Routledge
Taylor & Francis Group

Visual Cognition

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pvis20

Same set of visual pointers for biological and non-biological objects in working memory

Xinchi Yu & Ellen Lau

To cite this article: Xinchi Yu & Ellen Lau (11 Apr 2025): Same set of visual pointers for biological and non-biological objects in working memory, Visual Cognition, DOI: 10.1080/13506285.2025.2487860

To link to this article: https://doi.org/10.1080/13506285.2025.2487860

VISUAL COGNITION
https://doi.org/10.1080/13506285.2025.2487860

Same set of visual pointers for biological and non-biological objects in working memory

Xinchi Yu [a,b] and Ellen Lau [a,b]

[a]Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD, USA; [b]Department of Linguistics, University of Maryland, College Park, MD, USA

ABSTRACT
It has been well-established that visual working memory (VWM) is implemented by a limited set of ∼ 3 pointers/indexicals, binding features together. Most of these studies used non-biological objects (e.g., shapes) as stimuli. However, two separate lines of recent work on infants and adults have prompted a proposal that biological objects recruit a separate set of visual pointers in working memory, distinct from the classical ∼ 3 pointers for non-biological objects (the separate set hypothesis). The alternative possibility is that biological and non-biological objects share the same set of pointers (the same set hypothesis). In the current paper, we provide evidence supporting the same set hypothesis with a series of three experiments involving biological motions and complex shapes, as well as arguments reconciling evidence that seemed to support the separate set hypothesis.

ARTICLE HISTORY
Received 21 August 2024
Accepted 27 March 2025

KEYWORDS
Object pointers; animacy; working memory; feature binding

A ball moves behind an occluder. Now a book comes out from behind the same occluder, before returning behind it. Finally the occluder is removed, but there is only one object there! Although perplexing to adults, young preverbal infants seem not to be surprised by this sequence of events, as measured by looking time (Kibbe & Leslie, 2011; F. Xu & Carey, 1996), although they will be surprised if *no* objects are there. This is one striking example from a broad base of infant and adult evidence for the idea that a set of contentless visual pointers (or indexicals, indexes), pointing to features, supports human visual working memory (VWM) and visual object tracking (Baillargeon & Carey, 2012; Brody, 2020; Carey & Xu, 2001; Kahneman et al., 1992; Leslie et al., 1998; F. Xu, 1999; Pylyshyn, 1989, 2001; Quilty-Dunn et al., 2022; Yu & Lau, 2023). In this case, the now-standard interpretation of the infants' behavior is that infants represent the ball moving behind the occluder with an object pointer to "ball" conceptual features, and that when the book emerges, they represent this as the "same" object by changing the pointed features to those of "book" (Cheng & Kibbe, 2022; Kibbe & Leslie, 2013).

The pointers are hypothesized to be content-free and reassignable (Balaban et al., 2019, 2023; Thyer et al., 2022; Yu, 2024), i.e., that they are independent from the representation of features (which bind features together but do not represent features in themselves), and that they can be "remapped" to bind different features to the same object, or even represent different objects across time. Indeed, the existence of contentless visual pointers in working memory has also been supported by neural evidence in adults (Y. Xu & Chun, 2006; Balaban et al., 2019; Naughtin et al., 2016; Thyer et al., 2022; Y. Xu, 2009; Yu & Lau, 2025; for discussion see Yu & Lau, 2023). For example, Xu and Chun identified subregions in the posterior parietal cortex demonstrating load-dependent, capacity-limited responses during VWM retention with fMRI (functional magnetic resonance imaging); critically, this representation was agnostic to the number of features. This observation was subsequently corroborated by MEG (magnetoencephalography) evidence of a qualitatively similar response over the right posterior cortex (Yu & Lau, 2025). Complementary to the above univariate neural responses, recent work has explored multivariate responses tracking the number of VWM pointers as well (Jones et al., 2024a, 2024b; Thyer et al., 2022). Moreover, the number of pointers is limited to

CONTACT Xinchi Yu xcyu@umd.edu Department of Linguistics, 1401 Marie Mount Hall, University of Maryland, College Park, MD 20742, USA

around 3, as suggested by numerous infant and adult studies using VWM or object tracking tasks (Awh et al., 2007; Cowan, 2001; Feigenson et al., 2002; Luck & Vogel, 1997, 2013; Y. Xu & Chun, 2009). Apart from the limited capacity for pointers, the representation of features is also capacity-limited, as VWM is best modeled with a limited capacity for features *and* a limited set of pointers (Ngiam, 2023; Swan & Wyble, 2014; Van den Berg et al., 2014; Xie & Zhang, 2017).[1]

An intuitive hypothesis would be that these ~ 3 visual pointers support the representation of all kinds of objects, irrespective of e.g., their animacy. Somewhat surprisingly, some recent studies on infants and adults seem to offer evidence against this "same set" hypothesis, prompting the proposal that biological objects recruit a separate set of visual pointers other than those for non-biological objects (the separate set hypothesis; Liu et al., 2019; Murez & Smortchkova, 2014). Relevant infant and adult literatures are often not integrated; therefore here we first provide an overview of these two lines of evidence. We will show that the existing results can be accounted for by either hypothesis, and we then go on to introduce our current experiments designed to resolve the ambiguity in the adult studies.

Infant studies. In Xu and Carey's (1996) experiments, infants showed insensitivity to changes in the object behind an occluder, just as adults asked to visually track an object in motion are not disrupted by some featural changes (Zhou et al., 2010). However, the stimuli used in these infant studies were mostly inanimate. Subsequent studies observed that infants in a similar age range (or earlier) were surprised when the feature change is between biological and non-biological objects (Bonatti et al., 2002; Decarli et al., 2020; Kibbe & Leslie, 2019; Surian & Caldi, 2010). For example, Kibbe and Leslie observed a longer looking time when an inanimate ball changed into a doll head (or vice versa), compared to when the feature change did not change animacy status. This type of observation led Murez and Smortchkova (2014) to propose that biological objects recruit a separate set of object pointers (i.e., object files) from non-biological objects.

However, such evidence is not necessarily inconsistent with the same set hypothesis; one can simply assume that some kinds of object changes prompt a "reset" of visual pointers while others do not, and that these pointer resetting parameters are subject to change across the course of development. For example, in adult multiple object tracking studies, it has been found that some feature changes do not affect behavioral accuracy (suggesting the same pointer represents the object before and after the feature change), while other feature changes do hurt accuracy (Lyu et al., 2015; Zhao et al., 2020; Zhou et al., 2010). The process of pointer resetting is associated with a disruption in CDA (contralateral delay activity) in EEG (electroencephalography) which tracks the number of visual pointers in VWM (Balaban et al., 2023, 2024). In sum, infant evidence, although previously argued to support the separate set hypothesis, is in fact equally consistent with the same set hypothesis.

