Psychology and Aging

Age-Related Decline in Visual Working Memory: The Effect of Nontarget Objects During a Delayed Estimation Task

A. Caglar Tas, Matthew C. Costello, and Aaron T. Buss

Online First Publication, February 27, 2020. http://dx.doi.org/10.1037/pag0000450

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychology and Aging

ISSN: 0882-7974

http://dx.doi.org/10.1037/pag0000450

Age-Related Decline in Visual Working Memory: The Effect of Nontarget Objects During a Delayed Estimation Task

A. Caglar Tas
University of Tennessee, Knoxville

Matthew C. Costello
University of Hartford

Aaron T. Buss
University of Tennessee, Knoxville

Visual working memory (VWM) is an essential aspect of cognitive functioning that becomes compromised in older adults. A canonical probe of VWM is the change detection task in which participants compare a visually presented stimulus with items being maintained in VWM. Older adults show a decreased ability to detect changes between a stimulus and the contents of VWM compared with younger adults. Previously, we used a dynamic neural field (DNF) model to explore changes in neural connectivity that can explain this pattern of decline in performance. These simulations suggest that older adults have cortical interactions that are more diffuse compared to younger adults. In the current article, we examined the precision of representations in VWM using the delayed-estimation task. Participants are first presented with a memory array. After a delay, a location is cued, and participants click on a color wheel to indicate which color was at that location. The model predicted that older adults should show increased guessing rates and decreased precision (defined as the variability of color responses around the target location) relative to younger adults. The model also predicted that presenting the nontarget items during test should improve the precision of responses for older adults but not for younger adults. Results from two experiments supported these predictions of the model. These findings further advance an emerging theory of the neurocognitive decline of VWM and illustrate how older adults' VWM representations are influenced by the context in which information is being recalled.

Keywords: visual working memory, aging, delayed estimation task, dynamic field theory

Supplemental materials: http://dx.doi.org/10.1037/pag0000450.supp

Aging negatively affects many cognitive abilities, such as perception (e.g., Kline & Scialfa, 1997), attention (e.g., Zanto & Gazzaley, 2014), and executive function (e.g., Salthouse, Atkinson, & Berish, 2003). One of the most extensively studied areas for age-related cognitive decline is working memory (Park & Payer, 2006). Visual working memory (VWM) refers to a subset of the working memory system which allows short-term storage of visual information (for a review, see Luck & Vogel, 2013). In general, VWM is capacity-limited, that is, only a portion of the visual information is encoded into memory at a given time. VWM is essential during tasks which require keeping track of and updating of visual information, such as driving a car or completing a visual search. VWM peaks around age 20 and then declines steadily with advanced age (Brockmole & Logie, 2013). Further, performance in VWM tasks is strongly correlated with fluid intelligence (Fukuda, Vogel, Mayr, & Awh, 2010) as well as general cognitive abilities (M. K. Johnson et al., 2013), both of which have been shown to be compromised in older adults (Salthouse, 2010; Wecker, Kramer, Wisniewski, Delis, & Kaplan, 2000). Given its role in everyday life and general well-being, it is important to determine the mechanisms for why aging affects VWM.

Measuring the Capacity of Visual Working Memory

Traditionally, VWM capacity has been assessed with the change-detection task in which participants are presented with an array of visual stimuli to memorize. After a brief delay, they are presented with a second array of stimuli and asked to indicate whether all items remained the same or one of them has changed. Performance in this task is typically assessed by calculating mem-

A. Caglar Tas, Department of Psychology, University of Tennessee, Knoxville; Matthew C. Costello, Department of Psychology, University of Hartford; Aaron T. Buss, Department of Psychology, University of Tennessee, Knoxville.

The behavioral data in this study were presented at the 70th Annual Scientific Meeting of the Gerontological Society of America in November 2018. Part of this research was funded by NIH RO1HD092485 to Aaron T. Buss.

Correspondence concerning this article should be addressed to A. Caglar Tas, Department of Psychology, University of Tennessee, Knoxville, 1404 Circle Drive, Austin Peay Building, Knoxville, TN 37996. E-mail: caglartas@utk.edu

ory capacity using false alarms and hit rates. Previous studies which utilize the change detection task have found that average VWM capacity for younger adults is around three to four items (e.g., Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001). Although there is a debate about the nature of age-related decline in working memory (e.g., Naveh-Benjamin, 2000; Hasher & Zacks, 1988), studies using older adults have consistently shown that change detection performance significantly decreases with aging (e.g., Brockmole, Parra, Della Sala, & Logie, 2008; Brown & Brockmole, 2010; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Ko et al., 2014).

An alternative to the change-detection paradigm is the delayed-estimation task, in which participants are asked to estimate the feature value of an item across a range of possible values (e.g., Bays, Catalao, & Husain, 2009). For instance, to test color VWM, participants are first given an array of colors to encode. After a brief delay, the location of one of the array objects is cued. The task is to select the color of this cued object from an array of colors on a color wheel (Figure 1A). Color response distributions are then fitted with a maximum-likelihood estimate method to estimate the probability of reporting the correct color as well as the precision with which the item is recalled (see Results for more details of this method). Studies using this method have found that the standard deviation of the distributions increases as the number of objects in the study array increases, indicating that as one encodes more items in VWM the precision of each item decreases (Bays et al., 2009). Because the responses are collected from a continuous array, an important advantage of this method over the change detection task is that it is possible to estimate the source of error during recall. Specifically, this method allows us to estimate the probability of reporting a nontarget item (i.e., one of the noncued color values) and the probability of reporting a random color value that was not in the study array (i.e., random guesses). In addition, this method also allows us to determine not just how accurately a participant responds (whether they clicked near the correct feature value or not), but also how precise a representation is based on the width or variability of the response distribution. Using the delayed-estimation task, Peich, Husain, and Bays (2013) found two important effects of aging. First, older adults were more likely to report the features of one of the uncued memory items compared to younger participants, indicating more feature-location binding errors. Second, the precision of the recalled item was also significantly lower (i.e., higher variability for the correct target responses) for older adults than for younger adults, especially for larger memory arrays.

A Dynamic Neural Field Model of Age-Related Decline in VWM

To explore age-related differences in cognition, specifically in VWM, we have previously developed a dynamic neural field (DNF) model which allows us to examine the neural mechanisms of aging (Costello & Buss, 2018). DNF models are framed around the concept of an activation field that specifies the temporal activation dynamics within populations of neurons which are tuned to perceptual and/or motor dimensions. The basic idea is that cognitive states correspond to stable patterns of activation within populations of neurons that are tuned to perceptual or motor dimensions. By simulating the temporal evolution of neural activation, then, we can make inferences and predictions about cognition and behavior (for a review, see Johnson, Simmering, & Buss, 2014; Spencer & Schöner, 2003). Neural units in the model are connected through local-excitation and lateral-inhibition interactions that are defined by how strongly and how broadly neural

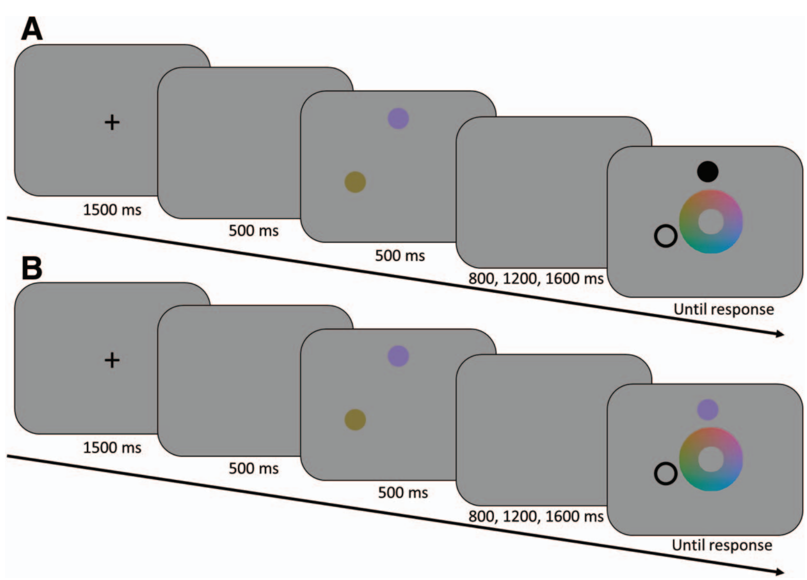

Figure 1. Sequence of events in a sample trial in Experiments 1 (Panel A) and 2 (Panel B). All colors were sampled from the CIE L*a*b color space. The color wheel showed the range of all possible color values in the color space. In both experiments, participants were asked to report the color for the cued (empty circle) item. In Experiment 2, the nontarget items were presented on the screen during test. See the online article for the color version of this figure.

output projects to other units. Units that are tuned to similar information share excitation with one another, and units that are tuned to different information inhibit one another. Through the balance of excitatory and inhibitory interactions, activation within the model can form stabilized "peaks" of activation that correspond to the representations of a stimulus value (e.g., a color hue) or motor plan (i.e., a motor plan to press a button). The equations and dynamics used in this modeling framework have been demonstrated to correspond to actual neural dynamics measured in awake, behaving animals (Jancke et al., 1999; Markounikau, Igel, Grinvald, & Jancke, 2010). Moreover, the equations that are used in DNF models can be reconstructed from neural population recordings using a distribution of population activation approach (Erlhagen, Bastian, Jancke, Riehle, & Schöner, 1999; for a more comprehensive review of the development and application of this modeling approach see Schöner, Spencer, & The DFT Research Group, 2015).

