

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/176208/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Cowan, Nelson and Guitard, Dominic 2025. Similar working memory outcomes with successive versus concurrent presentation of tones and colors. *Attention, Perception, & Psychophysics* 87 , pp. 884-898. 10.3758/s13414-025-03036-3

Publishers page: <https://doi.org/10.3758/s13414-025-03036-3>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



**Similar Working Memory Outcomes  
with Successive versus Concurrent Presentation of Tones and Colors**

Nelson Cowan

University of Missouri

Dominic Guitard

Cardiff University

Authors' Note

This research was supported by NIH Grant R01-HD21338 to Nelson Cowan. We thank Bret Glass for assistance. E-mail correspondence concerning this article should be addressed to Nelson Cowan at [cowann@missouri.edu](mailto:cowann@missouri.edu) or Dominic Guitard at [guitardd@cardiff.ac.uk](mailto:guitardd@cardiff.ac.uk).

Open Practices Statement

The computer program, data, and analyses are available on the Open Science Framework project page, <https://osf.io/5nb7e>. The research was not preregistered.

### Abstract

Performance in a short-term or working memory task may depend on the ability to construct and retain a representation of multiple items fitting a single, coherent scheme. Judging by a theoretical tradition based on selective attention tasks, there might be encoding of perceived items into working memory limited to a single coherent stream of information at a time. However, this notion of a single stream seems to conflict with another tradition in which several items can be absorbed into working memory and retained at once. In three large-scale experiments (N~120 per experiment), we examined working memory for series of colors and tones that were presented either successively, with all of one modality before the other, or concurrently, with one tone accompanying each color but uncorrelated with it. The single-stream account would suggest considerable difficulty for the concurrent presentation method. However, no such effect was obtained. There was no effect of the presentation method in the recognition of items (Experiments 1 and 2) and only a small modifying effect in memory for the serial position of items (Experiment 3). All experiments replicated earlier findings of a detrimental effect of having to retain two sets compared to one, and of an asymmetry in which the dual-set effect was larger for the visual modality. The extant evidence supports a theory in which multiple, separate episodic streams can be encoded into working memory concurrently, provided that the rate of information input does not exceed a rate of processing.

*Data availability:* The computer program, data, and analysis codes are available on the Open Science Framework project page, <https://osf.io/5nb7e>.

*Keywords:* Attention, Working Memory. Bimodal Memory, Audiovisual Attention, Encoding versus Maintenance

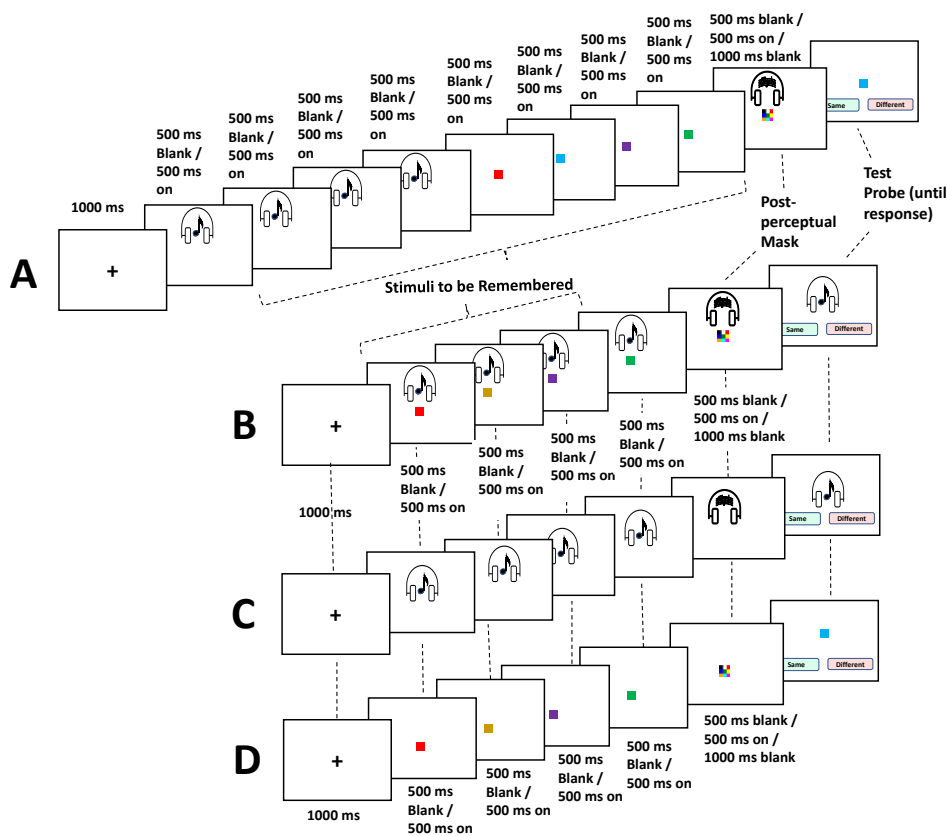
### **Similar Working Memory Outcomes with Successive versus Concurrent Presentation of Tones and Colors**

There are two often-repeated principles in cognitive psychology that at first glance appear at odds with one another, and we aim to address this apparent contradiction. First, there is a maxim that it is possible to attend to only one stream of information at a time (e.g., Broadbent, 1956, 1958; Cowan, 2012; Jones et al., 1996; MacDermid et al., 2023; Treisman, 1969; Wolfe & Robertson, 2012). Second, in apparent contrast, working memory can include multiple items at the same time (e.g., Adam et al., 2017; Cowan, 2001; Luck & Vogel, 1997; Miller, 1956). The question then arises as to the degree to which working memory's multiple-item capacity can compensate for attention's single-stream capacity. Presentation of two items at once, one visual and one acoustic, might be distributed into working memory in such a way that neither attention nor working memory are overwhelmed. Cowan and Guitard (2024) showed this in a developmental investigation of item memory, and here we replicate the finding, examine it in more detail than Cowan et al., and extend it to memory for serial order.

Cowan and Guitard (2024) examined memory in children in the elementary school years and in adults, using series of tones and series of colored squares presented at a rate of one event per second. In the *successive* condition, participants received four stimuli from one modality and then four stimuli from the other modality, always followed by a post-perceptual mask to cut short any lingering sensory memory. In the novel *concurrent* condition, participants received pairs of stimuli, with one tone concurrent with every color, until four pairs were presented, with a post-perceptual mask following that to interrupt any lingering sensory memory (Figure 1). In either case, recognition was tested with a single probe stimulus that was to be judged present in, or absent from, the stimuli that were to be remembered on the trial. Within a trial block attention

was to be paid to colors only, to tones only, or to both tones and colors, with the recognition test always occurring in an attended modality. Remarkably, the pattern of recognition (including single-task performance and dual-task costs), was nearly identical in the successive and concurrent conditions, in all age groups. Working memory appeared capable of absorbing the stimuli presented two at a time as well as one at a time, demonstrating that the capacity limit of working memory was not in the encoding phase but in the maintenance phase of the tasks. This finding is notable inasmuch as concurrent presentation might require a central attention system to shift back and forth between audition and vision to accommodate the concurrent stimuli (cf. Shomstein & Yantis, 2004).

Figure 1

*Conditions in Cowan and Guitard (2024)*

**Note.** Reproduced with permission (to be obtained) from Cowan and Guitard (2024), Figure 1, p. 4. **A**, successive presentation condition; either modality could come first and the auditory-visual order is illustrated here. **B**, concurrent presentation condition; **C**, auditory-only condition; **D**, visual-only condition. In the dual-modality conditions (A & B), attention was directed toward auditory stimuli only, visual stimuli only, or both modalities, depending on the trial block, and the probe item was in an attended modality. The task was to judge whether the probe item was present in the studied material.

Although the Cowan and Guitard (2024) result is remarkable in the degree of similarity between successive and concurrent methods of presentation of two sets to be retained, there are important uncertainties remaining. Most importantly, there was no test of memory for serial order. Guitard et al. (2022) and Guitard and Cowan (2023a, 2023b) have found that the retention of order information in working memory is more attention-demanding than the retention of item information, and Oberauer (2019) suggested that the capacity limit in working memory is based

on the maintenance of binding between items and their context (but that context can include the trial in which the item occurred: see Cowan, 2022). We know relatively little about the role of encoding in the limitations in the number of bindings that can be held in working memory, memory, such as the binding of items to their serial positions. If it takes more attention to encode order or position compared to encoding items, then an order or position test might show inferior retention of items from two modalities and more interference between them with concurrent as compared to successive presentation of items in two modalities. In contrast, if the limit in retention occurs only in the maintenance phase (Guitard & Cowan, 2023a), then there should be no difference between concurrent versus successive presentation. We examined this question.

We also improved on the methods of Cowan and Guitard (2024) in other ways. They included only 27 adults but a larger adult sample size is needed to examine serial position effects, an important benchmark of working memory tasks (Hurlstone, 2024; Oberauer et al., 2018). Different serial position effects for successive and concurrent presentation methods would indicate that they favor the entry of different items into working memory, even if the levels of overall performance do not differ. We replicated the item memory task in two slightly different experiments with 240 adults in all, and we carried out a third experiment (n=120) in which the test was one of serial position placement instead of simple item recognition.

### General Methods

In all three experiments, participants were either instructed to allocate their attention only to visual information, only to auditory information or both to visual and auditory information. In a *successive* presentation method, illustrated in Figure 1A, four visual items were presented sequentially before the sequential presentation of the four auditory items, or vice versa (e.g., *v,v,v,v,a,a,a,a* or *a,a,a,a,v,v,v,v*; where *v* corresponds to visual items and *a* corresponds to auditory items). In a *concurrent* method, illustrated in Figure 1B, we used a presentation in which

each visual item onset was simultaneous with one auditory item onset (which can be described as  $v\&a, v\&a, v\&a, v\&a$ ). Experiments 2 and 3 also included one-set trials, as illustrated in Figure 1C and 1D.

