

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/170356/>

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 2024. Encoding colors and tones into working memory concurrently: A developmental investigation. *Developmental Science* 27 (6) , e13552. 10.1111/desc.13552

Publishers page: <https://doi.org/10.1111/desc.13552>

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.



**Encoding Colors and Tones into Working Memory Concurrently:
A Developmental Investigation**

Nelson Cowan

Department of Psychological Sciences, University of Missouri

and

Dominic Guitard

School of Psychology, Cardiff University

Conflict of interest: The authors have no conflict of interest.

Data availability: The study was not preregistered. The materials, the data, and the analyses for this study are available at OSF through an anonymized link:

https://osf.io/jx7q5/?view_only=b26a140d24fc4f26bfea964b3d64458f

Acknowledgements: This research was funded by NIH Grant R01 – HD021338 to Cowan. We thank Bret Glass for a great job with data collection and curation.

Correspondence: Address correspondence to Nelson Cowan, Department of Psychological Sciences, University of Missouri, Columbia, Missouri 65211 USA. Email:

CowanN@missouri.edu and GuitardD@cardiff.ac.uk.

Research Highlights

- Memory for stimuli from multiple modalities is relevant to school performance and learning; here we investigate how attention is shared between remembering colors and tones.
- Participants received four colors and/or four tones for subsequent recognition on a trial, with dual modalities presented successively (0.5 seconds per stimulus) or concurrently (0.5 seconds per pair).
- Successive versus concurrent presentation had little effect on recognition, and the marked increase in memory performance with age did not come from dividing attention during encoding or maintenance.
- Children as young as first grade thus can encode and organize for later recognition colors and concurrently-presented, but unrelated, tones.

Abstract

Working memory serves as a means to accumulate information and reorganize it. Researchers have long assumed that the natural organization of information is one stream at a time. This logic leads to the expectation that, when two different series of stimuli are to be remembered, performance should be superior if the series are presented one before the other in succession, rather than concurrently. Moreover, different accounts of attentional limits lead to different expectations for the change in the ability to encode two sets across age groups in childhood. Testing children from first grade (6-7 years) to adulthood, we presented sequences of colored objects and tones in succession or concurrently (with one color accompanying an unrelated tone) and found that performance was equally good no matter which presentation method was used. The results for both presentation methods closely matched the intricate pattern of development observed by Cowan et al. (2018), who used successive presentation only. We found marked developmental improvement in the ability to retain materials in each modality without an increasing cost of attention-sharing between modalities. Humans at least from the elementary school years through young adulthood thus display an ability to accommodate and organize two concurrent streams of information.

Keywords: working memory development, short-term memory development, visual and auditory modalities, attention, concurrent and successive stimuli

Encoding Colors and Tones into Working Memory Concurrently:

A Developmental Investigation

Here, we explore a key issue from the field of cognitive development, namely how flexibly participants from the early elementary school years to young adulthood can take in (encode or consolidate) series of stimuli in two modalities and retain them for immediate recognition following the series. Immediate recognition refers to the use of working memory, the limited amount of information retained temporarily and used to carry out cognitive tasks (for variation in researchers' definitions of working memory, see Cowan, 2017a). We examine differences between the presentation of series of colors and of tones successively, in either order as in a past study (Cowan et al., 2018, Experiment 2), versus a new arrangement in which there is concurrent presentation, with one tone accompanying each color in the series as an audiovisual, multimodal presentation. The theoretical issue is whether there is stability or developmental change during the school years in how well concurrent information can be encoded into working memory, which requires attention-sharing or rapid attention-switching between modalities. The issue also is of practical importance, for example because multimodal classroom presentations might combine visual and acoustic sources of information into a concurrent, bimodal stream of information, though we do not yet understand how such information is processed.

In a great deal of adult work on both attention and working memory (for reviews see Cowan, 1995; Cowan et al., 2024) it is clear that there are two main factors causing a conflict between stimulus sets to be perceived or remembered: (1) specific interference that occurs when the features of the two sets are similar, and (2) general interference between any two sets that

occurs when an attention-limited process is overtaxed. It was to avoid the first of these two cases that Cowan et al. (2018) chose to examine the interference in working memory between two sets in different modalities, a choice that we follow here in examining the extension of this paradigm to an understanding of how well concurrent stimuli are processed.

It is clear that working memory has a capacity limit of several (3-5) items in adults that applies to a single set that is to be retained but cannot be rehearsed, such as an array of simple objects or a series of tones (Cowan, 2001; Li et al., 2013; Luck & Vogel, 1997). Yet, the same limit does not exactly hold when there are two distinct, very different sets to be remembered, in particular a verbal or acoustic list and a nonverbal, visual array, presented in either order (Cowan et al., 2014, 2018; Fougny et al., 2015; Uittenhove et al., 2019). In that case, there is interference between the sets, but the interference does not limit the total recalled to 3-5 items. Instead, participants seem to retain both sets as if they were presented alone, except for a loss of about one item between sets as the cost of retaining both sets in working memory. Cowan et al. (2014) therefore proposed that participants quickly memorize information from a set and off-load the information so that it no longer occupies the capacity-limited region of working memory (the region said to be the focus of attention). This frees up the capacity-limited region to encode the second set, and the relatively small amount of interference between sets was said to come from attention-demanding maintenance activities to keep the memory of the two sets active (for a review see Rhodes & Cowan, 2018). In a developmental study of children from the early elementary school years through early adulthood, Cowan et al. (2018) tested memory for tone series and arrays of colors and found that the small cost between sets did not change with age, but that more mature participants apparently could off-load much more information about each set.

Following up on Cowan et al. (2018), we ask whether participants as young as the early elementary school years would still be able to handle two sets with no more interference between sets (dual-set cost) than is found in adults if the sets were presented concurrently rather than one after the other. We present series of tones of different frequencies and spots of different colors, with the two series presented either (1) one after the other, or (2) concurrently, with one tone accompanying each color. Assessing the skill needed in these cases is important to understand the basic development of the role of attention in working memory. Apparently, participants from 7 years old to adulthood can attend to two sets presented one after another with a dual-set cost of only about one item. When the two modalities are presented concurrently, however, both sets would have to be encoded, for example by off-loading separate tone and color series to activated parts of long-term memory as newly-learned representations concurrently, or by saving bimodal objects, each with a color and a tone. The question of interest was whether there is development in the ability to do this encoding of concurrently-presented dual streams.

Even adults' ability to do concurrent encoding of two streams into working memory is in question. Broadbent (1956, Experiment 1) presented digits to the right eye and right ear concurrently at the rate of two pairs of digits per half second, for a total of three digits in each modality on a trial, and found an average of 4.62 digits recalled, indicating that bimodal attention was indeed possible. However, there has been little follow-up work of that nature. Dividing attention even in a simple task may be difficult for adults. For example, Cowan et al. (1990) required that participants monitor a slow stream of easily-discriminated, spoken syllables while reading silently and found that the syllabic detection rate was only 60%. Brain imaging research shows that when two series of stimuli are presented at once across the visual and auditory modalities, there is a working memory load from dividing attention that shows up in the

dorsolateral prefrontal cortex (Johnson et al., 2007) and the divided attention load shows up in a broader neural network as well (Santangelo, 2018). Yet, the extant research does not clarify whether the additional load on attention and working memory comes from the rapid rate of presentation of complex information typically used, or from the need to keep track of two series at once so that they can be encoded into working memory.

Developmental Literature

Despite a tradition of research on the development of attention (e.g., Ristic & Enns, 2015), we have been able to find little developmental research in which two streams of information are presented concurrently and both are relevant to the assigned task. Sinzig et al. (2008) required typically developing children and those with autism and/or attention deficit with hyperactivity to respond as quickly as possible when a square occurred among shapes and also when an alternating “high and deep tone was repeated.” This divided attention task was carried out most effectively by typically developing children. Boxhoorn et al. (2018) used the same task but found the deficit to be specific to children with autism spectrum disorder. McDougal et al. (2020) asked children to find matching pairs of spaceships among distractors, count tones, or do both tasks at once, a divided attention condition. Among typically developing children averaging 9 years of age, there were no correlates with divided attention but, among children with Autism Spectrum Disorder averaging 11 years of age, dividing attention was related to IQ, reading, math, and selective attention (but not sustained attention). Savage et al. (2006) administered two divided attention tasks to children 6-11 years old and found that they loaded on the central executive and visual-spatial components of working memory, as well as on reading and intellectual aptitude tasks.

None of these studies, however, examined the typical developmental trend. Talalay

(2024) examined one sort of divided attention developmentally by examining children's ability to search for two objects among a frame of nine objects. It was found that the ability improved markedly between 6 and 10 years, in comparison to near-perfect performance in searching for only one object. However, study of the developmental trend in a bimodal presentation situation could not be found in our literature search.

