Serial dependence tracks objects and scenes independently

Thérèse Collins
Integrative Neuroscience and Cognition Center
University of Paris & CNRS
45 rue des Saints-Pères 75006 Paris, France

Abstract

The visual world is made up of objects and scenes. Object perception requires both discriminating an individual object from others and binding together different perceptual samples of that object across time. Such binding manifests by serial dependence, the attraction of the current perception of a visual attribute towards values of that attribute seen in the recent past. Scene perception is subserved by global mechanisms like ensemble perception, the rapid extraction of the average feature value of a group of objects. The current study examined to what extent the perception of single objects in multi-object scenes depended on previous feature values of that object, or on the average previous attribute of all objects in the scene. Results show that serial dependence occurs independently on two simultaneously present objects, that ensemble perception depends on previous ensembles, and that serial dependence of an individual object occurs only on the features of that particular object. These results suggest that the temporal integration of successive perceptual samples operates simultaneously at independent levels of visual processing.

Introduction

The visual world as we perceive it is made up of objects and scenes. Perceiving individual objects requires both discriminating them from others and properly binding together different perceptual samples of the same object across viewpoints, retinal locations, changes in ambient lighting, and time. Temporal binding is thought to be subserved by a mechanism called a continuity field, a spatio-temporally tuned operator that integrates successive perceptual samples over a brief moment in time and across a restricted region of space. Continuity fields would promote perceptual stability across time by smoothing out spurious fluctuations in the proximal stimulus. Continuity fields manifest in the laboratory as a phenomenon called serial dependence, small misperceptions of currently presented visual attributes as pulled towards values of that attribute seen in the immediate past. Perceiving scenes is aided by mechanisms that extract summary statistics and average them into global representations of cluttered environments ((Alvarez, 2011)). Observers are remarkably good at extracting the average value of a group of objects, such as the average orientation of a group of Gabor patches, the average speed and direction of moving dots, or the average emotional expression in a crowd of faces (e.g. (Haberman & Whitney, 2007)). Importantly, the attributes of the individual objects that make up a group are often less well perceived than the average attribute ((Ariely, 2001)). Furthermore, the two effects may be contradictory: the attraction of object perception towards features values in its recent past versus the perceptual predominance of the average attribute. It remains unknown how individual object and ensemble representations are integrated across time.

Serial dependence has been examined mainly in the context of sequential presentations of a single object, and observer reports of a feature of that object. It is thus unknown to what extent the perception of individual objects in multi-object scenes depends on previous ensemble representations of the scene, or whether independent representations can be maintained for individual objects and global scene characteristics. Ensemble perception is subject to serial dependence: when asked to report the average attribute of a set of stimuli, subject responses were pulled towards the average of the previous set ((Manassi et al., 2017)). Manassi et al. (2017) also showed that serial dependence occurred between single objects and ensembles, and vice versa, but they did not contrast the effect of previous individual objects versus previous ensembles. This is the crucial question to determine how object and scene representations interact to ensure perceptual stability across time.

The main question of interest here is whether an individual object that has been averaged into an ensemble representation can still influence upcoming perception of the individual item. If that were the case, it would suggest that multiple continuity fields are maintained simultaneously for different objects and for the ensemble representation including those objects. This seems unlikely, given the literature on how representations of individual objects in visual short-term memory is biased by the summary statistics of the current scene ((Brady & Alvarez, 2011; Lew & Vul, 2015)). Furthermore, long-term memories may resemble ensemble representations ((Richards et al., 2014)) and scene information is a powerful cue for long-term object recall (e.g. (Brady et al., 2011; Hollingworth, 2006)).

This study examined to what extent the current perception of a single shape and the current perception of the average of a set of shapes depended on previously seen shapes or sets of shapes (Figure 1). The previous literature on ensemble perception and its influence on memory for individual objects suggests the following hypothesis: that both the current

perception of an individual shape within a scene and the current perception of the average shape will be pulled towards the average shape of previous scenes.

