# City Research Online

## City, University of London Institutional Repository

This is the published version of the paper.

This version of the publication may differ from the final published version.

# Reconstructing the recent visual past: Hierarchical knowledge-based effects in visual working memory

Marie Poirier[1] · Daniel Heussen[1] · Silvio Aldrovandi[2] · Lauren Daniel[1] · Saiyara Tasnim[1] · James A. Hampton[1]

**Abstract** This paper presents two experiments that examine the influence of multiple levels of knowledge on visual working memory (VWM). Experiment 1 focused on memory for faces. Faces were selected from continua that were constructed by morphing two face photographs in 100 steps; half of the continua morphed a famous face into an unfamiliar one, while the other half used two unfamiliar faces. Participants studied six sequentially presented faces each from a different continuum, and at test they had to locate one of these within its continuum. Experiment 2 examined immediate memory for object sizes. On each trial, six images were shown; these were either all vegetables or all random shapes. Immediately after each list, one item was presented again, in a new random size, and participants reproduced its studied size. Results suggested that two levels of knowledge influenced VWM. First, there was an overall central-tendency bias whereby items were remembered as being closer to the overall average or central tokens (averaged across items and trials) than they actually were. Second, when object knowledge was available for the to-be-remembered items (i.e., famous face or typical size of a vegetable) a further bias was introduced in responses. The results extend the findings of Hemmer and Steyvers (*Psychonomic Bulletin & Review, 16*, 80–87, 2009a) from episodic memory to VWM and contribute to the growing literature which illustrates the complexity and flexibility of the representations subtending VWM performance (e.g., Bae, Olkkonen, Allred, & Flombaum, *Journal of Experimental Psychology: General, 144*(4):744–63, 2015).

One of the most fundamental functions that memory performs is to enable the past to support our current interactions with the world. The research presented herein examines how prior knowledge affects our memory for recently encountered visual stimuli (visual working memory; VWM).

The intellectual lineage of our experiments can be traced to a seminal paper by Janellen Huttenlocher and her collaborators. Huttenlocher, Hedges and Vevea (2000) examined how the distribution of exemplars within a single dimensional category influenced stimulus judgment. Observers were presented with one stimulus at a time and after a brief 2-second pause, they were asked to reproduce one of its characteristics from memory. Across experiments, to-be-remembered features included the length of horizontal lines, the grayness of squares, and the "fatness" or width of schematic fish. The distributions from which these stimuli were sampled were varied in terms of their mean, dispersion, and form (e.g., uniform or normal distributions) and the influence of these variations on the remembered features was systematically explored.

The judgment tasks just described can easily be construed as one-item VWM tasks, so the reported findings inform us with respect to knowledge effects in VWM. The Huttenlocher et al. (2000) results strongly suggested that VWM is constructive as they showed that memory was biased towards the central values of the categories called upon. For instance, if a studied line was shorter than the overall average line length, participants remembered the line as being somewhat longer than the one actually studied – in other words they

✉ Marie Poirier
   M.Poirier@city.ac.uk

[1] Psychology Department, City University London, Northampton Square, EC1V 0HB London, UK

[2] Psychology Department, Birmingham City University, Birmingham, UK

remembered the line as being closer to the average than it actually was. The authors referred to this phenomenon as the central-tendency bias.

At the heart of the model that underpinned their predictions and conclusions was the idea that the central-tendency bias is adaptive. Over many trials, if estimates are biased towards the more prototypical exemplars, performance will be less error prone on average. For example, if I remember an extreme value for a given item – considering the fallibility and imprecision of memory – there is a good chance that the said memory is inaccurate; the actual value is likely to be closer to the mean of the relevant category. Hence, over time, the central-tendency bias should produce behavior that is beneficial rather than detrimental.

A number of studies have explored alternative explanations of this basic phenomenon while others have replicated and extended it. In 2005, Sailor and Antoine provided further evidence of a central-tendency bias for single item memory but also suggested that the bias could be explained through the influence of immediately preceding stimuli; if a stimulus from one end of a distribution is presented, the preceding stimulus is more likely to be a less extreme value. Sailor and Antoine (2005) showed that such sequential dependencies could produce a central-tendency bias. However, Duffy, Huttenlocher, Hedges, and Crawford (2010) directly tested this hypothesis against the central-tendency bias view; they reported two experiments that showed that participants adjust their estimates towards the mean of all the stimuli encountered previously rather than towards a smaller and more recently encountered subset. They note that these findings do not mean that there is never an influence of recent, prior stimuli; rather, their results imply that such an influence is generally far smaller than the influence of the entire distribution. Sailor and Antoine (2005), as well as DeCarlo and Cross (1990), reported evidence to the effect that both the distribution as a whole and the immediately preceding stimulus affected estimates, but the influence of the immediately preceding stimulus was minor, relative to the influence of the entire distribution. In summary, there is no strong evidence for an explanation of the central tendency bias as a memory distortion caused by a subset of immediately preceding stimuli.

Brady, Konkle, and Alvarez (2009) offered another illustration of how prior knowledge can be integrated with noisy representations to support VWM performance. In their experiments, observers were presented with displays consisting of a small number of circles which varied in color; they were asked to remember the colors as well as their locations. In their task design, covariance was introduced between colors in a display so that over trials some color pairs were more likely to appear than other color pairs. Their findings showed that these redundancies led to more efficient encoding – i.e., after being exposed to stimuli with these built-in regularities, observers can store more information in working memory.