The Present Study

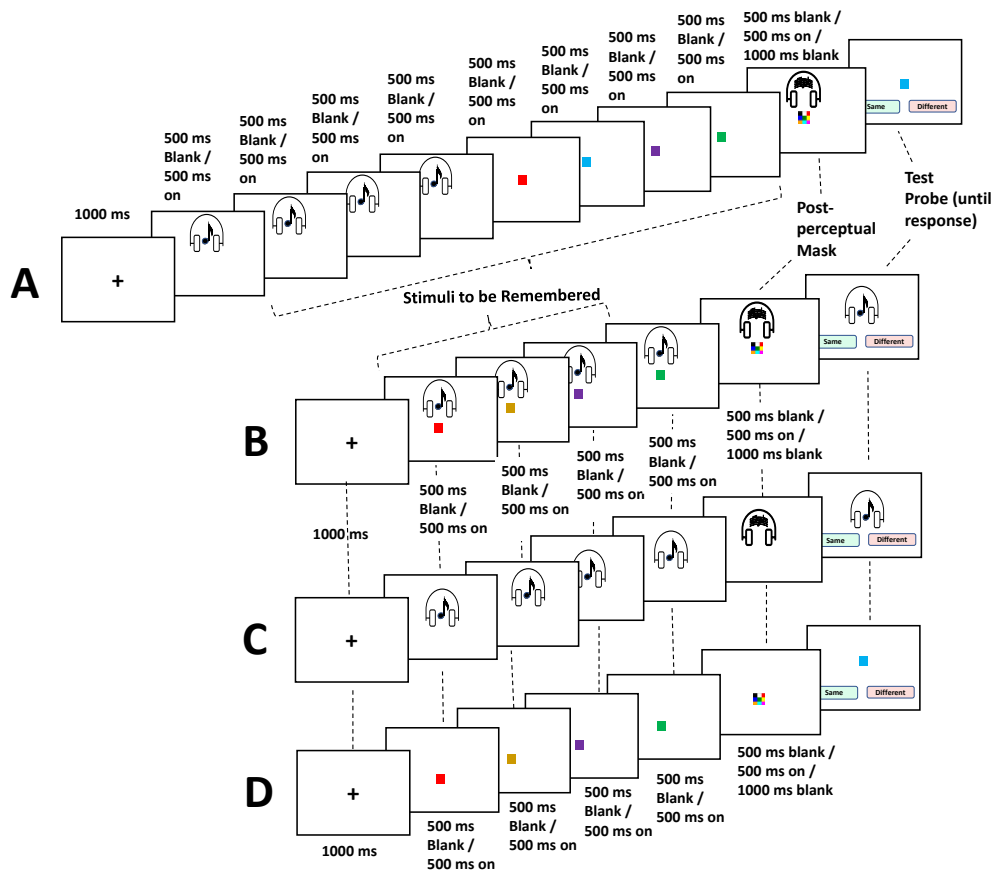
The method we use is an extension of methods used in a developmental study by Cowan et al. (2018, Experiment 2) and an aging study by Greene et al. (2020). Both studies presented series of tones and arrays of colored spots on every trial. On every trial, stimuli from one modality were presented first, followed by stimuli from the other modality. The stimuli were nonverbal to avoid developmental differences in verbal knowledge and comprised arrays of colored squares and series of tones. There was a recognition test, with the task of indicating whether a probe item was among in the stimuli to be remembered on that trial. When two sets had to be remembered (a color set and a tone set), the estimated number of items retained in working memory in total was higher than when only one of the sets had to be remembered, but it was not twice as high; that is, there was a small dual-set retention cost. There was a marked change in the total number of items retained over the age range studied, improving from about 1.8 items at 7 years to about 6.5 items in young adulthood (Cowan et al.) and then declining again from young adulthood to old age (Greene et al.). However, the dual-set retention cost remained similar at around 1 item or slightly less in all age groups.

Here, in both modalities, items are presented one at a time, but with two schedules of presentation: with one modality's stimuli presented after the other, like Cowan et al. (2018) and Greene et al. (2020), as shown in Figure 1A; or with the sets presented concurrently, with the

first tone presented along with the first color, the second tone with the second color, and so on (Figure 1B). The task was to indicate whether a probe item was included in one of the presented sets. Note that in either presentation method, there is no correlation between which colors and which tones were presented on a trial, i.e., no redundancy between modalities, so information from the two modalities must be maintained independently. Two main attention conditions were used with both successive and concurrent presentation of two sets: (1) the bimodal condition, in which participants knew that the probe could come from either set; and (2) the unimodal condition, in which participants knew in advance which set the probe would come from, so that the other set could be ignored during its presentation. In (3) one-set conditions, stimuli were presented in only one modality, to eliminate any interference from a second set.

Figure 1

Illustrations of the order of events for two-modality trials in the present study.



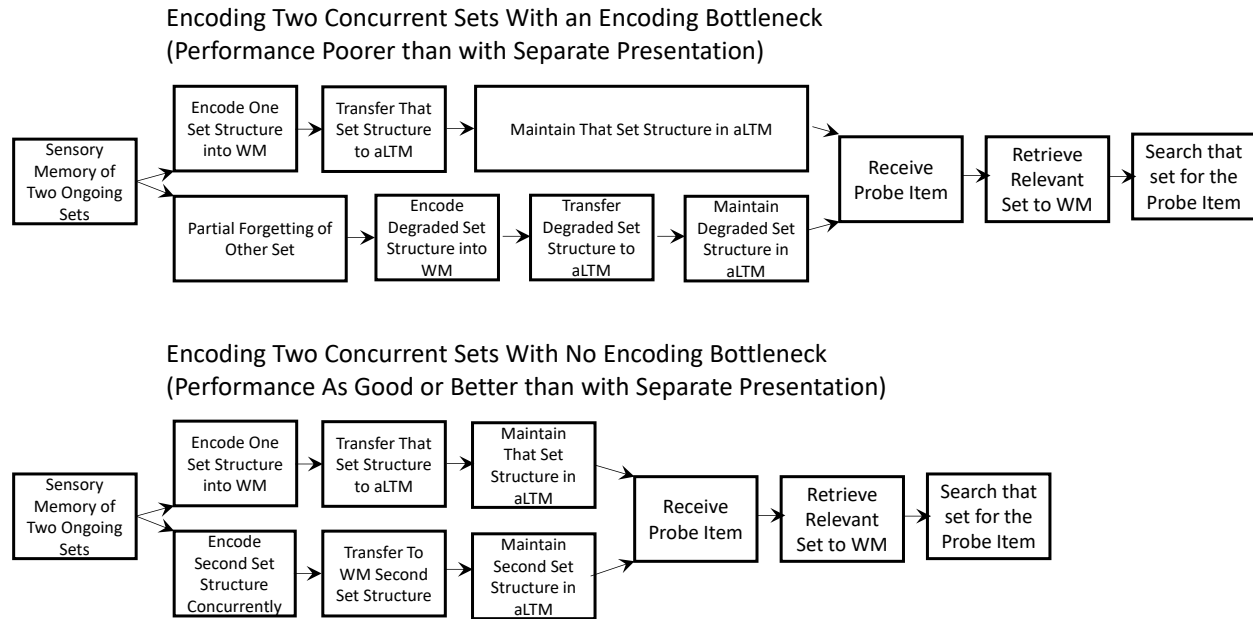
Note. **A:** The successive-presentation recognition procedure used in another experimental session; either modality could come first (as in Cowan et al., 2018). **B:** The novel, concurrent-presentation recognition procedure used in one experimental session. **C:** auditory-only trial. **D:** Visual-only trial. The timing was the same for all conditions but is reiterated several times in the figure. In dual-modality trials, either one or both modalities were to be attended, depending on the trial block. The probe question was whether it was included in the stimuli to be remembered. See Methods section for further details.

Figure 2 illustrates alternative processing mechanisms for concurrent presentation of two sets from different modalities. Concurrent presentation might (top panel of the figure) or might not (bottom) be more difficult than successive presentation of materials from the two sets. It is constructed within the framework of embedded processes (Cowan, 1988; Cowan et al., 2024), in which working memory consists of a temporarily activated portion of long-term memory and,

within it, a capacity-limited focus of attention. Both panels of the figure operate under the assumption that encoding information into working memory requires attention and that the process of encoding a series, either colors or tones, into working memory occurs. They both assume that, at first, encoding from sensory memory into working memory requires attention but that, soon afterward, the information is off-loaded from the focus of attention into the activated portion of long-term memory, freeing up the focus of attention for use on the other set (see Rhodes & Cowan, 2018). In the top panel of the figure, concurrent presentation should be more difficult because of an assumption that creating a series to be off-loaded can occur for only one set at a time. Therefore, there can be loss of sensory memory for the second set while the first set is in the process of being fully encoded and off-loaded. In the bottom panel, however, there is an assumption that the ongoing construction of a representation for both modalities can occur concurrently, at least for the simple item information that needs to be retained. If every time another color-tone item pair is presented both representations are formed, with no overloading of attention in the process, as in the bottom panel of the figure, then performance should be as good with concurrent presentation as with successive presentation of the two modalities. It is also possible that there would be a disadvantage of concurrent presentation for younger participants more than for older ones, if the process of encoding two streams at once improves with development, like various other aspects of attention (Ristic & Enns, 2015), a possibility that remains untested to our knowledge.