The architecture of the DNF model is composed of multiple populations of feature-sensitive (e.g., color or shape) neurons. In the architecture displayed in Figure 2, there are two main excitatory fields: the contrast field (CF) and working memory (WM) field. Stimuli are presented strongly to CF and weakly to WM. CF and WM both have self-excitatory connections. CF is involved with encoding perceptual information into the WM field and sends excitation to WM which, in turn, maintains active representations of stimuli that persist when the stimuli are no longer present. Inhibitory interactions within these fields are mediated by the Inhib field which receives activation from both the CF and WM fields. Inhibition sent to the CF serves to suppress encoding of stimuli that are activated in WM. Inhibition sent to the WM field serves to stabilize peaks of activation that represent the items stored in VWM and enables stable representations by preventing these peaks from shifting in location or spreading beyond the represented value.

This model architecture is capable of generating active same/different decisions in the change detection task and produces

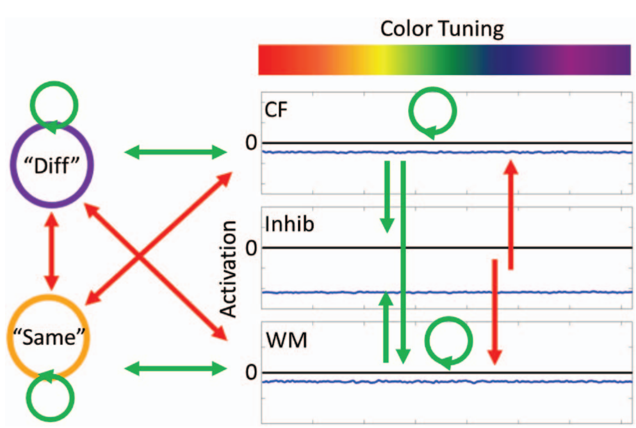

Figure 2. Architecture of DNF model of change detection task. Blue (wavy) lines show the activation fields. Black (solid) lines mark the activation threshold (at 0). Synaptic output is sent once activation crosses this activation threshold. Excitatory interactions are marked with green (light gray) arrows and inhibitory interactions are marked with red (dark gray) arrows. CF = contrast field; WM = working memory. See the online article for the color version of this figure.

variation in performance accuracy that mimics rates of performance by human participants (Costello & Buss, 2018; J. S. Johnson, Spencer, Luck, & Schöner, 2009; Simmering, 2016). The model generates these decisions by coupling the activation within the WM and CF layers to "same" and "different" decision nodes, respectively. Specifically, activation within the CF is integrated and sent to the "different" decision node, and activation within the WM field is integrated and sent to the "same" decision node. In this way, the "different" and "same" node compete to become activated once the test display is presented to the model (for more details of the model architecture, see Costello & Buss, 2018).

Costello and Buss (2018) used this model to explore whether alterations to neural interactions in the model could explain patterns of performance of younger and older adults in color and shape change detection tasks. The strength and width of the excitatory and inhibitory neural interactions were systematically manipulated to find which parameter values reproduced performance of older adults. Three different manipulations were able to explain age-related differences in performance between younger and older adults. To explore the mechanistic basis of these changes in performance, Costello and Buss (2018) examined the properties of representations within the model. Figure 3 shows the average number of peaks maintained within the model at the highest set size (SS5) as well as the number of neural units that, on average, participated in each representation within WM. In one model, the widths of excitatory and inhibitory interactions were increased (excitation and inhibition width model). Under this manipulation, neural interactions within the model were less precise and less stable. That is, the model that simulated older adult performance had more neural units participating in each representation and a decrease in the number of peaks that could be maintained. In a second model, the strength of input to the Inhib layer was decreased (to-inhibition strength model). Under this manipulation, inhibitory interactions were engaged more weakly and more slowly resulting in a shift in the balance of excitatory and inhibitory neural interactions. As a result of this manipulation, peaks actually became more stable (reflected by an increase in the average number of peaks) and wider (reflected by the increased number of neural units participating in the average representation within WM). In a third model, the width of only inhibitory interactions was increased (inhibitory width model). Under this manipulation, inhibitory interactions were more diffuse within WM and CF. Consequently, peaks not only became less stable but also narrower due to stronger lateral-inhibition near the point of activation.

The results from feature estimation studies can be informative as to which of these neural mechanisms best captures age-related differences in VWM. Specifically, the Peich et al. (2013) study reviewed above suggests that older adults have less precise representations as reflected by an increase in the standard deviation of the response distributions, and older adults also tend to forget items as reflected by an increase in the probability of random guessing. To map the model dynamics onto the response properties from feature-estimation tasks, we can assume that the guessing rate would be proportional to the likelihood of losing peaks before the test phase. Further, the width of the response distributions can be assumed to reflect the width of neural representations within the

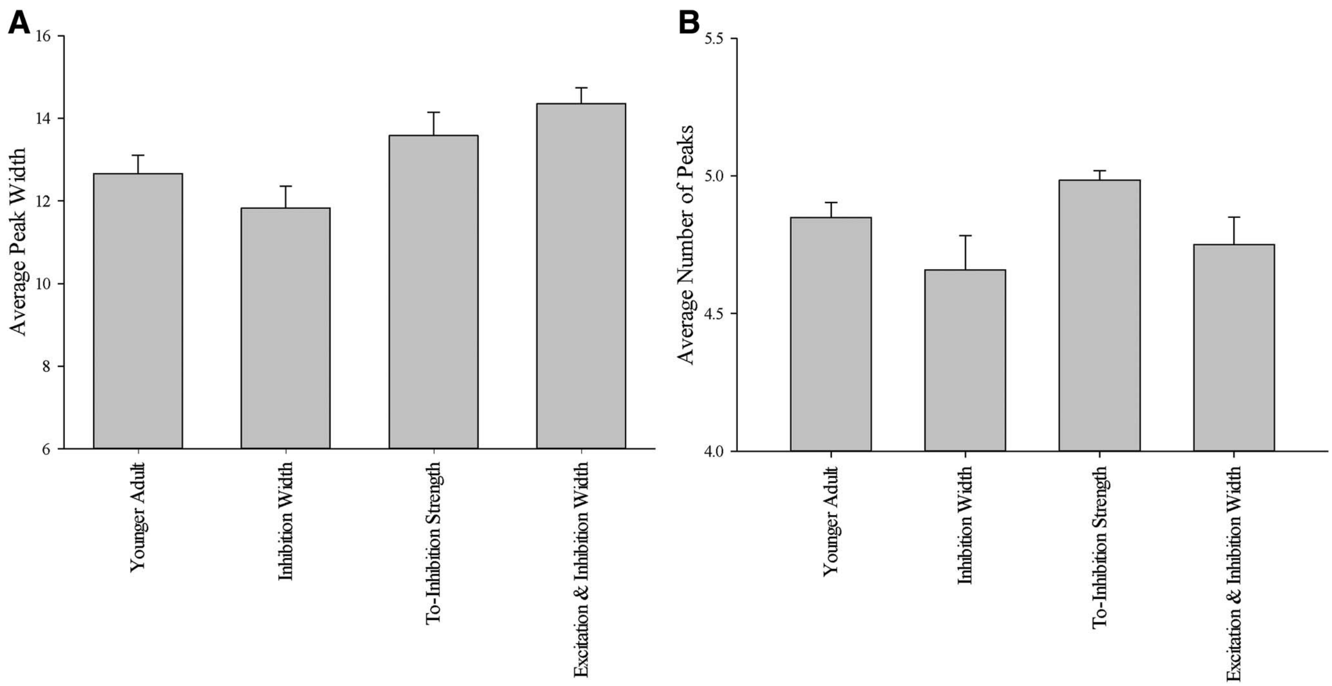

Figure 3. Average number of neural units participating in a peak (width, Panel A) and average number of peaks within the WM field (Panel B) for the different models used to explore age-related changes in VWM. Data are reproduced for set size 5 from Costello and Buss (2018).

WM field.[1] That is, at the time of color-report, participants' responses will reflect sampling from the distribution of neural units underlying the probed representation. In the model, the VWM representations will have sustained peaks of activation around the color value that is being remembered. When one of the memory locations is cued, these VWM peaks would send activation to the perceptual representation of the color wheel, building a response peak around the remembered color value. If the VWM peak is narrow, then it would project to a more restricted range of the color wheel, leading to more precise responses. A wider peak, on the other hand, would project activation to a broader range of the color wheel, leading to more variable responses. Although this method of estimating response distributions does not involve an active response mechanism, the aim is to be consistent with how responses might be actively generated by future adaptations of this model. Based on this conceptualization of the VWM representations, we have two distinctions between younger and older adult models. Older model should be able to maintain fewer VWM representations and have wider VWM peaks at test. Among the three models identified by Costello and Buss (2018), only the excitation and inhibition width model showed both a decrease in the number of peaks and an increase the width of peaks relative to the young adult model.