Insert Figure 1 here

Results

Subjects (n=15) viewed single shapes presented at screen center for 150 ms, then masked, and reported whether they had seen a circle or a square. This "single shape at center" condition was run as contiguous block of 400 trials. Performance as quantified by the psychometric function is shown in Figure 2 (blue curve). In the same experiment, this condition was followed by a randomly interleaved mix of single and double shapes (400 and 1000 trials, respectively; Figure 1). These trials were each preceded by a cue alerting subjects as to the number of shapes (single or double) and, in the case of a single shape, its location (left or right). This was done to ensure that any differences between the single-shape condition in blocked versus interleaved conditions could not be attributed to uncertainty about target location or reduced attention: indeed, the strength of serial dependence decreases in the absence of attention ((Fischer & Whitney, 2014)).

Figure 2 shows psychometric functions for the two single-shape conditions (blocked and interleaved with double-shape trials). There were no significant differences in performance between them, on either slope or point of subjective equality (PSE) as shown by the overlapping 95% confidence intervals (insets in Figure 2). Subjects were good at reporting shape and there was no cost of interleaving single and double shape trials.

Insert Figure 2 here

Serial dependence was quantified as a percentage by considering error trials: the number of trials in which the response erred towards the previously presented shape divided by the total number of errors. Serial dependence could vary between -50% (a repulsive effect) and +50% (an attractive effect), with no serial dependence at 0%. Figure 3a shows percent serial dependence for different temporal distances between current and previous trials (1-back to 10-back), and the (non-significant) percent serial dependence of current response on future shape. This analysis is a control, since of course there can be no influence of the future on the present. The shaded region corresponds to 95% confidence intervals, based on bootstrapping individual data, such that any condition for which the lower bound of the confidence interval is above zero is considered significant. The current response depends on the shape seen in the trial immediately preceding it (1-back). All confidence intervals overlap, arguing against any differences between the two single-trial conditions.

In the double-shape trials, subjects reported one of the two shapes. The psychometric function is in red in Figure 2: there were no significant differences with either of the two single-shape conditions, as shown by the overlapping confidence interval for PSE and slope (Figure 2, inset).

Figure 3b shows percent serial dependence for different temporal distances between

current and previous trials in the double-shape condition. Similar to the single-shape conditions, current response depended on the shape seen at the same location in the 1-back trial.

In the double-shape trials, it is possible to look at the influence of the preceding shape on the same side depending on whether that shape was cued or not. Figure 4 breaks down serial dependence in the double-shape trials as a function of whether the previous shape at the same location was cued or not. For ease of comparison, the single-shape conditions are given again – i.e. the same data as the 1-back trial in Figure 3a – as well as the overall serial dependence in double-shape trials, i.e. the influence of the preceding shape at the same location regardless of whether it was cued or not. Current response depended on previous shape at the same location (3.0%, 95% CI: [1.2–4.8%]; replicated in Exp. 2, see below). It did not mater very much if that shape had been cued or not (cued: 3.4%, [0.9–6.4%]; uncued: 3.78%, [0.6–6.8%]). However, the serial dependence of current response on previously cued shape was not significantly different from zero when previous shape was at the opposite location (0.6%, [-1.4–2.6%]). These results suggest that each object (or side) is monitored independently.

Insert Figure 3 here

Figure 4 also represents how current perception depends on the average shape seen in the previous trial. This analysis considers both single-shape and double-shape trials that were preceded by a double-shape trial. The amount of serial dependence was not significantly different from zero (2.3%, [-0.2–4.6%]) and this was replicated in Exp. 2 (see below). This suggests that ensemble perception is not obligatory: when subjects reported individual items, their responses were independent from the previous average visual attribute.

Insert Figure 4 here

To further explore serial dependence of individual versus ensemble features, Experiment 2 required a new group of participants (n=19) to report the average of the two shapes. Stimuli were identical to the first experiment except for the response cue which underlined both shapes when an average response was required (Figure 1). For comparison purposes, 1285 ensemble-shape trials were interleaved with 516 double-shape trials (29% of all trials), i.e. trials in which two shapes were present but subjects reported the shape of either the left or right shape, as instructed by a response cue, just like in the first experiment.

Figure 5 shows psychometric functions for single and ensemble responses in the second experiment, with bootstrapped mean and 95% confidence intervals in the inset. Overlapping confidence intervals show that there were no significant differences between the two conditions: subjects were as good at reporting one of two shapes as the average of two shapes. This confirms many previous reports of ensemble perception in the literature.