Figure 2

Two Alternative, Hypothetical Processing Mechanisms for the Concurrent Presentation Condition



Note. WM=working memory, aLTM=activated portion of long-term memory,

We examined the results taking into account the performance on probe-present trials (in which the probe item had been included in the target set to be remembered) and probe-absent trials (in which the probe item was not present in the target set). Specifically, we examined performance by condition using k , an estimate of the number of items held in working memory. After showing that the pattern of performance was closely similar using either of these measures, the k measure allowed a further investigation in which the development of shared-attention and modality-specific aspects of performance could be examined separately.

The technique for examining this question is illustrated in Figure 3. The first step for this further investigation is to estimate the number of items held in working memory, taking into account guessing (Cowan et al., 2013, Appendix, Model 2). This estimate, k , can be obtained for each modality in both the unimodal and bimodal attention conditions. It is based on the notion

that a correct response to a centrally presented probe item can be based either on detecting that the probe was not from the stimulus set when it was in fact absent, or on correctly guessing that the probe is new or old. The resulting estimate can be expressed as

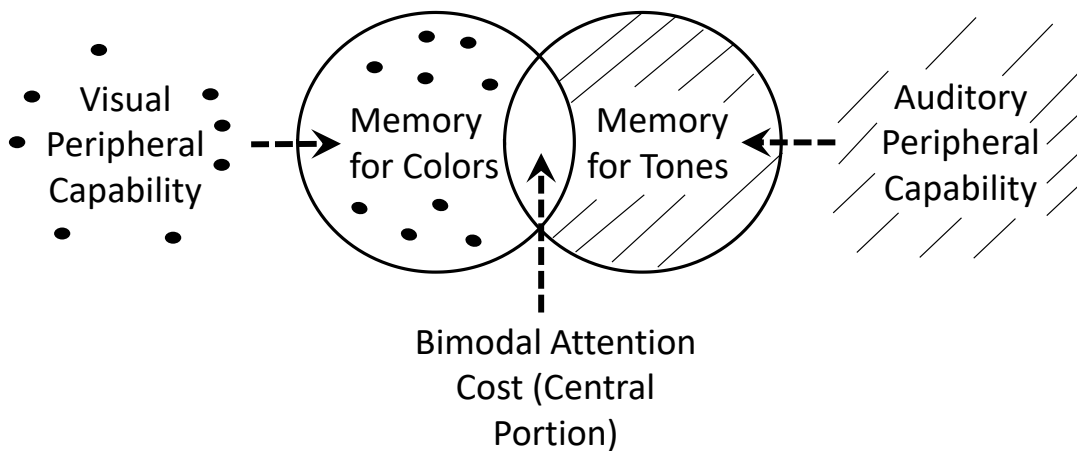
$$k = [p(\text{hits}) - p(\text{false alarms})] / p(\text{hits}) \quad (1)$$

where $p(\text{hits})$ is the proportion of new-probe trials in which the probe item was recognized, and $p(\text{false alarms})$ is the proportion of old-probe trials incorrectly called new. (Note that it would instead be possible to define hits as detection of an old item, in which case the denominator would change to $p(1 - \text{false alarms})$ but the outcome would be the same.) The unimodal attention condition k estimates are depicted by the two circles in the figure. The overlap between the circles reflects the use of the same attentional resource to maintain stimuli in two modalities at once. It can be estimated by adding the two unimodal-attention condition capacities and subtracting from them the two bimodal-attention condition capacities. The result is called a central attentional maintenance resource:

$$k_{\text{central resource}} = (k_{\text{unimodal visual}} + k_{\text{unimodal auditory}}) - (k_{\text{bimodal visual}} + k_{\text{bimodal auditory}}) \quad (2)$$

Figure 3

The logic of quantifying dual-set attention costs



The total items remembered summing across unimodal attention conditions (represented by the sum of the two circle areas considered separately) can be larger than the total items remembered summing across bimodal conditions (represented by the total area when the circles overlap). The decrease in total area caused by overlap of the circles matches the *central portion* of working memory Equation 2. This is the estimated number of items that can be allocated to either colors or tones or split between them, depending on task demands. The non-overlapping parts of the figure represent the items that can be retained in each modality without the central resource, called *peripheral portions*, which can be estimated by subtracting the central portion from each of the unimodal capacity estimates. It cannot be assumed that attention is uninvolved in maintaining these peripheral portions, but they could be formed via an “offloading” process in which attention is used to assist in memorizing the stimuli so that the commitment of attention for maintenance can decline (Rhodes & Cowan, 2018). The developmental finding of Cowan et al. (2018) was that the central portion does not increase with maturation, whereas the peripheral portions increase markedly. Participants improve in the ability to retain colors and tones without an increase in resource-sharing between the two when both have to be retained at once.

Here, when we presented stimuli in two modalities concurrently, one tone concurrent with each color (as in Figure 1B), we anticipated that the need to encode both modalities concurrently might increase the central resource estimate, if that resource is at some point taken up by an encoding process or the consolidation of information in working memory (e.g., Bayliss et al., 2015; Jolicoeur, 1999; Jolicoeur, P., & Dell’Acqua, 1998; Ricker et al., 2018; Vogel et al., 2006). We further expected that the need to encode both modalities concurrently might be especially problematic for the younger children, if their attentional resource is limited as is assumed from a neo-Piagetian theoretical view of development (e.g., Case, 1995; Cowan, 2016,

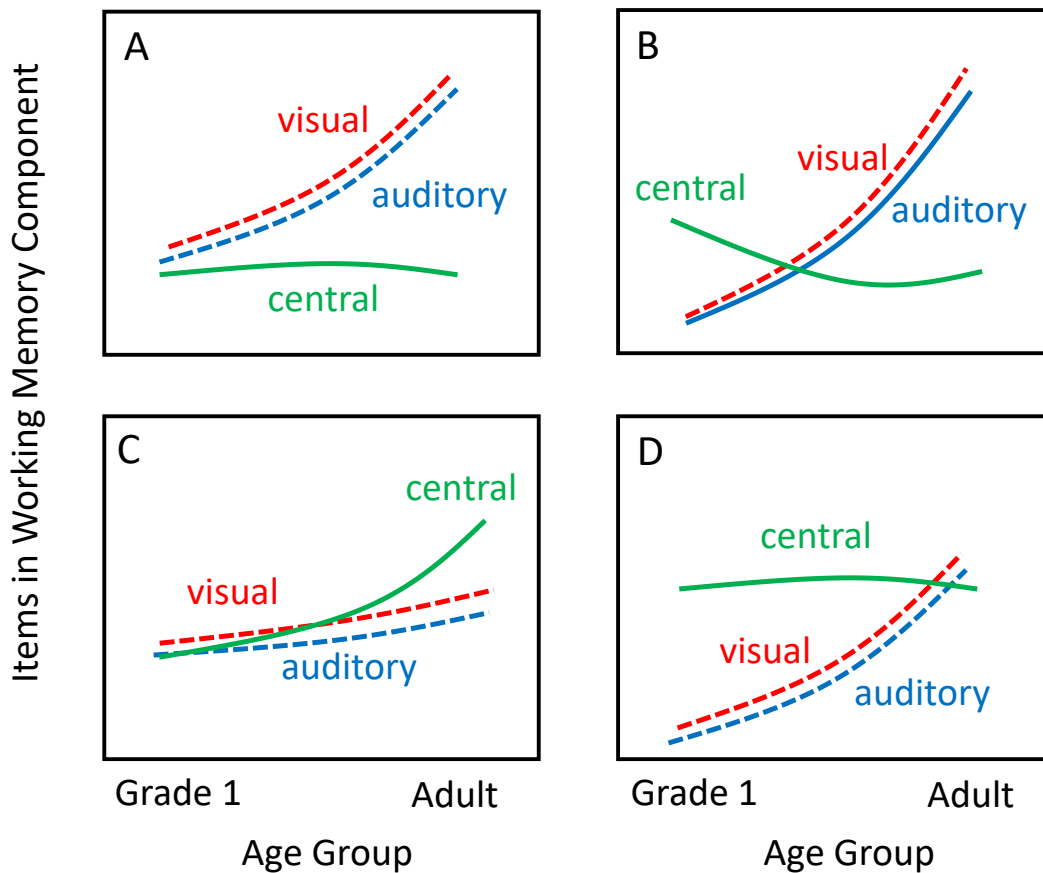
2017b, 2022; Pascual-Leone, 1970; Pascual-Leone & Johnson, 2021). If this is the case, then we should find a larger central portion for concurrent presentation of stimuli in the two modalities (Figure 1B), compared to successive presentation of the two modalities (Figure 1A), especially for younger children. On the other hand, these predictions do not hold if the resource limitation applies only to working memory maintenance, with little effect of encoding limits with a moderately-paced temporal schedules of presentation that we used.