The methods of presentation and testing differed for different groups in all three experiments, as shown in Table 1. In each experiment, one group received a successive presentation of the two sets to be remembered, whereas the other group received a concurrent presentation method. Experiments 1 and 2 examined memory for items as in the developmental study of Cowan and Guitard (2024), whereas Experiment 3 examined memory for serial position. Experiments 1 and 2 differed slightly in that Experiment 2 included trials in which only one set of four items (in one modality) was presented, like Cowan and Guitard, whereas these trials were omitted in Experiment 1 to determine whether they altered the task set for the dual-modality trials.

Table 1

*Methods in Six Experiments*

| Experiment | Presentation | Type of Test | Include One-Set Trials? |
|------------|--------------|--------------|-------------------------|
| 1a         | Successive   | Item         | No                      |
| 1b         | Concurrent   | Item         | No                      |
| 2a         | Successive   | Item         | Yes                     |
| 2b         | Concurrent   | Item         | Yes                     |
| 3a         | Successive   | Position     | Yes                     |
| 3b         | Concurrent   | Position     | Yes                     |

**Note.** N=60 per experiment.

## Materials

All materials were displayed on a uniform gray background (RGB = 192-192-192) and the experiment was programmed with PsyToolKit (Stoet, 2010, 2017). The material of this experiment was based on the materials of the second experiment of Cowan et al. (2018).

Specifically, the visual stimuli were color squares which we randomly sample without replacement from a pool of 9 possible colors presented at the center of the computer screen (black: RGB 0-0-0, white: RGB 255-255-255, red: RGB = 255-0-0, blue: RGB = 0-0-255, green: RGB = 0-128-0, yellow: RGB = 255-255-0, orange: RGB = 255-165-0, cyan: RGB: 0-255-255, magenta: RGB = 255-0-255). The visual mask was a multicolor square identical in size to the visual to-be-remembered items composed of all nine colors displayed at the center of the computer screen for 500 ms.

The auditory items were identical to those of Experiment 2 of Cowan et al. (2018). Precisely, the 9 discriminable tones developed by Li et al. (2013, Experiment 3) and truncated to the duration of 125 ms by Cowan et al. were used (Trumpet Section at 200 Hz; Smooth Clav at 262 Hz, Classic Rock Organ at 343 Hz, Negril Bass at 450 Hz, Tenor Sax at 589 Hz, Space Harpsichord at 772 Hz, Grand Piano at 1011 Hz, Live Pop Horns at 1324 Hz, Aurora Bell at 1735 Hz.) The auditory mask, like the visual mask, was presented for 500 ms and corresponds to a multi-tone composed of all nine tones presented sequentially (9 tones/500 ms). For additional details regarding the creation of the auditory tone see Li et al. (2013, Experiment 3) and Cowan et al. (2018, Experiment 2).

## **Design**

Each experiment was divided into three blocks, each corresponding to an attention condition. Participants were instructed to allocate their attention only to visual information (visual block) in one block, only to auditory information (auditory block) in a different block or both to visual and auditory information (bimodal block). We counterbalanced the order of the block across participants using a Latin square design.

## **Analyses**

**Overview.** Formulas were used to estimate the number of items in working memory taking into account guessing, under the simplifying assumption that an item is either in working memory sufficiently for that knowledge to be used, or out of working memory, in which case the participant is only able to guess randomly. Given that the pattern of responses for the successive presentation of visual and auditory materials has been examined already in previous work (Cowan et al., 2013, 2014, 2018; Cowan & Guitard, 2024), here what is of interest is a comparison of comparable conditions in the successive versus concurrent presentation conditions. These comparisons are made for each condition using Bayesian Mann-Whitney U tests (JASP Team, 2024), non-parametric tests that eliminate any concern about the actual distribution of data in each condition. The Bayesian Mann-Whitney U was also applied to proportion correct separately for each serial position in each condition to determine if the patterns differ between successive and concurrent conditions, and was applied to separate conditions equated across successive and concurrent presentation types for post-presentation delay. Bayesian analyses allow the possibility of providing evidence favoring the null hypothesis above the alternative. The sample size,  $n=60$  per experiment, is comparable to our previous work on item and order memory (Guitard et al., 2022; Guitard & Cowan, 2023a, 2023b) and allows  $N=120$  with comparable conditions except for the successive versus concurrent schedule of presentation, for a precise examination of that factor in each of three experiments.

**Models of items in working memory.** The estimation of items in working memory depends on the procedure used (Cowan et al., 2013). For item information (Experiments 1 and 2), the estimation is according to Model 2 in the Appendix of Cowan et al. A single, central probe must be compared to all list items and if no match is found, the participant has indecisive information and must guess. The guessing rate can change according to the set size and, for the current set size of 4 items, it can be assumed that the participant would be biased toward the

assumption that if no match was found, most likely none was present. Regardless of the guessing bias, this situation produces an estimation model in which, for our set size of 4,

$$k=4*\frac{p(\text{match correctly proclaimed})+p(\text{no-match correctly proclaimed})}{p(\text{no-match correctly proclaimed})}$$

where  $k$  is the estimated number of items in working memory. The only exception to this rule is that a restriction of  $k \geq 0$  was imposed. The value of  $k$  was estimated separately in the two-set trials for (1) bimodal attention trials with an auditory, tone probe, (2) bimodal attention trials with a visual, color probe, (3) unimodal, auditory-attention trials, and (4) unimodal, visual-attention trials.

For serial position information, in particular when the participant is asked to locate the probe among 4 serial positions (Experiment 3), the estimation of  $k$  can be determined as follows. The participant knows the correct answer on  $k/4$  of the trials and if it is not known (on the other  $1 - k/4$  of the trials), the participant must guess among the  $4 - k$  remaining choices. This situation results in

$$k=(4*\text{proportion correct})-1$$

where  $k$  is the number of item - serial position associations known. Once more, a restriction of  $k \geq 0$  was imposed. If  $k=3$  then the correct answer will be known on  $3/4$  of the trials and guessing on the other  $1/4$  of trials will occur out of only a single remaining location, resulting in perfect performance; yet people rarely achieved  $k=3$  for position information.

The  $k$  values in each experiment were used to estimate central and peripheral components of memory according to the method of Cowan et al. (2014). The central component reflects the

number of items that are recalled using a resource shared between auditory and visual presentations and is expressed as

$$\text{Central component} = k_{\text{unimodal auditory}} + k_{\text{unimodal visual}} - k_{\text{bimodal auditory}} - k_{\text{bimodal visual}}$$

where  $k_{\text{unimodal auditory}}$  refers to the number of items in working memory in the condition with unimodal attention and an auditory probe, and so on. When subtracting the bimodal condition capacities from the unimodal condition capacities, the difference is the resource that they share. To understand why, think of unimodal auditory and unimodal visual capacities as the areas of two circles, and think of bimodal trial capacity as represented by the area of these circles overlapping, with the shared or central resource reflected in the overlap. For example, suppose a participant remembered 2.3 auditory and 3.1 visual items with unimodal attention, for a total of 5.4 items across modalities based on separate trials for each modality held separately. Suppose also that the participant remembered 2.0 auditory and 2.5 visual items with bimodal attention, for a total of 4.5 items across modalities when both modalities must be held in working memory concurrently. Then the estimate of the central portion is  $5.4 - 4.5 = 0.9$  items.

The peripheral auditory and visual capacities, the portions that for any reason are not influenced by sharing attention between modalities, can be easily estimated by subtracting the central portion from the unimodal estimates:

$$\text{Peripheral auditory} = k_{\text{unimodal auditory}} - \text{Central component}$$

$$\text{Peripheral visual} = k_{\text{unimodal visual}} - \text{Central component}$$

In the hypothetical example above, the peripheral auditory capacity would be  $2.3 - 0.9 = 1.4$  item and the peripheral visual capacity would be  $3.1 - 0.9 = 2.2$  item.

## Experiment 1

### Method

#### *Successive Presentation Group*

In the successive presentation group in the first experiment, all four visual items were sequentially presented before the four auditory items for half of the trials while, for the other half of the trials, all auditory items were sequentially presented before the visual items (see Figure 1A).

**Participants.** Sixty participants were recruited from the online data collection agency, Prolific (<https://www.prolific.com/>) and were paid \$15.00 per hour (pro-rated) for their participation. Two participants were excluded and replaced for not following the instructions properly. All volunteers had to meet the following eligible criteria to participate in this study: participants were (1) native speaker of English, (2) British, American, or Canadian nationality, (3) had normal or corrected-to-normal vision, (4) had no cognitive impairment or dementia, (5) had no language-related disorders, (6) aged between 18 and 30 years old, and (7) had an approval rating of at least 90% on prior submissions at Prolific. The mean age of the participants was 25.28 ( $SD = 3.22$ , range 18–30); 40 self-identified as female, 18 as male, and 2 preferred not to specify their gender. This and all subsequent experiments were approved by the research ethics committee of the University of Missouri.

**Design.** Three subgroups of 20 participants each were included, who received the three attention blocks (Visual=V, Auditory=A, Bimodal=B) in different Latin square orders: VAB, ABV, and BVA. For both the visual and auditory unimodal attention blocks, participants completed 4 practice trials followed by 48 experimental trials. For the bimodal attention block,

participants completed 8 practice trials (4 visual trials, 4 auditory trials) followed by 96 trials (48 visual trials, 48 auditory trials). In the bimodal block, the modality of the test (visual or auditory) was randomized; therefore, participants had to be prepared for either kind of test. As mentioned, for half of the trials the visual items were presented before the auditory items while for the other half it was the opposite. For each block, there was an equal number of trials in which the test probe was a previously presented item (same probe) and an item not presented (different probe). Furthermore, each item position (1 to 4) was tested equally often in a given block. Lastly, the order of the trials conditions within a block was randomized for each participant.