Adult studies. A different way to evaluate these hypotheses in adults is to examine whether the working memory capacity for non-biological objects is shared with the capacity for biological objects, as predicted by the same set hypothesis. A number of studies using the delay-match-to-sample paradigm have observed that holding more biological motions (BMs) in working memory does not appear to interfere with the capacity limit for non-biological objects in working memory, which might seem to support the separate set hypothesis. We note, however, that several of these studies presented the sample objects sequentially at the same location (Liu et al., 2019, Experiment 4-5; Wood, 2007, 2011). Continuity of location is a strong cue for the objective continuation of the same object across time, and thus against pointer resetting (Bapat et al., 2017; Golomb et al., 2014). Therefore, sequential, same-location presentation may encourage participants to represent the input as feature change within a single object pointer, rather than as multiple distinct objects. In keeping with this, fMRI work has shown that responses associated with pointer maintenance only minimally increase with set size in sequential, same-location presentation, compared to simultaneous different-location presentation (Harrison et al., 2010; Y. Xu & Chun, 2006).

Perhaps the strongest current evidence for the separate-set hypothesis comes from two experiments by Liu et al. (2019) which used simultaneous different-location presentation (Exp. 2 and 3). In these two experiments, Liu and colleagues found that memorizing more BMs does not affect the capacity for color-shape bindings and vice versa. This result is consistent with the hypothesis that biological objects recruit a separate set of pointers from non-biological objects. However, another property of the experiment design provides an alternative explanation. In these experiments, the “foil” memory probes always used an object that had not appeared anywhere in the sample display (“object-change trials”) rather than using an “old” object that had been in the sample display but that was probed in the wrong location (“location-change” trials). This leaves open the possibility that participants could succeed at rejecting the foil trials without binding the BM features to object pointers at all, but simply by noting whether the probe’s BM feature had been detected recently (cf. Oberauer, 2017). In other words, in order to correctly respond to object-change trials, one does not need to represent BMs as bound objects, but just as a conceptual feature. And as the point-light displays (BMs) and color-shape bindings recruit very different feature spaces, confusion between the features themselves would be unlikely.

Fortunately, the use of location-change trials provides a straightforward way of probing the pointers for binding. If the probe is an old object (present in the sample) but at a wrong location previously occupied by another object (“location-change trials” hereafter), one cannot respond to these probes correctly by the conceptual features only; rather, correct response for location-change probes relies on the representation of conceptual feature-location *bindings*. The same set hypothesis predicts an interference effect when examining these location-change trials. In the current study, we put this to the test. In Experiment 1, we conceptually replicated the results of Liu et al. (2019) with another set of inanimate bound objects, complex shapes. In Experiment 2 and Experiment 3 (a replication), we introduced location-change trials to provide the crucial comparison. If biological and non-biological object representation relies on the same set of pointers, memory interference effects should be observed in these trials; the separate set hypothesis again predicts no interference.

Results

In Experiment 1 (Figure 1), we aimed to replicate previous results on biological and non-biological object maintenance using another set of bound inanimate objects: complex (multi-featured) shapes. Here we used the same object-change probes as previous

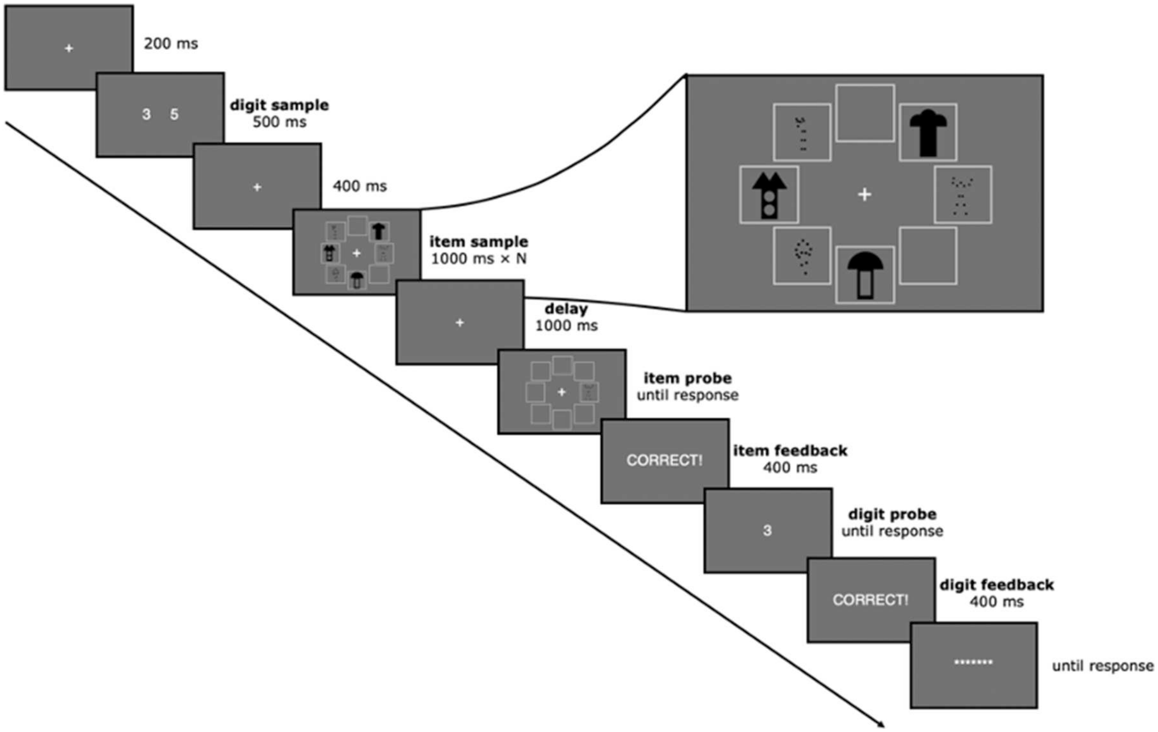


Figure 1. Illustration of a trial in Experiments 1-3.

experiments, but rather than asking participants to "judge whether the probe had appeared in the memory array" (Liu et al., 2019), we asked them to judge whether the probe matched the object that had been presented at that location, in order to encourage the representation of bound objects rather than free features.

Experiment 1

Subjects reached ceiling at the digit task (mean accuracy = 98%). In terms of Cowan's K, we observed a significant difference in K_{ACT} across the two conditions, W = 0, z = −4.3, $p < 0.001$ (K_{ACT} for 3OBJ + 1ACT: $M \pm SE = 0.83 \pm 0.04$; K_{ACT} for 3OBJ + 3ACT: $M \pm SE = 1.75 \pm 0.10$). The difference in K_{OBJ} across the two conditions was only marginally significant, W = 184.5, z = 1.8, $p = 0.06$ (K_{OBJ} for 3OBJ + 1ACT: $M \pm SE = 1.39 \pm 0.16$; K_{OBJ} for 3OBJ + 3ACT: $M \pm SE = 1.16 \pm 0.17$). See Figure 2.

In Experiment 1 we replicated Liu et al.'s (2019) observation of a non-interference effect, i.e., an increase in the Cowan's K for BMs held in working memory does not significantly discount the Cowan's K for complex shapes. However, this result may be driven by the fact that in both our Experiment 1 and Liu et al. (2019), the probe was always a new object (an object that was not in the sample). Although multi-featured complex shapes (our Experiment 1) and color-shape bindings (Liu et al., 2019) would be represented as bound objects, meaningful BMs can be represented as a conceptual *feature* only (action type, e.g., JUMPING, WAVING), without recruiting the pointers to bind multiple features together. Therefore, in Experiment 2 and Experiment 3, we added trials in which the probe is an old object presented at a wrong location previously occupied by another object (Figure 3). These trials cannot be correctly responded to by just holding the conceptual feature, enabling us to examine the Cowan's K for conceptual feature-location *bindings*. Note that the marginally significant effect for K_{OBJ} may reflect a true effect but with a small effect size. This is in line with our argument that biological and non-biological objects share the same set of VWM pointers.

