The role of motion in visual working memory for dynamic stimuli: more lagged but more precise representations of moving objects.

Yong Hoon Chung[1] Mark W. Schurgin[2] Timothy F. Brady[2]

[1] Department of Psychological and Brain Sciences, Dartmouth College
[2] Department of Psychology, University of California San Diego

Abstract

While most visual working memory studies use static stimuli with unchanging features, objects in the real-world are often dynamic, introducing significant differences in the surface feature information hitting the retina from the same object over time (e.g., changes in orientation, lighting, shadows). Previous research on dynamic stimuli has shown that change detection is improved if objects obey rules of physical motion, but it is unclear how memory for visual features interacts with object motion. In the current study, we investigated whether object motion facilitates greater temporal integration of continuously changing surface feature information. In a series of experiments, participants were asked to report the final color of continuously changing colored dots that were either moving or stationary on the screen. We found that the reported colors "lagged behind" the physical states of the dots when they were in motion. We also observed that the precision of memory responses was significantly higher for stimuli in the moving condition compared to the stationary condition. Together, these findings suggest that memory representation is improved – but lagged – for moving objects, consistent with the idea that object motion may facilitate integration of object information over longer intervals.

Introduction

Visual working memory is an active temporary storage system for manipulating and holding in mind visual information (Cowan, 2008; Schurgin, 2018). One of the main characteristics of visual working memory is its limited capacity, as we can only actively retain a relatively small amount of visual information for a short duration of time (e.g., Ma, Husain & Bays, 2014). Since the limited capacity of working memory is related to important cognitive functions such as fluid intelligence, reading comprehension, and attentional control (Alloway & Alloway, 2010; Daneman & Carpenter, 1980; Fukuda et al., 2010; Kane et al., 2001), understanding how visual working memory is limited and how it functions to hold in mind visual objects has been a major focus of research (Cowan, 2001; Ma, Husain & Bays, 2014; Luck & Vogel, 2013). In the real-world, visual objects are embedded in rich scene contexts, and many objects are not static: they move, and while they move the visual feature information hitting an observer's retina drastically changes. For example, as a car drives by in front of you, the color of the car's surface constantly changes as a result of varying viewing angles, lighting, and or shadows (Figure 1). However, each of these aspects of visual objects has generally been studied in isolation, without asking about how they interact or may support the formation of visual memory representations. In the current work, we ask how motion affects memory for continuously changing visual features (e.g., the color of the car as it moves past) and how this is affected by whether a single object is being tracked and remembered or more than one object is being tracked and remembered.

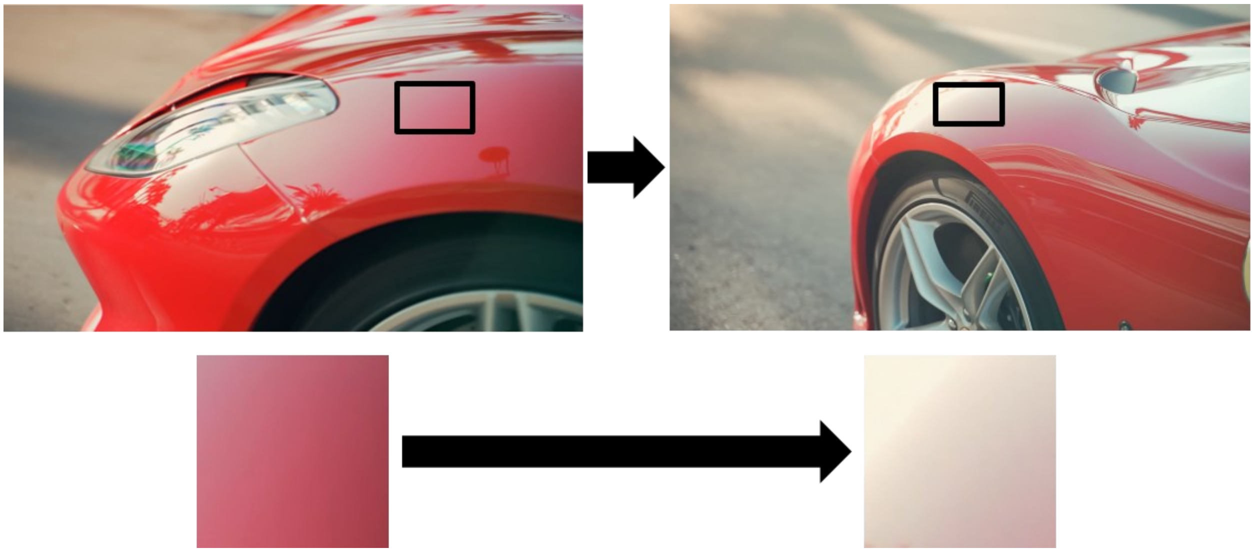

Figure 1. An illustration of varying object features over time and space. Even though the location of the black box on the car's surface remains the same, the color of the spot vastly changes around as the car drives by due to changes in viewpoints, lightings, or shadows. However, we are able to integrate the varying feature information into one single object representation.

Visual working memory for static stimuli

The majority of research on visual working memory has focused on extremely simplified situations, rather than more realistic, dynamic moments common in everyday life: studies primarily consider only stationary, non-changing simple features (e.g., colored circles, oriented bars; e.g., Luck & Vogel, 1997; Harrison & Tong, 2009, etc.). Even in the case of studies focusing on visual working memory for more realistic objects, most studies have focused on memory for stationary, non-changing objects (e.g., Hollingworth, 2004), often presented without any scene context on a white background (e.g., Brady & Stoermer, 2021; Starr, Srinivasan, & Bunge, 2020).

Studies using simple stimuli like colors and orientations have led to important conclusions regarding visual working memory capacity that are relevant to the real-world scenario where an object is both moving and changing in features. Continuous reproduction studies, where participants must report the exact color or orientation of a stimulus after a brief delay (e.g., Wilken & Ma, 2004; Zhang & Luck, 2008) have been particularly influential in understanding how visual objects are represented. Such studies show that the strength of working memory is strongly dependent on how many items are held in mind: participants can reproduce a color more accurately when only holding 1 color in mind than when holding 2 or 3 colors in mind (Zhang & Luck, 2008; Bays et al., 2009). Memory is also strongly dependent on delay: holding in mind information for longer results in weaker memories (Schurgin et al., 2020; Rademaker et al., 2018).