**Procedure.** All participants took part in one online experimental session lasting approximately 50 minutes. For all blocks, each trial progression was controlled by the participants, who initiated each trial by clicking on the screen. Before each trial, an encoding instruction was presented (remember only the colored squares, remember only the tones, remember the colored squares and the tones) with an image (eye for visual items, hear for auditory items, eye and hear for both visual and auditory items). Once participants initiated the trial, a trial began with the presentation of a fixation crossed (+) at the center of the computer screen for 1500 ms (1000 ms on, 500 ms off). After the presentation of the fixation crossed, 8 items (4 visual items, 4 auditory items) were presented sequentially at a rate of one item every 500 ms (250 ms on, 250 ms off). The visual items were presented at the center of the computer screen and the auditory items were presented via participants' personal audio system (headphone or speaker). 500 ms after the presentation of the last item the visual and auditory masks were presented simultaneously (500 ms on, 1000 ms off). Following the presentation of the masks, a probe was presented at the center of the computer screen. Participants had to identify if the probe was the same as a just presented item by clicking on the "same" green button or different from the just presented item by clicking on the "different" red button displayed below the probe. For

visual probe the item was displayed at the center of the computer screen while for auditory probe a question mark was displayed, and the probe was presented via participants' auditory system. After each response participants received feedback, smiley face with the word "correct" for correct response and a frowny face with the word "incorrect" for incorrect responses displayed for 1000 ms.

After participants completed the experimental session, they were asked a series of questions to better understand the strategy they used to remember the colors and the tones. They were two open-ended questions and twelve strategy questions that they had to respond via the following Likert scale: All the time, most of the time, some of the time, occasionally, never. The series of questions are presented in Appendix A.

### ***Concurrent Presentation Group***

For this group, the same method was used as for the successive presentation group except that the four visual and the four auditory items on a trial were presented using the concurrent presentation method shown in Figure 1B ( $v \& a$ ,  $v \& a$ ,  $v \& a$ ,  $v \& a$ , where  $v$  corresponds to visual items and  $a$  corresponds to auditory items).

**Participants.** Sixty different volunteers who did not take part in the successive presentation group were recruited from the online data collection agency, Prolific (<https://www.prolific.com/>). The inclusion criteria and the monetary compensation were the same as in the previous experiment and one participant was excluded and replaced for not following the instructions properly. The mean age of the participants was 23.87 ( $SD = 3.77$ , range 18–30); 37 self-identified as female, 22 as male, and 1 preferred not to specify their gender.

**Materials, design, and procedure.** The materials, the design, and the procedure were identical to those used for the successive presentation group except for the presentation of visual and auditory items in four bimodal pairs as shown in Figure 1B. The first visual item and the first

auditory item were presented concurrently, then the second visual item and the second auditory item were presented and so on until all four visual and auditory items were presented, at the rate shown in Figure 1B.

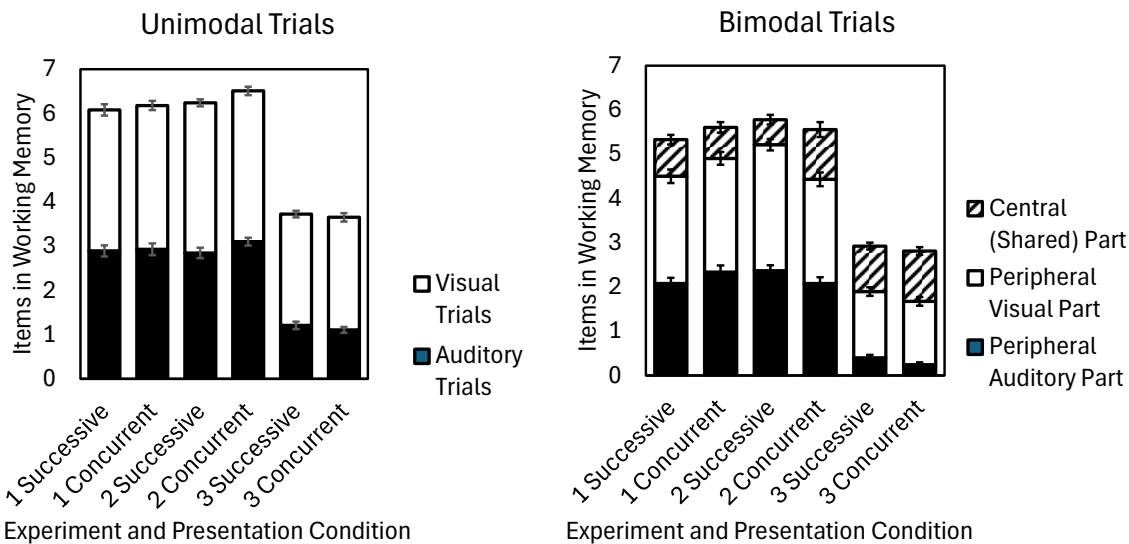
### **Experiment 1 Results and Discussion**

Figure 2 shows the performance in terms of number of items in working memory for the unimodal condition (left-hand panel) and the bimodal condition (right-hand panel). The first two bars in the left-hand panel show the Experiment 1 mean unimodal attention estimates for visual probe trials stacked on top of auditory probe trials. It is clear that the mean estimates are almost identical for the first and second bars, i.e., for the successive and concurrent presentation of the items in the two modalities.

Similarly, the first two bars on the right-hand side of Figure 2 show that the components of performance in bimodal attention trials were quite comparable for successive and concurrent presentation. The slight difference favors concurrent presentation, the opposite of what one would expect if there were a severe bottleneck for the input of information into working memory.

**Figure 2**

*For all experiments, estimates of items in working memory in each theoretical component*



**Note.** In the **left-hand panel** (for unimodal trials) the capacity for two-set, attend-one visual trials are stacked on top of the capacity for two-set, attend-one auditory trials. In the **right-hand panel** (for bimodal trials) the capacity estimates specific to auditory information, specific to visual information, and shared between the two modalities are stacked; see text for details about these estimates. 1 Successive = Experiment 1, successive presentation group; 1 Concurrent = Experiment 1, concurrent presentation group, etc. Error bars are standard errors of the mean for each section.

Table 2 shows the Bayes Factors for a comparison of presentation groups for every component in every experiment. For Experiment 1, the results for every component were less than 1.0, providing evidence suggesting no difference between components. The difference is sometimes in an indeterminate range but, along with Figure 2, it is clear that there was little if any difference between the successive and concurrent comparison groups in any component.

Table 2

*Bayes Factor for two-tailed, Mann-Whitney U tests comparing the successive and concurrent presentation groups' capacity component values in every experiment*

| <b><u>Component</u></b> | <b>Experiment</b> |                         |                 |
|-------------------------|-------------------|-------------------------|-----------------|
|                         | <b><u>1</u></b>   | <b><u>2</u></b>         | <b><u>3</u></b> |
| Unitary auditory        | 0.20              | 0.55                    | 0.28            |
| Unitary visual          | 0.21              | 0.20                    | 0.21            |
| Central                 | 0.48              | <b>4.75<sup>a</sup></b> | 0.27            |
| Peripheral visual       | 0.22              | 2.65                    | 0.24            |
| Peripheral auditory     | 0.70              | 0.34                    | 0.69            |

**Note.** Bayes Factor greater than 3.0 is shown in bold.

<sup>a</sup> Successive M=0.56, Concurrent M=1.12

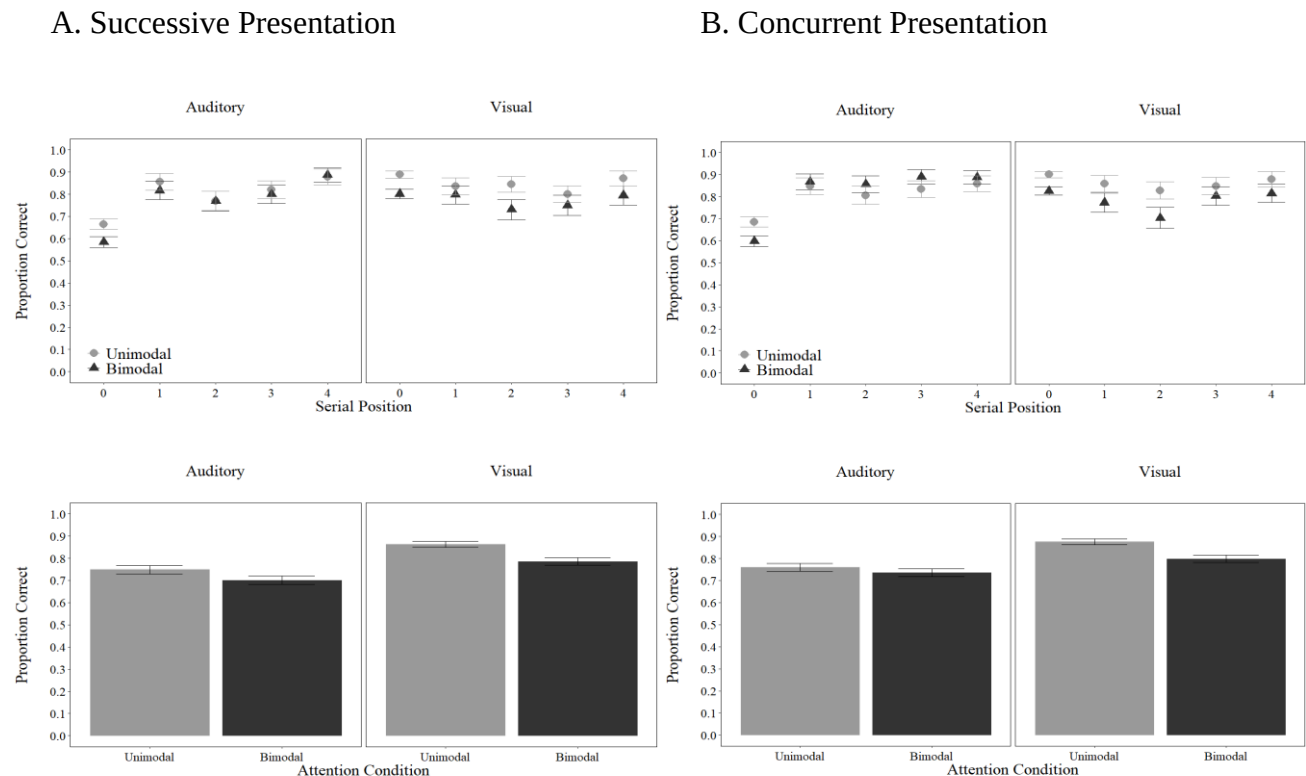
The top panels of Figure 3 shows the proportion correct in Experiment 1 according to the serial position of presentation, with Serial Position 0 indicating performance on probes that were absent from the list. In all conditions and serial positions, the pattern of performance looks quite similar for successive presentation (left-hand side of the figure) and concurrent presentation (right-hand side). The bottom panels of the figure show more starkly that the overall levels of performance and bimodal attention cost are very similar in successive and concurrent presentation groups.