The latter finding extended influential slot models of VWM which suggested that the capacity of VWM is limited to a fixed number of slots (e.g., Zhang & Luck, 2008). A number of extensions to these fixed capacity models have been proposed in order to account for additional factors that affect VWM performance (e.g., Bae, Olkkonen, Allred, & Flombaum, 2015; Bae, Olkkonen, Allred, Wilson & Flombaum , 2014; Bays, Catalao, & Husain, 2009; Bays, Wu, & Husain, 2011; van den Berg, Shin, Chou, George, & Ma, 2012). For instance, Bae et al. (2015) proposed a model of color VWM where memory for a very recently encountered color is significantly influenced by knowledge of color categories as well as by the specific color value encountered. Bae et al. (2015) reported a central tendency bias for color memory as well as evidence suggesting that the bias originated in perception (see also Sims, Ma, Allred, Lerch, & Flombaum, 2016).

The studies reviewed so far have called upon simple/ abstract stimuli and most have also examined the effects of knowledge developed over the course of the experiment. What of the knowledge that participants bring to the experiment, i.e. longer-term knowledge of more familiar and meaningful stimuli? As far as we are aware, there are no studies systematically examining the biasing effects of this type of long-term knowledge on VWM; this was one of the objectives of the work reported here. In effect, our aim was to test a series of hypotheses derived from a general Bayesian perspective (see Hemmer & Steyvers, 2009b) which predicts that multiple levels of knowledge impact performance. Our work calls upon novel strategies in the study of VWM and differs from previous work in a number of important ways. We systematically examine the influence of well-established knowledge for complex and meaningful stimuli on VWM. In doing so, we report the impact of hierarchical levels of knowledge, i.e. knowledge that relates to the category from which studied items are taken (e.g., fruit sizes) and one that relates to *item-specific knowledge* (e.g., typical apple size). This means the interplay of multiple levels of representations can be considered, i.e. the representation of the to-be-remembered item, the representation of the relevant ensemble statistics, as well as the relevant item-specific long-term knowledge that is brought to the experimental task. For example, if the task is to remember the size of the most recently encountered apple, the assumption is that the response will mainly be based on the representation of said apple. However, two further knowledge-based sources can play a role: one would be the knowledge of what the typical size of an apple is (item-specific categorical knowledge) and the other would be the average size of all the fruit encountered in the experiment (superordinate categorical knowledge).

The work reported here extends the recent findings on VWM (Bae et al., 2015; Brady et al., 2009, 2011; Duffy et al., 2010) by examining hierarchical knowledge-based

effects with concrete, familiar and complex stimuli. Finally, using familiar stimuli allowed us to test knowledge-based biases while being confident that the observed effects were the result of the knowledge brought to the experiment rather than an artifact of sequential dependencies (e.g., Sailor & Antoine, 2005).

Experiment 1 was designed to investigate whether prior knowledge can bias VWM for faces. Experiment 2 borrowed from Hemmer and Steyvers (2009a) and examined the effect of prior knowledge on VWM for the size of familiar and unfamiliar objects.

## Experiment 1

In Experiment 1, participants were asked to remember short series of photographs of six different faces. Each of these faces was taken from a set of "families" created by morphing two faces and generating a continuum of stimuli that went from one face to the other (see Fig. 1). To manipulate prior knowledge, half of the morph continua were created by going from a famous face to an unfamiliar face (famous continua) while for the control set both faces were unfamiliar (non-famous continua).

We made predictions based on the assumption that two sources of available knowledge combine with the most recent representations to produce a response. Although the specific faces called upon were unfamiliar at the outset, people have considerable expertise in processing faces generally. Also, each family of faces was encountered repeatedly across the experiment. We expected that summary representations of each continuum would develop – in a similar fashion to what is observed with item sizes in other studies; we assumed that this would include an average representation that corresponds approximately to the middle of the series. This experiment-based knowledge was predicted to lead to a central-tendency bias where reconstruction should be pulled towards the center of each continuum. For famous continua, we expected the same central-tendency bias but with the added influence of the knowledge brought to the experiment: The prediction was that these faces would be remembered as being somewhat closer to the famous face than they actually were.