The panels of Figure 4 illustrate different possible outcomes for the componential analysis, with age group represented along the x axis and the number of items held in each component represented in the y axis. Figure 4A resembles the result obtained by Cowan et al. (2018) with successive presentation. A small tradeoff between modalities is represented by the central component, and modality-specific gains across age groups presumably reflect improvements in the means of encoding stimuli within each modality. Assuming that we replicate this effect from Cowan et al., obtaining results like these also in the present concurrent-presentation condition would indicate that there is no additional cost of encoding and storing stimuli in both modalities at once. If instead, younger children show the most deficit due to concurrent encoding, the results of the concurrent condition could resemble instead Figure 4B. Another possibility is that more mature participants use attention more to encoding even one set, in which case dual-modality encoding could drain their attention more than younger participants, resulting in a pattern like Figure 4C. In support of this possibility Rosen and Engle (1997) found that among adults, those with higher working memory span relied more on attention for memory retrieval and were impeded more by dividing attention task during retrieval. Last, if concurrent presentation is equally demanding for all, and more demanding than successive presentation, then concurrent presentation could give rise to a pattern of results with a uniformly high central

portion, as in Figure 4D.

Figure 4

Alternative Predictions for Concurrent Presentation of Visual (Color) and Auditory (Tone) Stimuli



Note. Panel A: If there is no cost of concurrent presentation, the results could be comparable to successive presentation of the two sets, which should match the results of Cowan et al. (2018, Experiment 2), schematically depicted here. **Panel B:** A concurrent-presentation cost that is larger for young children should produce a result like this one. **Panel C:** If the concurrence-cost is especially large for more mature participants because they used attention for encoding more than younger participants, the results should resemble this figure. **Panel D:** If the concurrent-presentation cost is large for all participants, this configuration of results should emerge.

Some have suggested that the k measure can be biased if there is actually a continuum of

memory strength rather than an all-or-none working memory of each item that the k measure assumes (Williams et al., 2022). Williams et al. used d' as a measure of the sensitivity to a memory representation, i.e., memory strength, and reported that it did not yield the same results as k . The implication was that k cannot be used if the memory representations vary continuously in strength. To check this possibility, we calculated also a d' measure of sensitivity to a changed probe, for comparison with the k results.

Last, we carried out two auxiliary tasks, Raven Progressive Matrices (Raven, 2000) and running memory span (Cowan et al., 2005; Gray et al., 2017). The Raven task is a standard fluid intelligence task that can be used to assess whether our samples were typical and comparable for their age groups. Also, it can be used to examine whether there is a developmental factor of maturity beyond what can be gleaned from the age group, that plays a role in performance. Running memory span with spoken digit stimuli correlates well with aptitudes at any age group, better than standard digit span does in the older children and adults, and comparable to how well complex span does but with a much simpler procedure (Cowan et al., 2005). Cowan et al. considered running span to be an index of the capacity of the focus of attention, an interpretation that was strengthened by Gray et al., who found that running digit span in 9-year-olds was correlated with visual tasks, which are known to require attention, more strongly than with standard verbal span tasks, which allow verbal rehearsal and therefore tend to circumvent attention. Given the reliance of aptitudes and of running span on attention, it will be of interest to determine the degree to which they correlate with our central and peripheral components of working memory in a dual-task setting, which can reflect a role of attention.

Method

Participants

The study was approved by Institutional Review Board of the first author's university. The participants were tested one at a time, remotely, from locations throughout the United States, and received \$15.00 per hour of participation for two sessions. Our target was at least 24 per age group, with several additional pre-scheduled appointments also completed and included. Power-like analyses for Bayesian inference (Schönbrodt & Stefan, 2018; Schönbrodt & Wagenmakers, 2018) indicated that with an inferential cutoff of 3.0, we could detect an age group difference effect size of .85 with a probability of .81, whereas based on previous work (e.g., Cowan et al., 2018), we expected age differences with effect sizes of about 1.0 and higher. We refer to a Bayes Factor of 3, considered relatively lenient, so as not to ignore subtle trends in the data, but larger Bayes Factors should always be interpreted with more confidence. Children were assigned to groups 1-3 based on the grade in school that the parent indicated that the child was in currently or had just completed. In all, $N=102$. There were 24 children in Grades 1-2 (mean 7.37 years old, $SD=0.79$, 17 female, 9 male); 25 children in Grades 3-4 (for 24 of them, mean 9.40 years, $SD=0.65$, and a 10-year-old for whom an exact birthdate was unavailable; 9 female, 16 male); 26 children in Grades 5-7 (mean=12.03 years, $SD=1.20$; 14 female, 12 male). There also were 27 adults (mean=24.04 years, $SD=3.18$; 21 female, 6 male). Additional participants did not complete both sessions of the study (2 in Grades 1-2, 1 in Grades 3-4, 3 in Grades 5-7, and 4 adults; an additional child in Grades 3-4 may have completed the study but one session was not saved).

Materials, Stimuli, and Procedure

Children and adults were tested one at a time, remotely, using Zoom and monitored by

the experimenter during testing, and were paid for their participation. The experiment was programmed with PsyToolKit (Stoet, 2010, 2017) and the link to the program was given to the participant's parent or guardian, so that it ran on the computer that the child was to use. This method was used because we determined that if the program ran on the experimenter's computer, the display that the child received would not be consistently correct due to how the university handled outgoing internet signals. However, the experimenter was able to monitor the participant throughout the session.

Experimental Sessions

The experiment was conducted in two sessions, each lasting approximately 60 minutes, with a total of 20 practice trials and 96 test trials. In both sessions, the tone and color stimuli were identical to those used by Cowan et al. (2018) and the timing was intended to be like that study, but with changes in the timing of visual stimuli given that Cowan et al. placed them in spatial arrays, whereas we presented one visual item at a time.

In one session, a *successive* presentation of the four-tone series and four-color series for recognition was used, as in Cowan et al. (Figure 1A). On some trials, only four stimuli in one modality were presented. In this session, when both visual and auditory stimuli were presented, their order was counterbalanced across trials. The main procedure was followed by a running span test (Cowan et al., 2005). In the other session, *concurrent* presentation of four color-tone pairs was used (Figure 1B), although there also were trials with four stimuli in only one modality, as in the other session. This procedure was followed by Ravens Progressive Matrices (Raven, 2000). The order of these two sessions was counterbalanced across participants.

Tone and Color Recognition Test

Visual stimuli. The visual stimuli in this task were displayed on a gray background

(RGB = 192-192-192) with stimulus values matching Experiment 2 of Cowan et al. (2018). Colored squares (60 pixels x 60 pixels) presented for 250 ms (separated by 250-ms blank periods) were sampled without replacement from a pool of 9 possible colors presented at the center of the display with the following values: (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; and magenta, RGB = 255-0-255). The visual mask was a multicolored square the same size as the visual to-be-remembered items but comprising a grid of all nine colors in a 3x3 arrangement within the square presented for 500 ms.

Auditory stimuli. The auditory stimuli were the same as those developed by Li et al. (2013, Experiment 3) and used also by Cowan et al. (2018, Experiment 2). To allow them to have categorical distinctions as colors do, we used synthetic musical sounds created by GarageBand (Apple Inc., Cupertino, CA). We used 9 different musical sounds of 125 ms each, with different fundamental frequencies assigned to different instruments (trumpet section, 200 Hz; smooth clav, 262 Hz; classic rock organ, 343 Hz; Negril bass, 450 Hz; tenor sax, 589 Hz; space harpsichord, 772 Hz; grand piano, 1011 Hz; live pop horns, 1324 Hz; aurora bell, 1735 Hz.) Tones were separated by 375-ms periods, allowing the onset-to-onset time to be 500 ms, the same as for visual stimuli. One can confirm that there was a 31% increase in frequency from one tone to the next, an increase that allowed us to avoid familiar musical intervals. These intervals were such that no two tones in our set had a simple harmonic relation (for further details see Cowan et al., 2018). The auditory mask was presented for 500 ms, like the visual mask and concurrent with it; it was composed of short segments from all nine tones presented in rapid sequence with no break between them. No two stimuli to be remembered on a trial were identical.

Trial blocks within each session. The task included three blocks within each session

(i.e., in the successive and concurrent presentation sessions). Each of the three blocks was for a different attention condition, and the block order was counterbalanced in a Latin square design across participants (but with the same order for an individual in the two sessions). In the unimodal visual block, all tests were on the visual stimuli, and the auditory stimuli could be ignored. The block started with 2 practice trials in which there were only four visual stimuli and no auditory and, between them, 4 practice trials in which there were both visual and auditory stimuli, but the auditory stimuli could be ignored. The test session included, in random order, 8 one-set trials and 8 two-set trials. The trial block for auditory information was exactly comparable, except that the roles of auditory and visual stimuli were reversed so that participants now could ignore all visual stimuli.