Table 3 shows the Bayes Factor for a comparison of successive and concurrent presentation groups for every experiment in every condition and serial position. For Experiment 1 there was one case in which a difference may have occurred, specifically in the auditory probe for Serial Position 3 within the bimodal attention condition. In this case, the concurrent-presentation

performance ( $M=.89$ ) was higher than the successive-presentation performance ( $M=.80$ ), against what would be expected if concurrent presentation was problematic for the encoding process.

Figure 3

*For Experiment 1, performance in all conditions by the serial position of presentation*



In sum, this experiment shows that the theoretically important manipulation of the schedule of presentation of items in two modalities, either with the modalities in succession or with one auditory item concurrent with each visual item, makes little difference for performance. This pattern of results speaks clearly to the importance of storage rather than encoding as the typical limiting factor for working memory capacity.

Table 3

*Bayes Factor for two-tailed, Mann-Whitney U tests comparing the successive and concurrent presentation groups at every serial position in every experiment*

| <b><u>Condition</u></b> | <b>Experiment</b>       |                         |                         |
|-------------------------|-------------------------|-------------------------|-------------------------|
|                         | <b><u>1</u></b>         | <b><u>2</u></b>         | <b><u>3</u></b>         |
| AudBimSP0               | 0.22                    | 0.20                    | --                      |
| AudBimSP1               | 0.42                    | 0.25                    | <b>8.94<sup>c</sup></b> |
| AudBimSP2               | 1.69                    | 0.33                    | 0.23                    |
| AudBimSP3               | <b>4.17<sup>a</sup></b> | 0.24                    | 0.40                    |
| AudBimSP4               | 0.22                    | 0.25                    | 0.28                    |
| AudUniSP0               | 0.24                    | 0.28                    | --                      |
| AudUniSP1               | 0.21                    | 0.42                    | 0.55                    |
| AudUniSP2               | 0.28                    | 0.24                    | 0.20                    |
| AudUniSP3               | 0.24                    | 2.98                    | 0.88                    |
| AudUniSP4               | 0.27                    | 0.23                    | 0.36                    |
| AudOneSetSP0            | --                      | 0.23                    | --                      |
| AudOneSetSP1            | --                      | 0.72                    | 0.77                    |
| AudOneSetSP2            | --                      | 0.23                    | 0.40                    |
| AudOneSetSP3            | --                      | 0.31                    | 0.24                    |
| AudOneSetSP4            | --                      | 0.23                    | 0.21                    |
| VisBimSP0               | 0.29                    | 0.25                    | --                      |
| VisBimSP1               | 0.28                    | 0.28                    | 0.31                    |
| VisBimSP2               | 0.21                    | <b>4.80<sup>b</sup></b> | 0.21                    |
| VisBimSP3               | 0.48                    | 0.34                    | 0.33                    |
| VisBimSP4               | 0.21                    | 0.22                    | 0.21                    |
| VisUniSP0               | 0.29                    | 0.23                    | --                      |
| VisUniSP1               | 0.22                    | 0.21                    | 0.22                    |
| VisUniSP2               | 0.25                    | 0.24                    | 0.21                    |
| VisUniSP3               | 0.24                    | 0.23                    | 0.36                    |
| VisUniSP4               | 0.21                    | 0.25                    | 0.26                    |
| VisOneSetSP0            | --                      | 0.21                    | --                      |
| VisOneSetSP1            | --                      | 0.27                    | 0.23                    |
| VisOneSetSP2            | --                      | 0.23                    | 0.22                    |
| VisOneSetSP3            | --                      | 0.29                    | 0.26                    |
| VisOneSetSP4            | --                      | 0.25                    | 0.22                    |

---

**Note.** Aud=auditory probe, Vis=visual probe, Bim=bimodal attention, Uni=unimodal attention with two sets presented, OneSet=one set presented, SP=serial position. Dashes represent instances in which a condition was not included in an experiment. SP0 refers to probes that were absent from the list in the item-recognition experiments. Bayes Factors larger than 3.0 are shown in bold.

<sup>a</sup> Successive M=0.80, Concurrent M=0.89.

<sup>b</sup> Successive M=0.81, Concurrent M=0.68.

<sup>c</sup> Successive M=0.50, Concurrent M=0.37.

Within the successive presentation group, performance was slightly higher for the modality presented second compared to the modality presented first (Supplementary Figure S1). Apparently, however, the average across first and second presentations was similar to the average within the concurrent presentation group.

A final set of analyses of the proportion correct were carried out in order to compare only the second-presented set in the successive presentation condition to concurrent presentation. For these two types of testing, there was no intervening other set to delay testing. Mann-Whitney U tests showed no reliable difference between presentation types for the situation in which memory was bimodal and testing was auditory (successive Set 2, M=.76; concurrent, M=.74, BF=0.27), memory was unimodal and testing was auditory (M=.77 and .76, respectively; BF=0.22), memory was bimodal and testing was visual (M=.80 and .80, respectively; BF=0.21), or memory was unimodal and testing was visual (M=.87 and .88, respectively; BF=0.38).

Last, the open-ended questionnaire at the end of the experiment (see Appendix A) yielded a variety of responses regarding strategies participants used. For the 12 questions about strategies based on a Likert scale, Bayesian Mann-Whitney U tests showed no difference between presentation groups, with BF < 0.53 for every question.

## **Experiment 2**

The second experiment examines whether this pattern observed in Experiment 1 could be changed by including some trials in which only one modality is presented. This condition matches the method in the developmental investigation by Cowan and Guitard (2024) and theoretically could allow the refinement of a more modality-specific retention strategy than was observed in Experiment 1.

## Method

In the second experiment, we replicated and extend the procedure used in the first experiment. Specifically, for both the successive modality presentation group and the concurrent modality presentation group, trials were added in which only one kind of to-be-remembered material was presented (visual or auditory).

### *Successive Presentation Group*

**Participants.** Another group of sixty volunteers who did not take part in the previous experiment were recruited from Prolific (<https://www.prolific.com/>). The inclusion criteria and the monetary compensation were the same as in the previous experiment and two participants were excluded in this experiment for not following the instructions. The mean age of the participants was 25.55 ( $SD = 3.60$ , range 18–30); 39 self-identified as female, 19 as male, and 2 preferred not to specify their gender.

**Materials, design, and procedure.** The materials, the design, and the procedure were identical to the successive presentation group of Experiment 1 except for the following changes. In this experiment, to ensure participants were prepared to encode the second set of visual or auditory information, a fixation cross (+) was added before the presentation of the second set of items (1000 ms on, 500 ms off). Also, as mentioned, trials were added in which only visual or only auditory information was presented. When participants had to attend only to visual item, participants completed 6 practice trials (2 trials only one set of visual information, 4 trials with both visual and auditory information) and then 64 experimental trials (32 trials only one set of visual information, 32 trials with both visual and auditory information). Participants also completed 6 practice trials followed by 64 experimental trials in the auditory block, with the same distribution of trials with one set versus two sets of information as in the visual block. The bimodal attention block included 8 practice blocks (half with an auditory probe and half with a

visual probe) and then 64 test trials (again, half with an auditory probe and half with a visual probe. All conditions were balanced as in Experiment 1.

### ***Concurrent Presentation Group***

**Participants.** Sixty volunteers who did not take part in the previous experiment were recruited from Prolific (<https://www.prolific.com/>). The inclusion criteria and the monetary compensation were the same as in the previous experiment and no participant was excluded in this experiment. The mean age of the participants was 24.93 ( $SD = 3.30$ , range 19–30); 34 self-identified as female, 23 as male, and 3 preferred not to specify their gender.

**Materials, design, and procedure.** The method was identical to the successive presentation group in this experiment except that a concurrent method of presentation was used on two-set trials. (See Figure 1B.)

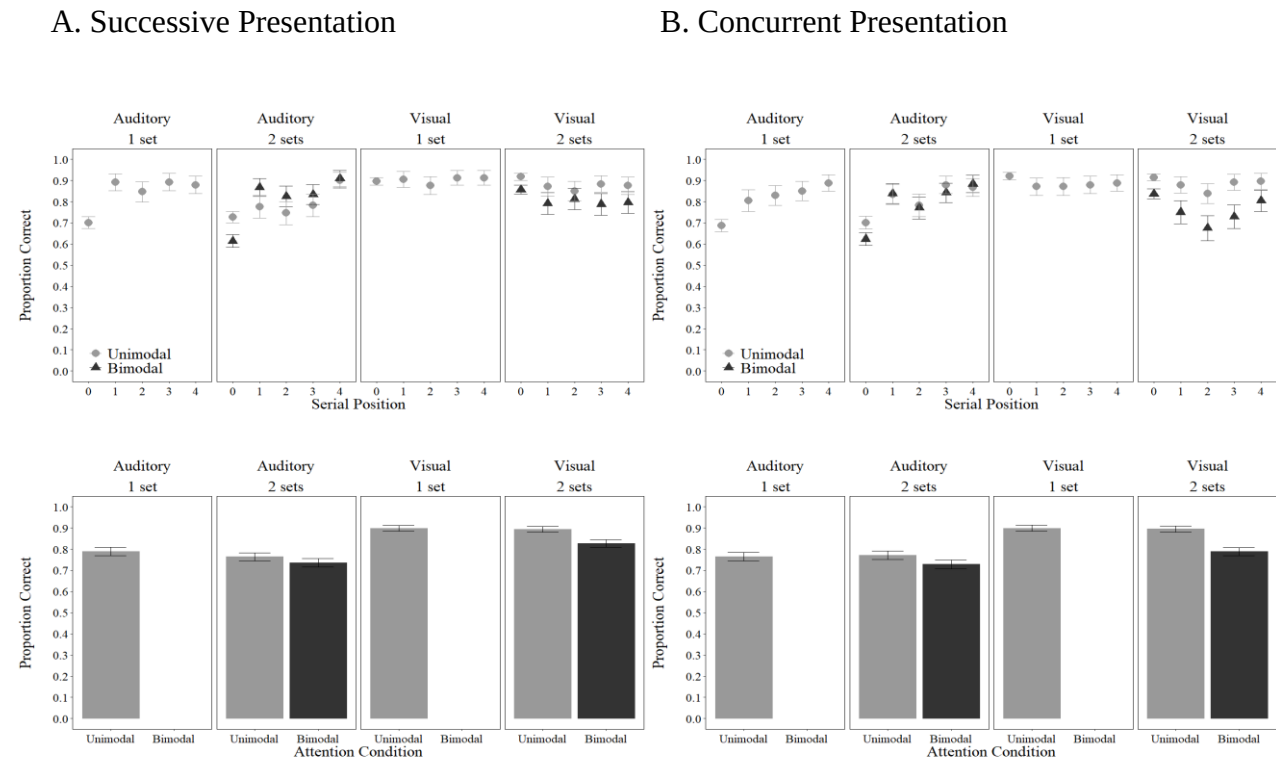
### **Experiment 2 Results and Discussion**

The middle two bars on the left side of Figure 2 show that once more, there was not much difference between successive and concurrent presentations in the unimodal attention condition, and that any slight advantage was for the concurrent presentation condition. The middle two bars on the right side of Figure 2 show that in the bimodal attention condition, there might have been a slight advantage for the successive presentation condition overall.