### Method

**Participants** Thirty psychology undergraduates took part in this study and received course credits for participating.
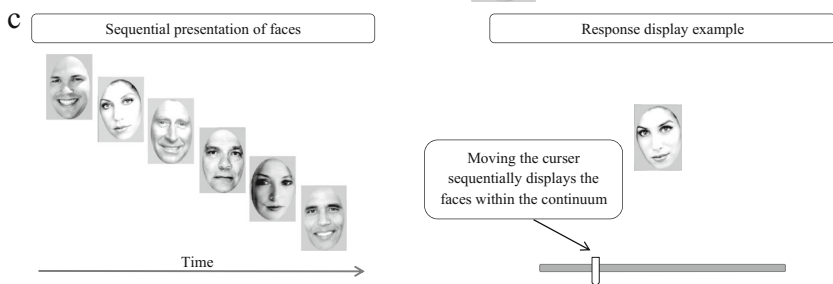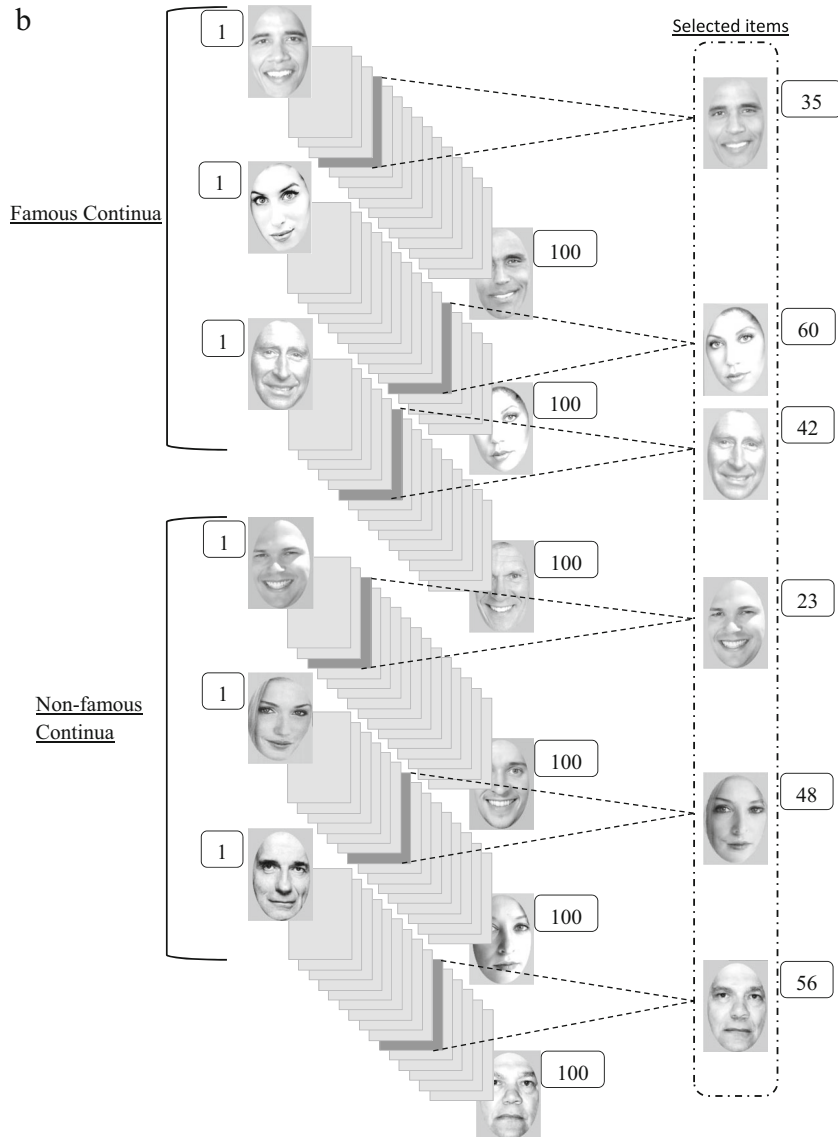
**Materials** Forty-eight grayscale images from Eimer, Gosling and Duchaine (2012) were used. As in Eimer et al., the faces were presented within an oval through which only the central features of each face were visible. These 48 images were organized into 24 pairs so that within-pair items had broadly similar characteristics; these included gender, approximate

age, facial expression, head orientation (or gaze direction), and other salient details (e.g., size of smile; see Fig. 1a). This matching allowed the morphing process to proceed more smoothly from one face to the other, i.e. each morph continuum was based on one of the matched face pairs. Of the 24 pairs, 12 contained a famous face while the other 12 did not. We therefore created 12 famous continua and 12 non-famous continua (using WinMorph 3.01). From each pair we obtained 100 image-steps; the image positions or numbers referred to below are related to those 100 steps. Figure 1 provides examples sampled from one famous and one non-famous continuum and illustrates the list construction process.

The procedure required six faces from different continua to be presented on each trial. To achieve this, the 24 face continua were randomly divided into four sets of six, each set containing three famous and three non-famous continua. This random division was performed 12 times, to create a total of 48 sets of six continua, each with the same property of having three famous and three non-famous continua. For each of these 48 sets of six, an individual face on each of the six continua was then selected for presentation by choosing an image at random from the range on the morphing scale of 20–79, subject to the constraint each half of the continuum had to be sampled from equally often. Figure 1b illustrates this process.

Only one of the six continua was tested on a given trial. Hence, from each of the 48 lists, a to-be-tested continuum was selected at random with two constraints: each of the 24 continua had to be tested twice across the experiment (each half of the continuum tested once) and each of the six study positions had to be tested equally often. One of the faces from the to-be-tested continuum had to be presented at the point of responding. The starting position of that test item was selected at random from position 10–89, with one constraint: the test image had to be a least ten steps away from the studied item. The testing range (10–89) was 20 images wider than the study range (20–79) as this allowed the test face to be at least ten steps either side of the studied face, even for the extreme morphs. The full range (from 1 to 100) of faces was not used as the first and last few images within each continuum did not have the slight blurriness that the other faces included due to the morphing process. Finally, at test, the relevant continuum was flipped on half the trials so that each end was to the left or right as often. Each face was presented at the center of a 15-in. monitor within a gray rectangle that was 6.5 cm high by 4.5 cm wide. Responses were provided using a mouse-controlled slider that made the displayed face change so that it travelled through the face continua under consideration. Figure 1c illustrates the study and test phase of a trial.

**Procedure** Participants were individually tested in a sound-attenuated room during a 30-min session. The experimenter first explained the task and answered questions; participants

a



b

Famous Continua

Non-famous
Continua

Selected items

1

1

1

100

100

100

35

60

42

23

48

56

1

1

1

100

100

100

c

| Sequential presentation of faces | Response display example |

Time

Moving the curser
sequentially displays the
faces within the continuum

**Fig. 1 a** The top row provides a sample of faces taken from within a famous face continuum (Obama); the bottom row presents a sample taken from a control continuum. **b** Illustration of the item selection process for each list; three famous and three non-famous continua were randomly selected; one image is then drawn from each to create a six item list. **c** Illustration of the study and test phases of a trial in Experiment 1

then provided consent. A reminder of the instructions was presented on screen followed by two practice and 48 experimental trials. On each trial, the six faces appeared sequentially for 1,500 msec each, with a blank of 500 msec after each image. After the sixth item, there was a blank screen, presented for 2 s, and then the test stimulus appeared along with a mouse-controlled horizontal slider bar used for responding. Participants could then move up and down the face continuum by using the mouse-controlled slider; they were instructed to identify the studied face and then click on a "Next" button to start the following trial. Upon completion of the experiment, participants were thanked and debriefed.