Representations of moving or changing objects

By contrast to such studies of static stimuli, our ability to concurrently track multiple moving objects has been extensively studied using the multiple object tracking paradigm (Pylyshyn & Storm, 1988; Scholl, 2009). Multiple object tracking, like visual working memory, is subject to a severe capacity limit (e.g., Scholl, 2009), but also reveals unique aspects of attempting to hold in mind stimuli in motion. Of particular relevance to the case of continuously changing stimuli, it has generally been found that when participants are required to track positions of moving stimuli on the screen, observers' position representations lag behind that of the current locations of those stimuli so that when asked to report the final position of the target, they generally would report a position slightly before the ending point (Howard et al., 2011). This perceptual lag in tracking is also found when observers are tracking multiple dynamically changing features of stimuli that are not in motion (i.e., continuously changing spatial frequencies or orientations): analogous to the spatial locations, participants are biased toward the past state when asked to report the final state of the target feature, particularly when tracking multiple items at once (Howard & Holcombe, 2008).

These results show that people are affected by past states of an object even when cued to report the object's final state. This could be because people are lagged in sampling objects, as it has generally been interpreted in the literature on multiple object tracking (e.g., as evidence of serial processing, Howard & Holcombe, 2008). However, such effects could also arise if participants' visual systems are purposefully integrating information over time. For example, in the case of a car driving past, our goal is indeed to integrate over time, as the car in fact has a single color and our goal is to extract this and maintain color constancy rather than to attempt to remember each feature separately at each moment in time (Shevell & Kingdom, 2008). When explicitly prompted to compute a perceptual average over time, participants are accurately able to do so (Albrecht & Scholl, 2009), suggesting that integration is a natural aspect of the way we deal with continuously changing stimuli. Even when not explicitly instructed, visual working memory representations can be biased through integration with other perceptual inputs, especially when they are similar (Fukuda et al., 2022). In theory, these accounts should be distinct in that integration should also come with increased precision by being able to integrate multiple representations: participants should not only be lagged but also more precise if they are integrating over time, compared to a situation with little integration over time (as in ensemble perception models: Alvarez, 2011).

At a broad level, this ability to treat spatiotemporally defined objects as unified entities in our mind, and integrate the information presented on them into one single object representation, has often been termed an object file and has been extensively studied (Scholl, 2001; Kahneman, Treisman & Gibbs, 1992; Spelke et al., 1995; Flombaum, Scholl & Santos, 2009; Flombaum & Scholl, 2006; Schurgin & Flombaum, 2017). By contrast to the multiple object tracking literature, however, nearly all object file studies present discrete information at two different moments in time (e.g., first one letter, then later another), asking about object-specific preview benefits, rather than exploring more naturalistic scenarios (e.g., the driving car) where motion and feature changes both occur smoothly over time. Abrupt onsets may ‘reset’ attention, resulting in different patterns of attentional engagement and in different memory biases than attempting to remember a particular moment in a smoothly varying stimulus (Callahan-Flintoft, Holcombe & Wyble, 2020).

The current studies

Given the current state of the literature, there remains a critical gap in our understanding of how object motion interacts with smooth changes in visual features. Specifically, do participants integrate visual information over time to a different extent when they are tracking 1 vs. 2 objects, or is the tendency to report past information a result of lag in sampling objects in a serial manner? Are participants lagged in their reports to the same extent when there are both object motion and visual feature changes occurring at once (as they tend to do in the natural world) as when the two are decoupled?

In order to investigate these questions, we ran a series of visual working memory experiments where participants had to track either 1 or 2 colored dot stimuli that were continuously changing their colors during each trial. After a short delay, we then asked participants to report the final color of one of the dots. To investigate how motion impacts the memoranda, these dots sometimes moved around the screen and sometimes remained stationary during the encoding time depending on the condition. We hypothesized that when objects are in motion, integration of changing features over time (e.g., into an 'object file') will be more prominent than when objects are stationary. This integration should make the memory representation more biased towards the past (i.e., more lagged behind the current state) as more past features are being incorporated, but at the same time, integration may make memories more precise, as many noisy samples can be averaged together, potentially canceling out some amount of noise (Figure 2).

Our results show that object motion can induce increased lag in representation but higher precision in memory report, demonstrating that participants are integrating information over time rather than simply lagged in their representation of the objects. Our results also point to a role for motion in visual working memory representations: in particular, they suggest an interaction between integrating features over time and object motion, where objects in motion are integrated over longer time windows.

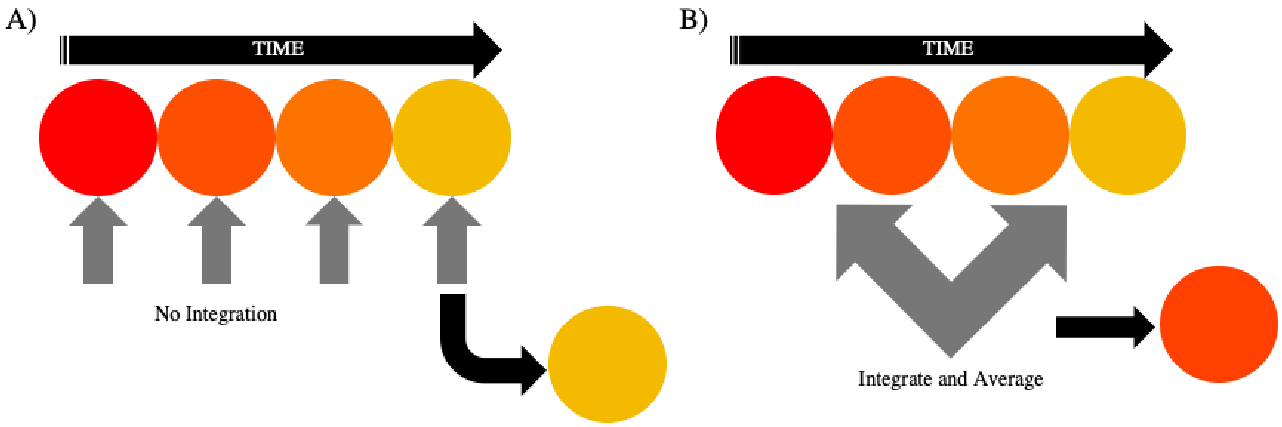

Figure 2. A demonstration of our hypothesis. A) When an object is stationary and therefore facilitates less integration of changing colors, the final color report will be less biased towards the previous colors. B) If the object is in motion, integration will happen over a longer time period, resulting in biased final color reports towards the past states.