Present Study

In the present study, we tested the predictions from the DNF model regarding how color reports are influenced by manipulations to the task. One aspect of the model that is distinct from previous accounts which have focused on encoding and maintenance stages is that the model is sensitive to the display characteristics during test. Feature-estimation tasks typically cue a single location and ask participants to report the feature that was at that location. We examined the effect of presenting the nontarget colors during recall in addition to the spatial cue indicating which item to report. To explore this manipulation in the model, we administered conditions which presented either a single cue during the test phase, or the nontargets along with the cue for the test item during the test phase at Set Size 3. We then compared the average width of peaks in the model during the test phase. For 200 runs of the "younger" model, there were no differences based on presenting the nontarget items during test, $t(198) = 0.51$, $p = .614$, $d = 0.07$.

[1] Note that other application of DNF models have used center-of-mass of peaks as a metric of precision of representations (Schutte & Spencer, 2009). Our task is different than those studies in two major ways. First, in the previous applications, the focus of the study was the drift of spatial memory over time, and the change in the center-of-mass was integral to the performance of the model. These drifts in memory were implemented by presenting stimuli that were spatially close. In our current study, however, our colors were at least 60° apart around the color wheel. Therefore, we can assume that the peaks were more stable in our model. Second, the mapping of a VWM representation to a response is more complex in our application of the DNF model. The spatial memory tasks used stimuli and responses that were directly aligned—the memory and the response were in the same frame of reference. In the color estimation task, however, a VWM representation of a color at a spatial location is cued, and this feature-space conjunction must then be recalled to select a spatial location on the color wheel. To be able to simulate both versions of the working memory tasks, we are currently expanding the model architecture to incorporate spatial-feature bindings and an active response mechanism that can build a peak in the response space defined by the color wheel.

For the "older" model, however, there was a significant reduction in the number of neural units participating in representations within the WM field when the nontarget items were presented compared with when only the probe item was presented, $t(198) = 2.56$, $p = .011$, $d = 0.36$. Thus, the model makes the qualitative prediction that older adults, but not younger adults, should show a decrease in the standard deviations of the response distribution (i.e., more precise memory representations) associated with color recall when nontarget items are presented during the response phase. Mechanistically, this decrease results from the increased inhibition associated with the presentation of more stimuli. Inhibition within the WM field is dependent upon the level of excitation that goes into the Inhib field. Stronger inhibition has the consequence of dampening excitation within the WM field and narrowing the width of peaks within this field, effectively sharpening the representations of the items in memory.

In Experiment 1, we first tested (a) whether older adults are more likely to have more random guesses than younger adults, and (b) whether older adults' color reports have higher standard deviation (i.e., less precise) than younger adults' color reports. Finding evidence for these two hypotheses would both replicate Peich et al. (2013) findings and provide evidence for the predictions made by the DNF model (Costello & Buss, 2018). More importantly, we tested a unique prediction made by the DNF model regarding the presence of nontarget colors during test display: Presenting nontargets during the test should improve the precision of older adults', but not younger adults', color reports. To achieve this, we ran Experiment 2 where we presented the nontargets during test and compared color report performance of older and younger participants across the two experiments.

Experiment 1

Method

Participants. Twenty-nine younger adults (nine male, 20 female; age range: 18–40, $M = 19.9$ years) and 31 older adults (nine male, 22 female; age range: 61–90, $M = 73.0$ years) participated in Experiment 1.[2] The younger group consisted of undergraduate students from the University of Hartford who received course credit for their participation. The older group were recruited from the greater Hartford area via local newspaper ads and received a gift certificate worth $15. All participants reported normal or corrected-to-normal vision. This project was approved by the Institutional Review Board of the University of Hartford, and participants in both experiments signed an informed consent form.

Stimuli and apparatus. Cognitive and perceptual measures: All participants first completed a battery of cognitive and perceptual tasks: the vocabulary subscale of The Wechsler Adult Intelligence Scale (WAIS-III), Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), forward and backward digit-span, Freiburg Visual Acuity Test (FrACT; Bach, 1996), and 17-plate version of the Pseudoisochromatic Plates color blindness test. These batteries lasted about 20 min. Table 1 displays participants' scores on these cognitive and perceptual tasks.

VWM task. To assess VWM, we used a delayed color estimation task. Stimuli consisted of colored circles, subtended 1 degree of visual angle (dva). The colors of the memory items were chosen from CIE $L^*a^*b^*$ color space where the L parameter of each color was set to 50. First, we randomly selected a color from a set of 360 possible colors equally distributed in the color space. Then, we created six possible colors, in steps of 60° in the clockwise direction from the initial color. Each trial could consist of one, two, or three memory items. For each set sizes, we randomly chose one (Set Size 1 [SS1]), two (Set Size 2 [SS2]), or three (Set Size 3 [SS3]) colors from the set of possible six colors. Therefore, the minimum separation between two colors was 60°. Spatial locations were sampled in the same way from a 360° space around an imaginary circle with a radius of 5dva. For the test cue, we used a black annulus with an outer radius of 1dva and an inner annulus of 0.5dva. All stimuli were presented against a neutral gray background. The color wheel used for collecting responses was an annulus with an outer radius of 8dva and an inner radius of 3dva. To eliminate any spatial biases, we created 36 color wheels rotated in steps of 10° and randomly selected one of the color wheels on each trial. Participants were seated approximately 53 cm from the monitor screen in a small quiet room. Head position was not restricted. Stimuli were presented on a 17″ LCD monitor with 60-Hz refresh rate. Stimulus presentation was controlled with E-prime 2.0 software (Schneider, Eschmann, & Zuccolotto, 2002).

Procedure. Each trial started with a fixation cross at the center of the screen which lasted for 1,500 ms (see Figure 1). After a delay of 500 ms, the memory array was presented for 500 ms which was followed by a retention interval. The memory array consisted of one (SS1), two (SS2), or three (SS3) colored circles. The retention interval was 800 ms, 1,200 ms, or 1,600 ms, randomly sampled in equal frequencies. At the end of each trial, participants were presented with the test array which stayed on the screen until response. For SS1, the test array consisted of the test cue at the location of the memory item and the color wheel at the center of the screen. For SS2 and SS3, one of the memory items was randomly chosen to be the test item. The location of this test item was cued with a black annulus, and the remaining memory items were replaced with black circles. Therefore, the test array in SS2 and SS3 trials consisted of the test cue, black placeholders, and the color wheel at the center of the screen (Figure 1A). The next trial started 1,500 ms after participants made a response. Participants completed a total of 15 blocks, five blocks of 15 trials for each set size, a total of 225 trials. Sessions lasted for approximately 30 min.

Data analyses. To determine the quality of the memory representations, each participant's response distributions were fitted with probabilistic mixture models formulated below:

$$p(x) = p_t \Phi_{\mu t, \kappa t}(x - \theta_t) + p_n \Phi_{\mu n, \kappa n}(x - \theta_n) + p_r / 2\Pi \qquad (1)$$

where x, θ_t, and θ_n refer to the reported color value, color value of the target item, and color value of the nontarget item, respectively (Bays et al., 2009). We calculated the probabilities of reporting the correct color value (p_t), the probability of reporting one of the nontarget items (p_n), and the probability of reporting a random color value (p_r). In the formula, $\phi_{i,k}$ refers to the probability

[2] We ran power analyses using MorePower 6.0.4 program (Campbell & Thompson, 2012) with the effect size calculated from the effect of aging on color memory precision in Peich et al. (2013) study (the effect of age for color in the high-load condition). The results showed that for our 3×2 mixed design, we need a minimum of 24 participants in each age group to achieve a minimum power of .90 for $\eta_p^2 = .241$.

Table 1
Mean and Standard Deviations of the Participant Demographics and Their Scores on the Cognitive and Perceptual Tasks in Experiments 1 and 2

Scale	Experiment 1			Experiment 2		
	Younger	Older	*p*	Younger	Older	*p*
Age (years)	19.90 (4.50)	73.03 (7.84)	<.001	18.83 (0.89)	71.60 (6.91)	<.001
Education (years)	13.41 (1.05)	16.21 (3.06)	<.001	13.21 (1.08)	17.90 (2.58)	<.001
WAIS-III (vocabulary subscale)	47.04 (6.72)	55.13 (6.93)	<.001	49.46 (6.40)	58.13 (4.74)	<.001
Forward digit span	7.04 (0.98)	6.84 (1.37)	.529	6.62 (0.86)	7.13 (1.17)	.061
Backward digit span	4.90 (1.08)	4.97 (1.35)	.823	4.76 (0.91)	4.97 (1.54)	.533
FRACT—Test 1	−0.19 (0.13)	0.08 (0.19)	<.001	−0.23 (0.13)	−0.001 (0.32)	.001
FRACT—Test 2	−0.16 (0.14)	0.07 (0.18)	<.001	−0.24 (0.11)	−0.03 (0.27)	<.001
Color blindness	17.00 (0)	16.61 (0.80)	.012	16.72 (0.65)	16.33 (1.24)	.137
MMSE	29.21 (1.01)	28.63 (0.80)	.103	28.83 (1.17)	29.30 (0.88)	.083

Note. WAIS-III = Wechsler Adult Intelligence Scale; FRACT = Freiburg Visual Acuity Test; MMSE = Mini-Mental State Examination. Standard deviations are presented in parentheses. FRACT scores are reported in logMar, with 0.0 equivalent to Snellen 20/20 (Bach, 1996).

density function of von Mises distribution with a mean of μ and a concentration of κ. With this model, we also estimated the standard deviation of the response distributions as a means of measuring the precision of the reports ($SD = \sqrt{1/\kappa}$). Because there was only one color on SS1 trials, we only included the target and random guessing distributions in the model to estimate the p_t and p_r ($p_t + p_r = 1$). For SS2, we used the formula depicted in Equation 1 where one of the colors was assigned as the target (p_t) and the other one was assigned as the nontarget (p_n; $p_t + p_n + p_r = 1$). Lastly, for SS3 we included two nontarget distributions in the equation. Therefore, in addition to the p_t and p_r, we estimated the probabilities of reporting two nontarget items (p_{n1} and p_{n2}; $p_t + p_{n1} + p_{n2} + p_r = 1$). We compared the probabilities of reporting the correct target values, the nontarget values, and random guesses between the younger and older adult groups. To determine whether aging affects the precision of the memory items, we also compared the standard deviations of the target distributions between the two age groups.