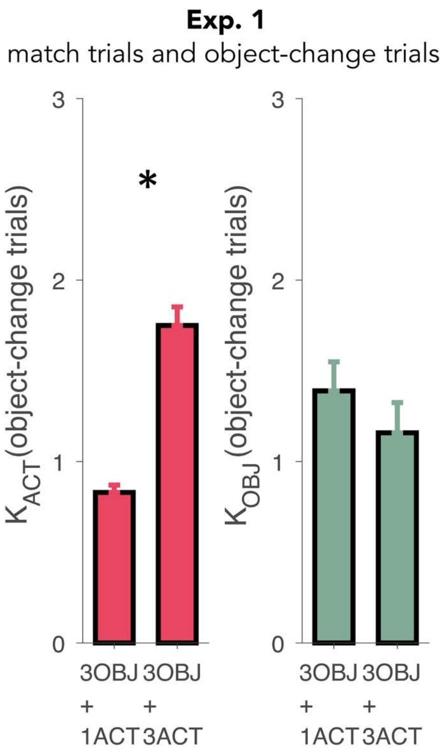


Figure 2. The results for our Experiment 1. K_{ACT} is the Cowan's K for BMs, and K_{OBJ} is the Cowan's K for static shapes. 3OBJ + 1ACT is the condition where 3 static shapes and 1 BM were in the sample, and 3OBJ + 3ACT is the condition where 3 static shapes and 3 BMs were in the sample. *: $p \leq 0.05$. Error bars stand for standard error.

Experiments 2 and 3 were very similar, both including the crucial location-change trials; Experiment 3 was essentially a replication of Experiment 2. The only difference between these experiments was the number of trials allocated to each condition, as described below.

Experiment 2

Subjects reached ceiling at the digit task (mean accuracy = 99%). For the match trials and the object-change trials, there was a significant difference between the K_{ACT} across the two conditions, W = 18, z = −3.8, $p < 0.001$ (K_{ACT} for 3OBJ + 1ACT: $M \pm SE = 0.79 \pm 0.05$; K_{ACT} for 3OBJ + 3ACT: $M \pm SE = 1.37 \pm 0.12$). Meanwhile, there wasn't a significant difference between the K_{OBJ} across the two conditions, W = 169, z = 1.4, $p = 0.17$ (K_{OBJ} for 3OBJ + 1ACT: $M \pm SE = 1.28 \pm 0.11$; K_{OBJ} for 3OBJ + 3ACT: $M \pm SE = 1.08 \pm 0.15$). See Figure 4.

For the match trials and the location-change trials, there was a significant difference between the K_{ACT} across the two conditions, W = 3, z = −4.2, $p < 0.001$ (K_{ACT} for 3OBJ + 1ACT: $M \pm SE = 0.91 \pm 0.04$; K_{ACT} for 3OBJ + 3ACT: $M \pm SE = 1.93 \pm 0.13$). There was also a

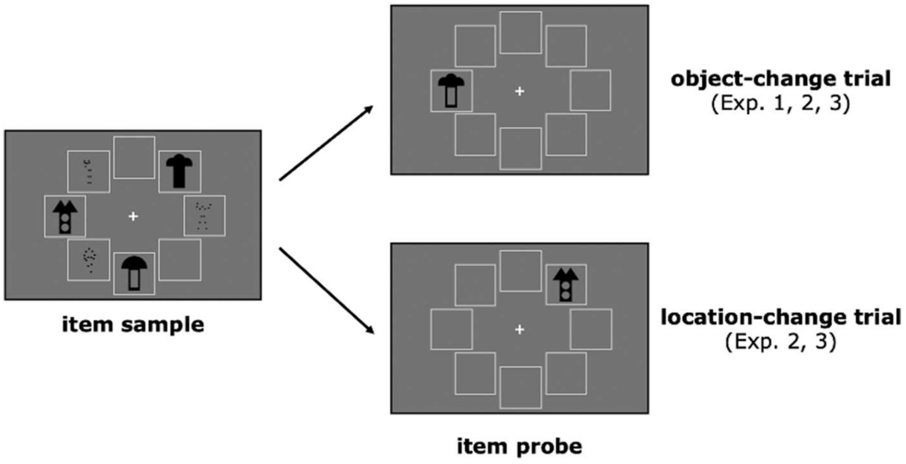


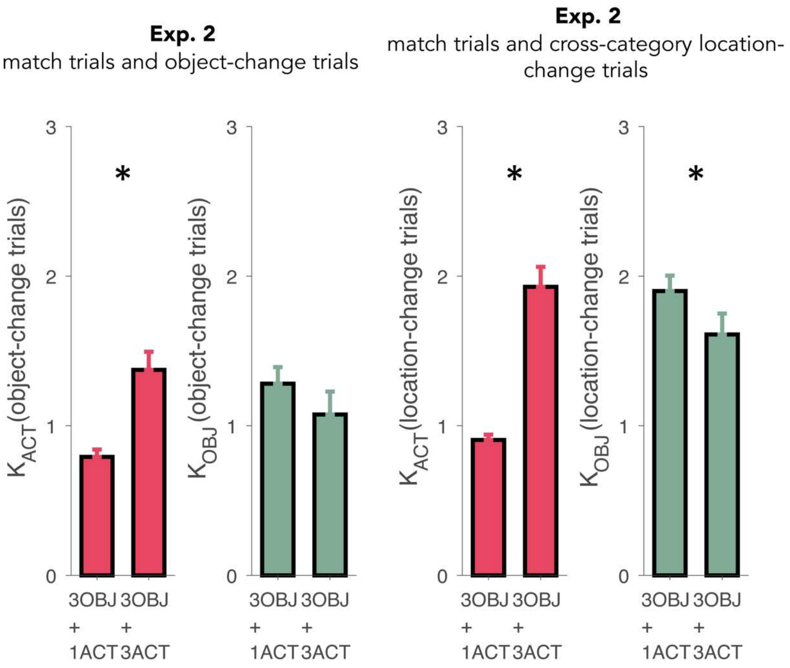
Figure 3. Illustration of the difference between an object-change trial and a location-change trial. In an object-change trial, the probe was a new object that didn't appear in the sample; in a location-change trials, the probe as an old object at a wrong location previously occupied by another sample object.

significant difference between the K_{OBJ} across the two conditions, W = 206, z = 2.6, $p = 0.01$ (K_{OBJ} for 3OBJ + 1ACT: $M \pm SE = 1.90 \pm 0.10$; K_{OBJ} for 3OBJ + 3ACT: $M \pm SE = 1.61 \pm 0.14$). See Figure 4.