Experiment 1

Methods

Participants. 30 undergraduate students from University of California San Diego were recruited to participate in the study in exchange for course credit. All participants gave informed consent and reported normal or corrected-to-normal color vision.

Stimuli. All colors were drawn from a fixed luminance circle in CIE L*a*b space (identical to that used by Suchow et al., 2013; Schurgin et al., 2020). The colored dots presented during the display started at a random point on the color wheel and continuously changed in a clockwise direction on the color wheel at a speed of 120 degrees-of-color per second. The dots appeared at random locations along preset half circular pathways around the fixation point to control eccentricity. In the moving condition, they then bounced back and forth along the half circular pathways at 2.7 radian per second. In the stationary condition, dots stayed at the same locations throughout the trial. In order to create some randomness in the motion path, the dots stochastically changed their direction at the ¼, ½, and ¾ intervals of the trial. This random motion was implemented to reduce predictability of association between dots' locations and colors.

Procedure. Participants were presented with two dots on the screen each trial, one on the left side and the other on the right side. At set size 1, one of the dots was presented black while the other had an initial color. At set size 2, both dots had initial colors. During the stimulus display, only the colored dots smoothly changed their colors while the black dots in the set size 1 trials remained constant and did not have to be remembered.

There were two different motion conditions: a moving condition where the dots moved along the preset circular paths and a stationary condition where the dots stayed at the same position throughout the trial. The display time was randomly jittered from 2.5 seconds to 3.5 seconds in order to prevent any temporal expectation for when the trial would end, since only the final color needed to be remembered. After the stimulus display, the dots disappeared from the screen, and there was a 1 second delay. After the delay, participants reported the final color of one of the dots using a continuous color wheel (For illustration, see Figure 3). To assess possible motion silencing of color change (Suchow & Alvarez, 2011), after the color report, participants were asked to judge how fast the colors were changing using a Likert scale from 1 (not changing much) to 4 (changing very fast). All conditions were counterbalanced and randomly intermixed. Participants were instructed to maintain eye fixation in the center of the screen throughout the experiment. The entire experiment consisted of 240 trials and each participant went through 10 practice trials beforehand to familiarize themselves with the experiment procedure.

[1]As we expected participants' memories would integrate information across multiple time points, model-based analysis (e.g., by fitting the TCC model: Schurgin et al., 2020) would not be appropriate, unless we knew the exact level of weighting of each lag of color in the final representation. For static stimuli, the circular standard deviation is strongly related to model-based memory strength (Schurgin et al., 2020 supplement), so we used this as our primary measure of memory.

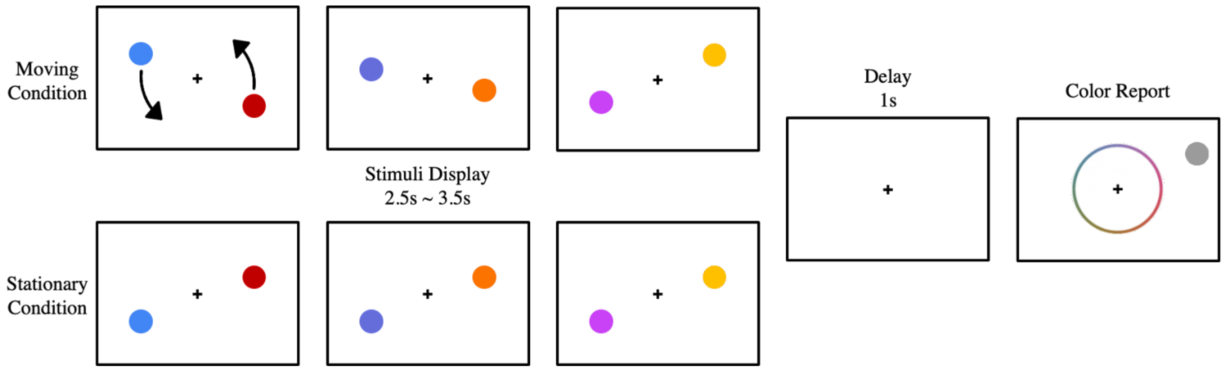

Figure 3. Experimental procedure. On each trial two dots were presented on the screen, one on the left and one on the right side. In set size 1, one of the dots was black and remained black during the entire duration of the trial while in set size 2 both dots had colors. In the stationary condition the dots did not move and in the moving condition the dots moved along preset circular paths around the fixation cross. After a 1 second delay, participants reported the final color of the probed dot by clicking their response on a color wheel.

Analysis. Color error was calculated by taking the angle in degrees that participants reported and subtracting the target color from the reported color using circular color distance: [Reported Color] - [Target Color]. For instance, if the target color was 45 degrees and the participant reported 30 degrees, the color error for that trial would be -15. Because the colors were always changing in a clockwise direction along the color wheel, negative color error value indicates a lag (participant reported the past color) while positive color error value indicates an extrapolation (participant reported the future color).

Precision of memory is analyzed by calculating the circular standard deviation of the error distribution. Smaller standard deviation value indicates higher precision of memory[1]. Precision specifically refers to the narrowness of the error distribution and higher precision means errors have less variance overall. Perceptual lag is indexed by looking at the overall mean shift of the distribution from the zero error mark (for illustration, see Figure 4). Again, negative shift indicates a lag towards past colors while positive shift indicates a possible future extrapolation.

[1]As we expected participants' memories would integrate information across multiple time points, model-based analysis (e.g., by fitting the TCC model: Schurgin et al., 2020) would not be appropriate, unless we knew the exact level of weighting of each lag of color in the final representation. For static stimuli, the circular standard deviation is strongly related to model-based memory strength (Schurgin et al., 2020 supplement), so we used this as our primary measure of memory.