Results

Probability of reporting the correct target value (p_t). Analyses revealed that delay did not impact any of the performance measures nor did it interact with any other factors, so all analyses below collapsed across this factor.[3] To test the role of set size and age on reporting the correct target color, we ran a 3 (set size: 1, 2, 3) × 2 (age: older, younger) ANOVA on the p_t data, where set size was entered as a repeated measure and age was a between-subjects variable. The results showed a significant main effect of set size, $F(2, 116) = 48.8$, $p < .001$, $\eta_p^2 = .457$ (Figure 4A). Bonferroni corrected pairwise comparisons showed that participants had significantly higher p_t values for SS1 than for SS2, $t(59) = 5.03$, $p < .001$, $d = 0.86$, for SS2 than for SS3, $t(59) = 4.88$, $p < .001$, $d = 0.52$, and also for SS1 than for SS3, $t(59) = 8.51$, $p < .001$, $d = 1.42$. We also found a significant main effect of age, $F(1, 58) = 17.1$, $p < .001$, $\eta_p^2 = .228$. Younger adults were more likely to report the correct target color compared with older adults. Importantly, there was a significant Set Size × Age interaction, $F(2, 116) = 9.61$, $p < .001$, $\eta_p^2 = .142$. Follow-up analyses showed that the probability of reporting the target color did not differ between older and younger participants for SS1, $t(58) = -1.15$, $p = .255$, $d = 0.30$. However, older participants had significantly lower p_t values compared with younger participants for both SS2, $t(58) = -2.86$, $p = .006$, $d = 0.74$, and for SS3, $t(58) = -4.55$, $p < .001$, $d = 1.18$.

Probability of reporting a nontarget value (p_n). We conducted a 2 (set size: 2, 3) × 2 (age: older, younger) mixed ANOVA on p_n values (Figure 4B). For SS3, we combined the reports for both nontarget colors. The results showed a significant main effect of set size, $F(1, 58) = 14.9$, $p < .001$, $\eta_p^2 = .204$, and a significant main effect of age, $F(1, 58) = 7.31$, $p = .009$, $\eta_p^2 = .112$. Participants were more likely to report a nontarget color for SS3 than for SS2. Further, older participants reported significantly more nontarget colors than the younger participants. Importantly, the Set Size × Age interaction was also significant, $F(1, 58) = 8.03$, $p = .006$, $\eta_p^2 = .122$. Follow-up analyses showed that p_n values did not differ between different age groups for SS2, $t < 1$. However, older participants reported significantly more nontarget colors for SS3 compared with the younger participants, $t(58) = 3.25$, $p = .002$, $d = 0.84$.

Probability of reporting a random color value/guessing (p_r). The first hypothesis the DNF model predicted was that older adults are more likely to randomly guess during an incorrect trial than younger adults. We investigated this hypothesis by comparing the p_r values using a 3 (set size: 1, 2, 3) × 2 (age: older, younger) mixed ANOVA (Figure 4C). We found a significant main effect of set size, $F(2, 116) = 48.8$, $p < .001$, $\eta_p^2 = .457$, a significant main effect of age, $F(1, 58) = 17.1$, $p < .001$, $\eta_p^2 = .228$, and a significant Set Size × Age interaction, $F(2, 116) = 9.61$, $p < .001$, $\eta_p^2 = .142$. In line with our first hypothesis, pairwise comparisons revealed that older participants made significantly more random guesses than younger participants. Further, participants made significantly more random guesses for SS2 than for SS1, $t(59) = 3.76$, $p < .001$, $d = 0.65$, for SS3 than for SS1, $t(59) = 6.72$, $p < .001$, $d = 1.14$, and for SS3 than for SS2, $t(59) = 3.58$, $p = .001$, $d = 0.36$. Lastly, p_r values were similar for older and younger participants for SS1, $t(58) = 1.15$, $p = .225$, $d = 0.30$. However, compared with the younger participants, older participants had

[3] The analyses and figures depicting data for different delay conditions can be found in the online supplementary materials.

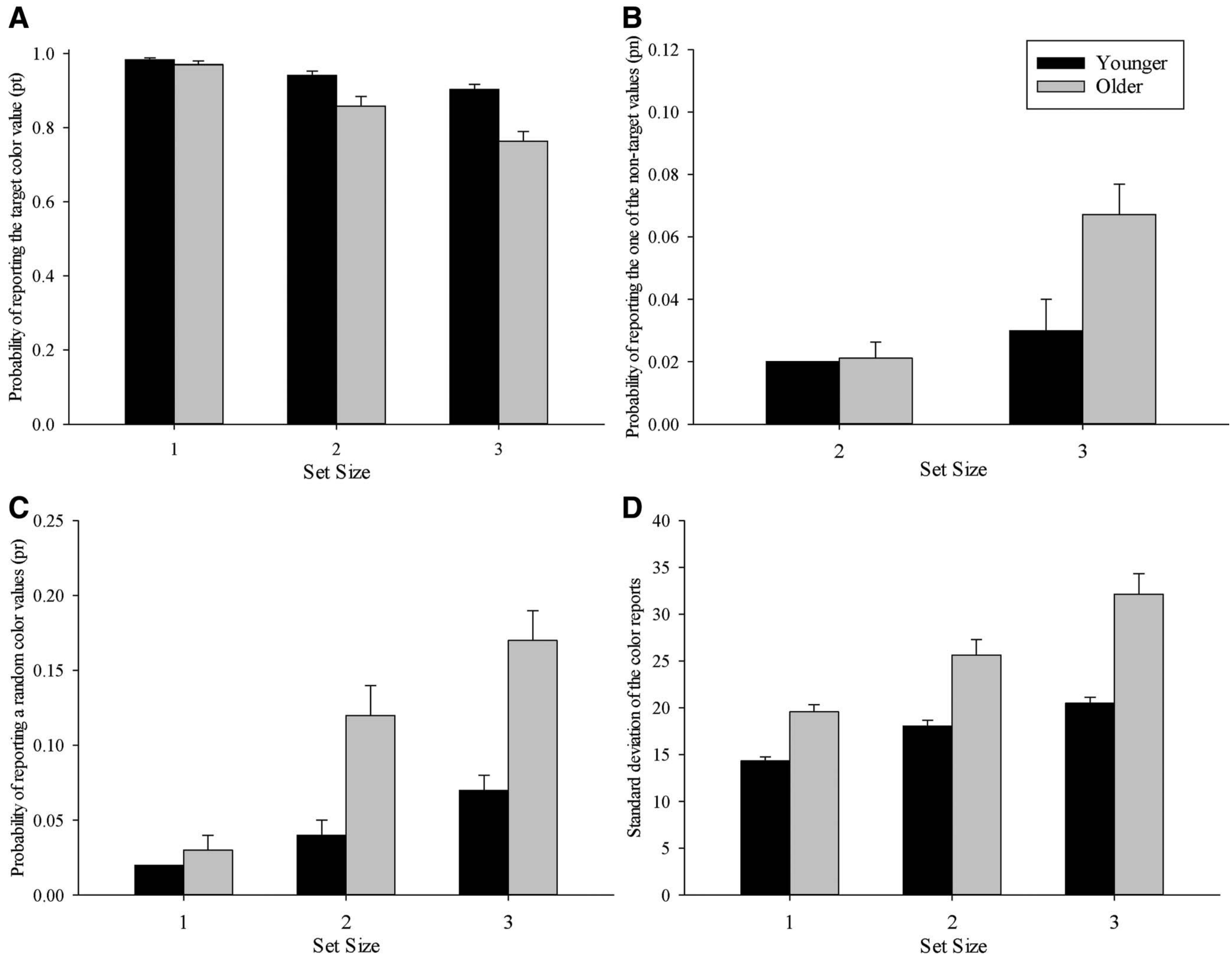

Figure 4. Experiment 1 results. Panel A shows the probability of reporting the target color value (p_t), Panel B shows the probability of reporting a nontarget value (p_n), Panel C shows the probability of reporting a random value (p_r), and Panel D shows the standard deviation of the response distribution for the target value across different set sizes and ages. Error bars represent 95% confidence intervals.

significantly higher p_r values for SS2, $t(58) = 3.05$, $p = .003$, $d = 0.79$, as well as for SS3, $t(58) = 4.10$, $p < .001$, $d = 1.06$.