Figure 2 and Table 2 show that in this experiment, there was a reliable difference between presentation conditions for the central component, which was higher in the concurrent presentation condition than in the successive presentation condition despite the bimodal performance overall being slightly lower in the concurrent presentation condition, though not to a reliable extent. An examination of the data by serial positions helps to explain why this difference may have occurred.

Figure 4

*For Experiment 2, performance in all conditions by the serial position of presentation*



**Note.** Error bars are standard errors of the mean. Serial Position 0 refers to probes that were absent from the list.

Once again, as in Experiment 1, the results examined by serial position, shown in Figure 4, suggests that there is little difference between successive and concurrent presentation methods. Figure 4 and Table 3 show, however, that there was a reliable disadvantage for concurrent presentation in the bimodal attention condition for visual probes at Serial Position 2. This one local difference could indicate to a larger need for attention for good performance with concurrent presentation in this one situation, helping to explain why the central component was reliably larger in the concurrent condition despite overall performance being slightly lower in that condition. This is the kind of effect we expected when we began this study but did not expect that it would emerge only in this limited situation.

Although not the main issue under investigation, Figure 4 shows as well that there was almost no difference between the one-set and two-set trials within unimodal attention sets; the presence of stimuli in the modality to be ignored had little or no influence on the outcome. This can be seen most easily on the bottom panel of the figure, which is averaged across serial positions.

As in Experiment 1, within the successive presentation group, performance was slightly higher for the modality presented second compared to the modality presented first (Supplementary Figure S2). Apparently, however, the average across first and second presentations was again similar to the average within the concurrent presentation group.

As in Experiment 1, final set of analyses of the proportion correct were carried out on trials with two sets of materials in order to compare only the second-presented set in the successive presentation condition to concurrent presentation. For these two types of testing, there was no intervening other set to delay testing. Mann-Whitney U tests showed a small, reliable difference between presentation types for the situation in which memory was bimodal and testing was auditory (successive Set 2,  $M=.79$ ; concurrent,  $M=.73$ ,  $BF=3.16$ ). The effect was not quite reliable when memory was unimodal and testing was auditory ( $M=.81$  and  $.77$ , respectively;  $BF=2.73$ ), or when memory was bimodal and testing was visual ( $M=.85$  and  $.79$ , respectively;  $BF=2.76$ ). It was unreliable when memory was unimodal and testing was visual ( $M=.92$  and  $.90$ , respectively;  $BF=0.76$ ).

Last, the post-experimental questionnaire (see Appendix A) showed no reliable difference between presentation groups for most conditions. The only exception occurred for the question, “I tried to find patterns among the tones.” This question produced a mean response of 3.58 ( $SEM=0.17$ ) in the successive presentation group, between “some of the time” and “occasionally,” whereas it produced a mean response of 2.78 ( $SEM=0.17$ ) in the concurrent

presentation group, between “most of the time” and “some of the time,”  $BF=7.79$ . It seems possible that the presentation of each modality alone in some conditions allowed participants enough resources to develop such a memorization strategy for the tones.

### **Experiment 3**

In the third experiment, we replicated and extended the procedure used in the second experiment to position memory. Participants were tested on the serial position of a probe item in the list for both the successive presentation group and the concurrent presentation group. This experiment allows us to investigate whether item and order information maintenance are based on common process or distinct processes. Cowan et al. (2013, 2014) already showed that memory capacity for binding information (e.g., binding a spoken item to a voice or a color to a shape) was considerably smaller than the capacity for items per se. We expected that binding items to serial positions would similarly decrease the capacity compared to the first two experiments. It is also reasonable to expect that the process of binding items to serial positions would be difficult if it had to be carried out for two modalities at once, as in the concurrent presentation. However, to anticipate, that was not the finding.

### **Method**

The method was the same as in Experiment 2 except that the probe item was always present in the list and a position (1-4) response was required instead of an item recognition (present-absent) response.

#### ***Successive Presentation Group***

This experiment was like the successive presentation group of Experiment 2, in that all visual items were presented before all auditory information or vice-versa. However, in this experiment, position memory was tested instead of item memory.

**Participants.** An additional group of sixty volunteers who did not take part in any of the previous experiments were recruited from Prolific (<https://www.prolific.com/>). The same inclusion criteria and monetary compensation previously described were used in this experiment. One participant was excluded in this experiment for not following the instructions. The mean age of the participants was 25.13 ( $SD = 3.59$ , range 18–30); 32 self-identified as female, 25 as male, and 3 preferred not to specify their gender.

**Materials, design, and procedure.** The materials, the design, and the procedure were like the successive group of Experiment 2 except for the following changes. In this experiment, after the presentation of the lists to be remembered, one of the presented items (a color square or tone) was presented. When the probe item was a tone, a question mark on the screen accompanied it. Above the color square probe or question mark was the instruction “Click on the position where the item was presented.” Below it, the four possible serial position answers were displayed within a 2 x 2 matrix (first row: 1, 2; second row: 3, 4). For both visual and auditory information, participants had to click on the position of where the item was presented within the list of the appropriate modality. Each position (1 to 4) was tested equally often in the experimental trials and the position tested was randomized within each block for each participant.

The number of practice trials were reduced to accommodate additional instructions for position memory. Specifically, there were 3 practice trial before the attend-visual and attend-auditory trial block sets and 4 practice trials before the attend-both block. The remaining details were as described for the successive group of Experiment 2.

### ***Concurrent Presentation Group***

**Participants.** An additional sixty participants who did not take part in any of the previous experiment were recruited from Prolific (<https://www.prolific.com/>) with the same criteria as before. In this experiment two participants were excluded for not following the instruction (e.g.,

not providing any responses). The mean age of the participants was 24.73 ( $SD = 3.67$ , range 18–30); 33 self-identified as female, 24 as male, and 3 preferred not to specify their gender.

**Materials, design, and procedure.** The materials, design, and procedure of this experiment were identical to the Successive Presentation Group except that two-set visual and auditory lists were presented according to the concurrent schedule ( $a\&v$ ,  $a\&v$ ,  $a\&v$ ,  $a\&v$  where  $a$  is an auditory item and  $v$  is a visual item; see Figure 1B).

### Experiment 3 Results and Discussion

The rightmost two bars in the left side of Figure 2 show that, indeed, there is little difference between performance in the successive versus concurrent presentation groups for unimodal attention conditions, and the rightmost two bars on the right side of the figure show the same thing for bimodal attention conditions. Table 2 verifies this finding; all of the Bayes Factors are smaller than 1.0, favoring no difference between presentation groups, and the value is less than 0.33 for four of the five components.

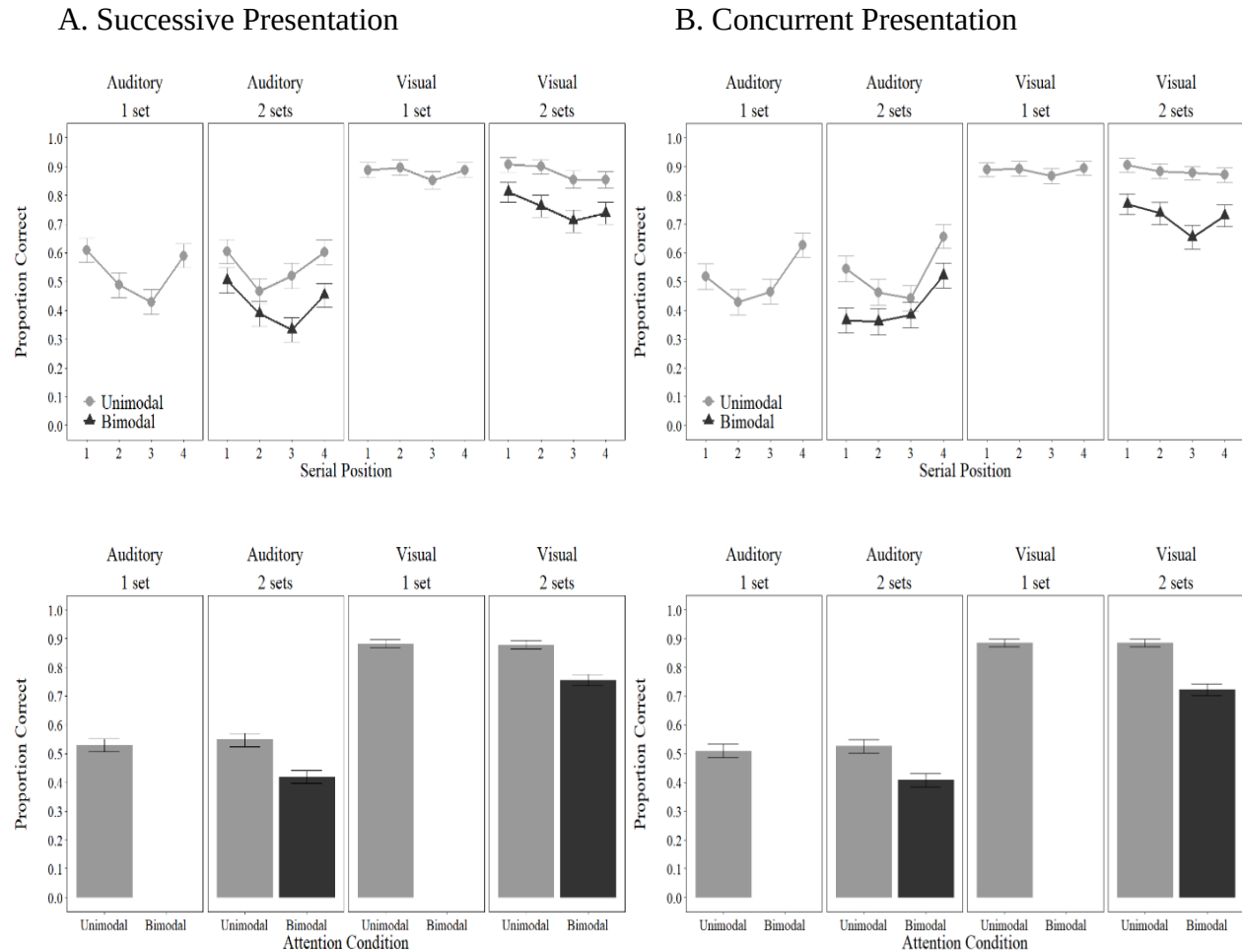
Figure 5 shows that, broken down by serial position, there is still little difference between the successive presentation group performance (left side of the figure) and the concurrent presentation group (right side of the figure). As seen in close inspection of the figure and verified in Table 3, however, there was one case in which a concurrent presentation did result in poorer performance, namely for an auditory probe in Serial Position 1 in the bimodal attention condition. Apparently, at the very beginning of a concurrent presentation when both modalities have to be attended, it is a little more difficult to tune in to the auditory channel and encode the first such item.