Standard deviation and mean shift values were computed for each participant and compared using 2x2 ANOVA with motion condition (stationary vs. moving) and set sizes (set size 1 vs. set size 2) as factors.

Color speed judgment was calculated separately for each condition by averaging the responses and compared using 2x2 ANOVA with motion condition and set sizes as factors.

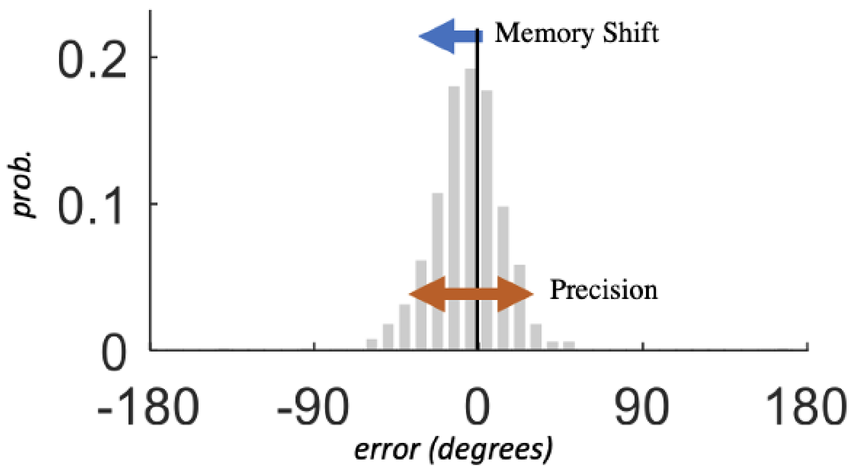

Figure 4. An illustration of the data analysis procedure. Color errors were calculated and plotted so that the error distribution is centered around 0 (where there was no error at all in the report) and negative error indicates bias towards the past colors. Precision is indexed by the circular standard deviation of the error distribution in a way that lower standard deviation reflects higher memory precision.

Results

An ANOVA on circular standard deviation — our measure of memory precision — yielded main effects of both motion vs. static, $F(1, 29) = 4.69$, $p = 0.039$, and set size 1 vs 2, $F(1, 29) = 51.29$, $p < 0.001$. For mean shifts, we found a main effect of motion, $F(1, 29) = 55.56$, $p < 0.001$, and an interaction between set size condition and motion condition, $F(1, 29) = 6.44$, $p = 0.017$. Overall, participants were both more precise and more lagged when the objects were moving, and this was particularly true at set size 2 (Figure 5).

An ANOVA on color speed judgment scores yielded a main effect of set size, $F(1, 29) = 14.07$, $p = 0.001$, and an interaction between set size condition and motion condition, $F(1, 29) = 7.45$, $p =$

0.01 (Figure 6). Overall, participants rated colors to be changing faster in moving dots than stationary dots, but this difference is more prominent at set size 1. This means that motion silencing is unlikely to be playing a role in the current paradigm as the results are against the prediction that motion should silence the perception of changing colors.

Follow up t-tests on the mean shifts for each of the conditions showed significant differences for the set size 1 moving condition ($t(29) = 4.95$, $p<0.001$) and set size 2 moving condition ($t(29) = 6.34$, $p<0.001$) compared to 0 (no mean shift at all), but no significant differences for the set size 1 stationary condition ($t(29) = 1.62$, $p = 0.116$) and set size 2 stationary condition ($t(29) = 1.13$, $p = 0.269$).

Overall, then, we found evidence in line with the idea that participants are integrating information over time in moving objects, rather than simply lagged in their representation of these objects: not only were they lagged in their reported colors, but these color representations were more precise, as you'd expect if people were integrating information across multiple time points. This also interacted with set size in a way that the mean shift was much more prominent for the set size 2 moving condition. Standard deviation was larger for set size two conditions than set size one conditions, as expected with increased set size.

Experiment 2

This experiment served as a replication of Experiment 1, with one minor change. In Experiment 1, participants could have inferred the color change direction from the response color wheel, as colors always changed in one direction along the color wheel. To address this, we varied the color change direction in Experiment 2: Colors could either smoothly change clockwise or counterclockwise along the color wheel.

Method

Participants. 30 undergraduate students from Dartmouth College were recruited to participate in the study in exchange for course credit. All participants gave informed consent and reported normal or corrected-to-normal color vision.

Stimuli. In half the trials colored dots changed their colors clockwise along the color wheel and in other half colors changed counterclockwise. All other stimuli set up were identical to Experiment 1.

Procedure. Other than color change directions, the experimental procedure was identical to Experiment 1.

Analysis. Color errors were calculated such that negative values indicate lag towards the past by flipping trials where the colors changed counterclockwise. All other data analysis was identical to Experiment 1. Color speed judgments were not collected for this experiment.

Results

An ANOVA on circular standard deviation, our measure of memory precision, yielded main effects of both motion condition, $F(1, 29) = 11.397$, $p = 0.002$, and set size, $F(1, 29) = 29.56$, $p < 0.001$. For mean shifts, we found a main effect of motion, $F(1, 29) = 126.36$, $p < 0.001$, but no significant interaction between set size condition and motion condition, $F(1, 29) = 0.63$, $p = 0.43$ (Figure 5).

One sample t-tests on the mean shifts for each of the conditions showed significant differences for the set size 1 moving condition ($t(29) = 11.10$, $p<0.001$), set size 2 moving condition ($t(29) = 7.47$, $p<0.001$), and set size 1 stationary condition ($t(29) = 3.30$, $p = 0.0026$) compared to 0, but no significant differences for the set size 2 stationary condition ($t(29) = 1.80$, $p = 0.08$).

Overall, this set of results provided a strong replication of Experiment 1: Mean shifts were significantly more negative while standard deviations were significantly smaller in the moving conditions compared to the stationary conditions. Mean shifts in stationary conditions were also not significant against the baseline.