Insert Figure 5 here

Percent serial dependence was calculated as before. Figure 3b (purple curve) shows that the temporal evolution of serial dependence was very similar to that observed on individual shapes in the first experiment: the current ensemble response depended on the ensemble seen in the 1-back trial. Figure 4 shows serial dependence in the 1-back trial, in red for the trials in which subjects reported one of the two shapes, and in purple when they reported the average. Several aspects of Experiment 2 replicate findings from Experiment 1: when subjects were asked to report one of the shapes, their response was significantly influenced by the shape seen on the same side in the previous trial (2.7% [0.4%-4.9%]), but their response was not significantly influenced by the average of the two previous shapes (0.8%, [-1.4–3.0%). However, responses were attracted towards the average of the two previous shapes when subjects were asked to report the average of two shapes (3.1%, [1.6%-4.7%]) – this is the same data as the 1-back trial in Figure 3b. Finally, when asked to report the average, responses did not depend significantly on the shape on the previously cued side (1.1%, [-0.7%-2.8%]).

Discussion

When participants were asked to report the shape of an individual object, their responses were attracted towards the shape of that object seen in the recent past. They were not influenced by a shape on the opposite side, even when that shape had been previously reported. This serial dependence between objects at the same location, but not across locations, may seem to contrast with results from other studies showing that serial dependence operates across locations (if they are not too far apart;(Collins, 2019)). However, in these previous studies, there was only ever one object, compared with the two in the current experiment. This result suggests that individual objects in scenes are monitored by independent continuity fields.

Participant reports of individual object shape were also not influenced by the average shape of the previous multi-shape scene, suggesting that ensemble perception does not obligatorily influence upcoming individual shape perception. The key difference with Manassi et al. (2017) is that they did not ask participants to report one object in a multi-object scene, and they did not examine to what extent single-object responses were attracted to the previous average or to the previous feature at the same location.

Finally, when participants were asked to report the average of two shapes, they were influenced by the previous ensemble representation (replicating Manassi et al., 2017). There was no influence of previously reported individual shapes on ensemble responses.

Taken together, these results show that the object representations that drive perception are influenced by multiple aspects of stimulus history, depending on the goal of current perception: individual object or global scene perception. The perception of individual objects in multi-object scenes is attracted towards previous features of that object, but not towards the ensemble representation of the previous scene, or by the previous features of a nearby object. Ensemble perception is influenced by previous ensembles, but not by the features of individual objects in a previously viewed scene. Because observers were not told whether they would have to report one of the objects or the average until after having viewed the scene, this means that single object representations and ensemble representations are maintained and integrated with previous history in parallel. In other words, multiple

continuity fields operate simultaneously on object and scene representations.

Materials and methods

Participants

Fifteen subjects participated in Experiment 1 (2 women, mean age 29 years old, range 18-65). Nineteen subjects participated in Experiment 2 (3 women, mean age 33 years old, range 18-56). The experiments were run on the online platform testable.org, and subjects were found through the Testable Minds subject pool. The number of subjects was determined by a power analysis based on reports of serial dependence effect sizes in the literature, which places the number at around 5 subjects for a power of 99% (i.e. estimated 1% probability of a false negative). The number of subjects was increased because it was practically feasible and ensured greater power, given that some of the conditions tested here were expected to have smaller serial dependence than what has been reported before. The relevant ethical information pertaining to the laboratory version of the study and approved by the French national ethics committee (CPP) in accordance with the declaration of Helsinki was provided at the onset of the experiment, and participants had to click to accept the consent form before proceeding with the experiment. Sex, age and screen characteristics, but no personally identifying information, were recorded for each participant.

Stimuli

Because the experiment was run on an online platform, exact control of stimulus size was not possible. Testable.org displays stimuli as images centered on observers' personal computer screens, adjusted for each participant's screen resolution but without distorting the image proportions (i.e. the same adjustment in width and height). Observers were instructed to sit 60 cm (arm's length) from their screen. All stimuli were 1024x768 pixel images; stimulus sizes below are given for a typical laptop screen, but because of individual screens the values can vary from 0.7 to 1.2 in proportion. Stimuli were generated using Psychtoolbox for Matlab ((Brainard, D., 1997; Kleiner, M et al., 2007; Pelli, G., 1997)) and saved as images.

