Expertise for upright faces improves the precision but not the capacity of visual working memory

**Elizabeth S. Lorenc · Michael S. Pratte ·
Christopher F. Angeloni · Frank Tong**

Abstract Considerable research has focused on how basic visual features are maintained in working memory, but little is currently known about the precision or capacity of visual working memory for complex objects. How precisely can an object be remembered, and to what extent might familiarity or perceptual expertise contribute to working memory performance? To address these questions, we developed a set of computer-generated face stimuli that varied continuously along the dimensions of age and gender, and we probed participants' memories using a method-of-adjustment reporting procedure. This paradigm allowed us to separately estimate the precision and capacity of working memory for individual faces, on the basis of the assumptions of a discrete capacity model, and to assess the impact of face inversion on memory performance. We found that observers could maintain up to four to five items on average, with equally good memory capacity for upright and upside-down faces. In contrast, memory precision was significantly impaired by face inversion at every set size tested. Our results demonstrate that the precision of visual working memory for a complex stimulus is not strictly fixed but, instead, can be modified by learning and experience. We find that perceptual expertise for upright faces leads to significant improvements in visual

E. S. Lorenc · M. S. Pratte · C. F. Angeloni · F. Tong
Vanderbilt University, Nashville, TN, USA

E. S. Lorenc · M. S. Pratte · C. F. Angeloni
Psychology Department, Vanderbilt University, Nashville, TN, USA

F. Tong
Psychology Department and Vanderbilt Vision Research Center,
Vanderbilt University, Nashville, TN, USA

E. S. Lorenc (✉)
UC Berkeley Neuroscience Department, D'Esposito Lab, 132 Barker
Hall #3190, Berkeley, CA 94720, USA
e-mail: elizabeth.lorenc@berkeley.edu

precision, without modifying the capacity of working memory.

Visual working memory is essential for our ability to maintain information about stimuli that are no longer in direct view. In daily life, we rely on this ability to maintain relevant information about objects in our environment and to facilitate the planning of actions toward such objects during everyday tasks (e.g., Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Hollingworth, 2004). Given that visual working memory is of such importance for normal functioning and yet has profound limitations, it has been the topic of considerable research and debate (e.g., Alvarez & Cavanagh, 2004; Bays & Husain, 2008; Luck & Vogel, 1997, 2013; van den Berg, Shin, Chou, George, & Ma, 2012; Zhang & Luck, 2008). Researchers have sought to define the mechanisms of successful encoding, maintenance, and retrieval, as well as the properties of working memory that account for its fundamental limits.

Whereas early studies relied on change detection paradigms to estimate the number of visual objects that can be maintained in working memory (Luck & Vogel, 1997 Phillips, 1974), more recent studies have adopted psychophysical procedures that provide separate estimates of memory precision and capacity (Zhang & Luck, 2008). Using the method of adjustment, one can estimate the precision of memory for basic visual features that vary along a continuum, such as color or orientation, by having participants manipulate a probe stimulus along a relevant feature dimension to indicate their memory of a previously viewed test stimulus. Although it is debated whether working memory is better described by a small number of discrete representations (Rouder et al., 2008;

Zhang & Luck, 2008) or by a continuous resource that can be flexibly subdivided to represent many items (Bays & Husain, 2008; van den Berg et al., 2012), research has demonstrated the utility of considering precision and memory capacity as separate psychological variables. For example, estimates of memory capacity, but not precision, are highly predictive of individual differences in fluid intelligence (Fukuda, Vogel, Mayr, & Awh, 2010) and further correlate with EEG measures of working memory maintenance (Anderson, Vogel, & Awh, 2011). For the purposes of the present study, we rely on the mixture-model analysis of Zhang and Luck and its assumption of a discrete capacity limit to estimate the precision and capacity of visual working memory.

Recent studies have successfully characterized the precision of working memory for basic visual features, but much less is known about the resolution and capacity of working memory for complex objects. Unlike basic features, which can be readily manipulated to vary along a continuous dimension, it is less obvious how one might devise a set of complex objects that smoothly vary along a continuum. Previous studies using change detection paradigms have suggested that working memory capacity may be more limited for complex visual objects (Alvarez & Cavanagh, 2004; Wheeler & Treisman, 2002; Xu, 2002). For example, Alvarez and Cavanagh found that participants could successfully maintain more than four colored squares in working memory, on average, but could maintain no more than two shaded cubes or random polygons. Similarly, Curby and Gauthier (2007) reported finding greater working memory capacity for upright than for inverted faces, which suggested that extensive perceptual training with upright faces led to an information-processing advantage and, thereby, reduced the load on working memory.