**Results and discussion**

To facilitate scoring and interpretation, the famous-face continua were re-organized to have the famous face always at the same end (zero/left) of the continuum and scores were corrected to reflect this. The relationship between studied and remembered positions on the continua was then examined. Figure 2 illustrates the findings; it presents the average remembered positions as a function of the studied positions.[1]

Two elements are noteworthy. First, a comparison of the slopes of the regression lines with the diagonal line representing perfect recall suggests that studied faces were remembered as being closer to the midpoint than they actually were. In essence, the slopes suggest a central-tendency bias. Assuming participants build a central representation for each continuum as the trials progress and that this knowledge is accessed to support reconstruction, then this tendency to regress towards the "best" representative of each continuum would be expected. As both functions appear to have very similar slopes, this bias seems equivalent for familiar and unfamiliar faces.

The second point of interest is the lower intercept obtained for famous face continua. When the target was a famous morph, there was an overall tendency to reconstruct more towards the zero end of the continuum, that is, towards the famous end of the continuum. Simply put, when studied at the same position as a non-famous face, a famous face will be reconstructed closer to the famous end of the continuum. This difference in intercept can be seen as a prior knowledge



**Fig. 2** Remembered face position as a function of study position on the morph continua. Each dot represents the average remembered position for a given studied position; the best fitting regression lines for famous (solid) and non-famous (dashed) stimuli are also provided. The diagonal dotted line going from the bottom left corner to the top right corner represents perfect performance

bias as its source is most probably the extra familiarity associated with the famous face that observers bring to the experiment.

The central-tendency and the influence of the famous faces were examined by running a series of per participant regression analyses where studied position was the predictor and remembered position was the dependent measure. We first determined if the central-tendency bias (slopes in Fig. 2) observed for the famous and non-famous continua were comparable. In order to do so, we ran separate regression analyses for the famous and non-famous conditions for each participant. The average slopes obtained for the famous (.35) and non-famous (.30) items were both significantly different from zero (famous faces: $t(29) = 7.1$, $p < .001$; non-famous faces: $t(29) = 6.4$, $p < .001$) but did not differ from each other ($t = -.90$, p $= .374$).

We then turned to the effect of the prior knowledge associated with the famous continua (intercept difference in Fig. 2). For each participant, we fitted a model with a single slope parameter and two intercepts (one for famous and one for non-famous stimuli) so there could be a test of the apparent difference in intercepts within the model. The famous or non-famous status was entered as a binary predictor in the regression model. Across participants, the mean slope was .33, and the average intercept values were 24.1 and 30.7 for the famous and non-famous data respectively. Hence, the average intercepts were ordered as predicted. T-tests confirmed that the average slope was different from zero, $t(29) = 8.1$, $p < .001$, and that the difference in intercepts was significant, $t(29) = 6.2$, $p < .001$.

---

[1] Note that because the study positions were randomly selected for each participant the number of observations per study position is not perfectly even; this is not an issue in the inferential statistics as the regressions were run independently for each participant.
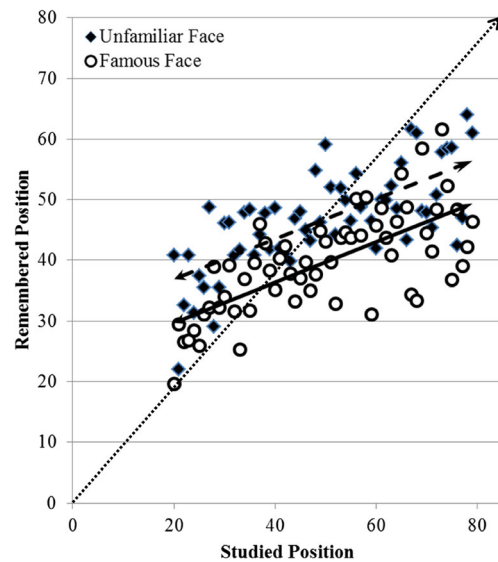
The aim of this study was to assess the influence of prior knowledge on VWM for photographs of faces. It was predicted that participants would be biased by knowledge in two ways. First, the familiarity with the stimuli developed during the experiment was expected to lead to a bias whereby remembered faces were drawn towards the center of the relevant continuum. Second, for the continua that involved a famous face, it was expected that prior knowledge would lead to a bias that would cause participants to remember the studied instance as being more like the famous face than it actually was. Both these predictions were born out.

It could be argued that faces are a unique type of stimulus (Wang, Fang, Tian, & Liu, 2012) and that these findings may not extend to other categories of objects. Experiment 2 called upon a different class of stimuli and also required reconstruction along another dimension: size.

## Experiment 2

Experiment 2 was based on prior work by Hemmer and Steyvers (2009a) who examined the impact of prior knowledge on episodic memory. In their work, Hemmer and Steyvers compared memory for the size of familiar items (fruit, vegetables) with memory for the size of unfamiliar items (random shapes). The task used a form of continuous recognition where participants were presented with 72 item lists. Study and test trials were randomly interleaved so that studied items were tested at random intervals within the list; on a test trial, participants were first asked if they recognized the item as having been studied before and were then asked to resize recognized items to their original studied size. The lag between study and test could vary between one and 24 trials; it follows that most lags would be outside what is typical in the study of immediate/working memory. Moreover, performance at all lags was averaged in the analyses. The results suggested that episodic memory of the studied items was affected by (a) fine-grained, item-specific representations and (b) two levels of categorical information. For both familiar and unfamiliar shapes, there was a central-tendency bias as the recalled size was systematically influenced by the mean size of the stimuli in the category. The results with familiar stimuli demonstrated the influence of a second categorical factor: item-level prior knowledge (e.g., the average size of apples).