Experiment 3

Subjects reached ceiling at the digit task (mean accuracy = 98%). For the match trials and the object-change trials, there was a significant difference in K_{ACT} across the two conditions, W = 21.5, z = −3.7, $p < 0.001$ (K_{ACT} for 3OBJ + 1ACT: $M \pm SE = 0.80 \pm 0.03$; K_{ACT} for 3OBJ + 3ACT: $M \pm SE = 1.33 \pm 0.12$). Meanwhile, there was no significant difference in K_{OBJ} across the two conditions, W = 155.5, z = 1.4, $p = 0.17$ (K_{OBJ} for 3OBJ + 1ACT: $M \pm SE = 1.07 \pm 0.12$; K_{OBJ} for 3OBJ + 3ACT: $M \pm SE = 0.91 \pm 0.12$). See Figure 5.

For match trials and location-change trials, there was a significant difference in K_{ACT} across the two conditions, W = 9, z = 4.0, $p < 0.001$ (K_{ACT} for 3OBJ + 1ACT: $M \pm SE = 0.86 \pm 0.03$; K_{ACT} for 3OBJ + 3ACT: $M \pm SE = 1.72 \pm 0.12$). There was also a significant difference in K_{OBJ} across the two conditions, W = 202, z = 1.9, $p = 0.05$ (K_{OBJ} for 3OBJ + 1ACT: $M \pm SE = 1.50 \pm$


Exp. 2
match trials and object-change trials

Exp. 2
match trials and cross-category location-change trials

Figure 4. The results for our Experiment 2. *: $p \leq 0.05$. Error bars stand for standard error.

0.16; K_{OBJ} for 3OBJ + 3ACT: $M \pm SE = 1.25 \pm 0.14$). See Figure 5.

GLME analysis for experiments 2 and 3

In order to implement a direct comparison between the significant and non-significant effects observed in Experiments 2 and 3 for K_{OBJ}, we then implemented a generalized linear mixed-effect (GLME) model (cf. Yu et al., 2024) for trial-by-trial results from Experiments 2 and 3, testing whether there was an interaction effect between the trial type and BM load on accuracy. Indeed, this interaction effect was statistically significant ($\beta = -0.33$, SE = 0.13, Z = −2.52, $p = 0.012$). The main effect of trial type was significant ($\beta = 1.10$, SE = 0.13, Z = 8.22, $p < 0.001$), and the main effect of the number of BMs was not ($\beta = -0.01$, SE = 0.05, Z = −0.21, $p = 0.83$). Based on the interaction effect, we separately analyzed object-change trials and location-change trials. We found that a main effect of BM load only manifested in location-change trials ($\beta = -0.38$, SE = 0.13, Z = −3.13, $p = 0.004$), but not in object-change trials ($\beta = -0.01$, SE = 0.05, Z = −0.22, $p = 0.83$). The overall accuracies for complex shapes across all conditions were 75.4% (1 BM, object-change trials), 75.1% (3 BM, object-change trials), 92.0% (1 BM, location-change trials), and 85.7% (3 BM, location-change trials). This complementary analysis strengthened the case that the interference across BMs and complex shapes was more pronounced for location-change trials than object-change trials.

Discussion

Across three experiments using a new set of complex shapes, we replicated the pattern observed previously: in object-change trials, the number of BMs to be remembered does not affect memory capacity for inanimate objects. However, under these conditions the task can be solved without using the limited set of pointers to bind BMs, by instead just coding the motions as unbound features of the scene (e.g., JUMPING, WAVING). Crucially, when we used location-change trials in Experiments 2 and 3 that preclude this solution by requiring feature-location *bindings*, we now observed an interference effect between the number of BMs and inanimate objects, exactly as the same set hypothesis would predict. A task-based explanation of the previous results can also help explain why some previous behavioral studies have observed interference between biological and non-biological objects, if variations in task parameters encourage binding vs.

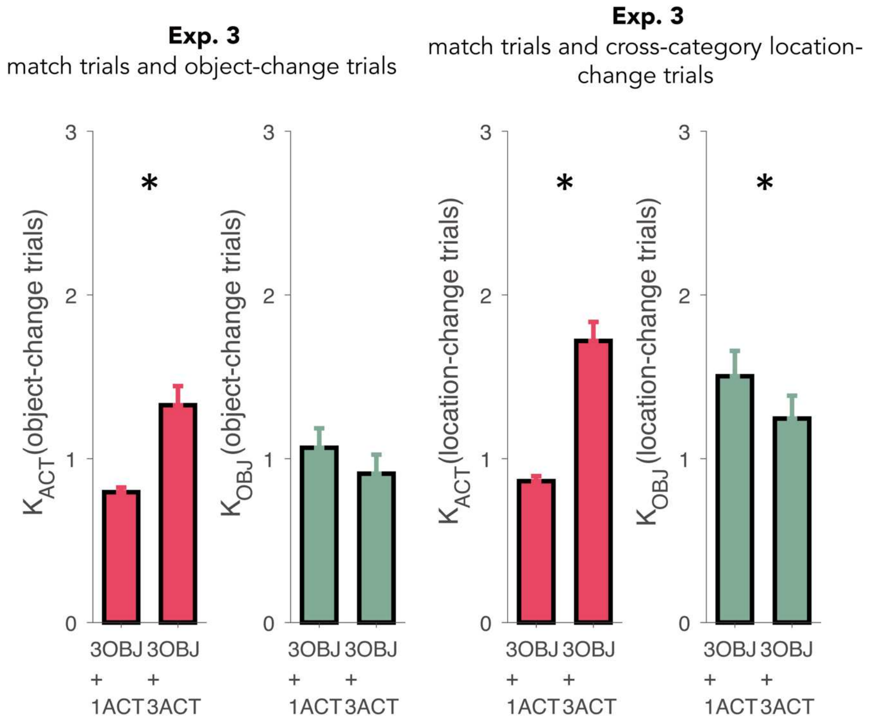


Figure 5. The results for our Experiment 3. *: $p \leq 0.05$. Error bars stand for standard error.

feature-based coding to different degrees (Li et al., 2023; Lu et al., 2019; Shen et al., 2014). Apart from task parameters (e.g., types of probes, and the proportion of trials with different types of probe), other factors including task instructions and individual preferences may also affect the extent to which binding is encouraged (Cao & Deouell, 2023; Qian et al., 2019; Vergauwe & Cowan, 2015); the contributions of these factors and their interactions remain to be fully established in the future.

In sum, the current results provide novel support for the hypothesis that biological and non-biological objects recruit the same set of visual pointers in working memory. This allows a simpler VWM architecture in which a single set of pointers can be used to bind and track visual objects regardless of their content.

Implications for shared and distinct representations for biological vs. non-biological objects

Our observation that a shared set of VWM pointers support the representation for both biological and non-biological objects is in interesting contrast with a line of research suggesting that biological and non-biological objects recruit rather different neural substrates. For example, it has recently been proposed that biological stimuli recruit a separate visual pathway (Küçük et al., 2024; Pitcher & Ungerleider, 2021), compared to non-biological ones. Our findings is not at odds with this line of research, since the *features* for biological and non-biological objects could indeed be represented by very different neural bases, given that these features appear to occupy rather different sub-spaces in the feature space, across visual and more conceptual features (Capitani et al., 2003; Carlson et al., 2013; Grootswagers et al., 2019; Hebart et al., 2020; Long et al., 2017; Spriet et al., 2022; R. Wang et al., 2022). In this context, the shared pointers allow for flexible binding of very distinct features to the same object: For example, if an object resembling a plane is approaching, and at a certain point as it comes closer you realize that it is actually Superman (Kahneman et al., 1992), one would be able to bind INANIMATE and ANIMATE sequentially to the same pointer without the need to treat the later-recognized animate object as a brand new object.