However, a potential concern in these change detection studies is that differences in memory precision might be confused for differences in capacity (Awh, Barton, & Vogel, 2007; Wilken & Ma, 2004). It is now known that memory precision decreases with increasing set size (Zhang & Luck, 2008); thus, working memory performance could be disproportionately impaired at large set sizes if participants were to have generally poorer memory precision for a particular object class. Awh et al. found that the estimated capacity of working memory across different object categories was strongly predicted by the visual similarity (or confusability) of the stimuli within each object class. Moreover, when participants were required to detect large between-category changes (e.g., from cube to polygon), rather than subtle within-category changes (e.g., polygon to polygon), estimates of memory capacity were comparable across the different object classes. In another study, Scolari, Vogel, and Awh (2008) found an upright face advantage for detecting within-category changes in facial identity (e.g., from one face to another), replicating the findings of Curby and Gauthier (2007). However, the researchers

found comparable performance for upright and inverted faces following between-category changes (e.g., from face to cube). The authors interpreted these findings to suggest that differences in memory precision may account for the differences in working memory performance previously reported for upright and inverted faces (Curby & Gauthier, 2007). Although these change detection studies suggest that memory precision is likely to be better for upright faces, the paradigm did not allow the researchers to directly quantify the precision of working memory for faces. Thus, it is difficult to specify how much more precisely participants can maintain upright, as compared with inverted, faces. Recent studies have demonstrated the importance of using continuous measures of memory for an item to obtain separate estimates of memory precision and capacity (Luck & Vogel, 2013), especially if the research goal is to determine whether an effect on working memory performance should be attributed to changes in memory precision, capacity, or both.

The purpose of our study was to directly evaluate the precision and capacity of visual working memory for the complex stimulus class of faces and to determine how perceptual expertise might enhance working memory performance. Faces constitute a very important and socially relevant class of visual stimuli for which people have acquired a lifetime of expertise. Evidence of expertise for faces comes from the well-documented *face inversion effect*: When a face is simply turned upside-down, both perceptual discrimination and recognition judgments are consistently impaired (McKone & Yovel, 2009; Tanaka & Sengco, 1997; Tong & Nakayama, 1999; Valentine, 1988; Yin, 1969). Stimulus inversion causes a disproportionate impairment in face recognition, when compared with the recognition of other types of objects (Yin, 1969), unless participants have acquired considerable expertise for that specific object class (Diamond & Carey, 1986; Gauthier & Tarr, 1997; but see also McKone, Kanwisher, & Duchaine, 2007).

In the present study, our motivation for using face stimuli to investigate working memory was twofold. First, faces vary along multiple dimensions, and people are adept at distinguishing subtle variations between individual faces in many of these dimensions (e.g., O'Toole, Vetter, Troje, & Bülthoff , 1997; Webster, Kaping, Mizokami, & Duhamel, 2004). By generating faces that varied along the dimensions of both gender and age, we created a smoothly changing circular face space (see Fig. 1). The circular nature of this space was critical, since it allowed participants to precisely report their memory of individual faces by method of adjustment and, therefore, yielded estimates of precision and capacity that were uncontaminated by response biases. Second, people have far greater perceptual expertise for upright than for inverted faces. Our goal was to test whether extensive experience with upright faces leads to improvements in the precision or capacity of working memory, on the basis of the

separate quantification of these parameters using the mixture model approach.

In Experiment 1, participants were presented with upright and inverted faces at set sizes of one, three, and five, and we obtained estimates of working memory precision and capacity for each participant and experimental condition. In Experiment 2, we tested a larger set size of six faces and implemented an articulatory suppression procedure to discourage the use of verbal encoding strategies. Across both experiments, we find consistently better memory precision for upright than for inverted faces, but no effect of face orientation on memory capacity.

Experiment 1

Method

Participants

A total of 8 participants (19–31 years of age, 6 female) participated in Experiment 1, which consisted of six 1-h experimental sessions. Participants included students or employees at Vanderbilt University, as well as members of the surrounding Nashville community. All participants had normal or corrected-to-normal vision. Participants were paid for their time, and they could earn additional pay for performing well on the task. Two of the 8 participants were removed from the analysis for failing to follow the task instructions (submitting mostly random responses even at set size 1). This study was approved by the Vanderbilt University Institutional Review Board, and all participants gave written informed consent.

Stimuli

A set of 80 grayscale 3-D face images were generated, using the FaceGen Modeller software package (Singular Inversions, Inc.), to form an approximately circular face space. First, we created eight 3-D-rendered faces that varied along the dimensions of age and gender, to form an octagonal space (Fig. 1). Next, each pair of neighboring faces was combined in varying proportions (10/90, 20/80, . . . 90/10) using a linear morphing procedure to create a total of 80 unique faces. For our analyses of memory performance, we consider these 80 face stimuli to span a 360˚ circular space and assume even spacing between neighboring faces, such that the difference between any two neighboring faces was equivalent to 4.5˚.