Although it is not the main factor under consideration, the figure also shows that the presence versus absence of a second modality within the unimodal attention conditions makes little or no difference. Notice that this was the case also in Experiment 2.



Figure 5

*For Experiment 3 (position recall experiment), performance in all conditions by the serial position of presentation*



**Note.** Error bars are standard errors.

As in the previous experiments, within the successive presentation group, performance was slightly higher for the modality presented second compared to the modality presented first (Supplementary Figure S3). Apparently, however, the average across first and second presentations was similar to the average within the concurrent presentation group.

As in the previous experiments, another analyses of the proportion correct were carried out on trials with two stimulus sets in order to compare only the second-presented set in the successive presentation condition to concurrent presentation. For these two types of testing, there

was no intervening other set to delay testing. Mann-Whitney U tests showed no reliable difference between presentation types for the situation in which memory was bimodal and testing was auditory (successive Set 2,  $M=.46$ ; concurrent,  $M=.41$ ,  $BF=0.50$ ), memory was unimodal and testing was auditory ( $M=.57$  and  $.53$ , respectively;  $BF=0.50$ ), memory was bimodal and testing was visual ( $M=.77$  and  $.72$ , respectively;  $BF=0.65$ ), or memory was unimodal and testing was visual ( $M=.90$  and  $.88$ , respectively;  $BF=0.24$ ).

Last, for this experiment on position information, the post-experimental questionnaire yielded one reliable difference between presentation groups that differed from what was found in the item-recognition experiments. Specifically, on the question, “I tried to ignore one set and just concentrate on the other, even when I was asked to remember both sets, both groups yielded a mean between occasionally and never but the response was more often “never” in the successive presentation group,  $M=3.97$  ( $SEM=0.14$ ) compared to the concurrent presentation group,  $M=3.40$  ( $SEM=0.15$ ),  $BF=3.19$ . Although the two presentation groups showed little overall difference in performance levels, this result suggests that there may have been somewhat more pressure or experienced difficulty for the concurrent presentation group.

### **General Discussion**

By many accounts, participants’ performance on working memory for items presented in series should depend on the ability to retrieve the required information (e.g., Oberauer et al., 2018), which was item information in the present Experiments 1 and 2 and serial position information in the present Experiment 3. In a developmental sample including a small sample of young adults, Cowan and Guitard (2024) previously found that successive and current presentation methods (refer to Figure 1) yielded similar outcomes in an item-recognition task. In both presentation conditions, there were comparable performance levels and comparable amounts of interference between series of colors and series of tones. The present results confirm this

surprising outcome in a larger sample of adults, with additional trials in which only one set of items is presented either excluded from the design (Experiment 1) or included in it (Experiment 2).

The aforementioned result is also extended here to a serial position recall task (Experiment 3). Similar to previous studies that examined color-shape or word-voice associations (Cowan et al., 2013, 2014), the performance levels for item-position associations (Experiment 3) were considerably lower than for recognition of individual items (in Experiments 1 & 2). Also similar to this past work, a componential analysis indicated that the difference was in peripheral auditory and visual memory. As in the past work, contributions of a central, attentional part of working memory shared between modalities was similar for item and associative information (see Figure 2).

In Experiment 2 (but not in the other experiments), performance on the second set in the successive condition was reliably higher than performance in the concurrent presentation condition. Taken together, the findings suggest that inclusion of one-set items may encourage a strategy in which participants focus on the last-presented set when only item information is needed. However, this advantage was balanced by slightly worse performance in the first-presented set, so that the method of presentation did not matter for the average result across sets in any of the experiments.

These results emphasize the importance of storage capacity in working memory and the flexibility of the process of storing or encoding information into working memory. In Experiment 3, especially, the need for serial position information means that a dual serial position coding must be used to establish the order of colors and, independently, tones. The fact that this can be accomplished in the concurrent condition, with one tone concomitant with each color, as well as it can in the successive condition, with one series followed by another, suggests that there is no

special difficulty in constructing two unrelated serial orders at once. This outcome might not be expected according to theories in which the capacity limit occurs in the process of constructing a coherent, object-oriented episodic record of the stimuli (e.g., Jones et al., 1996; MacDermid et al., 2023) or in a strict ability of attention to encode only one stream of information at a time (Broadbent, 1958; Cowan, 2012; Jones et al., 1996; MacDermid et al., 2023; Treisman, 1969; Wolfe & Robertson, 2012).

In an alternative interpretation of our results, in the concurrent presentation condition the two modalities might be joined to form conglomerate, multimodal objects that are held in working memory, though that interpretation seems more complex in requiring different mechanisms for the successive and concurrent presentation conditions despite similar results. Relevant to Experiment 3, the experiment in which we tested for serial order, models of serial order in working memory recently reviewed by Hurlstone (2024) include three potential, fundamental mechanisms of ordering (associative chaining, position marking, and a primacy gradient) and we see no reason why any of these mechanisms would be inconsistent with either the separate-streams or the multimodal-object interpretation of our concurrent-presentation results. Future modeling efforts could be directed at determining what theoretical models are consistent with our results, including our similar findings for successive and concurrent presentation of two modalities.

In our preferred embedded processes concept of working memory (Cowan, 1988; Cowan et al., 2024), items in the bimodal attention condition might be off-loaded from the focus of attention to an activated portion of long-term memory. With successive presentation, the participant must preserve one series of items, presumably by off-loading it to the activated portion of long-term memory and occasionally tending to it or refreshing it using attention (Barrouillet & Camos, 2015) while encoding the second set. With concurrent presentation, the

participant must add items on to two different series, one of colors and one of tones, when each color-tone pair arrives. Apparently, both processes are possible.

A more modular view of working memory such as the view of Baddeley and others (e.g., Baddeley et al., 2021) also is capable of understanding the current findings through a suggestion that visual and acoustic streams of information would be able to take in information separately no matter whether the information was input successively for each modality or for the two modalities concurrently. To account for the interference between these two sets, producing our central portion of working memory, a role of sharing attention-demanding processes between sets would be necessary but that is not a problem for the modular approach judging from recent work from this group on attention effects (e.g., Hitch et al., 2018).

Future research could show that the concurrent presentation method might be difficult if the two lists were in the same modality (resulting in more inter-list interference). In contrast to that possibility, though, one study suggests broadly equivalent performance and attentional demands for stimuli including inter-modal (color-shape) versus cross-modal (verbal-visual) materials (Allen et al., 2009), and another suggests comparable disruption from unimodal versus bimodal distractors (Turoman & Vergauwe, 2024). Future research might also show difficulty with the concurrent presentation method if the time between items were reduced. The near-equivalence of successive and concurrent presentation methods in Experiment 3, in which position tests were given, could be examined to determine whether items are associated with serial positions directly or with each other in a chaining process, given that both types of information are viable to contribute to performance (e.g., Logan & Cox, 2021). An advantage of item-to-position associations is that this is what participants were asked to provide, and the use of chaining would be indirect (e.g., one might retrieve a list order of colors, perhaps *red-blue-green-brown*, in order to answer that green was third in the list). On the other hand, a disadvantage of

item-to-position information is that it would suffer fan effects because each position must be associated to both a color and a tone. Another disadvantage in the case of tones is that a single tone is difficult to remember compared to the ease of remembering not the absolute pitch but the relation of tones to one another, which participants often indicated that they used in the post-test questionnaire in each experiment.

In sum, we have obtained results that might be considered surprising in that the entire pattern of item and position information encoded and maintained in working memory from sets of colors and tones is similar in two very different schedules of presentation, a successive presentation (one modality after the other) and a concurrent schedule (with one tone accompanying each color). The pattern of inter-set interference was similar to previous work (e.g., Cowan & Guitard, 2024). The study attests to the power of working memory to construct and maintain useful representations using a combination of attention-demanding and attention-free processes.

### **Declarations**

**Funding:** This research was supported by NIH Grant R01-HD21338 to Nelson Cowan.

**Conflicts of interest/Competing interests:** none.

**Ethics approval:** *The Development of Short-Term Memory for Speech Attributes* (JH-0012008-1, IRB, Project Number 2099340, IRB Review Number 400662) was reviewed and approved by the Institutional Review Board of the University of Missouri on December 20, 2023.