Connections with prior literature on feature binding

Our current study is consistent with prior observations suggesting that dual-task cost for working memory performance is more pronounced when feature binding instead of only the representation of features was probed (Wheeler & Treisman, 2002). For example, holding visual and auditory features simultaneously in VWM induces little inference with each other, while a more pronounced interference is observed when probing the capacity for bindings across visual/auditory features and location (Fougnie et al., 2015; Fougnie & Marois, 2011). Taken together, although the representation of features may be largely domain-specific and thus take up non-overlapping working memory resources (Fougnie et al., 2015; B. Wang et al., 2017), a domain-general binding resource (i.e., pointers) may bind features from a variety of representational systems. Whether or not we only have one set of pointers, and what each set of pointers covers (in terms of the types of features) remains to be fully charted (Yu, 2024; Yu & Lau, 2023). Our current study represents one attempt in this line.

Limitations and future directions

One limitation of our current study is that, for location-change trials, the interference effect across BMs and complex shapes may be explained by a limited VWM capacity for *locations*, instead of pointers. This in fact touches on two competing conceptions of the nature of VWM pointers. One conception is that VWM pointers are just locations (Treisman, 1998). The other conception is that VWM pointers are truly content-free, such that location is not qualitatively different from other features such as color and orientation, in the sense that they are all represented independently from pointers (Swan & Wyble, 2014; Wei et al., 2024; Yu, 2024).

While some studies have attempted to dissociate the contributions of location load and object load in VWM on neural responses (Jones et al., 2024a; Thyer et al., 2022; Y. Xu & Chun, 2006), this endeavor remains challenging given that location is perhaps one of the strongest cues toward objecthood. For example, flashing two objects sequentially at the same location (Jones et al., 2024a) could be mentally interpreted as one single object changing its features, or as one object replacing another. In the real world, an object at a certain location can indeed change its other features while remaining the same object (e.g., the same chameleon changing its color), or one object can take the place of another at the same location (e.g., the chameleon goes away and a parrot took its place). Therefore, whether two or one pointers are recruited when two objects are presented sequentially, likely depends on e.g., task parameters, including the type and extent of feature change (Park et al., 2020; Zhou et al., 2010), as well as individual differences.

Materials and methods

Experiment 1

Participants

24 subjects (17 female; age 18-30, $M \pm SD = 21.5 \pm 3.7$) participated in Experiment 1, matching the sample size in Liu et al. (2019). All subjects reported having normal or corrected-to-normal vision. Informed consent was obtained from all participants and they received monetary reimbursement for their participation. Procedures were approved by the UMCP IRB Office.

Stimuli

Following Liu et al. (2019), 7 point-light BMs from Vanrie and Verfaillie (2004) were used (waving, walking, spading, painting, jumping, cycling, chopping); the points were in black. For inanimate bound objects, 9 black complex shapes similar to those in Y. Xu and Chun (2006) were used.

Procedure

Each subject saw 168 trials in the main experiment. In 80 trials, the item sample consisted of 3 static shapes and 1 BM (3OBJ + 1ACT). In another 80 trials, the item sample consisted of 3 static shapes and 3 BMs (3OBJ + 1ACT). Of the 80 trials in each condition, 20 trials were probed with a static shape at the same location as in the sample (match), 20 trials were probed with a static shape not in the item sample (non-match), 20 trials were probed with a BM at the same location as in the sample (match), and 20 trials were probed with a BM not in the item sample (non-match). Non-match shape probes replaced the location of another static shape, and non-match biological probes replaced the location of another BM. An additional 8 filler trials consisted of 3 static shapes and 2 BMs. Each subject went through at least 12 practice trials before the experiment.

The experiment was run in MATLAB 2020b with Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007) using customized script on a 13-inch MacBook Pro. Like Liu et al. (2019), we included a digit task to occupy participants' phonological loop. At trial onset, a fixation cross was centrally presented center for 200 ms on a gray background. Then, two digits randomly sampled from 1 to 9 were centrally presented for 500 ms. Subjects were instructed to repeat the two digits aloud until a line of asterisks indicated the end of a trial. After another 400 ms fixation screen, the item sample appeared. Eight light grey squares (each 1°×1°) were equally distributed at 0°, 45°, 90°, … , 315° around an invisible circle with a radius of 1.5° centered on the screen. The squares were employed to separate the objects and discourage visual grouping. The static shapes and BMs were presented at the center of the squares, within an visual angle of 0.9°×0.9°. Following Liu et al. (2019) the duration of item sample was 1000 ms × N (the number of items) in order to ensure enough encoding time. Thus for the 3OBJ + 1ACT condition, the item sample was presented for 4000 ms; for the 3OBJ + 3ACT condition, the item sample was presented for 6000 ms. The BMs were presented on a loop with 12 frames across 1s. In a few rare trials (<0.5% of all experimental trials), a technical glitch resulted in an inaccurate duration of the item sample; these trials were excluded from further analysis. After a 1000 ms delay, a probe item (also within the range of 0.9°×0.9°) appeared. Subjects were asked to respond whether the probe item matched with the item presented before at the same location, by pressing m (match) or n (non-match) using their right hand. Feedback on their response (correct/incorrect) was presented for 400 ms. Then, a digit probe was presented at the center of the screen and participants responded whether it matched with one of the two digits presented before (50% possibility), pressing m (match) or n (non-match) using their right hand. Response feedback was then presented for 400 ms. A line of asterisks ended the trial; subjects pressed the space bar to proceed to the next trial. Note that neither the response to the item probe nor the digit probe was time-constrained, different from Liu et al. (2019). This is because the formula for Cowan's K does not take into account no-response trials (Cowan, 2001; Rouder et al., 2011). See Figure 1 for a schematic trial illustration.

Data analysis

Cowan's K. Memory task performance was quantified with Cowan's K (Cowan, 2001; Rouder et al., 2011), an indicator of the number of items held in working memory. The accuracy for the digit task was calculated as the proportion of correct responses among the rest of the 160 trials. Then for each subject, four Cowan's Ks were calculated: the Cowan's K for static shapes (K_{OBJ}) in the 3OBJ + 1ACT condition, the Cowan's K for BMs (K_{ACT}) in the 3OBJ + 1ACT condition, the K_{OBJ} in the 3OBJ + 3ACT condition, and the K_{ACT} in the 3OBJ + 3ACT condition. The Cowan's K formula for single probe experiments is $K = N \times (H-F)$, where N is the number of sample items, H is the hit rate for detecting the change in the probe on non-match trials, and F is the false alarm rate (erroneously "hallucinating" a change in the probe on match trials). For example, K_{OBJ} for the 3OBJ + 1ACT condition was calculated using shape-probe trials which had 3 static shapes in the sample. Thus, $N = 3$, H is the proportion of "non-match" responses among non-match trials, and F is the proportion of "non-match" responses in match trials. K_{ACT} for the 3OBJ + 1ACT condition was calculated using motion-probe trials which had 1 BM in the sample. Thus, $N = 1$, H is the proportion of "non-match" responses among non-match trials, and F is the proportion of "non-match" responses in match trials. The other two Ks were similarly calculated. Planned two-tailed paired-sample t tests were administered between the K_{OBJ} across the two conditions (i.e., 3OBJ + 1ACT vs. 3OBJ + 3ACT) and the K_{ACT} across the two conditions. Wilcoxon signed-rank tests were run in JASP 0.16.1 (JASP Team, 2022); a non-parametric test was employed because one distribution only marginally passed the Shapiro–Wilk normality test in this experiment ($p = 0.1$).