All face stimuli were normalized to equate for mean luminance and were displayed at a size of 5.2˚ × 6.5˚ of visual angle. In addition to the face stimuli, a set of 80 noise stimuli were constructed that shared the outline shape of a
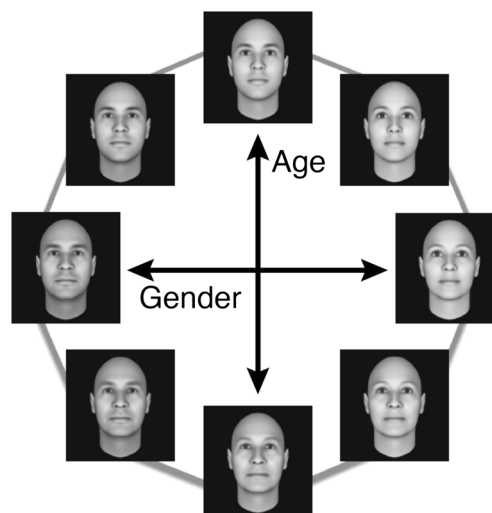

Fig. 1 Examples of face stimuli making up the continuous face space. A 3-D face-rendering software program was used to generate faces that varied along the separate dimensions of age and gender. After the above eight faces were created, intermediate faces were constructed by applying a linear morphing procedure between neighboring pairs of faces, to construct a total of 80 face stimuli

corresponding face stimulus but consisted of random noise (1/F noise) presented within the bounding mask.

Stimuli were presented on a 20-in. LCD monitor using MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Participants viewed the display at a distance of 44 cm, with head position stabilized by a forehead and chin rest.

Procedure

Each participant completed six 1-h sessions, consisting of 126 trials each, for a total of 756 trials per person. Working memory performance was measured for upright and inverted faces, shown in separate 21-trial blocks. Upright and inverted stimulus blocks were interleaved, and the inversion condition of the first block of the first session was counterbalanced across participants. On a given trial, participants were presented with one, three, or five faces chosen randomly from the stimulus space (see Fig. 2). The stimuli were presented one at a time at different spatial locations at an eccentricity of 6.5˚ relative to a central fixation dot. The polar angles between stimulus locations were chosen pseudorandomly on each trial, such that no two stimulus locations overlapped. Pilot testing indicated that the faces were too visually similar to allow for accurate discrimination in the periphery, so participants were instructed to fixate on the central dot at the beginning of each trial and then to look directly at each face as it was presented for 1,500 ms in its respective spatial location. This relatively long stimulus presentation duration was used to ensure that participants had enough time to saccade to each item and fully encode it into working memory, mitigating the possibility that

any observed face inversion effects would be due to differences in the speed of visual encoding.

On trials with fewer than five faces, noise stimuli were substituted to equate the total number of stimuli shown. The presentation of all five face/noise stimuli was followed by a 3,000-ms retention interval and then a 1,000-ms spatial cue (a circle with thicker outline) indicating which of the previously viewed faces would be tested. In the subsequent test period, one of the 80 possible faces was randomly selected and presented at the center of the screen. The participant used two keys on the keyboard to cycle through the face space, matching the appearance of the central face as precisely as possible to the probed face in memory, and made a keypress to submit his/her response. Trials were separated by a 1,000-ms intertrial interval. On inverted blocks, the face stimuli, noise stimuli, and test faces were all inverted.

At the beginning of the experiment, participants completed 12 practice trials (2 trials for each face orientation × set size condition). They were encouraged to remember as many stimuli as possible on each trial, as precisely as possible, and were told that they could earn bonus points by reporting the probed item with high accuracy. Participants received 2 bonus points for responding within 10 faces of the correct face (0°–45° difference) and 1 bonus point for responding between 11 and 20 faces (45°–90° difference) away from the correct face. Participants received an extra dollar for every 50 bonus points earned and could earn up to $5 in bonus pay per 1-h session in addition to the standard payment rate of $10/h. During the initial practice session, participants received feedback on every trial regarding the accuracy of their performance. During the main experiment, however, cumulative feedback was provided only at the end of each block of trials, during self-paced breaks. At the end of each experimental session, participants were informed of the cumulative amount of points earned and the corresponding dollar amount of bonus payment.

Data analysis

For each participant, we calculated response error as the angular difference between the reported face and the studied face. We applied Zhang and Luck's (2008) mixture model to the distributions of these errors separately for each participant, set size, and upright/inverted condition by maximizing the likelihood function. This model assumes that the responses are a mixture of pure guesses, which are distributed uniformly across the face space, and responses based on noisy memory representations, which form a von Mises distribution centered on the studied face. This model has two free parameters: (1) The standard deviation of the von Mises reflects the precision (SD) of faces stored in working memory, and (2) the proportion of responses arising from the von Mises distribution, rather than the uniform guessing distribution, reflects the probability that a face was stored in memory. This proportion is then transformed into an

estimate of working memory capacity (K) by multiplying the probability of successful memory by the set size in each condition. Within-subjects analyses of variance (ANOVAs) and planned contrasts were performed to test for reliable differences in memory precision and capacity across the manipulations of face orientation and set size.

Results

We estimated the precision and capacity of visual working memory for each participant and condition, on the basis of the observed differences between the reported faces and the studied faces. Figure 3 shows the distributions of response errors of a representative participant and the corresponding fits of the mixture model. The distribution of response errors was binned for the purposes of illustration, but each data point was considered individually in fitting the model using maximum likelihood estimation. The entire face space is assumed to span 360°, and the error sizes and standard deviations are reported in degrees.