**Consent to participate (include appropriate statements):** The form for consent to participate was approved by the IRB and administered over Prolific electronically, with approval obtained by the participant clicking an approval button online. The form included an indication that the participant could quit at any time and it indicated the rate of payment.

**Consent for publication (include appropriate statements):** The consent form did indicate that the de-identified data would be used in publications without our obtaining further consent.

**Availability of data and materials (data transparency):** The computer program, data, and analyses are available on the Open Science Framework project page, <https://osf.io/5nb7e>.

**Code availability (software application or custom code):** The computer program, data, and analysis codes are available on the Open Science Framework project page, <https://osf.io/5nb7e>.

**Authors' contributions (optional: please review the submission guidelines from the journal whether statements are mandatory):** Both authors were involved in all phases of the research except that Dominic Guitard wrote the computer program to carry out the study and Nelson Cowan obtained funding. The first draft of the Methods section was written by Dominic Guitard and the first drafts of other sections were written by Nelson Cowan, but both authors edited all sections.

### References

- Adam, K. C. S., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology*, 97, 79-97. doi.org/10.1016/j.cogpsych.2017.07.001
- Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2009). Cross-modal binding and working memory. *Visual Cognition*, 17(1-2), 83–102. <https://doi.org/10.1080/13506280802281386>
- Baddeley, A.D., Hitch, G., & Allen, R. (2021). A multicomponent model of working memory. In R.H. Logie, V. Camos, and N. Cowan (eds.), *Working Memory: State of the Science*, Oxford University Press. (pp. 10-43)
- Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. London, UK: Psychology Press.
- Broadbent, D.E. (1956). Successive responses to simultaneous stimuli, *Quarterly*

*Journal of Experimental Psychology*, 8, 145-152, DOI: 10.1080/17470215608416814

Broadbent, D.E. (1958). *Perception and communication*. New York: Pergamon Press.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163-191. <https://doi.org/10.1037/0033-2909.104.2.163>

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185. DOI: <https://doi.org/10.1017/S0140525X01003922>

Cowan, N. (2012). Focused and divided attention to the eyes and ears: A research journey. In J.M. Wolfe & L. Robertson, *From perception to consciousness: Searching with Anne Treisman*. Oxford, U.K.: Oxford University Press.

Cowan, N. (2022). Item-position binding capacity limits and word limits in working memory: A reanalysis of Oberauer (2019). *Journal of Cognition*, 5(1), 3. DOI: <http://doi.org/10.5334/joc.193>

Cowan, N. & Guitard, D. (2024) Encoding colors and tones into working memory concurrently: A developmental investigation. *Developmental Science*, e13552. <https://doi.org/10.1111/desc.13552>

Cowan, N., Bao, C., Bishop-Chrzanowski, B.M., Costa, A.N., Greene, N.R., Guitard, D., Li, C., Musich, M.L., & Ünal, Z.E. (2024). The relation between attention and memory. *Annual Review of Psychology*, 75, 183-214. <https://doi.org/10.1146/annurev-psych-040723-012736>

Cowan, N., Blume, C.L., & Saults, J.S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 731-747. <https://doi.org/10.1037/a0029687>

- Cowan, N., Sauls, J.S., & Blume, C.L. (2014). Central and peripheral components of working memory storage. *Journal of Experimental Psychology: General*, 143, 1806-1836. DOI: 10.1037/a0036814
- Cowan, N., Li, Y., Glass, B. A., & Sauls, J.S. (2018). Development of the ability to combine visual and acoustic information in working memory. *Developmental Science*, 21, 1-14. <https://doi.org/10.1111/desc.12635>
- Guitard, D., & Cowan, N. (2023a). The tradeoff between item and order information in short-term memory does not depend on encoding time. *Journal of Experimental Psychology: Human Perception and Performance*, 49, 51–70.
- Guitard, D., & Cowan, N. (2023b). Attention allocation between item and order information in short-term memory. *Quarterly Journal of Experimental Psychology* 76(6): 1391–1409. DOI: 10.1177/17470218221118451
- Guitard, D., Saint-Aubin, J., & Cowan, N. (2022). Tradeoffs between item and order information in short-term memory. *Journal of Memory and Language*, 122, 104300. <https://doi.org/10.1016/j.jml.2021.104300>
- Hitch, G.J., Hu, Y., Allen, R.J., & Baddeley, A.D. (2018). Competition for the focus of attention in visual working memory: perceptual recency versus executive control. *Annals of the New York Academy of Science*, 1424, 64–75. <https://doi.org/10.1111/nyas.13631>
- Hurlstone, M.J. (2024). Serial recall. In M.J. Kahana & A.D. Wagner (Eds.), *The Oxford Handbook of Human Memory*. Oxford University Press. (pp. 799-830)
- JASP Team (2024). JASP (Version 0.19.0)[Computer software].
- Jones, D. M., Beaman, P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 209-237). Hove, U.K.: Erlbaum.

Li, D., Cowan, N., & Sauls, J.S. (2013). Estimating working memory capacity for lists of nonverbal sounds. *Attention, Perception, and Psychophysics*, 75, 145-160.

<https://doi.org/10.3758/s13414-012-0383-z>

Logan, G. D., & Cox, G. E. (2021). Serial memory: Putting chains and position codes in context. *Psychological Review*, 128, 1197–1205. <http://dx.doi.org/10.1037/rev0000327>

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281. DOI: 10.1038/36846

MacDermid, A.E., Duggan, V.A., Miller, B.L., Neath, I., & Surprenant, A.M. (2023). Irrelevant speech, changing state, and order information. *Memory & Cognition*, 51, 1836-1848. DOI: 10.3758/s13421-023-01437-z

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97. doi: 10.1037/0033-295x.101.2.343.

Oberauer, K. (2019). Working memory capacity limits memory for bindings. *Journal of Cognition*, 2, 1–13. DOI: <https://doi.org/10.5334/joc.86>

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G.D.A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G.J., Hurlstone, M., Ma, W.J., Morey, C.C., Nee, D.E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of working memory. *Psychological Bulletin*, 144(9), 885-958. doi: 10.1037/bul0000153

Shomstein, S., & Yantis, S. (2004). Control of Attention Shifts between Vision and Audition in Human Cortex. *The Journal of Neuroscience*, 24(47), 10702–10706.  
<https://doi.org/10.1523/JNEUROSCI.2939-04.2004>

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42, 1096-1104.

<https://doi.org/10.3758/BRM.42.4.1096>

Stoet, G. (2017). PsyToolKit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44, 24-31.

<https://doi.org/10.1177/0098628316677643>

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, 76, 282-299.

Turoman, N., & Vergauwe, E. (2024). The effect of multisensory distraction on working memory: A role for task relevance? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 50(8), 1220–1248. <https://doi.org/10.1037/xlm0001323>

Wolfe, J.M., & Robertson, L. (2012). *From perception to consciousness: Searching with Anne Treisman*. Oxford, U.K.: Oxford University Press.

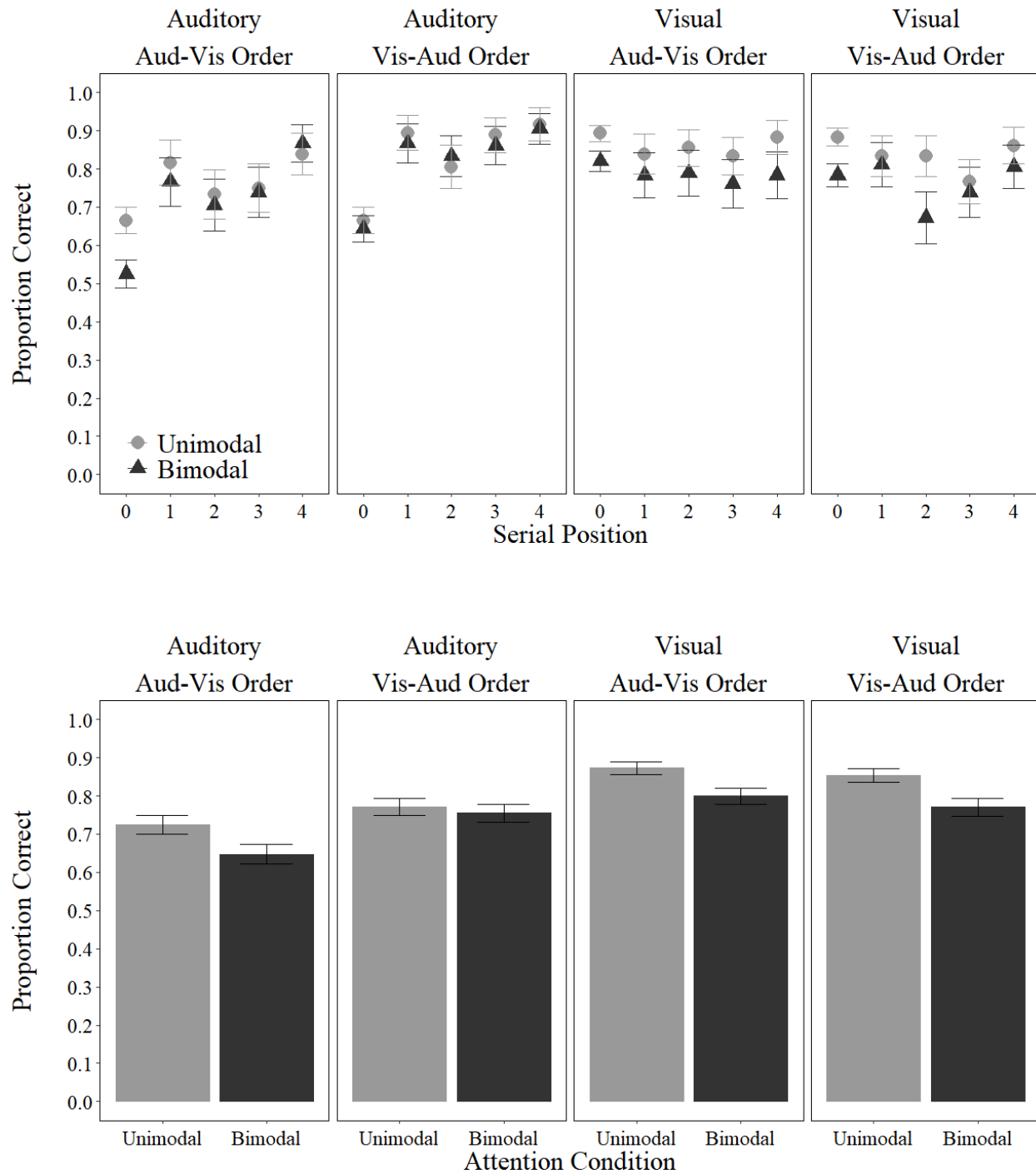
**Appendix A**  
**Post Experiment Strategy Survey**