In the bimodal trial block, there were 8 practice trials and 32 test trials. In both of these, half of the trials ended in a visual test and half ended in an auditory test, so the participant had to attend to all stimuli.

In all trial blocks, there were several variables that were presented an equal number of times across trials of the block, but in random order: (1) the presentation of the auditory or the visual stimuli first, (2) the presentation of a probe stimulus that did or did not match one of the stimuli to be remembered, and (3) when the probe did match one of the stimuli to be remembered, the match to Positions 1-4 in the sequence.

Procedure. Before each trial, an encoding instruction was verbally presented by the experimenter and an appropriate icon was shown (an eye, an ear, or both, depending on the modalities to be attended). The participant initiated the trial by clicking the screen. Then the trial began with a fixation cross at the center of the screen for 1000 ms, followed by a 500-ms blank period. Next, items (or, in the concurrent session, color-tone pairs) were presented at a rate of 1

item every 500 ms. After the last item's off period, there was another 500-ms blank period and then the visual and auditory masks were presented simultaneously for 500 ms and followed by a 1000-ms off period. (For a fair comparison of trials, the mask was always bimodal.) Next, the probe item was presented (a visual item in the center of the screen or a 125-ms sound accompanied and followed by a question mark, depending on the condition, lasting until the response). The participant was to indicate whether the probe was an item present within the stimuli to be remembered by clicking a green button for "same" or "old" and a red button for "different" or "new." After each response, as feedback, for 1000 ms a smiley face was accompanied by the word "correct" or a frowny face was accompanied by the word "incorrect," depending on the participant's response accuracy.

After participants completed this task, they were asked a series of questions to better understand the strategies used to remember the colors and tones. There were two open-ended questions and twelve strategy questions on a Likert scale. These appear in the Open Science Framework page along with the rest of the data,

https://osf.io/jx7q5/?view_only=b26a140d24fc4f26bfea964b3d64458f.

Running Memory Span

Cowan et al. (2005) described a version of a running span task in which spoken digits were digitally compressed to 250 ms per digit and were presented in series of 12-20 digits. The participants could not group the series because the endpoint of each trial is unknown. In each series, within a window of six digits, no two are identical. When the series ends, the participant spoke as many digits as they could remember from the end of the list. The experimenter typed these digits into the computer. These responses were scored according to items recalled in the correct serial position relative to the end of the list.

Raven Progressive Matrices

Raven Progressive Matrices is a standardized test of abstract reasoning in which a visual-spatial series of complex objects is shown and the participant must select the choice that best completes the pattern that the series follows. There are traditional age norms collected in 1979 (Raven, 2000) and newer norms (Pind et al., 2003) with which the scores can be compared.

Results

Appendix A (with Table A1) provides accuracy means for every condition in the experiment. The results section will focus on developmental trends, but also with some key supporting effects across age groups that provide cues to the nature of processing in our task. The results are divided into analyses of three types. These include (1) d-prime analyses to compare successive and concurrent presentation methods; (2) comparable analyses using k , an estimate of the number of items held in working memory; (3) central and peripheral components of working memory in a dual-set situation, in line with Cowan et al. (2018); (4) potential effects of presentation order within the successive sets; and (5) auxiliary tasks that help describe the participants, namely running memory span and Ravens Progressive Matrices. The data and analyses are available at https://osf.io/jx7q5/?view_only=b26a140d24fc4f26bfea964b3d64458f.

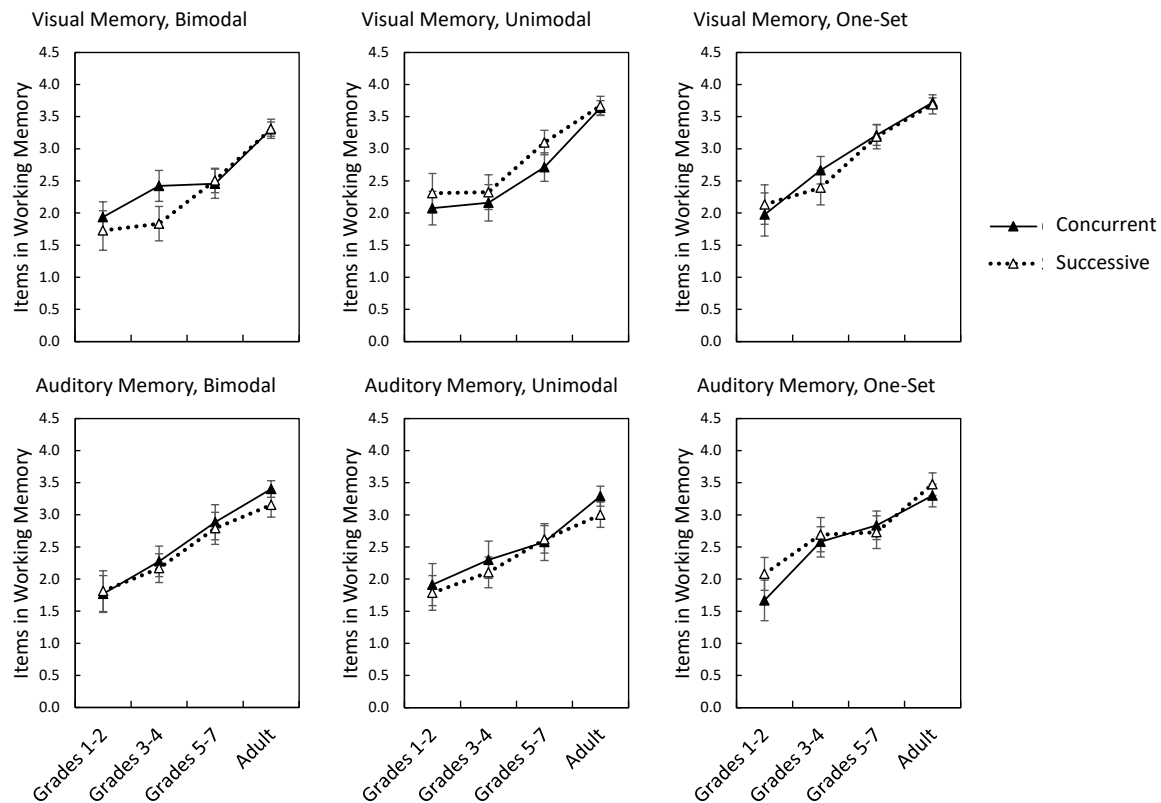
Inferential statistics were performed in a Bayesian framework (JASP Team, 2023) because that framework allows for the conclusion that the null model fits better than the model with a particular effect. A Bayes Factors (BF) for including an effect, compared to matched models omitting that effect, of $BF > 3$ is taken as evidence that the effect is probably not null, whereas $BF < 0.33$ is taken as evidence that the null model fits best. In either direction, more extreme values count as stronger evidence for or against the existence of the effect.

Items Held in Working Memory

Figure 6 shows the data in terms of the number of items in working memory. Statistically, there was a large effect of age group, $BF > 149$ or higher. There was no main effect of presentation condition, although the effect was in the indeterminate range for visual unimodal presentations (top middle panel of Figure 6), $BF = 0.87$. For the other conditions, the evidence favored no effect, $BF < 0.32$ or lower in each case, or at least 3.1 to 1 in favor of the null. There was no interaction in any condition, $BF < 0.18$ or lower in each case, or at least 5.56 to 1 in favor of the null.

Figure 6

Development of Memory for Visual and Auditory Stimuli Presented Concurrently vs. Successively Using a Measure of Items in Working Memory



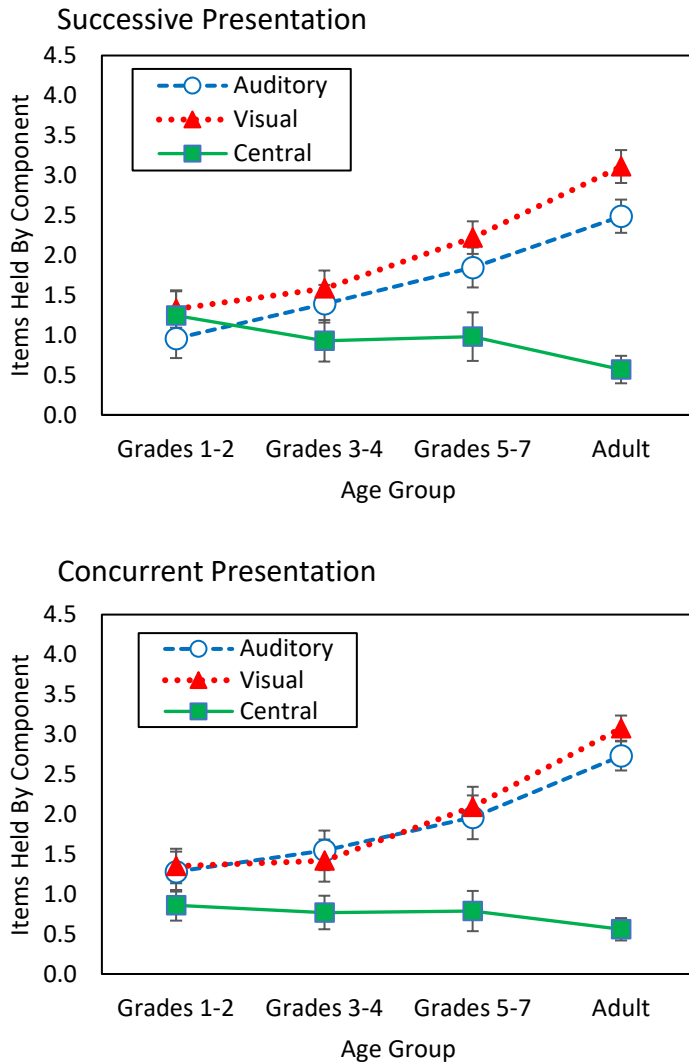
Note. Error bars are standard errors.