In general, our participants performed the task very well, as indicated by the group-averaged results showing the estimates of memory capacity (Fig. 4a) and precision (Fig. 4b). Participants were able to retain up to approximately 4.5 items on average, regardless of face orientation. A within-subjects ANOVA confirmed that the number of successfully maintained items increased as a function of set size, $F(2, 10) = 515.9$, $p \leq .0001$. More important, this analysis revealed no effect of face orientation on the number of items that were successfully maintained, $F(1, 5) = 0.56$, $p = .49$, and no evidence of an interaction between face orientation and set size, $F(2, 10) = 0.64$, $p = .55$. These results indicate that the capacity of visual working memory was not reliably impaired by face inversion.

In contrast, we found that face inversion led to consistently poorer memory precision, as indicated by larger standard deviations for inverted faces than for upright faces across all set sizes. An ANOVA revealed that memory precision was significantly worse for inverted than for upright faces, $F(1, 5) = 27.1$, $p = .003$, with no evidence of an interaction effect between face orientation and set size, $F(2, 10) = 0.44$, $p = .66$. Planned contrasts revealed a significant difference in precision at each of the three set sizes (paired t-tests, all ps < .02). Thus, we found a consistent advantage for upright faces in memory precision, even when only a single face was maintained and the demands on working memory capacity were very modest.

Further analyses indicated that memory precision declined significantly with increases in set size, $F(2, 10) = 4.918$, $p = .03$. However, precision did not decline as rapidly as one might expect on the basis of Zhang and Luck's (2008) slots-plus-averaging model. This model assumes that when participants are given only a single item to remember, multiple slots
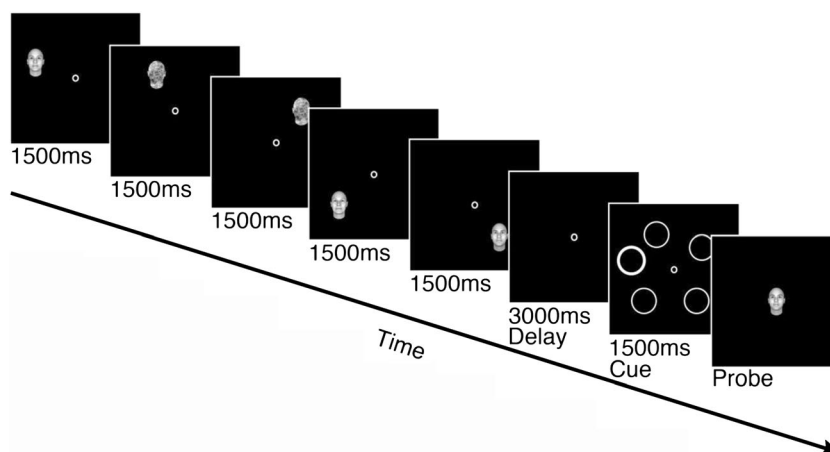
Fig. 2 Example of a working memory trial in Experiment 1. A total of five face or noise stimuli were presented sequentially at different locations, and following a 3s retention interval, a spatial cue indicated which face to report from memory

can be used to maintain independent representations of that single item, leading to improvements in memory precision as would be predicted by signal averaging. For example, if a working memory system with four slots were used to encode a single item, precision for that single item should improve by a factor of two, or the square root of the number of slots used for simultaneous encoding. In our study, average memory capacity exceeded a value of four, yet the improvement in precision at set size 1 was very modest when compared with memory precision at set size 5. Due to our use of sequential presentation, participants could not anticipate exactly how many faces

would appear on a given trial; as a consequence, they might not have been able to optimize their encoding strategy on set size 1 trials. Regardless, the results of Experiment 1 indicate a consistent advantage in memory precision for upright faces.

Experiment 2

We performed a second experiment in order to confirm the results from Experiment 1 and rule out potential alternative accounts of the finding that face inversion impairs precision
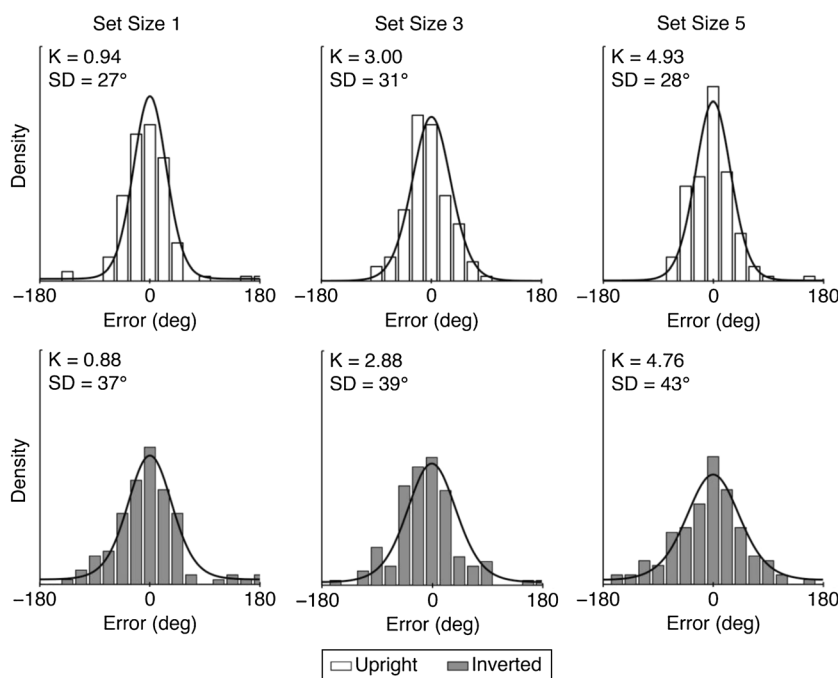