| Order | Questions                                                                                                                                | Question Type |
|-------|------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| Q1    | Could you explain in about 30 words or less something about your strategy for remembering the colors and tones?                          | Open-Ended    |
| Q2    | I tried to find patterns among the colors                                                                                                | Likert        |
| Q3    | I tried to find patterns among the tones                                                                                                 | Likert        |
| Q4    | When I had to remember two sets, I tried to memorize one set so I would not have to keep attending to it while waiting for the test item | Likert        |
| Q5    | I tried to hum the tones to myself while waiting for the test item                                                                       | Likert        |
| Q6    | I tried to imagine the tones while waiting for the test item                                                                             | Likert        |
| Q7    | I tried to imagine the colors while waiting for the test item                                                                            | Likert        |
| Q8    | I tried to relate the two sets to one another, one color for each tone                                                                   | Likert        |
| Q9    | I tried to ignore one set and just concentrate on the other, even when I was asked to remember both sets                                 | Likert        |
| Q10   | I used a different strategy not shown here – which I have explained in the open-ended question                                           | Likert        |
| Q11   | I used a different strategy not shown here – but forgot to mention it in the open-ended question                                         | Likert        |
| Q12   | I just remembered passively and then answered, without a strategy                                                                        | Likert        |
| Q13   | I just gave up and guessed                                                                                                               | Likert        |
| Q14   | Is there anything about your strategy for remembering that you did not say before, that would be helpful to tell us?                     | Open-Ended    |

*Note.* For the Likert questions participants had to use the following scale to answer: all the time, most of the time, some of the time, occasionally, never.

**Figure S1**

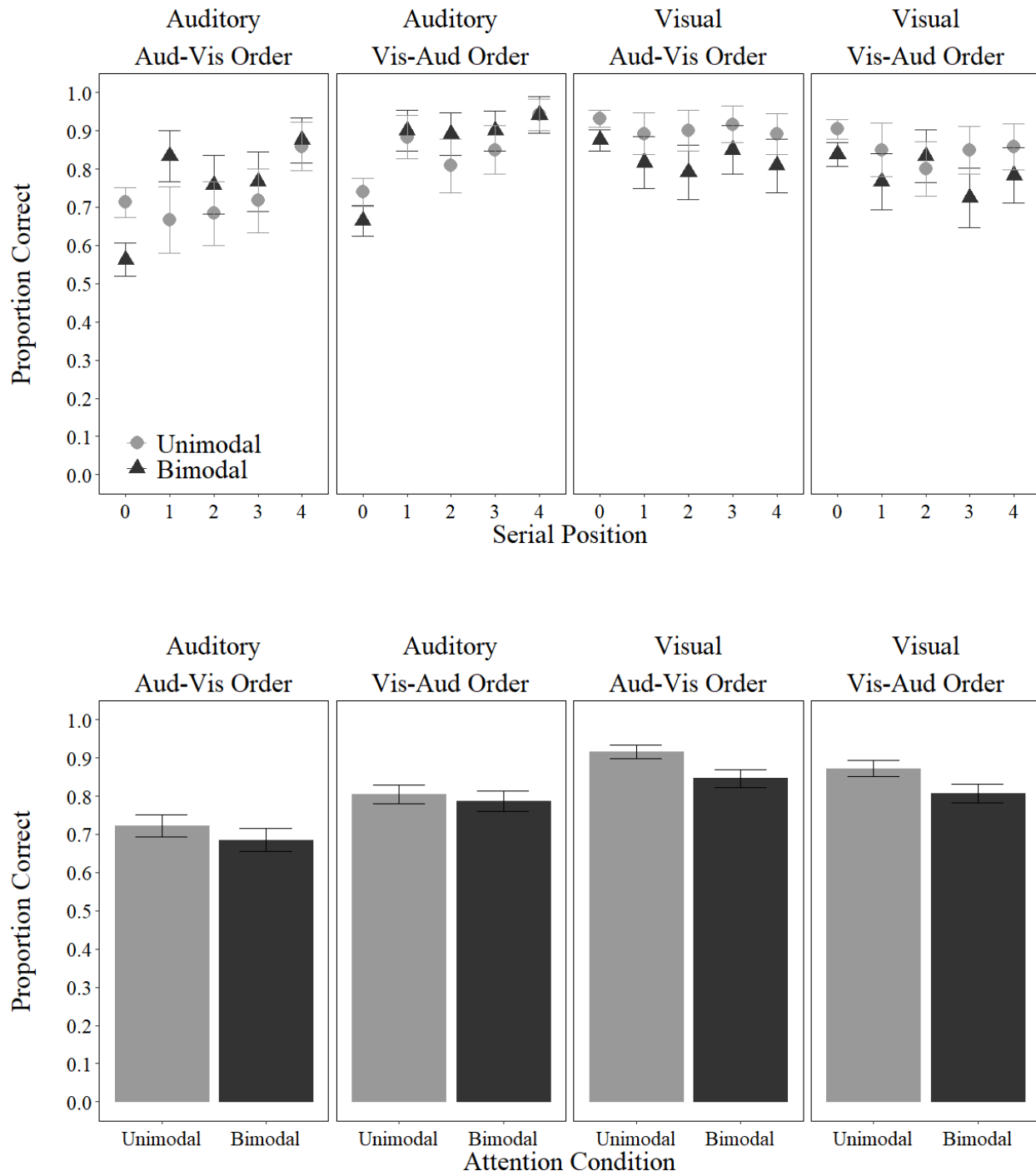
*For the successive presentation group in Experiment 1, performance in both presentation orders*



**Note.** Error bars are standard errors.

**Figure S2**

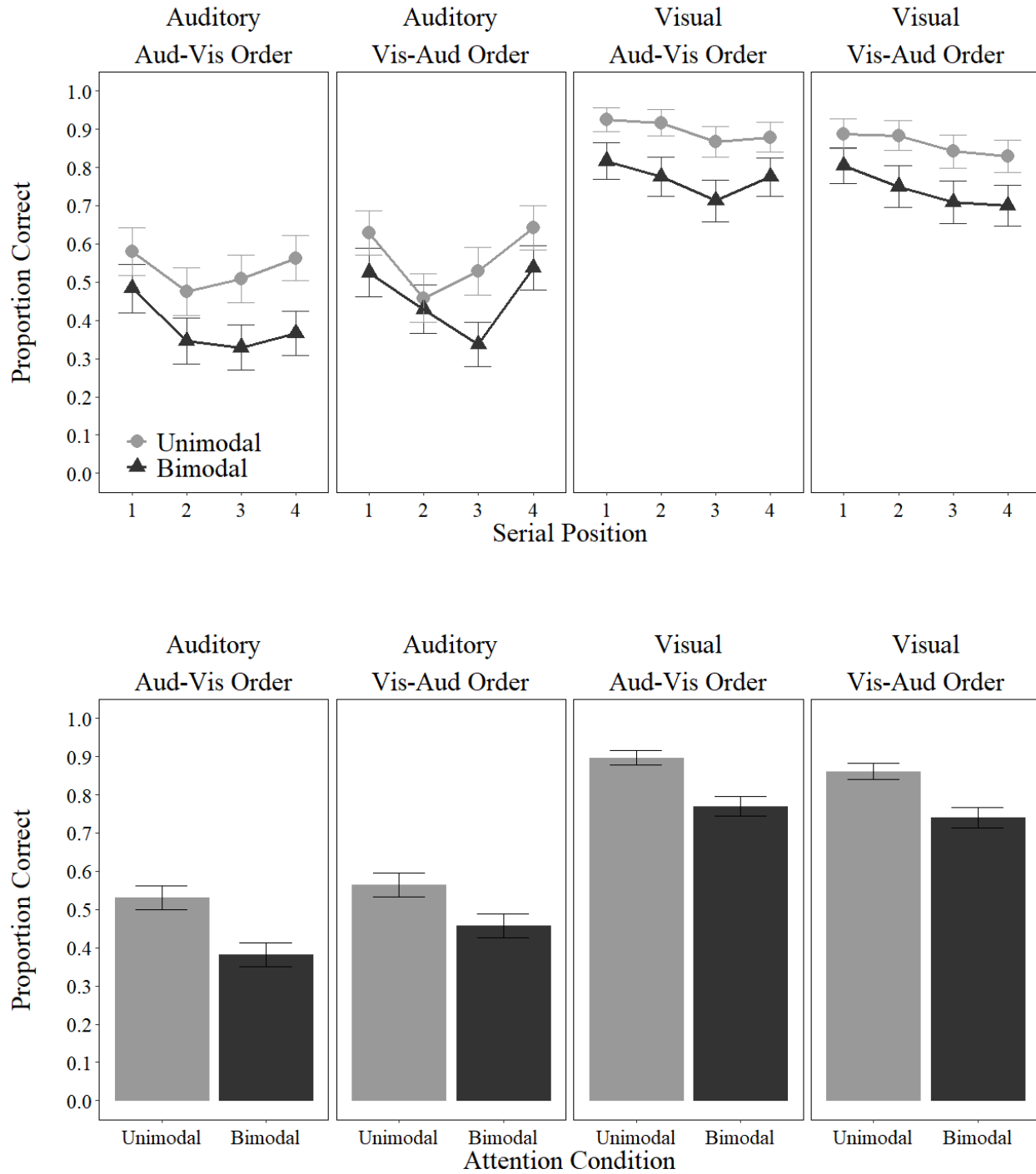
*For the successive presentation group in Experiment 2, performance in both presentation orders*



**Note.** Error bars are standard errors.

**Figure S3**

*For the successive presentation group in Experiment 3, performance in both presentation orders*



**Note.** Error bars are standard errors.