A black dot, 0.25 degrees of visual angle (dva) in diameter and presented at screen center, served as a fixation point. Test stimuli were 99 geometrical shapes that ranged from a circle to a square. For the double-shape trials, there were thus 9801 possible combinations; a random subset of 1000 trials was selected. Shapes were 2.7 dva wide (circle diameter or square side). The transition from one shape to the other was made by placing, at each corner of a square, arcs defined by a central angle from 45° (circle) to 0° (square), in steps of 0.45°. For ease of interpretation, each shape was given a value from 0 (circle) to 1 (square); the fully ambiguous shape (value of 0.50) was not included because there is no correct response. Shapes were embedded in pixel noise. Masks were screens of pixel noise alone.

Procedure

Experiment 1 had two parts. The first 400 trials of the experiment were single shape trials. On each trial, a shape appeared at screen center for 150 ms, followed by a mask which also served as the response screen. The second part of the experiment was the double shape session, composed of 1000 double shape trials and 400 single shape trials (200 shapes on the left, 200 shapes on the right), randomly interleaved. The subject was cued about the number

of shapes and, if there was only one, about its location (left or right). In double shape trials, two shapes appeared to the left and right of fixation for 150 ms, followed by the mask. A response cue appeared simultaneously with the mask and indicated the shape the subject was to report (left or right). Half of the trials called for a response about the shape on the left, half about the shape on the right. Subjects responded by pressing one of two buttons on the keyboard (one for "circle" and another for "square"), followed immediately by the cue for the next trial. In single shape trials, the procedure was identical except that only one shape appeared, on the left or right, and the response cue was always on the valid side.

Experiment 2 was similar except that participants were asked to report one of the two shapes, identically to the double-shape trials in Experiment 1, or to report the average of the two shapes. In this case, the response cue underlined both shapes. There were 400 single-report trials and 1400 ensemble-report trials.

Data analysis

To quantify performance, psychometric functions were fit on individual response data using the quickpsy package developed for R by ((Linares & López-Moliner, 2016)) . Two parameters are of interest: the slope of the psychometric function, and the point of subjective equality (PSE), i.e. the stimulus level for which subjects selected each response option 50% of the time. Average psychometric functions shown in figures were obtained by bootstrapping individual fits.

Serial dependence was calculated as a percentage by considering error trials (i.e. trials in which the shape value was less than 0.5 in which observers reported seeing a square, and trials in which the shape value was greater than 0.5 in which observers reported seeing a circle). The number of trials in which the response erred towards the previously presented shape was divided by the total number of errors. If there was no relationship between current response and previous shape, 50% of the errors should be preceded by more circle-like shapes, and 50% by more square-like shapes. For ease of interpretation, 50% was thus subtracted from the percent serial dependence. Thus, no serial dependence is 0%, 50% serial dependence is if all erroneous responses erred towards the previous shape, and -50% serial dependence is if all erroneous responses erred away from the previous shape (a repulsive effect).

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. https://doi.org/10.1016/j.tics.2011.01.003

Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science*, *12*(2), 157–162. https://doi.org/10.1111/1467-9280.00327

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. https://doi.org/10.1177/0956797610397956

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), 4–4. https://doi.org/10.1167/11.5.4

Collins, T. (2019). The perceptual continuity field is retinotopic. *Scientific Reports*, *9*(1), 18841. https://doi.org/10.1038/s41598-019-55134-6

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743. https://doi.org/10.1038/nn.3689

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751–R753. https://doi.org/10.1016/j.cub.2007.06.039

Hollingworth, A. (2006). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*, *14*(4–8), 781–807. https://doi.org/10.1080/13506280500193818

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

Lew, T. F., & Vul, E. (2015). Ensemble clustering in visual working memory biases location memories and reduces the Weber noise of relative positions. *Journal of Vision*, *15*(4), 10. https://doi.org/10.1167/15.4.10

Linares, D., & López-Moliner, J. (2016). quickpsy: An R Package to Fit Psychometric Functions for Multiple Groups. *The R Journal*, *8*(1), 122. https://doi.org/10.32614/RJ-2016-008

Manassi, M., Liberman, A., Chaney, W., & Whitney, D. (2017). The perceived stability of scenes: Serial dependence in ensemble representations. *Scientific Reports*, *7*(1), 1971. https://doi.org/10.1038/s41598-017-02201-5

Richards, B. A., Xia, F., Santoro, A., Husse, J., Woodin, M. A., Josselyn, S. A., & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature Neuroscience*, *17*(7), 981–986. https://doi.org/10.1038/nn.3736