Fig. 3 Distribution of errors for faces reported from memory for a representative participant. Individual plots show the distribution of differences between the actual face and reported face for each set size and face orientation condition. The mixture model was fit to the data (lines denote predicted distributions), providing estimates of memory capacity (K) and precision (SD). The estimates of these parameters for this participant are shown for each condition
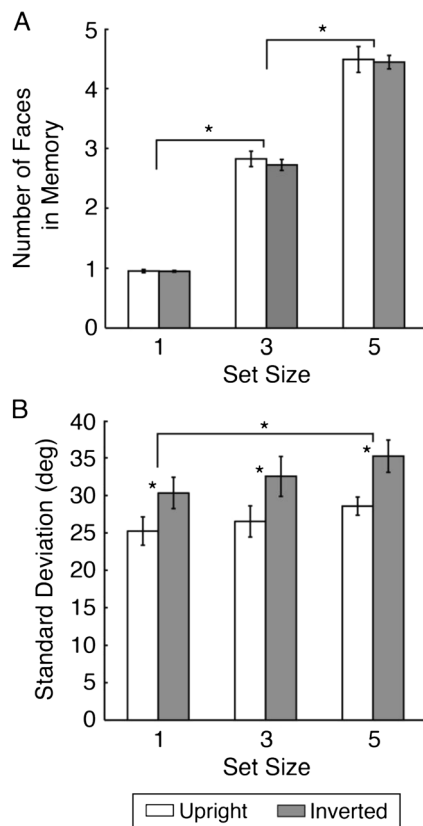
1980

Fig. 4 Results of Experiment 1: Estimates of memory capacity (**a**) and precision (**b**) for upright and inverted faces at set sizes of one, three, and five. Planned contrasts indicated significantly better precision for upright than for inverted faces at each set size tested and significant main effects of set size on both capacity and precision (significant differences indicated with asterisks). However, there was no interaction between face orientation and set size on either capacity or precision. Error bars represent ±1 standard error of the mean

but not capacity. Because the average capacity estimates in Experiment 1 were quite high, approximately 4.5 items out of 5, it is possible that a ceiling effect might have obscured our ability to detect an effect of face inversion on memory capacity in this experiment. In Experiment 2, we used a set size of six on every trial to ensure that participants' working memory capacities were reliably exceeded when testing for differential effects of face orientation on capacity. In addition, we were concerned that some participants might employ a verbal encoding strategy to remember the face stimuli. To discourage the use of verbal encoding strategies, the second experiment included a concurrent articulatory suppression task (Murray, 1967), which is commonly used to prevent verbal rehearsal during memory tasks (Baddeley & Hitch, 1975).

Method

Participants

A total of 14 participants (18–32 years of age, 7 female) participated in Experiment 2, and each had normal or corrected-to-normal vision. All participants were members of the Vanderbilt/Nashville community and provided written informed consent. Participants were paid for their time, with additional pay for accurate performance on the working memory task, as described in Experiment 1. None of the participants in Experiment 2 had previously participated in Experiment 1.

Stimuli and procedure

The stimuli and design of Experiment 2 were identical to those of Experiment 1, with the exception that six face stimuli were presented on every trial and participants performed a concurrent articulatory suppression task. Each trial began with the auditory presentation of two randomly chosen digits, followed by a fixation period (total of 3 s). Participants were instructed to overtly rehearse these digits out loud throughout the subsequent memory task. An experimenter monitored compliance with this task by either remaining in the room or listening remotely via a microphone. After participants made their responses on the memory task, they were prompted to enter the digits they had spoken throughout the trial. The accuracy of this report provided a quantitative measure of how well participants performed the articulatory suppression task.

Participants completed four 1-h sessions, each beginning with 6 practice trials that provided posttrial feedback regarding bonus points earned on the working memory task and the accuracy of performance on the digit report task. Each session included eight blocks of 8 experimental trials each, and upright and inverted faces were presented in alternating blocks, as in Experiment 1. Feedback about the number of bonus points earned was provided at the end of each block. Each participant completed a total of 256 trials, providing 128 trials per face orientation condition.

Results

Performance on the concurrent articulatory suppression task was highly accurate, with all participants entering at least one of the two digits correctly on every trial. Out of the total 768 digits to be rehearsed across all four sessions, participants had a mean accuracy above 99%, suggesting that the secondary task was performed diligently throughout the session and across conditions.