Experiment 3

While Experiment 1 and 2 both showed consistent data patterns, they both used identical motion and color change speeds. In order to ask if the effects generalize beyond these particular parameters, in Experiment 3 we significantly varied the stimuli motion and color change speeds and asked if the same patterns arose

Method

Participants. 30 undergraduate students from University of California San Diego were recruited to participate in the study in exchange for course credit. All participants gave informed consent and reported normal or corrected-to-normal color vision.

Stimuli. The color changed more slowly than previous experiments at 60 degrees per second along the color wheel (half the speed of the previous experiments). For the motion condition, the speed of dots randomly changed between 1.5 radian per second and 3.9 radian per second in effort to make the motion less coherent and predictable than it had been in Experiments 1 and 2. All other stimuli set up were identical to Experiment 1.

Procedure. Experimental procedure was identical to Experiment 1.

Analysis. Data analysis was identical to Experiment 1.

Results

An ANOVA on circular standard deviation, our measure of memory precision, showed a main effect of set size, $F(1, 29) = 94.65$, $p < 0.001$, and a significant interaction between set size and conditions, $F(1, 29) = 7.61$, $p < 0.05$, but not a significant main effect of motion condition, $F(1, 29) = 2.60$, $p = 0.117$. For mean shifts, we found the main effect of motion, $F(1, 29) = 72.81$, $p < 0.001$, but no significant interaction was found between set size condition and motion condition, $F(1, 29) = 3.80$, $p = 0.06$ (Figure 5).

ANOVA on color speed judgment scores yielded a main effect for set sizes, $F(1, 29) = 10.44$, $p = 0.003$, and an interaction between set size condition and motion condition, $F(1, 29) = 17.89$, $p < 0.001$ (Figure 6). Again, a similar pattern of data was observed as Experiment 1 where color changing speed was reported to be faster in moving condition for set size 1 but not in set size 2.

One sample t-tests on the mean shifts for each of the conditions showed significant differences for the set size 1 moving condition ($t(29) = 6.27$, $p<0.001$), set size 2 moving condition ($t(29) = 6.78$, $p<0.001$), set size 1 stationary condition ($t(29) = 2.55$, $p=0.016$), and the set size 2 stationary condition ($t(29) = 2.40$, $p = 0.02$) compared to 0.

Overall, Experiment 3 mostly replicates the previous experiments' results in both mean shifts and standard deviations. Here, there was no significant main effect of motion condition in standard deviations, but a significant interaction was found where set size 1 moving condition yielded a significantly higher precision than set size 1 stationary condition.

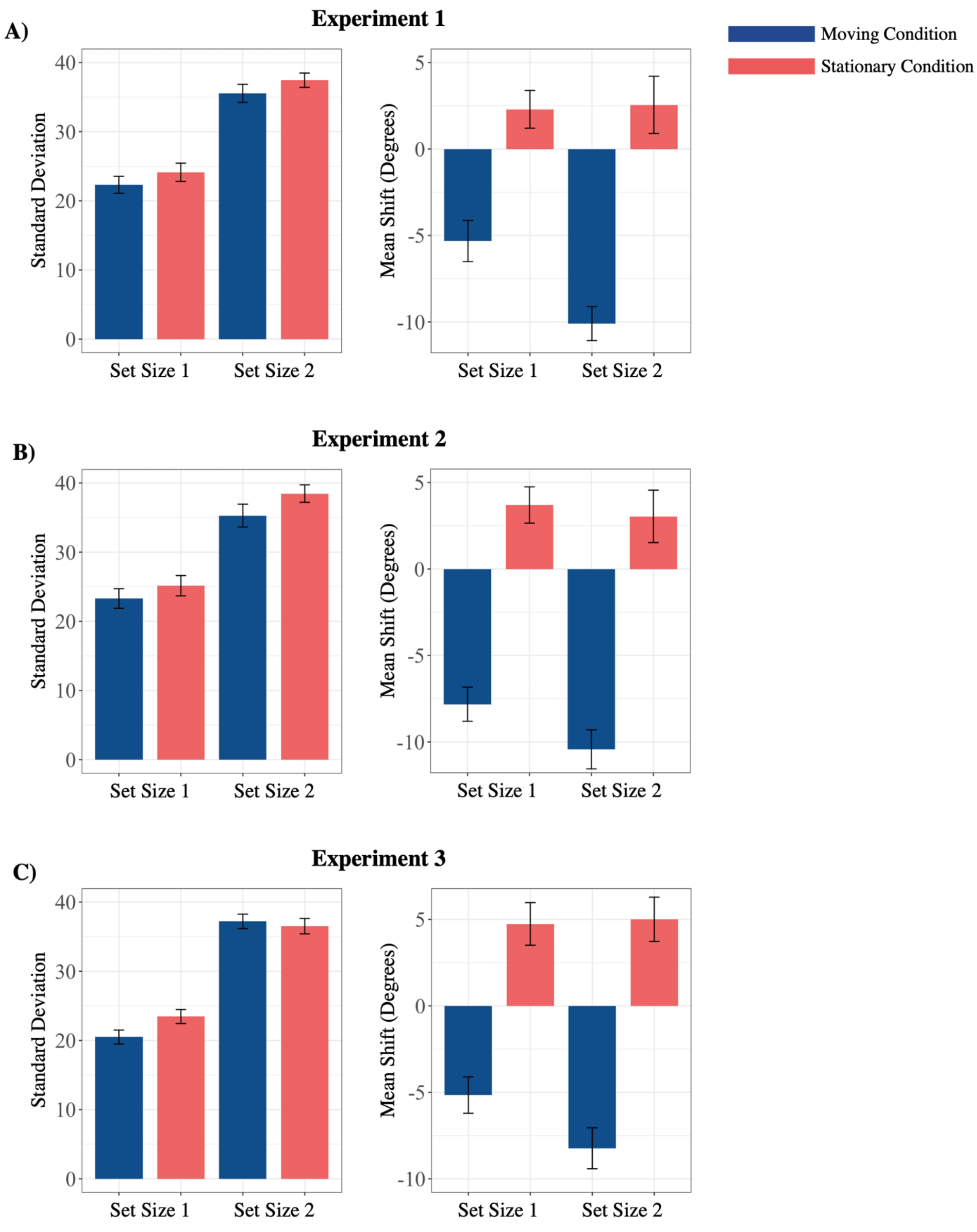

Figure 5. Mean shift and circular standard deviation data for each experiment.