In Experiment 2, we asked if the findings of Hemmer and Steyvers (2009a) could also be found in a VWM task. We used lists containing familiar items (photographs of vegetables) or unfamiliar ones (random shapes). As before, six items were sequentially presented, but in this case, at test, participants were to reconstruct the size of one of the studied objects.

From Hemmer and Steyvers (2009a) normative data were available for the familiar items; these included the normative average size (norm hereafter) for each item as well as the

largest and smallest realistic sizes. We assumed these norms were reasonable approximations of the knowledge participants brought to the experiment regarding familiar item sizes. With the help of these data, items could be presented either above or below the norm. This made it possible to predict the direction of any knowledge-based bias at the item level. Specifically, we expected that the remembered size of a familiar object (i.e., the just-seen apple) would drift towards the object's norm (i.e., the average apple size). Moreover, as before, we expected a central-tendency bias for both *familiar* and *unfamiliar* items whereby small items (a mushroom or a small shape) and large items (a cabbage or a large shape) would drift slightly towards the average size within the category. In essence, we tested predictions relating to two levels of knowledge: (1) for the familiar items, an object-level bias, where the size of each item is remembered as being slightly closer to its prototypical size and (2) for both types of items, a central-tendency bias where memory is influenced by the overall mean of item sizes presented within the experiment. Figure 3 summarizes the assumed influence of knowledge at both object and experiment levels.

### Method

**Participants** Forty-two undergraduate students volunteered for the study. Some received course credits for their participation.

**Materials** Stimuli were taken from Hemmer and Steyvers (2009a) and consisted of 24 high-resolution color photographs of vegetables against a white background as well as 24 images of random blue shapes.

These images were used to create 48 six-item lists, 24 lists of familiar items and 24 lists of unfamiliar ones. The familiar and unfamiliar items were yoked such that the presentation size of shapes was matched to that of the vegetables.

Study sizes of familiar items were determined as follows. In each list, two items were presented at their normative mean size, two items were larger than their normative mean size, and two items were smaller than said mean. All items were studied as often in all three sizes; however, tested items were always studied smaller or larger than their normative mean.

The sizes that were "larger" and "smaller" than the norm were obtained as follows. Recall that normative data contained three estimates: a normative mean size, a normative "smallest reasonable size" (e.g., the smallest realistic size for a radish), and a normative "largest reasonable size" (e.g., the largest realistic size for a radish). For each item, the range from the mean to the smallest reasonable size and the range from the mean to the largest reasonable size were calculated. Items presented smaller than their normative mean were presented at the size that was at 0.6 of the range from the mean to the smallest realistic size. So, if
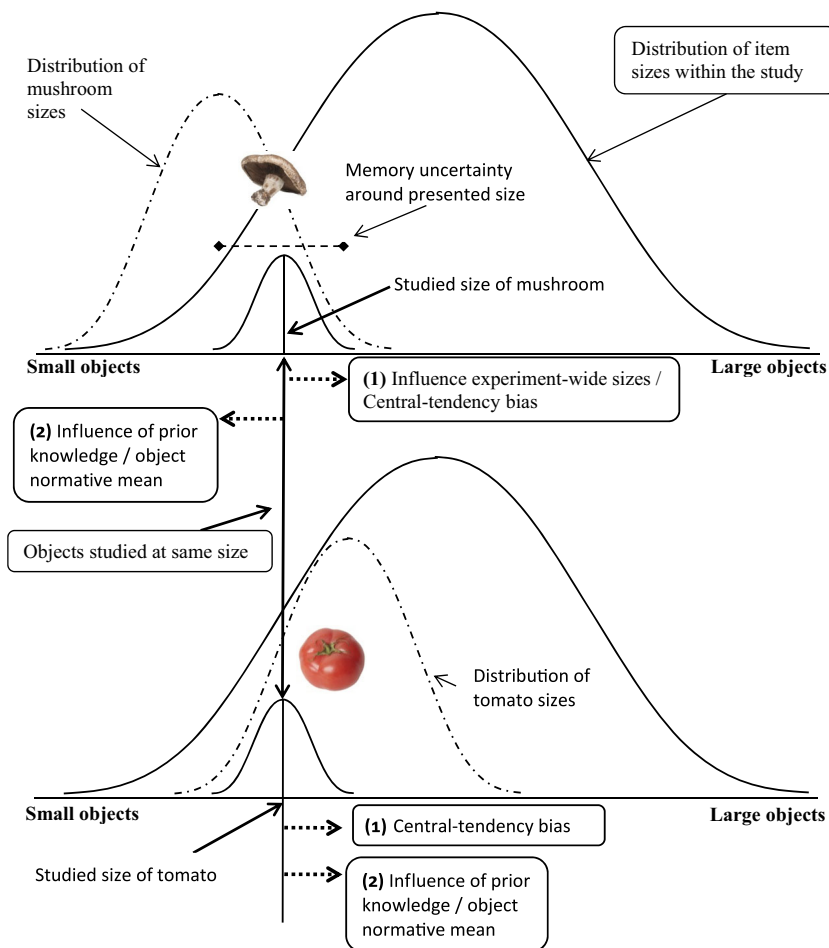
**Fig. 3** Schematic representation of the levels of information hypothetically contributing to the reconstruction of *familiar* object sizes. Top and bottom halves represent two different items studied at the same physical size. However, the item in the top half is studied at a size that is larger than its normative mean size while the item in the bottom half is studied at a size that is smaller than its normative mean size. The pull of the overall sizes (1) presented within the experiment will be similar in both cases, but the effect of prior knowledge about typical object sizes (2) will influence reconstruction in opposite directions. For *unfamiliar* items, the same predictions hold, except the distribution based on prior knowledge of object sizes is removed

the mean size for a beetroot was 0.25 and the smallest realistic size for a beetroot was 0.15, then the range was 0.10, and a small beetroot would be presented at 0.19, that is $[0.25 - (0.6 \times 0.10)]$. Likewise, the size of an item studied larger than its normative mean was set to be at 0.6 of the range between the mean and the largest realistic size for said item.