Standard deviations of the target distributions (precision of correct reports). The second hypothesis of the DNF model was that aging also negatively affects the precision of the color reports. To test this prediction, the same 3 (set size: 1, 2, 3) × 2 (age: older, younger) mixed ANOVA was run on the standard deviations of the color report distributions (Figure 4D). We found a significant main effect of set size, $F(2, 116) = 65.9$, $p < .001$, $\eta_p^2 = .532$, suggesting that color reports were more precise for SS1 than for SS2, $t(59) = -7.28$, $p < .001$, $d = 0.77$, for SS1 than for SS3, $t(59) = -9.36$, $p < .001$, $d = 1.14$, and for SS2 than for SS3, $t(59) = -5.26$, $p < .001$, $d = 0.48$. Importantly, supporting our second hypothesis, we found a significant main effect of age, $F(1, 58) = 28.4$, $p < .001$, $\eta_p^2 = .328$, where older participants' reports were significantly less precise than younger participants' reports. Lastly, we found a significant Set Size × Age interaction, $F(2, 116) = 7.88$, $p = .001$, $\eta_p^2 = .120$. Follow-up analyses showed that older participants' memory reports were significantly less precise than the younger participants' reports for all set sizes, but this difference became larger as set size increased, $t(58) = 5.69$, $p < .001$, $d = 1.47$ for SS1, $t(58) = 4.23$, $p < .001$, $d = 1.09$ for SS2, and $t(58) = 4.91$, $p < .001$, $d = 1.27$ for SS3.

Discussion

In Experiment 1, we tested the role of aging on the capacity and the precision of VWM representations. First, we found no effect of delay on any of the parameters, meaning that longer delays do not affect the properties of VWM representations. Second, we found an effect of set size on all parameters. Specifically, as the set size increases the probability of reporting the correct color decreases, precision of the memory representation decreases, and the probability of reporting the nontargets and random guessing increase.

More importantly, we found a significant effect of age on all parameters. Older participants were more likely to report one of the nontarget items or a random color compared with younger participants. Older participants' memory representations were also less precise than younger participants' representations. Finally, these age differences became larger as set size increased. Specifically, older participants were less likely to correctly report the target color when there was more than one color to remember, and also more likely to report a nontarget color if there was more than one nontarget color. These results both replicate previous findings (Peich et al., 2013) and support the predictions of the DNF model (Costello & Buss, 2018). In Experiment 2, we tested the last prediction of the DNF model that presenting the nontarget colors during recall will increase the precision of the memory representations for older adults, but not for younger adults.

Experiment 2

Method

Participants. Thirty-one younger adults and 30 older adults participated in Experiment 2. None of the participants took part in Experiment 1. Data from one participant from the younger group were excluded from the analyses, because they had very low color blindness test scores. The final sample consisted of 29 younger adults (11 male, 18 female; age range: 18–21, $M = 18.8$ years) and 30 older adults (four male, 26 female; age range: 60–82, $M = 71.6$ years). The younger group consisted of undergraduate students from the University of Hartford who received course credit for their participation. The older group were recruited from the greater Hartford area via local newspaper ads and received a gift certificate worth $15. All participants reported normal or corrected-to-normal vision.

Stimuli and apparatus. Stimuli and apparatus were the same as in Experiment 1. As in Experiment 1, all participants completed the battery of cognitive tests prior the experiment which took about 20 min to complete (see Table 1).

Procedure. The procedure was the same as in Experiment 1, except that the uncued memory items were presented on the screen during the test array (Figure 1B). Therefore, the test array consisted of the test cue at the location of one of the memory items, the other memory items, and the color wheel at the center of the screen. Participants completed a total of 15 blocks, five blocks of 15 trials for each set size, a total of 225 trials. Sessions lasted approximately 30 min.

Data analysis. The same probabilistic mixture modeling used in Experiment 1 was used to fit each participants' data, with one important difference. In Experiment 2, all nontarget items were presented on the screen during the test array. Therefore, we did not include a distribution for nontarget color values and used the formula in Equation 2 for every set size. Note that, there is only one Gaussian distribution for the target reports and a unimodal distribution for the random guessing. Thus, for the following analyses $p_t + p_r = 1$.

$$p(x) = p_t \Phi_{\mu t, \kappa t}(x - \theta_t) + p_r / 2\Pi \qquad (2)$$

Results

Probability of reporting the correct target value (p_t). We again found that delay did not impact any of the performance measures. All analyses below collapsed across this factor. We first tested the effects of set size and age on reporting the correct target color with a 3 (set size: 1, 2, 3) × 2 (age: older, younger) mixed ANOVA (Figure 5A). Replicating Experiment 1, we found a significant main effect of set size, $F(2, 114) = 80.4$, $p < .001$, $\eta_p^2 = .585$. Pairwise comparisons showed that participants were significantly more accurate for SS1 than for SS2, $t(58) = 3.68$, $p = .001$, $d = 0.51$, for SS2 than for SS3, $t(58) = 8.44$, $p < .001$, $d = 1.21$, and for SS1 than for SS3, $t(58) = 8.74$, $p < .001$, $d = 1.50$.

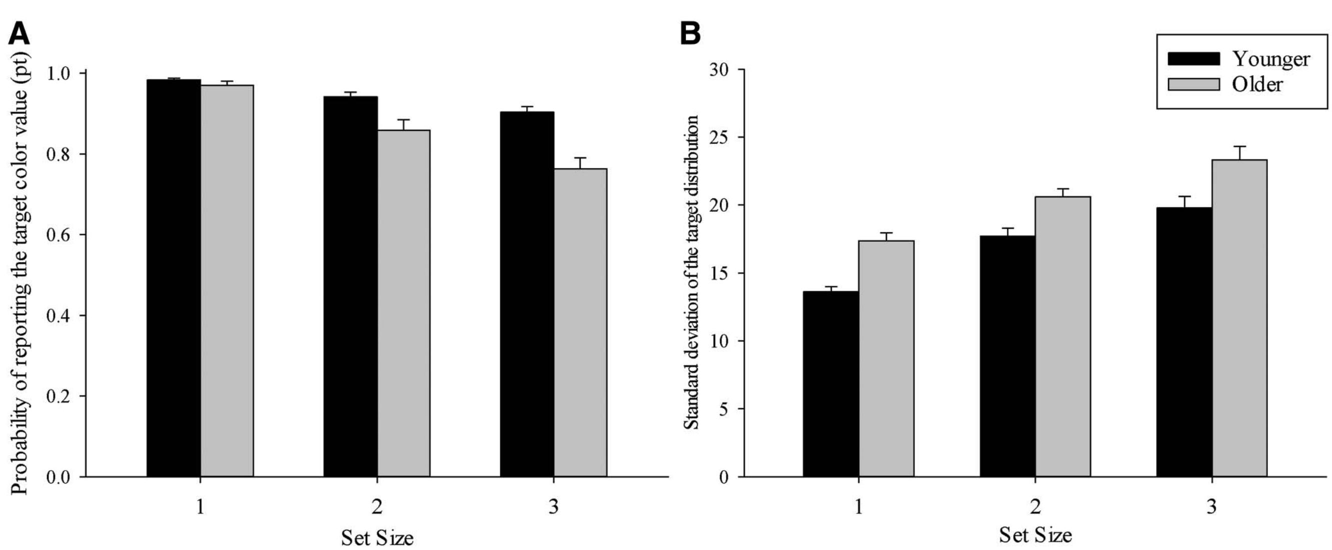

Figure 5. Experiment 2 results. Panel A shows the probability of reporting the target color value (p_t) and Panel B shows the standard deviation of the response distribution for the target value across different set sizes and ages. Error bars represent 95% confidence intervals.

As in Experiment 1, we also found a significant Set Size × Age interaction, $F(2, 114) = 11.9$, $p < .001$, $\eta_p^2 = .173$. We found no significant difference between the older and younger groups for the p_t parameter for SS1, $t(57) = 1.93$, $p = .058$, $d = 0.51$, and for SS2, $t < 1$. However, older participants had significantly lower p_t values compared with younger participants for SS3, $t(57) = 2.81$, $p = .007$, $d = 0.75$. The main effect of age was not significant, $F(1, 57) = 2.07$, $p = .156$, $\eta_p^2 = .035$. In this experiment $p_t + p_r = 1$; thus, the results are the same for the p_r variable, suggesting that older participants had significantly more random guesses than the younger participants for SS3, replicating our findings in Experiment 1 and further supporting our first hypothesis.

Standard deviations of the target distributions (precision of correct reports). To test whether age affects the precision of the color reports, we ran the same mixed ANOVA on the standard deviation of the color report distribution (Figure 5B). There was a significant main effect of set size, $F(2, 114) = 71.5$, $p < .001$, $\eta_p^2 = .556$, suggesting that color reports were less variable for SS1 than for SS2, $t(58) = -8.65$, $p < .001$, $d = 1.09$, for SS1 than for SS3, $t(58) = -10.7$, $p < .001$, $d = 1.39$, and for SS2 than for SS3, $t(58) = -4.57$, $p < .001$, $d = 0.54$. Importantly, we also found a significant main effect of age, $F(1, 57) = 18.7$, $p < .001$, $\eta_p^2 = .247$, where older participants' color reports were significantly less precise than younger participants' color reports, replicating our findings in Experiment 1 and further supporting our second hypothesis. The interaction between age and set size was not significant, $F < 1$.