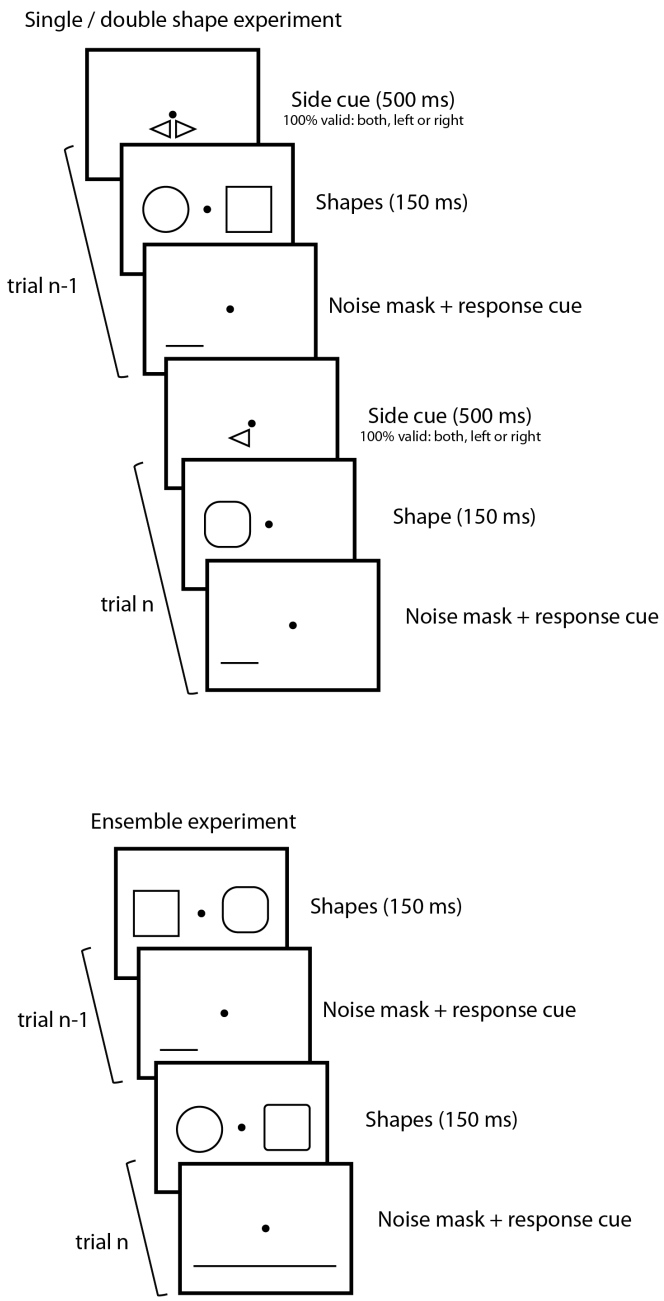

Figure 1. Procedure. Top panel: single / double shape experiment (Exp. 1). Two trials are pictured. In each, a 500-ms cue indicates number of shapes (one or two) and, in the case of one shape, its side (left or right). Shapes are then presented for 150 ms, followed by a noise mask and response cue until subject response. Bottom panel: ensemble experiment (Exp. 2). There were always two shapes, presented for 150 ms, followed by the noise mask and response cue until subject response.

