.78</b>	<b>2.94</b>	<b>2.69</b>
SEM	0.01	0.02	0.02	0.02	0.01	0.03	0.03	0.03	0.25	0.23	0.22

Note. Parameter values that are above zero by  $t$  test,  $p < .05$ , are shown in **bold**.