Experiments 2 and 3

Participants

Another 24 subjects (18 female; age 18-29, $M \pm SD = 20.9 \pm 3.1$) participated in Experiment 2; yet another 24 subjects (14 female, age 18-27, $M \pm SD = 20.2 \pm 2.4$) participated in Experiment 3. The sample size of Experiment 3 was sufficient to detect the K_{OBJ} difference for location-change trials observed in Experiment 2 with > 80% power ($\alpha = 0.05$; G*Power 3.1, Faul et al., 2007). All subjects reported having normal or corrected-to-normal vision. Informed consent was obtained from all participants and they received monetary reimbursement or course credit for participation. Procedures were approved by the UMCP IRB Office.

Stimuli

Same as Experiment 1.

Procedure

In addition to the 168 trials in Experiment 1, we added 80 non-match trials where the probe had appeared in the sample but in a different location (40 trials in each sample condition; half shape probes and half motion probes). The only other change from Experiment 1 was that this time our instructions explicitly requested subjects to use the digit task to help them avoid verbally labeling and rehearsing the items, in order to further reduce the contribution of the articulatory loop.

The only difference between Experiment 2 and 3 is the allocation of different types of trials within the 80 non-match trials. In Experiment 2, 40 non-match probes were old biological action, and 40 were old complex shape, presented at an old but wrong location. Because we selected that wrong location randomly, and because in the 1ACT + 3OBJ condition there were more object locations than action locations, the number of cross-category location-change trials to calculate K_{ACT} varied between 2 and 10 for the 1ACT + 3OBJ condition; for three subjects there were only a few (2 or 3) cross-category location-change trials with which to calculate K_{ACT} for the 1ACT + 3OBJ condition.[2] Therefore, in Experiment 3 we replicated Experiment 2 but this time fixed the number of different types of trials based on the expectancy for random swapping. For the 1ACT + 3OBJ condition, the number of location-change trials from a complex shape to a biological action was 10, the number of location-change trials from a biological action to a complex shape was 10, and the number of within-category location-change trials for complex shapes was 20. For the 3ACT +

3OBJ condition, the number of location-change trials from a complex shape to a biological action condition was 12, the number of location-change trials from a biological action to a complex shape was 12, the number of within-category location-change trials for complex shapes was 8, and the number of within-category location-change trials for BMs was 8.

Data analysis: experiment 2

Cowan's K. As in Experiment 1, we excluded from analysis a small number of trials with inaccurate timing (<0.5% experimental trials across the entire dataset). For Cowan's K, we first calculated K_{ACT} and K_{OCT} with the match trials and the object-change trials as in Experiment 1. Then, we calculated K_{ACT} and K_{OCT} with the match trials and the newly-added location-change trials. The absolute value of Cowan's K can be affected by the similarity between the objects before and after change (Awh et al., 2007; Brady & Alvarez, 2015; Fukuda et al., 2010; Gao et al., 2013); therefore, in order to match for this across Ks, we only used location-change trials where the change is cross-category (i.e., a BM changing into a complex shape or vice versa). K_{OBJ} was calculated with match trials where the probe was a complex shape, and location-change trials where a complex shape changed to a BM (3OBJ + 1ACT: 20 trials/subject; 3OBJ + 3ACT: on average 11.8 trials/subject). K_{ACT} was calculated with match trials where the probe was a BM, and location-change trials where a BM changed to a complex shape (3OBJ + 1ACT: 6.2 trials/subject; 3OBJ + 3ACT: 11.7 trials/subject).

GLME (generalized linear mixed-effect) model analysis. Complementarily, we used a GLME model to fit trial-level data to predict participants' trial-by-trial VWM performance (coded 1 for correct, and 0 for incorrect responses), for data in Experiments 2 and 3 (where both object-change trials and location-change trials were present) collectively. In this model, participants' overall likelihood of recognition success () across trials can be directly modeled based on a linear combination of trial-level predictors with a logistic link,

$$\ln\left(\frac{P_{success}}{1-P_{success}}\right) \sim \beta_0 + \beta_{trial\_type} + \beta_{BM\_load} + \beta_{trial\_type\times BM\_load} + 1|(subject:experiment)$$

where β_{trial_type} captures the variance accounted for by trial type (object-change trial vs. location-change trial), β_{BM_load} captures the variance accounted for by BM load (1 or 3). $\beta_{trial\_type\times BM\_load}$ stands for the interaction between trial type and BM load. Subject and Experiment (Experiment 2 vs. Experiment 3) were treated as random effects. This analysis was performed on the same trials that entered the Cowan's K analyses. BM load was standardized before regression. Based on the significant interaction effect between trial type and BM load, we conducted separate regression analyses for data from each trial type (i.e., object-change trials and location change trials). Statistical significance was evaluated using Wald's Z test, and all reported *p*-values are two-tailed.

Data analysis: experiment 3

Cowan's K. Same as Experiment 2. As in Experiments 1 and 2, we excluded from analysis a small number of trials with inaccurate timing (0.5% experimental trials across the entire dataset).

GLME model analysis. Detailed in the previous section.

Notes

1. This isn't to say that this pointer system is the only format of VWM; evidence also suggests a complementary system for numerosity estimation and summary statistics (Feigenson, 2011; Odic & Starr, 2018; Knowlton & Gomes, 2022; Kutter et al., 2023).
2. Removing these three subjects did not qualitatively change the results.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by NSF under grant [#1749407] (to E.L.).

Author contributions

Xinchi Yu: Conceptualization, data curation, investigation, formal analysis, methodology, software, visualization, writing-original draft, writing-review & editing. **Ellen Lau**: Conceptualization, funding acquisition, methodology, resources, supervision, writing-original draft, writing-review & editing.

Data availability statement

Raw data and analysis code are openly shared via OSF (https://osf.io/2stpw/).

ORCID

Xinchi Yu http://orcid.org/0000-0002-6165-7909
Ellen Lau http://orcid.org/0009-0007-0830-2205

References

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Baillargeon, R., & Carey, S. (2012). Core cognition and beyond: The acquisition of physical and numerical knowledge. In S. M. Pauen (Ed.), *Early childhood development and later outcome* (pp. 33–65). Cambridge University Press.