As for list composition, the 24 familiar items (and their yoked unfamiliar shapes) were divided into two groups based on their normative size; one group contained the 12 largest items while the other held the 12 smallest items. Forty-eight six-item lists were constructed so that: (a) three items were from the large group and three were from the small group. Also, items were divided into two sets of 12 items, matched for size. To improve experimental control, one set was used as targets for half the participants and the other set was used for the other half; this strategy allowed us to test the same item twice for each participant, once in a size above and once in a size below its normative mean; in essence, each item could be its own control and across participants, all items were used as targets. On each trial, a single item was selected from the list of six for testing. Each sequential position was tested equally often. Lists of unfamiliar items mirrored the construction of the familiar lists. There were also two practice trials created from the same stimuli that had the same structure.

**Procedure** The procedure was as in Experiment 1 except for the following. Images were presented for 1,000 msec each, followed by a 500-msec blank screen. Following the last item of a list, there was a further 1.5 s with a blank screen and then one of the items was presented again in a new size, randomly set to .2 (i.e., at a size corresponding to 20% of the display), .4, .6, or .8 of the presentation window. To reconstruct target sizes, participants moved a cursor placed in the center of a horizontal sliding bar (at bottom of the screen). Moving the

mouse-controlled cursor to the left made the target smaller and moving it to the right made it larger.

## Results and discussion

As in Hemmer and Steyvers (2009a), the presentation size was subtracted from the remembered size to obtain reconstruction error. A positive error indicates the item was remembered larger than studied; a negative error indicates the reverse. Figure 4 (a and b) presents the mean reconstruction error associated with each studied size separately for items studied larger than their norm, and items studied smaller than their norm. (For the unfamiliar shapes, the norm was taken to be the norm of the vegetable with which they were yoked). The left panel shows the data for familiar items (vegetables) and the right panel shows the data for the unfamiliar colored shape items. As expected, negative slopes were obtained in all conditions; small items were reconstructed larger and large items smaller, consistent with a central-tendency bias. For familiar items there were two distinct regression lines, corresponding to the items studied larger or smaller than their norm. In other words, two items studied at the same objective size can be remembered differently. If the studied size of one item was smaller than its norm, then it tended to be remembered as larger than it actually was. If the size of the corresponding item studied was larger than its normative size, it tended to be remembered as being slightly smaller than it was at study. For the unfamiliar items, the two regression lines were superimposed. As familiar and unfamiliar items were yoked in size, the difference must be due to the knowledge associated with the familiar items.

As before, we ran per participant regressions to analyze these findings. We first compared the two slopes obtained for the familiar items as well as those obtained for the unfamiliar items. We ran separate per participant regressions for the familiar items studied smaller than the norm and for those studied larger than the norm as well as the corresponding analyses for the unfamiliar conditions; the dependent variable was the error score and the predictor was the studied size. We then compared the slopes in a 2 (relative size, lager / smaller) × 2 (familiarity, vegetables / shapes) repeated measures ANOVA. There was no effect of relative size ($F(1,41)<1$, p=.61), a significant effect of category ($F(1,41)= 56.1$, p<0.001), and these factors did not interact ($F(1,41)<1$, p=.70). As Fig. 4 suggests, the slopes for each category (familiar / unfamiliar) are similar for each relative size; however, the mean slope for the unfamiliar items (−.58) is steeper than the mean slope for the familiar items (−.27). The slopes for both vegetables, $t(41) = −8.6$, $p < .001$, and shapes, $t(41) = −15.6$, $p < .001$) were significantly different from zero.

Following Hemmer and Steyvers (2009a), we then tested for the expected interaction between category (familiar/unfamiliar) and relative size (smaller/larger) for the intercepts. The hypothesis was that familiar items would show a knowledge-based bias through a difference in intercept as illustrated in Fig. 4a. As the unfamiliar items cannot benefit from equivalent knowledge, there should be no difference in intercept in this case, as suggested in Fig. 4b. Further per participant regressions were run for the familiar and unfamiliar items with the error score as the predicted variable. The predictors were the studied size along with a binary variable corresponding to whether an item was smaller or larger than its normative size.
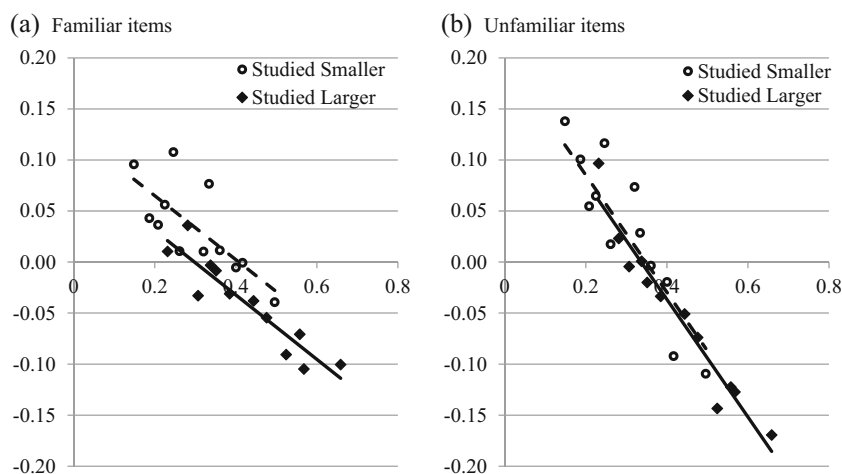


**Fig. 4** Mean reconstruction error as a function of the studied size for each item, item type [(**a**) familiar / (**b**) unfamiliar], and relative study size (i.e., larger than normative mean/smaller than normative mean). The latter is notional for the unfamiliar items, but as they were yoked to the familiar items, the distinction is maintained for comparison and analyses purposes

When averaged across participants, for the familiar items the two average intercept values were 0.12 and 0.16 respectively for the items studied larger and smaller than their normative means. For the unfamiliar items, the corresponding intercept values were .20 and .21. The means slopes were as above.