The results of the working memory task were very consistent with the findings in Experiment 1. As in the first experiment, we found that the mixture model fit the data well (see Fig. 5a for model fits to a representative participant's data). We also found that estimates of memory capacity were almost identical for upright and inverted faces (Fig. 5b), with no evidence of a reliable statistical difference [paired samples t-test, $t(13) = 0.37$, $p = .72$]. In contrast, the precision of working memory (Fig. 5c) was significantly better for upright faces
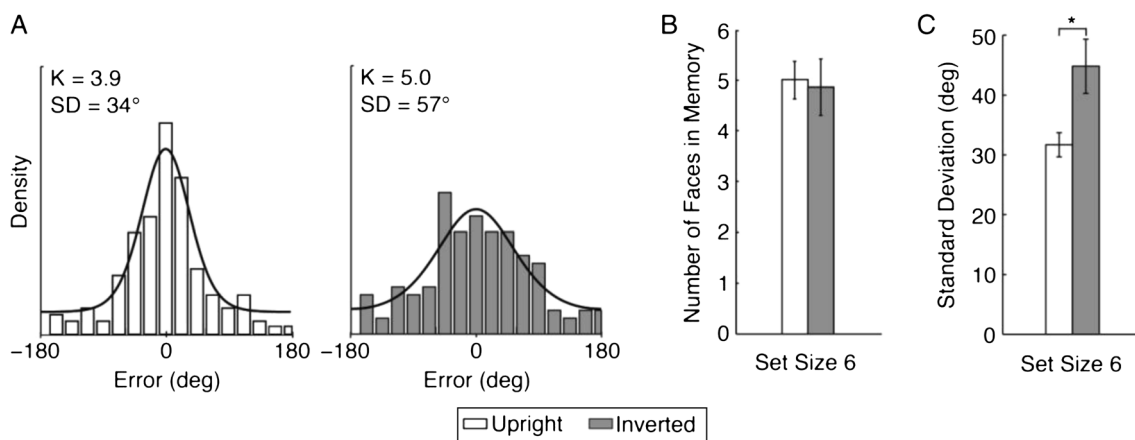
Fig. 5 Results of Experiment 2. **a** Distribution of response errors for a representative participant. The left and right panels show errors for upright and inverted faces, respectively, and the lines and parameter estimates denote mixture model predictions for this participant. **b** Estimates of capacity for upright and inverted faces across all participants. **c** Estimates of precision, which differed significantly between upright and inverted faces. Error bars represent ±1 standard error of the mean

(mean $SD = 38.8°$) than for inverted faces (mean $SD = 49.1°$), as indicated by a t-test, $t(13) = -5.11$, $p < .001$.

Capacity estimates for participants in Experiment 2 were reliably below the full set size of six, mitigating the possibility that a ceiling effect might have impeded our ability to observe an effect of face inversion on capacity estimates. Nonetheless, we conducted an additional analysis to address whether the lack of a difference in memory capacity might have resulted from a subset of participants who exhibited high capacity estimates. We repeated the original analysis, excluding any participants with mean capacity estimates greater than five items. With 6 participants remaining in the analysis, we still observed no effect of face inversion on memory capacity, $t(5) = 0.4331$, $p = .683$.

We conducted an additional analysis to address whether the null effect of memory capacity was likely to reflect a type II error—that is, a failure to reject the null hypothesis when it is false. Significance tests like the t-test are not designed to control for type II error rates, so they are not well suited for firmly accepting the null hypothesis. In contrast, Bayesian hypothesis tests, known as Bayes factors, provide an assessment of the evidence both for and against the null hypothesis. Specifically, the Bayes factor equivalent to the t-test (Rouder, Speckman, Sun, Morey, & Iverson, 2009) provides the likelihood that the null hypothesis is true, as well as the likelihood of the alternative hypothesis that there exists a true difference between conditions. We applied this analysis to our data and obtained a Bayes factor indicating that the hypothesized face inversion effect on precision was 77 times more likely than the null hypothesis of there being no such effect, in agreement with the t-test reported above. However, our analysis of capacity estimates indicated that the *null* hypothesis of no inversion effect was 4.7 times more likely to be true than the alternative hypothesis of an effect of face inversion on capacity. Bayes factors greater than 3.0 are considered to be "substantial" evidence for a hypothesis (Jeffreys, 1998).

Consequently, these results suggest that face inversion significantly affected precision and was unlikely to have affected capacity. In addition, for the subset of participants who had overall estimated capacities less than 5.0, the null hypothesis of no inversion effect on capacity was 3.2 times more likely than the alternative hypothesis. Taken together, these results suggest that our failure to observe an effect of face inversion on memory capacity is attributable neither to a ceiling effect nor to a lack of statistical power.