The effect of nontargets on reporting the correct target value (p_t). To examine differences in performance between experiments, we conducted a series of statistical tests that used experiment as a between-subjects factor. Note that we focused analyses on SS2 and SS3 because SS1 was identical between the two experiments. To test whether presenting participants with the nontarget colors during the test phase increased the proportion of reporting the target value (p_t), especially in the older group, we combined the data from both experiments and ran a 2 (set size: 2, 3) × 2 (experiment: 1, 2) × 2 (age: older, younger) mixed ANOVA on the p_t data, where the set size was entered as a repeated measure, and experiment and age were entered as between-subjects factors. Replicating the previous findings, a main effect of set size revealed that participants were more accurate for SS2 than for SS3, $F(1, 115) = 85.8$, $p < .001$, $\eta_p^2 = .427$. Further, younger participants were significantly more accurate than older participants, revealed by a significant main effect of age, $F(1, 115) = 21.9$, $p < .001$, $\eta_p^2 = .160$. We also found a significant main effect of experiment with participants in Experiment 2 being significantly more accurate than participants in Experiment 1, $F(1, 115) = 10.03$, $p = .002$, $\eta_p^2 = .080$. The results also revealed a significant Set Size × Age interaction, $F(1, 115) = 13.4$, $p < .001$, $\eta_p^2 = .104$, and a significant Experiment × Age interaction, $F(1, 115) = 6.37$, $p = .013$, $\eta_p^2 = .052$. Set Size × Experiment interaction, $F(1, 115) = 1.45$, $p = .231$, $\eta_p^2 = .012$, and the three-way interaction, $F < 1$, were not significant.

To further investigate the Set Size × Age interaction, we ran an independent samples t test and found that younger participants were significantly more accurate than older participants for both SS2, $t(117) = 2.56$, $p = .012$, $d = 0.47$ and SS3, $t(117) = 5.21$, $p < .001$, $d = 0.96$. To interpret the Experiment × Age interaction, we collapsed data across different set sizes (Figure 6A). We found that younger participants' p_t values did not differ between different experiments, $t < 1$. In contrast, older participants' p_t values were significantly better in Experiment 2 than in Experiment 1, $t(59) = 3.60$, $p = .001$, $d = 0.94$. Specifically, presenting nontarget items on the screen significantly improved older participants' color estimation performance.

The effect of nontargets on random color reports/guessing (p_r). Next, we compared p_r values across different experiments to test whether presenting nontargets also decreased the proportion of random guesses (Figure 6B). To compare p_r values in SS2 and SS3, we ran a 2 (experiment: 1, 2) × 2 (age: older, younger) × 2 (set size: 2, 3) mixed ANOVA where the experiment and age were entered as between-subjects and set size was within-subjects. The main effects of age, $F(1, 115) = 19.9$, $p < .001$, $\eta_p^2 = .147$, and set size were significant, $F(1, 115) = 73.6$, $p < .001$, $\eta_p^2 = .390$. Older participants had significantly higher guessing rates than younger participants, and the guessing rates were in general higher for SS3 than for SS2. The Experiment × Age, $F(1, 115) = 4.47$, $p = .037$, $\eta_p^2 = .037$, the Experiment × Set Size, $F(1, 115) = 10.2$, $p = .002$, $\eta_p^2 = .081$, and the Set Size × Age interactions, $F(1, 115) = 8.06$, $p = .005$, $\eta_p^2 = .066$, were also significant. The main effect of experiment, $F < 1$, and the three-way interaction, $F(1, 115) = 2.37$, $p = .126$, $\eta_p^2 = .020$, did not reach significance. To further investigate the interactions, we ran follow-up analyses. First, aging significantly negatively affected the guessing rates in both experiments, but this aging effect became smaller in Experiment 2 when the nontargets were presented during the test phase, $t(58) = 3.92$, $p < .001$, $d = 1.03$ in Experiment 1 and $t(57) = 2.18$, $p = .034$, $d = 0.57$ in Experiment 2. Second, presenting the nontargets during test significantly improved guessing for SS2, $t(117) = -2.44$, $p = .016$, $d = 0.45$, but not for SS3, $t < 1$, indicating that as the task became more difficult the effect of nontargets decreased. Lastly, we found that although aging increased the guessing rates for both SS2, $t(117) = 2.78$, $p = .006$, $d = 0.52$, and SS3, $t(117) = 4.93$, $p < .001$, $d = 0.91$, this age effect became larger as set size increased.

The effect of nontargets on the precision of the target distributions. As discussed in the Introduction, a unique hypothesis the DNF model predicts is that presenting the nontarget items the during recall phase should improve precision of color reports for older adults, but not for younger adults. To test this third hypothesis, we ran the same mixed ANOVA on the standard deviation parameter (Figure 6C). A significant main effect of set size revealed that reports were significantly more precise for SS2 than for SS3, $F(1, 115) = 48.5$, $p < .001$, $\eta_p^2 = .297$. We also found a significant main effect of age, indicating that younger participants' color reports were significantly more precise than older participants' color reports, $F(1, 115) = 34.9$, $p < .001$, $\eta_p^2 = .233$. Further, we found a significant main effect of experiment, with color reports in Experiment 2 being significantly more precise than color reports in Experiment 1, $F(1, 115) = 11.8$, $p = .001$, $\eta_p^2 = .093$. We also observed significant Set Size × Experiment interaction, $F(1, 115) = 4.38$, $p = .039$, $\eta_p^2 = .037$, Set Size × Age interaction, $F(1, 115) = 5.73$, $p = .018$, $\eta_p^2 = .047$, and Experiment × Age interaction, $F(1, 115) = 8.65$, $p = .004$, $\eta_p^2 = .070$. The three-way interaction was not significant, $F(1, 115) = 2.99$, $p = .086$, $\eta_p^2 = .025$. Follow-up analyses showed that color reports in Experiment 2 were significantly more precise than color reports in Experiment 1 for both SS2, $t(117) = 2.50$, $p = .014$, $d = 0.46$, and SS3, $t(117) = 3.14$, $p = .002$, $d = 0.58$. Second, older participants

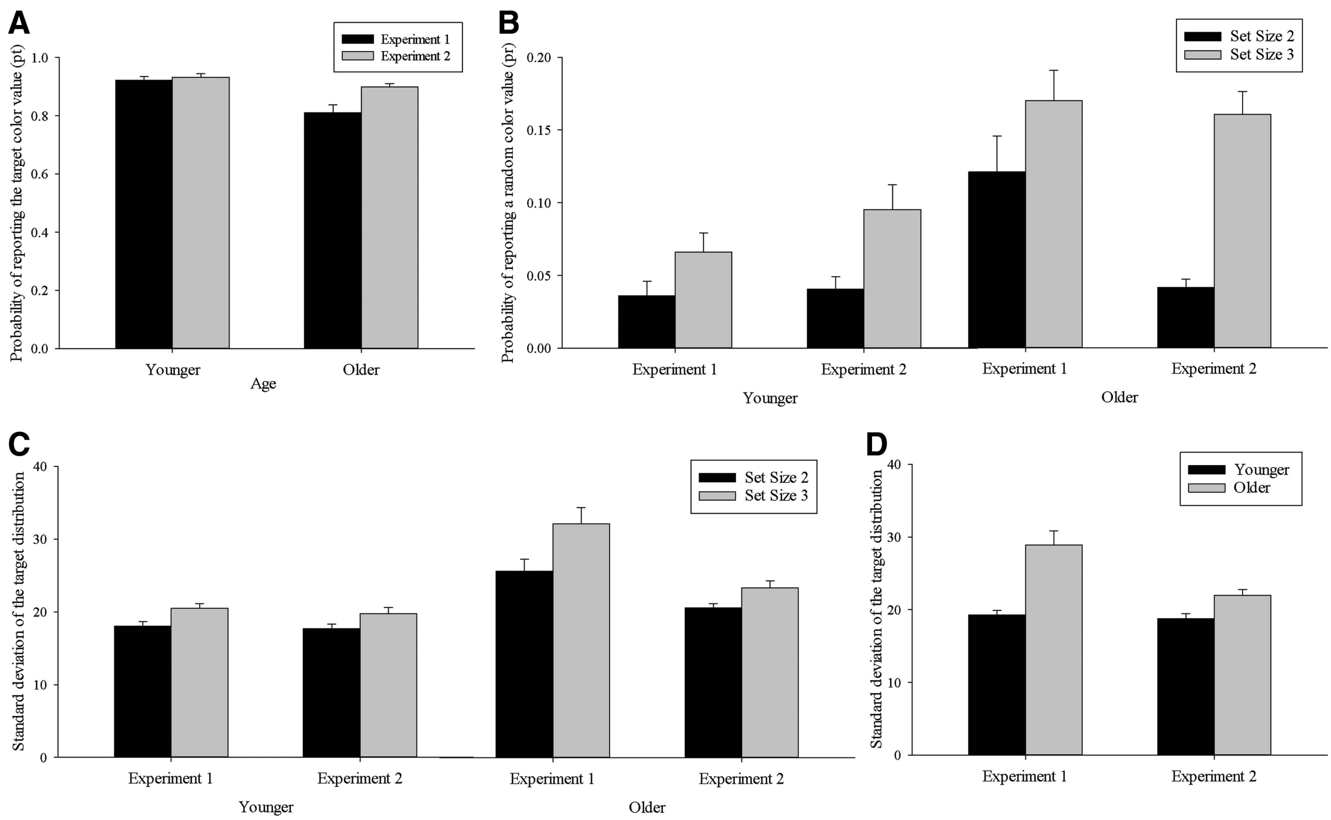

Figure 6. Comparison of Experiments 1 and 2. Panel A shows the probability of reporting the target color value (p_t) collapsed across different set sizes, Panel B shows the probability of reporting a random value (p_r,) Panel C shows the standard deviation of the target distribution, and Panel D shows the standard deviation of the target distribution collapsed across different set sizes. Error bars represent 95% confidence intervals.

reported significantly less precise colors compared with the younger participants for both SS2, $t(117) = 5.08$, $p < .001$, $d = 0.94$ and for SS3 $t(117) = 5.21$, $p < .001$, $d = 0.96$. In addition, this difference became larger as set size increased. Importantly, although precision of younger participants' color reports was not significantly affected by the presentation of the nontargets, $t < 1$, older participants in Experiment 2 reported significantly more precise colors than older participants in Experiment 1, $t(59) = -3.56$, $p = .001$, $d = 0.90$ (Figure 6D). This last finding suggests that only older participants benefited from the presentation of the nontarget colors during the test, as predicted by the DNF model.