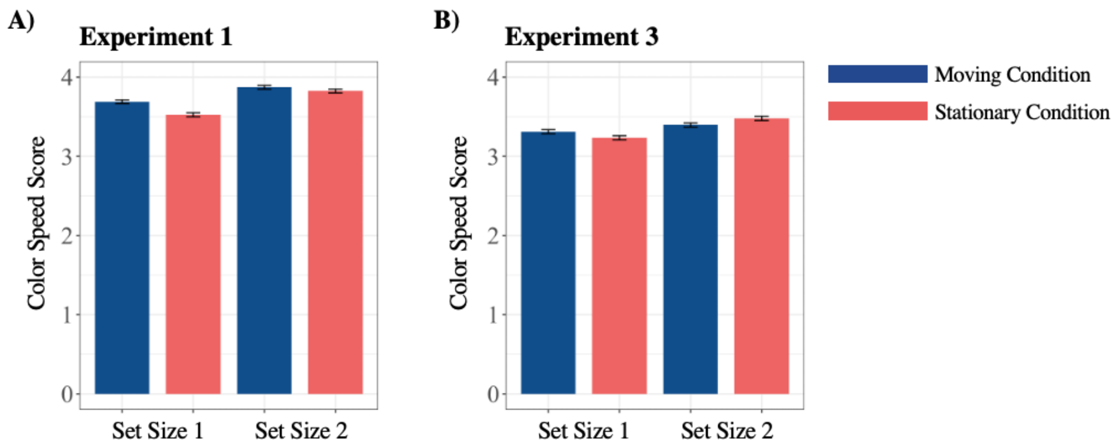

Figure 6. Color change speed judgment scores for A) Experiment 1 and B) Experiment 3. Participants rated how fast the colors were changing using a Likert scale 1~4 with 4 being the fastest.

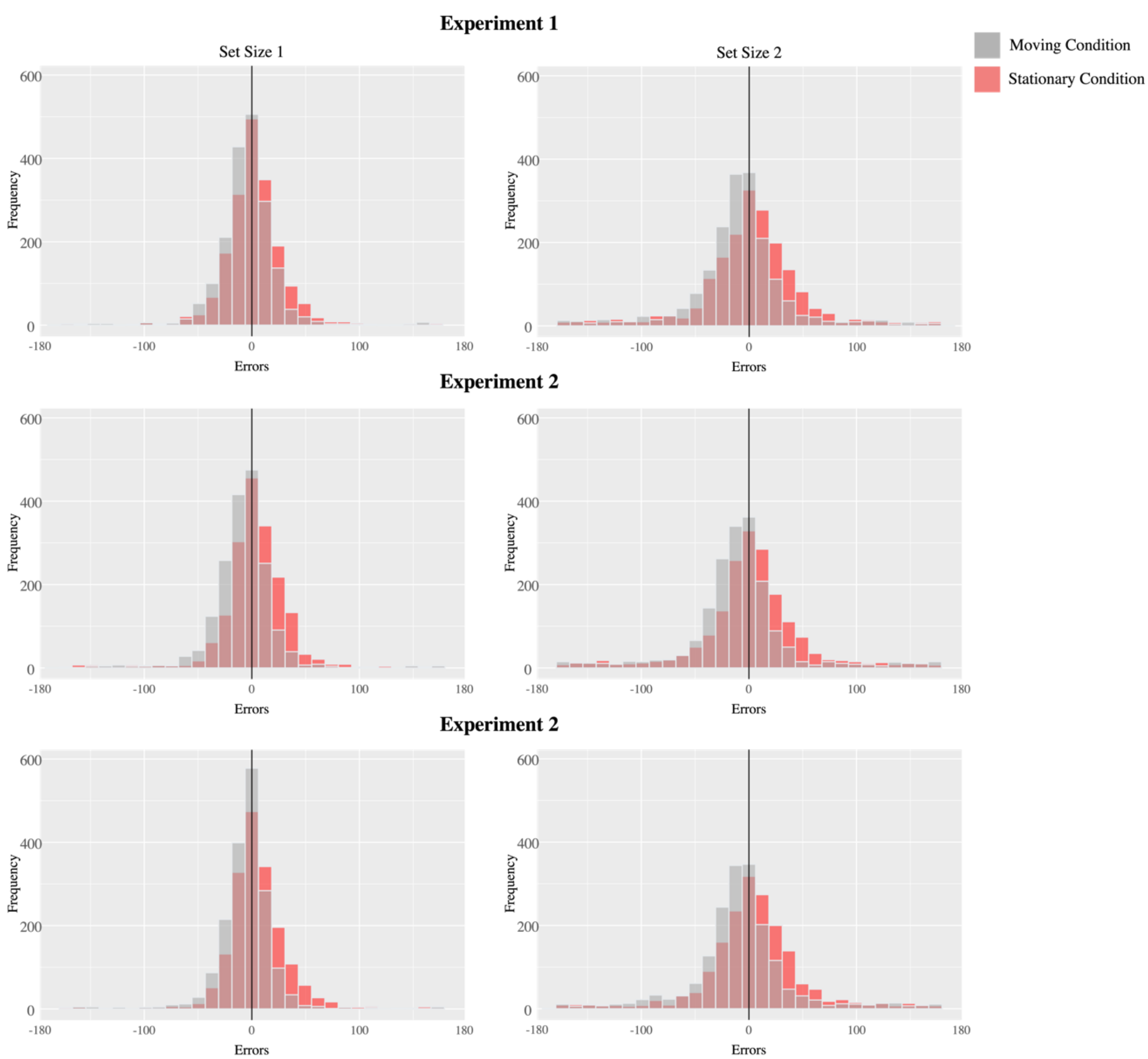

Figure 7. Error distribution of each experiment by condition.