Analysis into Components

Given the confidence in the k measure for the present data set (Figure 6), we proceeded to calculate the capacities of central and peripheral portions of working memory on each individual as explained above in Equations 1 and 2. (Two of the participants in the youngest group were omitted from this analysis because of a divide-by-zero error for Equation 1 in one condition; in both cases, 0% detection of a new auditory probe with concurrent presentation and bimodal attention.) The results, shown in Figure 7, are remarkably similar across successive presentation (top panel) and concurrent presentation (bottom panel) methods and closely resemble what Cowan et al. (2018, Experiment 2) obtained with successive presentation, matching the theoretical prediction of Figure 4, Panel A. These figures show that although there is a small contribution of shared attention between modalities, the bimodal attention condition did not detract more from younger participants. The fact that this was the case even with concurrent presentation indicates that even the younger children have the capability of encoding visual and auditory stimuli concurrently without draining attention in the process.

Figure 7

Analysis of Central and Peripheral Components with Successive (Top) and Concurrent (Bottom) Presentation Methods



Note. Error bars are standard errors of the mean.

Confirming this pattern statistically, a Bayesian ANOVA of the central portion of working memory, shared across modalities, was examined with age group between participants and the presentation method (successive versus concurrent) within participants. It yielded no effects of the age group, $BF=0.17$, or 5.89 to 1 in favor of the null; of the presentation method, $BF=0.30$, or

3.33 to 1 in favor of the null; or of the interaction, $BF=0.08$, or 12.5 to 1 in favor of the null.

In contrast to the central component, the peripheral auditory component yielded a large effect of age group, $BF=12,259$; the presentation method in the indeterminate range, $BF=0.41$; and the interaction favoring the null, $BF=0.06$ or 16.67 to 1 in favor of the null. Similarly, an analysis of the peripheral visual component yielded a large effect of age group, $BF>100,000$; no effect of the presentation method, $BF=0.18$ or 5.56 to 1 in favor of the null; and no interaction, $BF=0.06$ or 16.67 to 1 in favor of the null.

Presentation Order Effects

The means for the two-set trials within the successive presentation condition can be separated according to whether the presentation involved auditory or visual items first (Appendix A). It is possible that there was a benefit of receiving one set at a time but that it was cancelled out by the longer wait before being tested on the first set, compared to concurrent presentation. We examine this on the basis of the proportion of hits and proportion of correct rejections (when the probe was absent from the memoranda) to provide useful detail in the result and to avoid recalculating a derived measure (k or d') based on fewer trials than previously. Table 1 shows the results. It suggests that for target-present trials, targets presented within the first set tend to be forgotten compared to targets presented in the second set, so that it indeed may be the case that there is a disadvantage for Set 1 based most likely on forgetting, and a relative advantage for Set 2.

Table 1*Effect of Presentation Order on Accuracy*

Attention	Modality	Presented First	Presented Second
Target Present			
Bimodal	Auditory	0.71	0.82
	Visual	0.69	0.73
Unimodal	Auditory	0.67	0.78
	Visual	0.73	0.82
Target Absent			
Bimodal	Auditory	0.61	0.63
	Visual	0.76	0.78
Unimodal	Auditory	0.71	0.71
	Visual	0.84	0.87

Note. Each mean accuracy shown was averaged across the averages for the four age groups. The pooled standard error of the mean is .02 for each cell.

Four separate Bayesian ANOVAs were carried out for trials in the successive presentation condition, for auditory probes in target-present and target-absent trials separately, and the same for visual probes. This means of analysis allowed us to assess the role of order as a main effect, and it helped to maintain homogeneity of variance within an analysis. Each analysis included the attention condition (bimodal and unimodal) and the set presentation order (auditory-visual, or visual-auditory) within participants and age group between participants. Both analyses of target-present trials yielded main effects of the presentation order (auditory trials, $BF=4,078$; visual trials, $BF=5.68$). This was not the case for the target-absent trials (auditory trials, $BF=0.16$, or 6.25 to 1 in favor of the null; visual trials, $BF=0.39$, indeterminate but tending toward favoring the null). These results conform to what is seen in Table 1.

Among these analyses, there was one interaction involving age group, in the auditory

target-absent trials: the three-way interaction between attention condition (bimodal or unimodal), audiovisual presentation order, and age group, $BF=3.27$. It is described here for the sake of completeness of our reporting but does not alter our theoretical interpretation. The basis of this interaction was that there was less age difference for unimodal trials with an auditory-visual presentation order (across age groups, $M=0.69, 0.67, 0.69$, and 0.78) than for the other three conditions (bimodal with auditory-visual presentation, $M=0.52, 0.60, 0.66$, and 0.65 ; bimodal with visual-auditory presentation, $M=0.56, 0.55, 0.61$, and 0.80 ; and unimodal with a visual-auditory order, $0.55, 0.74, 0.75$, and 0.82). We speculate that the unimodal attention to sound made young children especially confident that, if the probe tone had been presented in the list, they would have remembered it despite its being followed by visual items. This finding would conform to the observation of Forsberg et al. (2021) that young children tend to be overconfident in their working memories, which helps when the probe was not present in the list.

A d' Measure of Memory Strength

Some have questioned the validity of the k measure on the grounds that the strength of memory (or sensitivity of working memory to an item's presence) might vary continuously, as in a d' measure (Williams et al., 2022), whereas the k measure rests on the assumption that an item is more simply either in working memory or not in working memory. Williams et al. argue that the k measure can distort the results. To address this issue, we calculated d' for every condition as Williams et al. did, for comparison with the k measure results. The results of this analysis are reported in Appendix B (with Figure B1) and show that the d' measure yielded results closely comparable to the k results shown in Figure 6. This result provides reassurance that the k measure and its all-or-none memory assumption about each item was adequate for the present purposes.

Further Details of the Accuracy Results

Further details help to establish the pattern of results already noted. We examine accuracy across serial positions and across trials.

Serial Position Functions

It is possible that the results for successive versus concurrent presentations might differ by serial positions, which could point to different mechanisms in the two cases. We carried out a Bayesian ANOVA in the dual-modality memory condition for auditory stimuli, and another for visual stimuli. Each analysis included the Age Group between participants and two within-participant factors: the order of presentation (auditory-visual, visual-auditory, or concurrent) and the serial position (probe absent and then probe at Positions 1, 2, 3, and 4). This analysis provided no evidence of an interaction of the order of presentation and the serial position, either for the auditory stimuli, $BF=.03$ or 33.33 to 1 in favor of the null, or for visual stimuli, $BF=.07$ or 14.29 to 1 in favor of the null. Serial position means are shown in Appendix C (within Table C1). In the auditory probe analysis, there was no interaction of the serial position with age group and no group $\times$ presentation order $\times$ serial position three-way interaction, with Bayes Factors of at least 500 to 1 in favor of the null. The same was true for the visual probe analysis, with Bayes Factors of at least 125 to 1 in favor of the null. Thus, successive and concurrent presentations not only yield comparable overall results, but also give rise to reasonably similar serial position functions.

Trial-by-Trial Accuracy

It is theoretically possible that the age effect comes from younger participants becoming tired toward the later portions of each session. To examine this possibility, we plotted mean performance for each trial in the main task of the experiment, including practice trials, averaged

across individuals in each age group. Different participants would receive trials of different orders but these differences average out across participants, leaving primarily trial-order-related mean performance functions as shown in Appendix D. Figure D1 within that appendix suggests that there was not much fatigue in any age groups.

Auxiliary Tasks

The mean running span in number of items correct increased steadily with age group (Grades 1-2, $M=2.13$, $SD=0.49$; Grades 3-4, $M=2.56$, $SD=0.56$; Grades 5-7, $M=2.83$, $SD=0.43$; Adult, $M=3.15$, $SD=0.62$). These scores are comparable with previous results (e.g., Cowan et al., 2005). The same was true of Ravens Progressive Matrices (Grades 1-2, $M=34.25$, $SD=7.12$; Grades 3-4, $M=42.88$, $SD=5.62$; Grades 5-7, $M=44.85$, $SD=6.91$; Adult, $M=52.04$, $SD=4.26$). For these scores, age norms (Pind et al., 2003; Raven, 2000) suggest that all groups were at or slightly below the 75th percentile of the general population.