Each participant completed four experimental sessions on separate days, and we investigated whether memory performance changed with practice over the course of these sessions. There was not sufficient data to fit the mixture model to an individual's data for a single session (32 trials in each condition per session), so we simply considered the root mean squared error (RMSE) between the studied face and the reported face as an overall measure of memory performance. The results are shown in Fig. 6a. An ANOVA performed on these error scores revealed a consistent main effect of both face inversion across sessions, $F(1, 13) = 33.6$, $p < .001$ and a main effect of experimental session due to gradual improvement with prolonged training, $F(3, 39) = 4.34$, $p = .01$. However, we did not find evidence of an interaction between experimental session and face orientation, $F(3, 39) = .37$, $p = .78$, suggesting that extended practice with these stimuli led to similar improvements in working memory performance for upright and inverted faces.

Recent studies of working memory for visual orientation have found evidence of a recency effect, with superior precision found for items that appear toward the end of the sequence of items (Gorgoraptis, Catalao, Bays, & Husain, 2011). Because our stimuli were presented individually, we could also explore whether memory performance for faces varied depending on the temporal position of the face that was probed for recall. Again, there was insufficient data to fit the mixture model to the data from individual probed locations, so
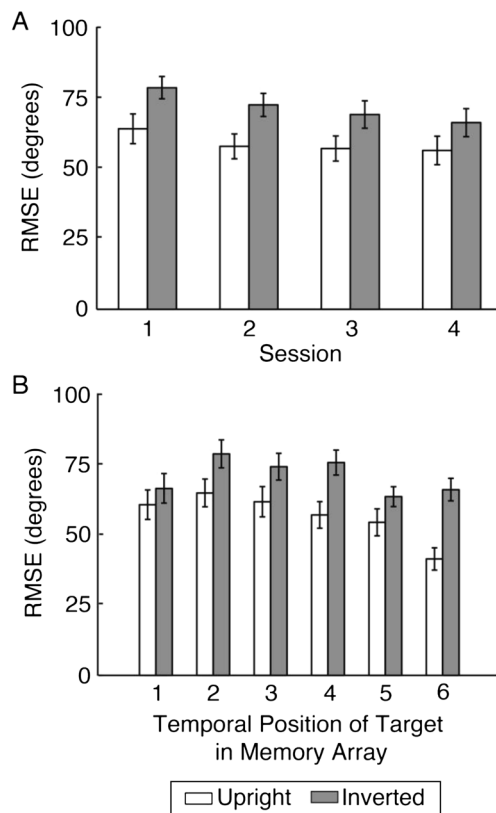
Fig. 6 **a** Memory performance (RMSE) as a function of experimental session. Performance was worse for inverted than for upright faces across all experimental sessions and improved slightly across sessions for both upright and inverted faces. **b** Overall memory performance (RMSE) for upright and inverted faces in Experiment 2, for each of the six study locations in the memory array. Performance was worse for inverted faces across all temporal positions. In addition, performance for both upright and inverted faces was slightly higher for items presented at the beginning and end of the array. Error bars represent ±1 standard error of the mean

we considered RMSE as an overall measure of memory performance (see Fig. 6b). A repeated measures ANOVA revealed a significant main effect of the temporal position of the target stimulus on overall performance, $F(5, 65) = 8.60$, $p < .001$, which appeared to reflect the presence of both primacy and recency effects. There was also a main effect of face orientation, $F(1, 13) = 29.24$, $p < .001$, but a nonsignificant interaction between temporal position and face orientation, $F(5, 65) = 2.02$, $p = .09$. Although this interaction effect was marginal, for both upright and inverted faces there appears to be a trend of better memory performance for the most recent items. It should be noted that the temporal position of the probed face was randomized from trial to trial, such that the mixture model results reported above are not sensitive to differences in performance across temporal position. However, these analyses do demonstrate the robustness of the face inversion effect on memory performance: Even with practice over multiple days or intervening stimuli in a sequential memory array, participants' memories are consistently impaired for inverted faces.

Discussion

In the present study, we developed a set of continuously varying faces to measure the precision and capacity of visual working memory for complex stimuli. We found that simply turning a set of faces upside-down leads to a significant loss of memory precision, even though face inversion fails to affect the number of items that can be actively maintained in memory. Our results demonstrate that the precision of working memory for complex stimuli is not strictly fixed but, rather, can be modified by learning and visual experience.

Our findings are relevant to current discussions regarding the properties and limits of visual working memory. One prominent view is that complex objects place greater information-processing demands on working memory and, thereby, lead to a reduced memory capacity (Alvarez & Cavanagh, 2004; Brady, Konkle, & Alvarez, 2011). Given that inverted faces require more intensive part-based processing and upright faces allow for more efficient holistic processing (McKone & Yovel, 2009; Tanaka & Sengco, 1997), one might expect that face inversion should lead to diminished memory capacity. Instead, we found that stimulus inversion led to a selective impairment in the precision of visual working memory, without impairing capacity. These findings support and clarify the results of previous studies involving change detection with faces (Curby & Gauthier, 2007; Scolari et al., 2008). Furthermore, our study raises the possibility that the presumed impact of object complexity on working memory capacity may, instead, be attributable to differences in visual precision (see also Awh et al., 2007). Previous studies relied on change detection paradigms to estimate memory capacity for complex objects and were unable to separately measure memory precision and capacity. It would be interesting for future studies to develop continuously varying stimuli in other complex object domains, to evaluate the generality of the present findings.