General Discussion

While many visual working memory studies use static, unchanging stimuli, objects in the real world are often constantly changing in both their locations and features. Here we found that visual working memory representations of dynamically changing stimuli are more biased towards the past when the objects are moving compared to when they are stationary. This aligns with our hypothesis that when objects are in motion, their feature information is more naturally integrated over time into a single object file, and that this increased incorporation of past features makes the memory representations lag behind the current state. Critically, we also found that object motion resulted in more precise memory reports compared to static objects. This again aligns with our hypothesis that increased feature integration for moving objects would result in reduced noise in memory representations as multiple noisy representations of features get averaged, by contrast to existing hypotheses that such lags represent evidence of serial processing and delayed sampling. Moreover, such effects were robust over low-level visual changes in stimuli such as more vs. less random motion patterns and slower vs. faster color changes. The combination of larger memory lag and higher memory precision suggest the role of object motion is that of facilitating more feature integration of dynamically changing stimuli.

Previous studies using dynamic stimuli have found that tracking the changing features result in perceptual lags where observers' representation lags behind the current state of the stimuli (Howard & Holcombe, 2008; Howard et al., 2011). One suggested mechanism of this result was that serial processing of each stimulus is contributing to delays in observers' ability to update the continuously changing feature state, resulting in lagged reports (Howard & Holcombe, 2008). However, such serial processing cannot account for our findings in increased precision when the dynamic stimuli are in motion. However, this precision gain *can* be explained by increased temporal integration of the features in the moving condition. Prior works in ensemble perception has shown that observers can accurately judge the average information of multiple items (e.g., Ariely, 2001; Chong & Treisman, 2005), and that the uncorrelated noises can be cancelled out through the process of averaging multiple noisy inputs, leading to more precise ensemble representation than individual item representation (e.g., Alverez, 2011; Brezis et al., 2015; Baek & Chong, 2020). Similarly, more temporal integration in the moving condition in our task could also result in more "samples" of each individual feature representation as part of the final object representation, with noise cancelling out, resulting in higher precision observed in our current data.

Unlike the previous investigations that showed perceptual lags when tracking dynamic stimuli, here we also find that stationary dynamic colored dots resulted in a slight, non-significant positive bias towards the direction of the color rotation. While this may seem contradictory to previous results, Howard & Holcombe (2008) also report that the perceptual lag was not equally strong for all types of features. For instance, tracking changing spatial periods resulted in much larger perceptual lag than tracking rotating orientations. Another study by Callahan-Flintoft et al. (2020) used stationary dynamic color dots very similar to our paradigm and showed that when observers are asked to report the color at a specific cued time reports are biased towards the colors that follow the cue. Although our paradigms did not have a cue but rather asked observers to report the final state, it could be possible that in the stationary conditions participants inferred from the smooth color changing and reported colors that are slightly biased towards the future colors in the direction of the color change. Such biases are completely gone when the stimuli are in motion and representation lags become a lot more prominent.

Why is temporal integration so much more prominent when stimuli are in motion? Theories based on the idea of "object files" may be useful to helping explain such findings. In real-world vision, we encounter an overwhelming amount of visual information that is changing around from moment to moment. Nevertheless, we are able to maintain a stable visual percept in such a noisy environment. One strategy our visual system takes to do this is to simplify the world by forming enduring representations of a few objects and integrating the information about these objects over time, a process often referred to as an 'object file' (e.g., Kahneman et al., 1992). Representations that are formed into an object file have certain privileges such as allocation of attention and priming effects (Blaser, Pylyshyn & Holcombe, 2000). This can also be advantageous when establishing invariant object perception. For instance, we are able to recognize objects despite dramatic variations (Li & DiCarlo, 2008; Cox et al., 2005). Similarly, when features of the object (i.e., colors) change due to varying environmental aspects of the object such as viewpoints or lighting, we are able to integrate the changing features into the single object file. Previous works have shown that spatiotemporal continuity has an important role in creating such enduring object files (e.g., Spelke et al., 1995; Gao & Scholl, 2008) and may also be important for creating robust, invariant long-term memories of real-world objects (Schurgin & Flombaum, 2017). This explains that motion in our experiments induced more enduring object files for the dynamic color dots, ultimately resulting in color integration over a longer range of time periods than the static color dots. Other works have shown that such continuous motion has been found to improve color constancy (Wener, 2007) and also result in more bias towards past states of stimuli (Liberman et al., 2016). This is consistent with the finding in our experiment that the effects were less robust in Experiment 3, where both motion direction and speed change more sporadically, interrupting the continuity that helps facilitate the representation of a stable object file.

One concern in current experiments is that the stationary condition and motion condition are inherently different in their low-level aspects and motion may naturally induce additional cognitive processes during the task. For example, moving stimuli are known to induce representational momentum where observers project the representation forward towards future (e.g., Kelly & Freyd, 1987), potentially influencing how colors and object positions are bound in the memory representation. However, such perceptual differences between moving and stationary stimuli would not predict higher precision in memory reports for the moving stimuli. Rather, tracking the moving objects while concurrently tracking the changing features would require more attentional resources and therefore would predict an overall impairment on the working memory task. Generally, motion is known to make object recognition harder as it limits the viewing time of the object at a specific location. Many other works investigating multiple identity tracking have shown that identity information (i.e., objects' features) is not well encoded when tracking their motion (e.g., Scholl et al., 1999; Horowitz et al., 2007; Bahrami, 2003). While these studies are largely dealing with much more cognitively demanding situations than ours (i.e., tracking more than two items), another work by Saiki (2003) has shown that color change detection suffers when stimuli are are moving even at set size 2. By contrast to such a prediction, we find that memory precision was significantly higher when the dynamically changing stimuli were in motion. Future investigations could aim to control for the low-level differences between the two conditions by employing other manipulations that can effectively break the continuous object perception. For instance, other studies have shown that past stimuli features have less influence on the later features if the object features are changing randomly without any continuous trajectory (Callahan-Flintoft et al., 2020) or when the stimuli do not follow a continuous motion trajectory (Liberman et al., 2016). Similarly, we expect that the main effects observed in our study would be reduced if either the stimuli features or the motion is less continuous.