Balaban, H., Drew, T., & Luria, R. (2019). Neural evidence for an object-based pointer system underlying working memory. *Cortex*, *119*, 362–372. https://doi.org/10.1016/j.cortex.2019.05.008

Balaban, H., Drew, T., & Luria, R. (2023). Dissociable online integration processes in visual working memory. *Cerebral Cortex*, *33*(23), 11420–11430. https://doi.org/10.1016/j.cortex.2019.05.008

Balaban, H., Smith, K. A., Tenenbaum, J. B., & Ullman, T. D. (2024). Electrophysiology reveals that intuitive physics guides visual tracking and working memory. *Open Mind*, *8*, 1425–1446. https://doi.org/10.1162/opmi_a_00174

Bapat, A. N., Shafer-Skelton, A., Kupitz, C. N., & Golomb, J. D. (2017). Binding object features to locations: Does the "spatial congruency bias" update with object movement? *Attention, Perception, & Psychophysics*, *79*(6), 1682–1694. https://doi.org/10.3758/s13414-017-1350-5

Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, *44*(4), 388–426. https://doi.org/10.1006/cogp.2002.0779

Brady, T. F., & Alvarez, G. A. (2015). No evidence for a fixed object limit in working memory: Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 921. https://doi.org/10.1037/xlm0000075

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brody, G. (2020). Indexing objects in vision and communication. [Unpublished doctoral dissertation]. Central European University.

Cao, R., & Deouell, L. Y. (2023). Task-Dependent representation of feature binding within the same dimension in visual working memory. *bioRxiv*. https://doi.org/10.1101/2023.12.19.572444

Capitani, E., Laiacona, M., Mahon, B., & Caramazza, A. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology*, *20*(3-6), 213–261. https://doi.org/10.1080/02643290244000266

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1-2), 179–213. https://doi.org/10.1016/s0010-0277(00)00154-2

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*(10), 1–1. https://doi.org/10.1167/13.10.1

Cheng, C., & Kibbe, M. M. (2022). Development of updating in working memory in 4–7-year-old children. *Developmental Psychology*, *58*(5), 902. https://doi.org/10.1037/dev0001337

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/s0140525 (01003922

Decarli, G., Franchin, L., Piazza, M., & Surian, L. (2020). Infants' use of motion cues in object individuation processes. *Journal of Experimental Child Psychology*, *197*, 104868. https://doi.org/10.1016/j.jecp.2020.104868

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Feigenson, L. (2011). Objects, sets, and ensembles. In S. Dehaene & E. M. Brannon (Eds.), *Space, time and number in the brain* (pp. 13–22). Academic Press. https://doi.org/10.1016/b978-0-12-385948-8.00002-5.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science*, *13*(2), 150–156. https://doi.org/10.1111/1467-9280.00427

Fougnie, D., & Marois, R. (2011). What limits working memory capacity? Evidence for modality-specific sources to the simultaneous storage of visual and auditory arrays. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1329. https://doi.org/10.1037/a0024834

Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. *Journal of Experimental Psychology: General*, *144*(1), 30. https://doi.org/10.1037/a0038211

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*(5), 673–679. https://doi.org/10.1037/e520562012-095

Gao, Z., Ding, X., Yang, T., Liang, J., & Shui, R. (2013). Coarse-to-fine construction for high-resolution representation in visual working memory. *PLoS One*, *8*(2), e57913. https://doi.org/10.1371/journal.pone.0057913

Golomb, J. D., Kupitz, C. N., & Thiemann, C. T. (2014). The influence of object location on identity: A "spatial congruency bias". *Journal of Experimental Psychology: General*, *143*(6), 2262. https://doi.org/10.1037/xge0000017

Grootswagers, T., Robinson, A. K., Shatek, S. M., & Carlson, T. A. (2019). Untangling featural and conceptual object representations. *NeuroImage*, *202*, 116083. https://doi.org/10.1016/j.neuroimage.2019.116083

Harrison, A., Jolicoeur, P., & Marois, R. (2010). "What" and "where" in the intraparietal sulcus: An fMRI study of object identity and location in visual short-term memory. *Cerebral Cortex*, *20*(10), 2478–2485. https://doi.org/10.1093/cercor/bhp314

Hebart, M. N., Zheng, C. Y., Pereira, F., & Baker, C. I. (2020). Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nature Human Behaviour*, *4*(11), 1173–1185. https://doi.org/10.1038/s41562-020-00951-3

JASP Team. (2022). *JASP 0.16.1*.

Jones, H. M., Diaz, G. K., Ngiam, W. X., & Awh, E. (2024a). Electroencephalogram decoding reveals distinct processes for directing spatial attention and encoding into working memory. *Psychological Science*, *35*(10), 1108–1138. https://doi.org/10.1177/09567976241263

Jones, H. M., Thyer, W. S., Suplica, D., & Awh, E. (2024b). Cortically disparate visual features evoke content-independent load signals during storage in working memory. *Journal of Neuroscience*, *44*(44), e0448242024. https://doi.org/10.1523/JNEUROSCI.0448-24.2024.

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

Kibbe, M. M., & Leslie, A. M. (2011). What do infants remember when they forget? Location and identity in 6-month-olds' memory for objects. *Psychological Science*, *22*(12), 1500–1505. https://doi.org/10.1177/0956797611420165

Kibbe, M. M., & Leslie, A. M. (2013). What's the object of object working memory in infancy? Unraveling 'what' and 'how many'. *Cognitive Psychology*, *66*(4), 380–404. https://doi.org/10.1016/j.cogpsych.2013.05.001

Kibbe, M. M., & Leslie, A. M. (2019). Conceptually rich, perceptually sparse: Object representations in 6-month-old infants' working memory. *Psychological Science*, *30*(3), 362–375. https://doi.org/10.1177/0956797618817754

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Knowlton, T., & Gomes, V. (2022). Linguistic and non-linguistic cues to acquiring the strong distributivity of each. *Proceedings of the Linguistic Society of America*, *7*(1), 5236–5236. https://doi.org/10.3765/plsa.v7i1.5236

Küçük, E., Foxwell, M., Kaiser, D., & Pitcher, D. (2024). Moving and static faces, bodies, objects, and scenes are differentially represented across the three visual pathways. *Journal of Cognitive Neuroscience*, *36*(12), 2639–2651. https://doi.org/10.1162/jocn_a_02139.

Kutter, E. F., Dehnen, G., Borger, V., Surges, R., Mormann, F., & Nieder, A. (2023). Distinct neuronal representation of small and large numbers in the human medial temporal lobe. *Nature Human Behaviour*, *7*(11), 1998–2007. https://doi.org/10.1038/s41562-023-01709-3

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing 'what' and 'where' systems. *Trends in Cognitive Sciences*, *2*(1), 10–18. https://doi.org/10.1016/s1364-6613(97)01113-3

Li, C., Tian, W., He, Y., Wang, C., Wang, X., Xu, X., … Wu, S. (2023). How are patterned movements stored in working memory? *Frontiers in Psychology*, *14*, 1074520. https://doi.org/10.3389/fpsyg.2023.1074520

Liu, Y., Lu, X., Wu, F., Shen, M., & Gao, Z. (2019). Biological motion is stored independently from bound representation in working memory. *Visual Cognition*, *27*(9-10), 701–713. https://doi.org/10.1080/13506285.2019.1638479

Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision*, *17*(6), 20–20. https://doi.org/10.1167/17.6.20