This study is also relevant to the prominent slot-plus-averaging model proposed by Zhang and Luck (2008). On the basis of their studies of memory for basic features, they proposed that visual working memory consists of a discrete number of slots, each of which can retain information about a single item with a fixed degree of visual precision. The fact that we find comparable capacity estimates for upright and inverted faces is consistent with the predictions of this discrete capacity model. That said, it should be acknowledged that our capacity estimates for face stimuli ranged from 4.5 to 4.9 items across the two experiments, a value that is somewhat higher than those reported in studies of working memory for basic visual features (Anderson et al., 2011; Rademaker, Tredway, & Tong, 2012; Zhang & Luck, 2008).

Whereas memory capacity is the same for upright and inverted faces, we find that the precision of visual working memory for a complex stimulus is not strictly fixed and, instead,

depends on one's level of familiarity and expertise with that item. Recent studies have shown that experimental manipulations of set size, attentional allocation, encoding duration, and temporal position can all influence working memory precision (Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011; Gorgoraptis et al., 2011 Zhang & Luck, 2008). In addition, there appears to be some natural stochastic variability in the precision of visual working memory from trial to trial (Fougnie, Suchow & Alvarez, 2012; Rademaker et al., 2012). However, much less is known about the extent to which perceptual training might modify the precision of visual representations that are called upon during working memory tasks. Studies have demonstrated improvements in working memory performance for an object class following visual training (e.g., Curby, Glazek, & Gauthier, 2009; Moore, Cohen, & Ranganath, 2006), but these change detection studies could not distinguish between effects of memory capacity and precision. In the present study, we observed a consistent advantage in visual precision, even at set size 1 when the capacity of working memory was far from being fully taxed. Such a result cannot be readily explained in terms of a modified capacity limit for the trained object category but, rather, appears to reflect an improvement in the ability to encode and maintain a precise representation of the trained stimulus.

These findings lead us to propose that acquired expertise for a stimulus class, such as upright faces, leads to the development of more precise perceptual representations that can be called upon to support visual working memory. The role of early perceptual areas in visual working memory has been demonstrated in recent fMRI decoding studies (e.g.,; Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Emrich, Riggall, LaRocque, & Postle, 2013; Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009), suggesting that perception and working memory rely on common visual representations. Given that perceptual learning is known to modify the representations of both basic features (Sasaki, Nanez, & Watanabe, 2009) and complex objects (Palmeri, Wong, & Gauthier, 2004), we predict that extensive perceptual training with other types of stimuli should also lead to improvements in the precision of visual working memory.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. *Current Biology, 23*(15), 1427–1431.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106–111.

Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *The Journal of Neuroscience, 31*(3), 1128–1138.

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.

Baddeley, A. D., & Hitch, G. J. (1975). Working memory. *The psychology of learning and motivation, 8* (47-89).

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision, 11*(10):6, 1-15.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*(5890), 851–854.

Brady, T. F., Konkle, T. & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision.* 11(5):4, 1-4.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin & Review, 14*(4), 620–628.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 94–107.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107–117.

Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *The Journal of Neuroscience, 33*(15), 6516–6523.

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications, 3,* 1229.

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.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research, 37*(12), 1673–1682.

Gorgoraptis, N., Catalao, R. F., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *The Journal of Neuroscience, 31*(23), 8502–8511.

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

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision, 3*(1), 49–63.

Hollingworth, A. (2004). Constructing visual representations of natural scenes: the roles of short-and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 519–537.

Jeffreys, H. (1998). *The theory of probability.* Oxford University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences, 17*(8), 391–400.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*(1), 8–15.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review, 16*(5), 778–797.

Moore, C. D., Cohen, M. X., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *The Journal of Neuroscience, 26*(43), 11187–11196.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology. Revue Canadienne de Ppsychologie, 21*(3), 263–276.

O'Toole, A. J., Vetter, T., Troje, N. F., & Bülthoff, H. H. (1997). Sex classification is better with three-dimensional head structure than with image intensity information. *Perception, 26,* 75–84.

Palmeri, T. J., Wong, A. C., & Gauthier, I. (2004). Computational approaches to the development of perceptual expertise. *Trends in Cognitive Sciences, 8*(8), 378–386.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*(2), 283–290.

Rademaker, R. L., Tredway, C. H., & Tong, F. (2012). Introspective judgments predict the precision and likelihood of successful maintenance of visual working memory. *Journal of Vision, 12*(13), 1–13.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences, 105*(16), 5975–5979.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review, 16*(2), 225–237.

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2009). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience, 11*(1), 53–60.

Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic Bulletin & Review, 15*(1), 215–222.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20*(2), 207–214.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*(5), 583–592.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1016–1035.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*(4), 471–491.

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.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*(6982), 557–561.

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

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12), 1120–1135.

Xu, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 458.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233–235.