Additionally, we collected participants' subjective ratings of the color changing speeds for Experiments 1 and 3 to investigate whether the "motion silencing" effect was present in our study. Motion silencing refers to an illusion where changing features of stimuli do not appear changing when they are in motion (Suchow & Alvarez, 2011). While further investigations on motion silencing have shown that visual crowding is required (Turi & Burr, 2013), we wanted to ensure that the current results did not arise due to observers not seeing the color changes in the motion conditions. Our results showed that observers rated colors to be changing faster in set size 2 compared to set size 1. For set size 1 conditions, observers found moving colored dots to be changing faster, but this pattern was not found for set size 2 conditions. It may be that motion silencing was present for set size 2 conditions and thus resulted in elimination of motion induced color speed bias observed in the set size 1 conditions. Nevertheless, overall our data shows that the motion silencing effect is not likely to be the main driver of the perceptual lag and increased precision we found in the motion condition.

Taken together, we find that motion in dynamically changing stimuli can facilitate more enduring object files, resulting in increased integration of the changing features. Such integration process is reflected by more lagged but also more precise memory representations for moving stimuli. This can support our ability to maintain stable memory representations in drastically changing visual environments. Overall, these results suggest the role of motion and object files for supporting the creation of robust memory representations.

Open Science Statement

All data and an example analysis code are available online at https://osf.io/z7hk8/.

References

Albrecht, A. R., & Scholl, B. J. (2010). Perceptually averaging in a continuous Visual World: Extracting statistical summary representations over time. *Psychological Science*, *21*(4), 560–567. https://doi.org/10.1177/0956797610363543

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

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), 1–11. https://doi.org/10.1167/9.10.7

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature*, *408*(6809), 196–199. https://doi.org/10.1038/35041567

Callahan-Flintoft, C., Holcombe, A. O., & Wyble, B. (2020). A delay in sampling information from temporally autocorrelated visual stimuli. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-020-15675-1

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research* (Vol. 169, Issue 07). https://doi.org/10.1016/S0079-6123(07)00020-9

Daneman, M., & Carpenter, P. (1980). Inidividual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 840–853. https://doi.org/10.1037/0096-1523.32.4.840

Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2012). Spatiotemporal priority as a fundamental principle of object persistence. *The Origins of Object Knowledge*, 135–164. https://doi.org/10.1093/acprof:oso/9780199216895.003.0006

Fukuda, K., Pereira, A. E., Saito, J. M., Tang, T. Y., Tsubomi, H., & Bae, G. Y. (2022). Working Memory Content Is Distorted by Its Use in Perceptual Comparisons. *Psychological Science*, *33*(5), 816–829. https://doi.org/10.1177/09567976211055375

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin and Review*, *17*(5), 673–679. https://doi.org/10.3758/17.5.673

Gao, T., & Scholl, B. J. (2010). Are objects required for object-files? Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition, 18*(1), 82–109.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632–635. https://doi.org/10.1038/nature07832

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(3), 519–537. https://doi.org/10.1037/0096-1523.30.3.519

Howard, C. J., & Holcombe, A. O. (2008). Tracking the changing features of multiple objects: Progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Research, 48*(9), 1164-1180.

Howard, C. J., Masom, D., & Holcombe, A. O. (2011). Position representations lag behind targets in multiple object tracking. *Vision Research*, *51*(17), 1907–1919. https://doi.org/10.1016/j.visres.2011.07.001

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O

Kane, M. J., Conway, A. R. A., Bleckley, M. K., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169–183. https://doi.org/10.1037/0096-3445.130.2.169

Kelly, M. & Freyd, J. (1987) Explorations of representational momentum. Cognitive Psychology 19(3) 369-401

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(1996), 279–281.

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

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. https://doi.org/10.1038/nn.3655

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179–197.

Rademaker, R. L., Park, Y. E., Sack, A. T., & Tong, F. (2018). Evidence of Gradual Loss of Precision for Simple Features and Complex Objects in Visual Working Memory. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(6), 925–940. https://doi.org/10.1037/xhp0000491.supp

Saiki, J. (2003). Spatiotemporal characteristics of dynamic feature binding in visual working memory. *Vision Research*, *43*(20), 2107–2123. https://doi.org/10.1016/S0042-6989(03)00331-6

Shevell, S. K., & Kingdom, F. A. A. (2008). Color in complex scenes. *Annual Review of Psychology*, *59*, 143–166. https://doi.org/10.1146/annurev.psych.59.103006.093619

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1–2), 1–46. https://doi.org/10.1016/S0010-0277(00)00152-9

Scholl, B. J. (2019). What Have We Learned about Attention from Multiple-Object Tracking (and Vice Versa)? *Computation, Cognition, and Pylyshyn*, *2*, 49–78. https://doi.org/10.7551/mitpress/8135.003.0005

Schurgin, M. W., & Flombaum, J. I. (2017). Exploiting core knowledge for visual object recognition. *Journal of Experimental Psychology: General, 146*(3), 362–375. https://doi.org/10.1037/xge0000270

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, *4*(11), 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Schurgin, M. W. (2018). Visual memory, the long and the short of it: A review of visual working memory and long-term memory. *Attention, Perception, and Psychophysics*, *80*(5), 1035–1056. https://doi.org/10.3758/s13414-018-1522-y

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of*

Developmental Psychology, 13(2), 113–142. https://doi.org/10.1111/j.2044-835X.1995.tb00669.x

Starr, A., Srinivasan, M., & Bunge, S. A. (2020). Semantic knowledge influences visual working memory in adults and children. *PLoS ONE*, *15*(11 November), 1–12. https://doi.org/10.1371/journal.pone.0241110

Stoermer, V., & Brady, T. (2021). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Vision*, *21*(9), 2644. https://doi.org/10.1167/jov.21.9.2644

Suchow, J. W., & Alvarez, G. A. (2011). Motion silences awareness of visual change. *Current Biology*, *21*(2), 140–143. https://doi.org/10.1016/j.cub.2010.12.019

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10), 1–8. https://doi.org/10.1167/13.10.9

Turi, M., & Burr, D. (2013). The "motion silencing" illusion results from global motion and crowding. *Journal of Vision, 13*(5), Article 14. https://doi.org/10.1167/13.5.14

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 1120–1135. https://doi.org/10.1167/4.12.11

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860