Lu, X., Ma, X., Zhao, Y., Gao, Z., & Shen, M. (2019). Retaining event files in working memory requires extra object-based attention than the constituent elements. *Quarterly Journal of Experimental Psychology*, *72*(9), 2225–2239. https://doi.org/10.1177/1747021819839993

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://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*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Lyu, C., Hu, S., Wei, L., Zhang, X., & Talhelm, T. (2015). Brain activation of identity switching in multiple identity tracking task. *PLoS One*, *10*(12), e0145489. https://doi.org/10.1371/journal.pone.0145489

Murez, M., & Smortchkova, J. (2014). Singular thought: Object-files, person-files, and the sortal person. *Topics in Cognitive Science*, *6*(4), 632–646. https://doi.org/10.1111/tops.12110

Naughtin, C. K., Mattingley, J. B., & Dux, P. E. (2016). Distributed and overlapping neural substrates for object individuation and identification in visual short-term memory. *Cerebral Cortex*, *26*(2), 566–575. https://doi.org/10.1093/cercor/bhu212

Ngiam, W. X. Q. (2023). Mapping visual working memory models to a theoretical framework. *Psychonomic Bulletin & Review*, *31*(2), 442–459. https://doi.org/10.3758/s13423-023-02356-5.

Oberauer, K. (2017). What is working memory capacity?/¿ Qué es la capacidad de la memoria de trabajo? *Studies in*

Psychology, *38*(2), 338–384. https://doi.org/10.1080/02109395.2017.1295579

Odic, D., & Starr, A. (2018). An introduction to the approximate number system. *Child Development Perspectives*, *12*(4), 223–229. https://doi.org/10.1111/cdep.12288

Park, B., Walther, D. B., & Fukuda, K. (2020). Dynamic representations in visual working memory. *Journal of Vision*, *20*(11), 900–900. https://doi.org/10.1167/jov.20.11.900

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. *Trends in Cognitive Sciences*, *25*(2), 100–110. https://doi.org/10.1016/j.tics.2020.11.006

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*(1), 65–97. https://doi.org/10.1016/0010-0277(89)90014-0

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*(1-2), 127–158. https://doi.org/10.1016/s0010-0277(00)00156-6

Qian, J., Zhang, K., Liu, S., & Lei, Q. (2019). The transition from feature to object: Storage unit in visual working memory depends on task difficulty. *Memory & Cognition*, *47*(8), 1498–1514. https://doi.org/10.3758/s13421-019-00956-y

Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, *46*, e261.

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*, 324–330. https://doi.org/10.3758/s13423-011-0055-3

Shen, M., Gao, Z., Ding, X., Zhou, B., & Huang, X. (2014). Holding biological motion information in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1332. https://doi.org/10.1037/a0036839

Spriet, C., Abassi, E., Hochmann, J. R., & Papeo, L. (2022). Visual object categorization in infancy. *Proceedings of the National Academy of Sciences*, *119*(8), e2105866119. https://doi.org/10.1073/pnas.2105866119

Surian, L., & Caldi, S. (2010). Infants' individuation of agents and inert objects. *Developmental Science*, *13*(1), 143–150. https://doi.org/10.1111/j.1467-7687.2009.00873.x

Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, & Psychophysics*, *76*(7), 2136–2157. https://doi.org/10.3758/s13414-014-0633-3

Thyer, W., Adam, K. C., Diaz, G. K., Velazquez Sanchez, I. N., Vogel, E. K., & Awh, E. (2022). Storage in visual working memory recruits a content-independent pointer system. *Psychological Science*, *33*(10), 1680–1694. https://doi.org/10.1177/09567976221090923

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1373), 1295–1306. https://doi.org/10.1098/rstb.1998.0284

Van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, *121*(1), 124. https://doi.org/10.1037/a0035234

Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 625–629. https://doi.org/10.3758/bf03206542

Vergauwe, E., & Cowan, N. (2015). Working memory units are all in your head: Factors that influence whether features or objects are the favored units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1404. https://doi.org/10.1037/xlm0000108

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 226. https://doi.org/10.1037/xlm0000295

Wang, R., Janini, D., & Konkle, T. (2022). Mid-level feature differences support early animacy and object size distinctions: Evidence from electroencephalography decoding. *Journal of Cognitive Neuroscience*, *34*(9), 1670–1680. https://doi.org/10.1162/jocn_a_01883

Wei, N., Song, J., Zhang, H., & Zhou, T. (2024). Unravelling the object-based nature of visual working memory: Insight from pointers. *Memory & Cognition*. https://doi.org/10.3758/s13421-024-01643-3

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64. https://doi.org/10.1037/0096-3445.131.1.48

Wood, J. N. (2007). Visual working memory for observed actions. *Journal of Experimental Psychology: General*, *136*(4), 639. https://doi.org/10.1037/0096-3445.136.4.639

Wood, J. N. (2011). A core knowledge architecture of visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 357. https://doi.org/10.1037/a0023751

Xie, W., & Zhang, W. (2017). Dissociations of the number and precision of visual short-term memory representations in change detection. *Memory & Cognition*, *45*(8), 1423–1437. https://doi.org/10.3758/s13421-017-0739-7

Xu, F. (1999). Object individuation and object identity in infancy: The role of spatiotemporal information, object property information, and language. *Acta Psychologica*, *102*(2-3), 113–136. https://doi.org/10.1016/s0001-6918(99)00029-3

Xu, Y. (2009). Distinctive neural mechanisms supporting visual object individuation and identification. *Journal of Cognitive Neuroscience*, *21*(3), 511–518. https://doi.org/10.1162/jocn.2008.21024

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, *30*(2), 111–153. https://doi.org/10.1006/cogp.1996.0005

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95. https://doi.org/10.1038/nature04262

Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*(4), 167–174. https://doi.org/10.1016/j.tics.2009.01.008

Yu, X. (2024). Mapping the neural taxonomy of mental objects in moment-to-moment cognition. *PsyArXiv*, https://doi.org/10.31234/osf.io/3awv9

Yu, X., & Lau, E. (2023). The binding problem 2.0: Beyond perceptual features. *Cognitive Science*, *47*(2), e13244. https://doi.org/10.1111/cogs.13244

Yu, X., & Lau, E. (2025). A finite set of content-free pointers in visual working memory: magnetoencephalography (MEG) evidence. *NeuroReport*, *36*(3), 153–160. https://doi.org/10.1097/WNR.0000000000002132.

Yu, X., Thakurdesai, S. P., & Xie, W. (2024). Associating everything with everything else, all at once: Semantic associations facilitate visual working memory formation for real-world objects. *PsyArXiv*, https://doi.org/10.31234/osf.io/4fqbh

Zhao, C., Hu, L., Wei, L., Wang, C., Li, X., Hu, B., & Zhang, X. (2020). How do humans perform in multiple object tracking with unstable features. *Frontiers in Psychology*, *11*, 1940. https://doi.org/10.3389/fpsyg.2020.01940

Zhou, K., Luo, H., Zhou, T., Zhuo, Y., & Chen, L. (2010). Topological change disturbs object continuity in attentive tracking. *Proceedings of the National Academy of Sciences*, *107*(50), 21920-21924. https://doi.org/10.1073/pnas.1010919108