A 2 (familiar/shapes) × 2 (relative study size: smaller/larger than norm) ANOVA on the intercepts produced a significant effect of familiarity, $F(1,41)= 20.2$, $p < .001$, study size, $F(1,41)= 42.1$, $p < .001$ and more importantly the two factors interacted, $F(1,41)= 7.7$, $p= .008$. T-tests showed a significant difference between the intercepts observed for the familiar objects, $t(41)= -6.0$, $p < .001$, but not for shapes, $t(41)= -1.8$, $p=.08$. Thus, for the familiar items, objects studied at the same size, but respectively larger and smaller than the norm were not remembered in the same way. Items presented larger than the norm tended to be underestimated while items that were smaller than the norm were overestimated.

## General discussion

This paper tested the predictions of a general Bayesian view (see Fig. 3) which predicted that multiple levels of knowledge would impact VWM performance. In Experiment 1, morph continua were created between two faces; in half of the cases, a famous face was used as one of the faces of the morphed pair. After studying six faces, participants located one of the studied items along its relevant morph continuum based on their memory of the studied item. The results showed that responses were influenced by two levels of knowledge. On the one hand, a central-tendency bias meant that responses were pulled towards the center of the continua. On the other hand, when the studied item was from a continuum involving a famous face at one end, participants tended to remember the studied face as being more like the famous face than it actually was.

Experiment 2 called upon a different type of stimulus. Participants were presented with six photographs of realistically sized vegetables or six pictures of unfamiliar shapes. Their memory for the size of one of the items was tested by asking them to reconstruct the studied size of the item. Here also there was a clear impact of two levels of knowledge. Responses for both vegetables and unfamiliar shapes were influenced by a central-tendency bias as small items were remembered as being somewhat larger than they actually were and large items were remembered as being somewhat smaller than they were. In addition, when the size of the studied vegetables deviated from their respective normative mean size, the remembered size tended to drift towards the norm. The results extend the findings reported by Hemmer and Steyvers (2009a) to VWM.

## Central-tendency bias: recently developed knowledge?

When reviewing their findings, Hemmer and Steyvers (2009a) discussed the central-tendency bias that they observed as originating from knowledge of the average size of all the items within the categories called upon (in their case fruit and vegetables). One can ask if this is knowledge that is brought to the experiment or if participants develop a representation of the relevant central values over the course of the experiment. In the case of unfamiliar items such as the random shapes used here, one has to assume that the mean representation that leads to the central tendency bias develops over the course of the experiment as participants did not have prior knowledge of the specific stimuli called upon. The same would be true in other studies calling upon simple / abstract stimuli such as line length and greyness of squares (e.g., Huttenlocher et al., 2000). In the case of the familiar items however, both within-experiment knowledge and prior knowledge about the category could play a role. In Experiment 1, the similarity in the central-tendency bias for the familiar and unfamiliar items (slopes) makes it tempting to conclude that both originate from a common source – i.e., from statistics computed over the course of the experiment. However, two considerations suggest that this could be a hasty conclusion. First, in Experiment 1, the famous faces although quickly recognisable were most likely new instances for participants; also, generally speaking, people have a high level of expertise when it comes to processing faces – whether they are familiar or new. Taken together, this may have reduced any differences in the impact of prior knowledge about the faces used in the experiment. Second, in Experiment 2, the central-tendency bias obtained for familiar items (vegetables) was significantly smaller than the bias observed for unfamiliar items (random shapes); this difference suggests that the knowledge brought to the experiment can reduce the central-tendency bias. As we did not include any manipulations or specific conditions that could disentangle these potential sources, future research is needed to clarify the issue.

## Knowledge-based support and biases

In both the reported experiments there was a systematic impact of hierarchical levels of knowledge. To explain these findings, we suggest that performance involved (a) the development of a specific representation of the target's relevant

features, (b) knowledge from prior experience with similar items (when available), and (c) a representation of the ensemble central value – all of which inform the reconstruction of the most recently encountered instance. Our findings hence highlight that retrieval within a VWM paradigm can be simultaneously sensitive to superordinate category knowledge (e.g., mean size of items in the category) as well as prior, well-established knowledge – e.g., famous face or typical item size.

In reporting these knowledge effects we have insisted on the biases or errors that knowledge introduces in performance. However, as mentioned in the introduction, it is thought that knowledge-based biases originate from a process that is adaptive / helpful overall. These biases can be considered a relatively minor cost generated by a processing strategy that produces significant benefits over time (Brady & Alvarez, 2011; Hemmer & Steyvers, 2009a, 2009b; Huttenlocher et al., 2000). The suggestion is that our memory systems are designed to auto-correct; when the memory system produces an extreme value, there is an increased likelihood that this includes some error and so an adjustment towards more central values is adaptive. There is evidence that this auto-correction is also observed – and perhaps originates – in perception (e.g., Bae et al., 2014; Bae et al., 2015; Allred & Flombaum, 2014, 2016).