General Discussion

The present study investigated the effect of aging on color VWM representations. Our results are consistent with previous research that has explored precision of VWM in the context of feature-binding (Peich et al., 2013). Specifically, in both Experiments 1 and 2, we found that older participants overall had significantly smaller proportions of correctly reported target (p_t) values compared with younger participants. This age-related effect increased as set size increased. Experiment 1 also revealed that older adults were more likely to report the color of one of the nontarget items especially for the largest set size. Further, we found that older adults were more likely to randomly guess than younger participants. Importantly, our results are also consistent with the DNF model of age-related decline in VWM (Costello & Buss, 2018). Specifically, smaller p_t values are reflected in the model that simulates older adults' change detection performance by losing more peaks within the WM field. In addition, we found that older adults' color reports were significantly more variable than younger adults' color reports (Figure 6C-D) which is consistent with the simulation results showing that representations in the older adult model use more neural units than the younger adult model. It should be noted that our younger and older adult groups had similar general working memory scores based on the forward and backward digit span tasks (see Table 1). Our older adult sample was self-selective, as in many other cognitive aging studies, and were generally healthier and more educated than the general population. There is evidence that sampling bias affects cognitive aging studies (cf. Brodaty et al., 2014), although the specific degrees of differences between convenience versus random sampling are mixed (Hultsch, MacDonald, Hunter, Maitland, & Dixon, 2002). Further, digit span tasks are typically used to measure verbal working memory while we employed a visual working memory task. Therefore, it is possible that the digit span task and delayed estimation task engage different working memory systems. In fact, there is evidence that adding a verbal memory load does not impair visual working memory as measured with the change-detection task (Luck & Vogel, 1997). Replicating those earlier findings, even with similar general verbal working memory abilities, our design was able to capture the significant difference in VWM performance between older and younger adults in both Experiments 1 and 2. Future work could extend the present findings in a random sample.

The DNF model predicted that presenting nontarget items would increase the precision of the memory items in older adults. Our results supported this prediction: Presenting nontarget items during recall in Experiment 2 improved the precision of correct color reports compared with Experiment 1. Importantly, this improvement was only present for older participants. Although older adults' color reports were significantly less precise than younger adults' color reports in both Experiments 1 and 2, their color reports were more precise when provided with the nontarget colors. Younger participants' color report performance was already almost at ceiling levels. As can be seen in Figures 6A and 6D, the presentation of the nontarget items in Experiment 2 led to older participants' p_t levels and precision levels to reach to similar levels as younger participants. In the DNF model, the increase in precision was driven by increased inhibition within the WM field that results from the presentation of more stimuli. Specifically, based on the coupling between the WM and Inhib fields of the model, the amount of inhibition and excitation within the WM field are proportional. With more inputs, there is more excitation which results in stronger inhibition. As a result, stronger inhibition serves to make the individual representations more precise. Taken together, these results suggest that emerging age-related declines in VWM can be partially ameliorated by providing nontarget stimuli during recall.

This work highlights some important limitations of the DNF model which can be addressed in future work. First, the model does not have an active response mechanism for feature estimation reports. The model has been developed in the context of the change detection task which requires making a "same/different" decision. Future work can expand the model architecture to be able to select a color from a location in a color wheel input (see Schutte & Spencer, 2009; Schutte, Spencer, & Schöner, 2003 for examples of models which make responses in a continuous space). Without such a mechanism, the model is unable to simulate the quantitative details of responses such as the ones presented in the analyses. Currently, the model can only speak to qualitative aspects of performance such as those explored here which predicted a decrease in standard deviation specifically for the older adult age group.

It should be noted that there is a possible alternative explanation for our findings because of how we selected color stimuli on each trial. Specifically, the colors in the present study were at least 60° apart on the color wheel. It is possible that the participants may have been implicitly influenced by this regularity and used an alternative response strategy in Experiment 2. That is, when presented with nontargets during test, based on the color of those nontarget items, participants could rule out large regions of the color space when responding. Although possible, we believe this explanation is unlikely for two reasons. First, the benefit of presenting nontargets at test was only present for older adults. If this were a general response strategy, then we should have seen similar improvements for younger adults as well. One can argue that our younger adults were already performing at the ceiling level and did not have room for improvement. This argument may be true for the probability of reporting the target parameter; however, the standard deviations of the reports significantly increased as set size increased, suggesting that younger adults' color reports became less precise with increased set size. Second, if the older participants indeed selected the target color value by a process of elimination of the color space, rather than actually remembering the correct color, then we should have seen either (a) a higher amount of guessing or (b) wider response distributions for SS2 than SS3. This is because the possible remaining color space to pick a response for SS2 is bigger than the remaining color space for SS3. The fact that we found almost no guessing for older adults in SS2 (and a decreased performance in SS3 compared to SS2), as well as a significant increase in precision as set size increased speaks against this alternative.

Related to these points, the color selection method in the present study was different than previous studies which used similar tasks (e.g., Peich et al., 2013). The main reason for this decision was to ensure that the task administered to our participants could be mimicked as closely as possible in the model. The model does not have a mechanism to bind features to spatial locations; thus, it has no way of representing two items that have the same color or very similar colors. Therefore, in our study, we randomly selected a set of six colors at the beginning of each trial which were 60° apart. The memory and test colors were then randomly selected among those six colors. This ensured, first, that colors would not repeat on a given trial. Second, this separation also ensured that the colors would not interfere with each other which may distort memory representations (J. S. Johnson et al., 2009).

There are two influential accounts of age-related decline in VWM performance. The binding account proposes that aging affects binding of different features to form the object, resulting in increased difficulty encoding the stimuli set (Cowan et al., 2006; Naveh-Benjamin, 2000). On the other hand, the inhibition account proposes that aging affects attentional control, resulting in trouble with filtering out irrelevant information during encoding (Gazzaley et al., 2008; Hasher & Zacks, 1988). Both accounts hypothesize that the difficulty arises during encoding of stimuli. Although we replicated the general finding that aging negatively affects the retention of the target properties, our broader pattern of results is not consistent with either the binding or the inhibition accounts. Our Experiments 1 and 2 were identical in their encoding stage but only differed during test. Therefore, as they stand, neither the binding nor the inhibition account can fully explain why presenting nontarget items during test would aid performance in older adults. Each account can, however, explain the findings if we attribute the source of error to the maintenance and decision stages, rather than encoding. First, it is possible that presenting nontarget items during recall would release the need for inhibiting the irrelevant nontargets which in turn would help older adults to retrieve the relevant target color from memory. In fact, the nontargets are not task-irrelevant until the test phase. Therefore, inhibiting them during encoding or maintenance would actually impair performance. However, these nontargets must be inhibited during the test phase in order to accurately report the correct target information. If they no longer require attentional resources to be suppressed, as in our Experiment 2, then attention can fully be focused on the target properties. This would be consistent with the explanation offered by the DNF model which illustrates how providing additional perceptual structure during the recall phase of the trial leads to stronger inhibition within the WM field. However, in the DNF model, the primary mechanism by which the performance of older adults was facilitated in Experiment 2 goes beyond suppressing attention to the nontarget items and arises from lateral inhibition sharpening the representation of the target item.

Second, if older adults' poorer memory performance is a result of incorrectly binding the feature-location information, but not necessarily a result of retaining the correct features, then presenting them with nontarget feature-location bindings should make it easier to report the correct feature. Although the binding account can explain why older adults' overall performance improved, it cannot explain the finding that older adults were also less likely to randomly guess when the nontargets were on screen during test. The finding that participants' random guesses also decreased with the aid of nontargets shows a general improvement in performance which is not specific to distractor effects caused by the nontargets. Therefore, it is unlikely that release from inhibition or resolving misbinding can account for the decrease in guessing. Moreover, the misbinding problem does not explain why the precision of responses would be improved by presenting nontarget items during recall. That is, misbinding may account for the finding that participants select one of the distractor colors but not for why precision of responses changes depending on nontarget presence or absence.