We explored whether Ravens Progressive Matrices and running span could serve as markers of maturation such that central, peripheral auditory, and peripheral visual components of working memory could vary within an age group in a manner that reflects participants' relative levels of maturation within that age group. We averaged the components from the main task across concurrent and successive presentations and carried out a Bayesian analysis of covariance for the Ravens measure in which the age group was entered as a fixed factor and the three components were added as covariates. This analysis yielded a large effect of age group, $BF>100,000$, and also an added effect of the peripheral auditory component, $BF=4.49$, but not the visual component, $BF=0.56$, or the central component, $BF=0.75$, both in the indeterminate range. The model accounted for 52% of the Ravens variance but without the auditory component it was still 50% (and the same with only age group in the model). Thus, most of the variance in

the Ravens was absorbed by age group. For a comparable analysis of running span, there was again an effect of age group, $BF=157.20$, but on top of that no clear effect of auditory ($BF=0.96$), vision ($BF=0.74$), or central ($BF=0.31$) components. The model accounted for 28% of the variance in running span. For the most part, then, our components reflect the age group but not a fine-grained view of maturation within an age group. Without considering age group, the auditory peripheral component correlates with Ravens, $r=.51$, $BF>100,000$, and running span, $r=.40$, $BF=474$. Similarly, the visual component correlates with Ravens, $r=.50$, $BF>100,000$, and running span, $r=.41$, $BF=1,321$, whereas the central component is not related to the Ravens, $r=-.11$, $BF=0.22$, or running span, $r=-.11$, $BF=0.23$. The central component is relatively fixed across age groups, and the improvement in performance with age appears to come from how auditory and visual streams can be encoded into memory, rather than from how much attention can be shared between modalities.

Discussion

We set out to investigate whether there is a bottleneck that applies to the encoding of two sets of information, one visual and one acoustic, into working memory concurrently with a moderate rate of presentation. Would the working memory representation be inferior to when one set is presented after the other? With concurrent presentation of the two sets, would an analysis of the components of working memory show a greater use of a general capacity-limited attention to maintain information for some age groups than for others? The answer to both questions was no; the results for concurrent presentation were quite similar to successive presentation, in keeping with the prediction of Figure 4, Panel A. For both methods of presentation there was a small cost of retaining two sets that did not change with age, despite a marked increase in the number of items of both types retained in working memory with age; the patterns of

development were quite comparable for successive and concurrent presentation, as shown in Figure 7. Individual differences in aptitude and attention indexed by the Raven Progressive Matrices (Raven, 2000) and by running memory span with spoken digit stimuli (Cowan et al., 2005; Gray et al., 2017), respectively, reflected age group as the main source of variance in ability, apparently large in comparison to within-age variation.

There has been a large variety of research showing developmental improvement in working memory in childhood (e.g., Cowan, 2022; Gathercole et al., 2004; Henry, 2012; Shimi et al., 2014) and developmental improvement in attention, which should be needed to encode stimuli in two sensory modalities (e.g., Ristic & Enns, 2015). None of this research, to our knowledge, indicates that even young children should be as good as adults at encoding into working memory separate visual and auditory stimuli concurrently. Yet, the present study shows that with our moderate presentation rate, it hardly mattered whether series of colors and tones to be remembered in an immediate recognition task were arranged with all of one modality presented before all of the other modality (Figure 1A, successive presentation) or with one color coincident with each tone (Figure 1B, concurrent presentation). The absence of consequence of the arrangement of items could not have been guessed from the extant literature reviewed in the introduction. The common wisdom is still that participants can attend to only one channel of information at a time (e.g., Broadbent, 1958; Cowan et al., 2024), but it was never clear whether that result obtained because the rate of information occurring in both channels was fast, or whether the act of constructing two separate streams of representations at once was too much of a burden on working memory. The present results indicate that even in the early elementary school years, constructing two streams of representations at once is entirely do-able within the overall working memory capacity limit of the individual, which increases markedly with

development.

We allowed 500 ms per pair of items in the concurrent condition and, apparently, encoding and maintenance were as good as they were when the modalities were presented one at a time. We know that there is a limit to the speed of encoding of items into working memory, about 250 ms per item in adults (Vogel et al., 2021) but it does not seem to be a problem to encode two very distinct channels within the constraints of this time window. Or, possibly there is a difficulty of concurrent encoding that is exactly balanced out by the difficulty in the successive-presentation condition of maintaining the first sequence while encoding the second one into working memory. Our results examining the performance levels for the first and second sets in the successive condition support that interpretation, showing better performance for the second set across all age groups.

The concurrent and successive presentation orders of the present study also both replicate what was obtained previously (Cowan et al., 2018) in a successive-presentation condition using the same stimuli as the present study (Figure 3). With maturation, there is no evidence that children start storing more information concurrently in the focus of attention, a form of “central” memory that must be allocated to one or both modalities depending on the task. Instead, with development, there is a dramatic increase in the number of tones and colors remembered in a manner that is impervious to a dual-task cost. The dual-task cost remains the same across age groups, whereas the modality-specific or what we have termed peripheral components of working memory, the parts impervious to dual-task interference, increase markedly with age (and decrease again with adult aging: Greene et al., 2020). This is not to say that attention is uninvolved in the development of working memory; indeed, like Cowan et al., we believe that attention becomes more adept at forming patterns that can be retained with a reduced investment

of attention for maintenance. It presumably occurs by off-loading newly-learned sets of items, which become temporarily activated parts of long-term memory, freeing up the capacity-limited, attention-demanding focus of attention for other activities (Cowan et al., 2024; Rhodes & Cowan, 2018). These may include both encoding of another set into working memory, and intermittent attentional restoration or improvement of the memory representations (Camos & Barrouillet, 2018; Ricker & Vergauwe, 2022) that we take to be held in the activated portion of long-term memory. Assessing this interpretation is a priority for further research.

One limitation of the previous study is that although the rate of stimulus presentation is equated between conditions, the concurrent presentation condition required maintenance of one of the modalities for a shorter period than the first-presented modality in the successive presentation condition. In future work it would be possible to equate the two conditions for the total duration of the presentation period (cf. Ricker et al., 2014), or to insert an unfilled delay between the stimuli to be remembered and the test. Although this kind of experiment theoretically could show superior performance for the successive presentation condition, the present results show that one can speed up the presentation of two modalities by making them concurrent without loss of working memory, regardless of the participant's age from early elementary school to adulthood. That is a result of potential theoretical importance for the theory of cognitive development, and of potential practical importance for teaching methods.

When an individual is confronted with two concurrent stimuli to be encoded into working memory, as in the concurrent presentation condition of the present study, there are likely to be several ways that the task can be carried out. According to a parallel processing method, both stimuli (auditory and visual) can be encoded into working memory concurrently. According to a serial processing method, however, participants could encode one modality, then switch attention

to the sensory memory of the stimulus in the other modality (Cowan, 1988, 1995) to encode it. Previous research suggests that it is quite difficult to distinguish between the serial versus parallel nature of processing (Moran et al., 2016; Townsend, 1990; Treisman, 1988) or between the sharing of resources between two events or switching attention between them (Hitch et al., 2001). Another choice for stimulus presentation could have been to alternate modalities, which would prevent parallel processing of the two modalities but, given that it would impose no more demand than the concurrent presentation of modalities and no less demand than successive presentation of the modalities, we would expect comparable results. Participants should be able to form separate streams for two modalities in alternation (Broadbent, 1956) though this is still an important prediction for future research.

Whatever the best explanation of the findings, they provide preliminary evidence that a multisensory presentation in an educational setting may not be overly taxing to the working memory encoding capabilities of children even early in the elementary school years. In contrast, the data confirm that a key limitation is the total number of separate items to be held concurrently in working memory (Cowan et al., 2018), no matter whether they were presented concurrently or successively.