### Representations subtending performance

Allred and Flombaum (2016), Bae et al. (2015), and Brady and Alvarez (2011) have highlighted the importance of adequately characterizing the representations that underlie performance in VWM tasks. They noted that past research has largely focused on the nature of underlying limits that restrict the amount and quality of content that the system can store and that the nature of the content itself has had less attention. There is clearly some agreement about the importance of this issue as there is a growing VWM literature illustrating the complexities of the representations and processing involved (Bae, et al., 2014, Bae et al. 2015; Brady, Konkle, Alvarez, & Oliva, 2013; Oberauer & Eichenberger, 2013; Vergauwe, & Cowan, 2015).

Our predictions were based on a general Bayesian view where a number of representational levels interact in the reconstruction process; we now turn to a consideration of the mechanisms that might be involved – although admittedly this is speculative. In the type of task considered here, the most recently encountered target item, or at least its most relevant features – needs to be represented, bound, and probably linked to context, in order to have an identity and be sufficiently distinctive. If this was not the case, participants would simply recall average/prototypical sizes or not recognize the items. Moreover, there has to be a process through which the ensemble statistics about the items across trials are computed, i.e., some system has to support the computation of the dimensions such as average size (or whatever else is relevant). Finally, a

mechanism for the input from existing knowledge (i.e., the typical size of an item) needs also to be identified. Further, we know there must be associations developed between the current exemplar and the context of its presentation (Cowan, 2009). Perhaps the binding of items to general context can provide a means of grouping the relevant items and features so that summary statistics relating to the current task can be computed. There has to be some means to isolate the group of items that are relevant to the computation of said statistics; shared context could be a candidate for this mechanism.

The complexities relating to representation that are being brought to the fore by recent research on VWM could possibly benefit from borrowing from current theories of conceptual representation that suggest that instances of a concept are represented through reliance on distributed and flexible brain networks (see Barsalou, In press, a and b; see Hemmer & Persaud, 2014 for a similar idea in relation to episodic memory). One could for example, assume that processing each exemplar encountered involves a categorization process where the object is identified, activating a distributed representation of relevant features. A further assumption would be that instance-specific features such as current size, color, location, and general context are also activated and bound together perhaps through attentional processes (see Cowan, 1999; Cowan & Chen, 2009). At the point of retrieval, the constructed response would be mainly influenced by the focal encoded instance, but also by the knowledge embedded in these multiple levels of representation.

## Conclusion

The two studies in this paper tested the predictions of a general Bayesian view of how knowledge combines with the most recent representations to influence/bias VWM retrieval. Using very different types of stimuli (faces, objects) and two different dimensions (resemblance, size), the results showed the effect of various types and levels of knowledge: memory was affected not only by the features of the most recently studied instance, but by experiment-wide item statistics, category knowledge, as well as by the item-level knowledge that the participants brought to the experiment. The results extend the findings of Hemmer and Steyvers (2009a) to VWM and contribute to the growing literature which illustrates the complexity and flexibility of the representations subtending VWM performance.

# References

Bae, G. Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General, 144*(4), 744–63.

Bae, G. Y., Olkkonen, M., Allred, S. R., Wilson, C., & Flombaum, J. I. (2014). Stimulus-specific variability in color working memory with delayed estimation. *Journal of Vision, 14*(4), 7.

Barsalou, L.W. (In Press-a). On staying grounded and avoiding Quixotic dead ends. *Psychonomic Bulletin & Review.*

Barsalou, L.W. (In Press-b). Cognitively plausible theories of concept composition. In Y. Winter & J. A. Hampton (Eds.), *Compositionality and concepts in linguistics and psychology.* London: Springer Publishing

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 7.1–711.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia, 49*(6), 1622–1631.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science, 22*(3), 384–392.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: Independent forgetting of different object details from visual memory. *Journal of Experimental Psychology: General, 142*(3), 791–808. doi:10.1037/a0029649

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press.

Cowan, N., & Chen, Z. (2009). How chunks form in long-term memory and affect short-term memory limits. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 86–101). Hove: Psychology Press.

DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General, 119,* 375–396. doi:10.1037/0096-3445.119.4.375

Eimer, M., Gosling, A., & Duchaine, B. (2012). Electrophysiological markers of covert face recognition in developmental prosopagnosia. *Brain: A Journal of Neurology, 135,* 542–554. doi:10.1093/brain/awr347

Hemmer, P., & Persaud, K. (2014). Interactions between categorical knowledge and episodic memory across domains. *Frontiers in Psychology, 5,* 584. doi:10.3389/fpsyg.2014.00584

Hemmer, P., & Steyvers, M. (2009a). Integrating episodic memories and prior knowledge at multiple levels of abstraction. *Psychonomic Bulletin & Review, 16,* 80–87. doi:10.3758/PBR.16.1.80

Hemmer, P., & Steyvers, M. (2009b). A bayesian account of reconstructive memory. *Topics in Cognitive Science, 1,* 189–202. doi:10.1111/j.1756-8765.2008.01010.x

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General, 129*(2), 220–241. doi:10.1037/0096-3445.129.2.220

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition, 41*(8), 1212–1227. doi:10.3758/s13421-013-0333-6

Sailor, K. M., & Antoine, M. (2005). Is memory for stimulus magnitude Bayesian? *Memory & Cognition, 33,* 840–851.

Sims, C. R., Ma, Z., Allred, S. R., Lerch, R. A., & Flombaum, J. I. (2016). *Exploring the cost function in color perception and memory: An information-theoretic model of categorical effects in color matching. Proceedings of the 38th annual conference of the cognitive science society* (pp. 2273–2278). Austin: Cognitive Science Society.

van den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, 109*(22), 8780–8785.

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), 1404–1416. doi:10.1037/xlm0000108

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science, 23*(2), 169–177.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 452,* 233–235.