It should be noted that the present results challenge some of the earlier findings reported in a similar delayed estimation study (Peich et al., 2013). First, Peich et al. (2013) did not find any significant difference between younger and older adults' guessing rates. Second, they found that incorrect reports for nontarget features were the main source of error in older adults. The authors used these two findings as strong evidence for the binding account. Specifically, older adults were able to encode and recall the features but had trouble to correctly identify which feature goes with which object, resulting in misbinding errors. Contrary to that study, our results showed significantly higher guessing in older adults compared to younger adults, even for set sizes as small as two (Experiment 1). Although we also found that older adults had higher levels of misbinding errors (i.e., higher reports of a nontarget color in Experiment 1), it is curious as to why Peich et al. (2013) did not find higher guessing rates. An important difference between these two studies is that participants in Peich et al. (2013) study were asked to retain a color-orientation-location conjunction information and report two features of an object whereas we only asked them to retain color-location binding. Previous studies have shown that multifeature memory tasks result in poorer estimates of VWM compared with single-feature memory tasks (e.g., Alvarez & Cavanagh, 2004; Olson & Jiang, 2002). Therefore, it is possible that the Peich et al. (2013) task was simply more attention-demanding than ours which may be why they did not find differences in guessing rates between the two age groups.

Interestingly, neither of our experiments showed a significant effect of delay or any interactions with delay manipulation. This finding is inconsistent with the results of Pertzov, Manohar, and Husain (2017) who showed that delay alone has a negative effect on memory performance of younger participants, especially for larger set sizes. There are two major differences between our design and Peich et al.'s (2013) study: They used longer delay times (100 ms, 1 s, 2 s, 3 s) as well as larger set sizes (1, 2, 4, and 6) compared with our experiments. In addition, they reported that the effect of delay was not significant for Set Size 1, and for Set Size 2 the delay seems to have a significant effect only when it was at least 3 s long. Therefore, it is possible that our delay manipulation was not strong enough (even for older adults) or that we did not use large enough set sizes to show a significant decline in performance with increased delay.

Conclusions

The present study provides further evidence regarding age-related differences in VWM between younger and older adults. In general, older adults' recall of color features was less precise and less accurate compared to older adults. Further, presenting the nontargets during recall improved both the precision of older adults' color recall, as well as the probability of reporting the correct color. We offered an explanation grounded in a DNF model: Providing nontarget items during recall increases the precision of color estimation reports by increasing the strength of neural inhibition within the VWM system. The present findings point to the need to address how perceptual structure during test can influence measures of VWM which has not been addressed by previous theories of VWM. Using model-based explanations to aging can inform us about how changes in neural dynamics can give rise to age-related differences in VWM.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106–111. http://dx.doi.org/10.1111/j.0963-7214.2004.01502006.x

Bach, M. (1996). The Freiburg Visual Acuity test—Automatic measurement of visual acuity. *Optometry and Vision Science, 73,* 49–53. http://dx.doi.org/10.1097/00006324-199601000-00008

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 7.

Brockmole, J. R., & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55,753 participants aged 8–75. *Frontiers in Psychology, 4,* 12. http://dx.doi.org/10.3389/fpsyg.2013.00012

Brockmole, J. R., Parra, M. A., Della Sala, S., & Logie, R. H. (2008). Do binding deficits account for age-related decline in visual working memory? *Psychonomic Bulletin & Review, 15,* 543–547. http://dx.doi.org/10.3758/PBR.15.3.543

Brodaty, H., Mothakunnel, A., de Vel-Palumbo, M., Ames, D., Ellis, K. A., Reppermund, S., . . . Sachdev, P. S. (2014). Influence of population versus convenience sampling on sample characteristics in studies of cognitive aging. *Annals of Epidemiology, 24,* 63–71. http://dx.doi.org/10.1016/j.annepidem.2013.10.005

Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 63,* 2067–2079. http://dx.doi.org/10.1080/17470211003721675

Campbell, J. I. D., & Thompson, V. A. T. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods, 44,* 1255–1265. http://dx.doi.org/10.3758/s13428-012-0186-0

Costello, M. C., & Buss, A. T. (2018). Age-related decline of visual working memory: Behavioral results simulated with a dynamic neural field model. *Journal of Cognitive Neuroscience, 30,* 1532–1548. http://dx.doi.org/10.1162/jocn_a_01293

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology, 42,* 1089–1102. http://dx.doi.org/10.1037/0012-1649.42.6.1089

Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., & Schöner, G. (1999). The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations. *Journal of Neuroscience Methods, 94,* 53–66. http://dx.doi.org/10.1016/S0165-0270(99)00125-9

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review, 17,* 673–679. http://dx.doi.org/10.3758/17.5.673

Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., & D'Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 13122–13126. http://dx.doi.org/10.1073/pnas.0806074105

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation, 22,* 193–225. http://dx.doi.org/10.1016/S0079-7421(08)60041-9

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Maitland, S. B., & Dixon, R. A. (2002). Sampling and generalisability in developmental research: Comparison of random and convenience samples of older adults. *International Journal of Behavioral Development, 26,* 345–359. http://dx.doi.org/10.1080/01650250143000247

Jancke, D., Erlhagen, W., Dinse, H. R., Akhavan, A. C., Giese, M., Steinhage, A., & Schöner, G. (1999). Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex. *The Journal of Neuroscience, 19,* 9016–9028. http://dx.doi.org/10.1523/JNEUROSCI.19-20-09016.1999

Johnson, J. S., Simmering, V. R., & Buss, A. T. (2014). Beyond slots and resources: Grounding cognitive concepts in neural dynamics. *Attention, Perception & Psychophysics, 76,* 1630–1654. http://dx.doi.org/10.3758/s13414-013-0596-9

Johnson, J. S., Spencer, J. P., Luck, S. J., & Schöner, G. (2009). A dynamic neural field model of visual working memory and change detection. *Psychological Science, 20,* 568–577. http://dx.doi.org/10.1111/j.1467-9280.2009.02329.x

Johnson, M. K., McMahon, R. P., Robinson, B. M., Harvey, A. N., Hahn, B., Leonard, C. J., . . . Gold, J. M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology, 27,* 220–229. http://dx.doi.org/10.1037/a0032060

Kline, D. W., & Scialfa, C. T. (1997). Sensory and perceptual functioning: Basic research and human factors implications. In A. D. Fisk & W. A. Rogers (Eds.), *Handbook of human factors and the older adult* (pp. 27–54). San Diego, CA: Academic Press.

Ko, P. C., Duda, B., Hussey, E., Mason, E., Molitor, R. J., Woodman, G. F., & Ally, B. A. (2014). Understanding age-related reductions in visual working memory capacity: Examining the stages of change detection. *Attention, Perception & Psychophysics, 76,* 2015–2030. http://dx.doi.org/10.3758/s13414-013-0585-z

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

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

Markounikau, V., Igel, C., Grinvald, A., & Jancke, D. (2010). A dynamic neural field model of mesoscopic cortical activity captured with voltage-sensitive dye imaging. *PLoS Computational Biology, 6*(9), e1000919. http://dx.doi.org/10.1371/journal.pcbi.1000919

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1170–1187. http://dx.doi.org/10.1037/0278-7393.26.5.1170

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception & Psychophysics, 64,* 1055–1067. http://dx.doi.org/10.3758/BF03194756

Park, D. C., & Payer, D. (2006). Working memory across the adult lifespan. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 128–142). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195169539.003.0009

Peich, M.-C., Husain, M., & Bays, P. M. (2013). Age-related decline of precision and binding in visual working memory. *Psychology and Aging, 28,* 729–743. http://dx.doi.org/10.1037/a0033236

Pertzov, Y., Manohar, S., & Husain, M. (2017). Rapid forgetting results from competition over time between items in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 528–536. http://dx.doi.org/10.1037/xlm0000328

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16,* 754–760. http://dx.doi.org/10.1017/S1355617710000706

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132,* 566–594. http://dx.doi.org/10.1037/0096-3445.132.4.566

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

Schöner, G., Spencer, J. P., & The DFT Research Group. (2015). *Dynamic thinking: A primer on dynamic field theory*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199300563.001.0001

Schutte, A. R., & Spencer, J. P. (2009). Tests of the dynamic field theory and the spatial precision hypothesis: Capturing a qualitative developmental transition in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1698–1725. http://dx.doi.org/10.1037/a0015794

Schutte, A. R., Spencer, J. P., & Schöner, G. (2003). Testing the dynamic field theory: Working memory for locations becomes more spatially precise over development. *Child Development, 74,* 1393–1417. http://dx.doi.org/10.1111/1467-8624.00614

Simmering, V. R. (2016). I. Working memory capacity in context: Modeling dynamic processes of behavior, memory, and development. *Monographs of the Society for Research in Child Development, 81,* 7–24. http://dx.doi.org/10.1111/mono.12249

Spencer, J. P., & Schöner, G. (2003). Bridging the representational gap in the dynamic systems approach to development. *Developmental Science, 6,* 392–412. http://dx.doi.org/10.1111/1467-7687.00295

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92–114. http://dx.doi.org/10.1037/0096-1523.27.1.92

Wecker, N. S., Kramer, J. H., Wisniewski, A., Delis, D. C., & Kaplan, E. (2000). Age effects on executive ability. *Neuropsychology, 14,* 409–414. http://dx.doi.org/10.1037/0894-4105.14.3.409

Zanto, T. P., & Gazzaley, A. (2014). Attention and ageing. In A. C. Nobre & S. Kastner (Eds.), *Oxford library of psychology. The Oxford handbook of attention* (pp. 927–971). New York, NY: Oxford University Press.

Received May 3, 2019
Revision received January 24, 2020
Accepted January 25, 2020 ■