References

- Bayliss, D.M., Bogdanovsa, J., & Jarrold, C. (2015). Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *Journal of Memory and Language*, 81, 34–50.
Doi:10.1016/j.jml.2014.12.004
- Boxhoorn, S., Lopez, E., Schmidt, C., Schulze, D., Hänig, S., & Freitag, C.M. (2018). Attention profiles in autism spectrum disorder and subtypes of attention-deficit/hyperactivity disorder. *European Child and Adolescent Psychiatry*, 27, 1433–1447.
<https://doi.org/10.1007/s00787-018-1138-8>
- Broadbent, D.E. (1954). The role of auditory localization and attention in memory span. *Journal of Experimental Psychology*, 47, 191-196. Doi:10.1037/h0054182
- 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.
- Camos, V., & Barrouillet, P. (2018). *Working memory in development*. Routledge.
- Case, R. (1995). Capacity based explanations of working memory growth: A brief history and reevaluation. In F.E. Weinert & W. Schneider (eds.), *Memory performance and competencies: Issues in growth and development*. Mahwah, NJ: Erlbaum. (pp. 23 44)
- 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. Doi:10.1037/0033-2909.104.2.163
- Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series, No. 26. New York: Oxford University Press. (Paperback edition: 1997)

- 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. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, 11, 239-264. DOI: 10.1177/1745691615621279
- Cowan, N. (2017a). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, 24, 1158–1170. DOI: 10.3758/s13423-016-1191-6
- Cowan, N. (2017b). Mental objects in working memory: Development of basic capacity or of cognitive completion? *Advances in Child Development and Behavior*, 52, 81-104. doi: 10.1016/bs.acdb.2016.12.001
- Cowan, N. (2022). Working memory development: A 50-year assessment of research and underlying theories. *Cognition*, 224, 105075. doi.org/10.1016/j.cognition.2022.105075
- 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., Saults, 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., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S., Hismjatullina, A., & Conway, A.R.A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51, 42-100. doi: 10.1016/j.cogpsych.2004.12.001

- Cowan, N., Li, Y., Glass, B., & Sauls, J.S. (2018). Development of the ability to combine visual and acoustic information in working memory. *Developmental Science*, 21, e12635, 1-14. doi: 10.1111/desc.12635.
- Cowan, N., Lichty, W., & Grove, T.R. (1990). Properties of memory for unattended spoken syllables. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 258-269. Doi: 10.1037//0278-7393.16.2.258
- Forsberg, A., Blume, C., and Cowan, N. (2021). The development of metacognitive accuracy in working memory across childhood. *Developmental Psychology*, 57, 1297-1317. DOI: 10.1037/dev0001213
- Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. *Journal of Experimental Psychology: General*, 144, 30–47. DOI: 10.1037/a0038211
- Gathercole, S.E., Pickering, S.J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40, 177 190. doi: 10.1037/0012-1649.40.2.17
- Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young school-age children and its relation to intelligence. *Journal of Memory and Language*, 92, 183-201. DOI: 10.1016/j.jml.2016.06.004
- Greene, N.R., Naveh-Benjamin, M., & Cowan, N. (2020). Adult age differences in working memory capacity: Spared central storage but deficits in ability to maximize peripheral storage. *Psychology and Aging*, 35, 866-880. doi.org/10.1037/pag0000476
- Henry, L. (2012). *The development of working memory in children*. London, UK: Sage.
- Hitch, G. J., Towse, J. N., & Hutton, U. (2001). What limits children's working memory span?

- Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General*, 130, 184–198. <https://doi.org/10.1037/0096-3445.130.2.184>
- JASP Team (2023). JASP (Version 0.17.1) [Computer software].
- Johnson, J.A., Strafella, A.P., & Zatorre, R.J. (2007). The role of the dorsolateral prefrontal cortex in bimodal divided attention: two transcranial magnetic stimulation studies. *Journal of Cognitive Neuroscience*, 19, 907-920. doi: 10.1162/jocn.2007.19.6.907
- Jolicoeur, P. (1999). Restricted attentional capacity between sensory modalities. *Psychonomic Bulletin & Review*, 6, 87-92. DOI: 10.1038/42947
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36, 138-202. Doi:10.1006/cogp.1998.0684
- Li, D., Cowan, N., & Sauls, J.S. (2013). Estimating working memory capacity for lists of nonverbal sounds. *Attention, Perception, & Psychophysics*, 75, 145-160. doi: 10.3758/s13414-012-0383-z
- Li, Y., & Cowan, N. (2021). Attention effects in working memory that are asymmetric across sensory modalities. *Memory & Cognition*, 49, 1050-1065. doi.org/10.3758/s13421-021-01142-9
- 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
- McDougal, D.M., Riby, D.M., & Hanley, M. (2020). Profiles of academic achievement and attention in children with and without Autism Spectrum Disorder. *Research in Developmental Disabilities*, 106, 103749, <https://doi.org/10.1016/j.ridd.2020.103749>
- 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.
- Moran, R., Zehetleitner, M., Liesefeld, H.R., Müller, H.J., & Usher, M. (2016). Serial vs. parallel models of attention in visual search: accounting for benchmark RT-distributions. *Psychonomic Bulletin & Review*, 23, 1300–1315. <https://doi.org/10.3758/s13423-015-0978-1>
- Morey, C. C., Morey, R. D., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*, 69, 324–348. doi.org/10.1016/j.jml.2013.04.004
- Pascual-Leone J. & Johnson, J. M. (2021). *The working mind: Meaning and mental attention in human development*. Cambridge, MA: MIT Press.
- Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica*, 32, 301-345. Doi:10.1016/0001-6918(70)90108-3
- Penney, C.G. (1989). Modality effects and the structure of short term verbal memory. *Memory & Cognition*, 17, 398 422. Doi:10.3758/BF03202613
- Pind, J., Gunnarsdóttir, E.K., & Jóhannesson, H.S. (2003). Raven's Standard Progressive Matrices: new school age norms and a study of the test's validity. *Personality and Individual Differences*, 34, 375-386. Doi:10.1016/S0191-8869(02)00058-2
- Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology*, 41, 1–48. Doi:10.1006/cogp.1999.0735
- Ricker, T.J., Nieuwenstein, M.R., Bayliss, D.M., & Barrouillet, P. (2018). Working memory consolidation: insights from studies on attention and working memory. *Annals of the New*

- York Academy of Sciences*, 1424, 8-18. doi: 10.1111/nyas.13633.
- Rhodes, S., & Cowan, N. (2018). Attention in working memory: Attention is needed but it yearns to be free. *Annals of the New York Academy of Science*, 1424, 52-63. doi: 10.1111/nyas.13652
- Ricker, T.J., & Cowan, N. (2014). Differences between presentation methods in working memory procedures: A matter of working memory consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 417-428. doi: 10.1037/a0034301
- Ricker, T.J., & Vergauwe, E. (2022). Boundary conditions for observing cognitive load effects in visual working memory. *Memory & Cognition*, 50, 1169–1185.
<https://doi.org/10.3758/s13421-022-01320-3>
- Ristic, J., & Enns, J.T. (2015). The changing face of attentional development. *Current Directions in Psychological Science*, 24, 24-31. DOI: 10.1177/0963721414551165
- Rosen, V.M., & Engle, R.W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, 126, 211-227. doi: 10.1037//0096-3445.126.3.211.
- Santangelo, V. (2018). Large-Scale Brain Networks Supporting Divided Attention across Spatial Locations and Sensory Modalities. *Frontiers in Integrative Neuroscience*, 12,
- Savage, R., Cornish, K., Manly, T., & Hollis, C. (2006). Cognitive processes in children's reading and attention: The role of working memory, divided attention, and response inhibition. *British Journal of Psychology*, 97, 365–385. doi: 10.1348/000712605X81370
- Schönbrodt, F. D., & Stefan, A. M. (2018). BFDA: An R package for Bayes Factor Design Analysis, version 0.4.0. <https://github.com/nicebread/BFDA>
- Schönbrodt, F. D. & Wagenmakers, E.-J. (2018). Bayes Factor Design Analysis: Planning for

- compelling evidence. *Psychonomic Bulletin & Review*, 25, 128-142. doi:10.3758/s13423-017-1230-y.
- Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting attention within visual short-term memory: Development and mechanisms. *Child development*, 85, 578-592. Doi: 10.1111/cdev.12150
- Sinzig, J., Bruning, N., Morsch, D., & Lehmkuhl, G. (2008). Attention profiles in autistic children with and without comorbid hyperactivity and attention problems. *Acta Neuropsychiatrica*, 20, 207–215. doi:10.1111/j.1601-5215.2008.00292.x
- Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42, 1096-1104. DOI: 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. Doi:10.1177/0098628316677643
- Talalay, I. V. (2024). The development of sustained, selective, and divided attention in school-age children. *Psychology in the Schools*. <https://doi.org/10.1002/pits.23160>
- Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, 1, 46–54. doi:10.1111/j.1467-9280.1990.tb00067.x
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. The *Quarterly Journal of Experimental Psychology*, 40, 201–237. DOI: 10.1080/02724988843000104
- Uittenhove, K., Chaabi, L., Camos, V., & Barrouillet, P. (2019). Is working memory storage intrinsically domain-specific? *Journal of Experimental Psychology: General*, 148, 2027–

2057. DOI: 10.1037/xge0000566

Vogel, E.K., Woodman, G.F., & Luck, S.J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1436-1451. DOI:10.1037/0096-1523.32.6.1436

Williams, J.R., Robinson, M.M., Schurgin, M.W., Wixted, J.T., & Brady, T.F. (2022). You cannot "count" how many items people remember in visual working memory: The importance of signal detection-based measures for understanding change detection performance. *Journal of Experimental Psychology: Human Perception and Performance*, 48, 1390-1409. doi: 10.1037/xhp0001055.