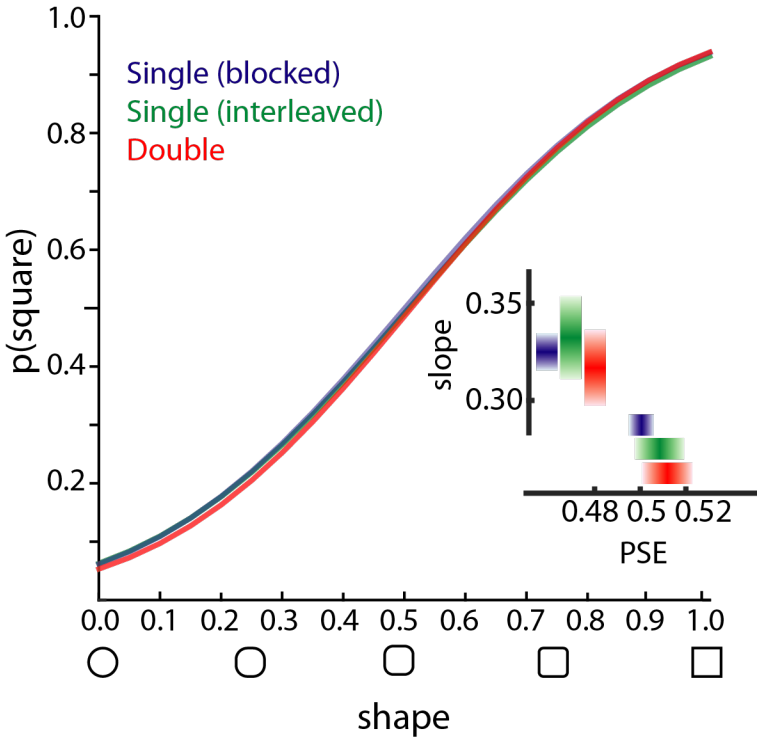

Figure 2. Psychometric functions (probability of responding "square" as a function of physical shape value), in the single/double shape experiment (Exp. 1). Blue: single shape trials, blocked design. Green: single shape trials, interleaved with double shape trials (in red). Inset: Average slope and PSE (± bootstrapped 95% confidence intervals).

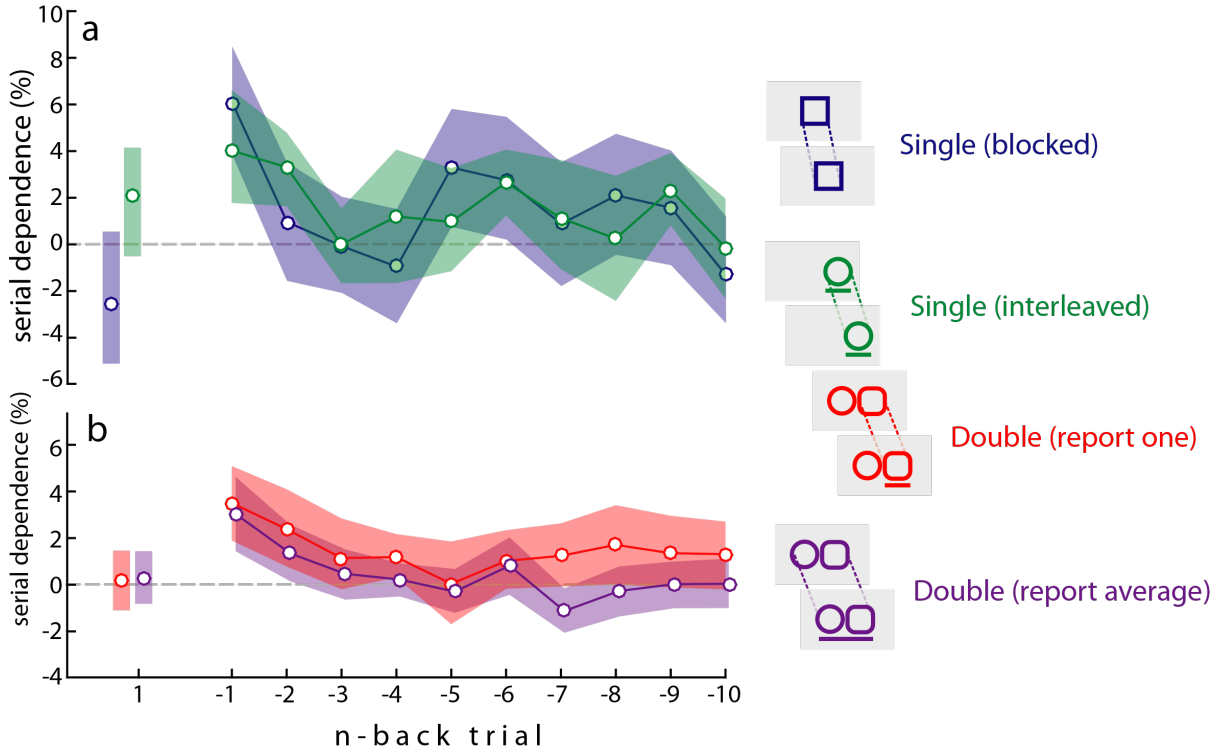

Figure 3. Percent serial dependence. **(a)** Single shape trials (Experiment 1), blocked (blue) or interleaved (green) design. **(b)** Double-shape trials, in which subjects had to report one of the two (red, Experiment 1) or the average (purple, Experiment 2).

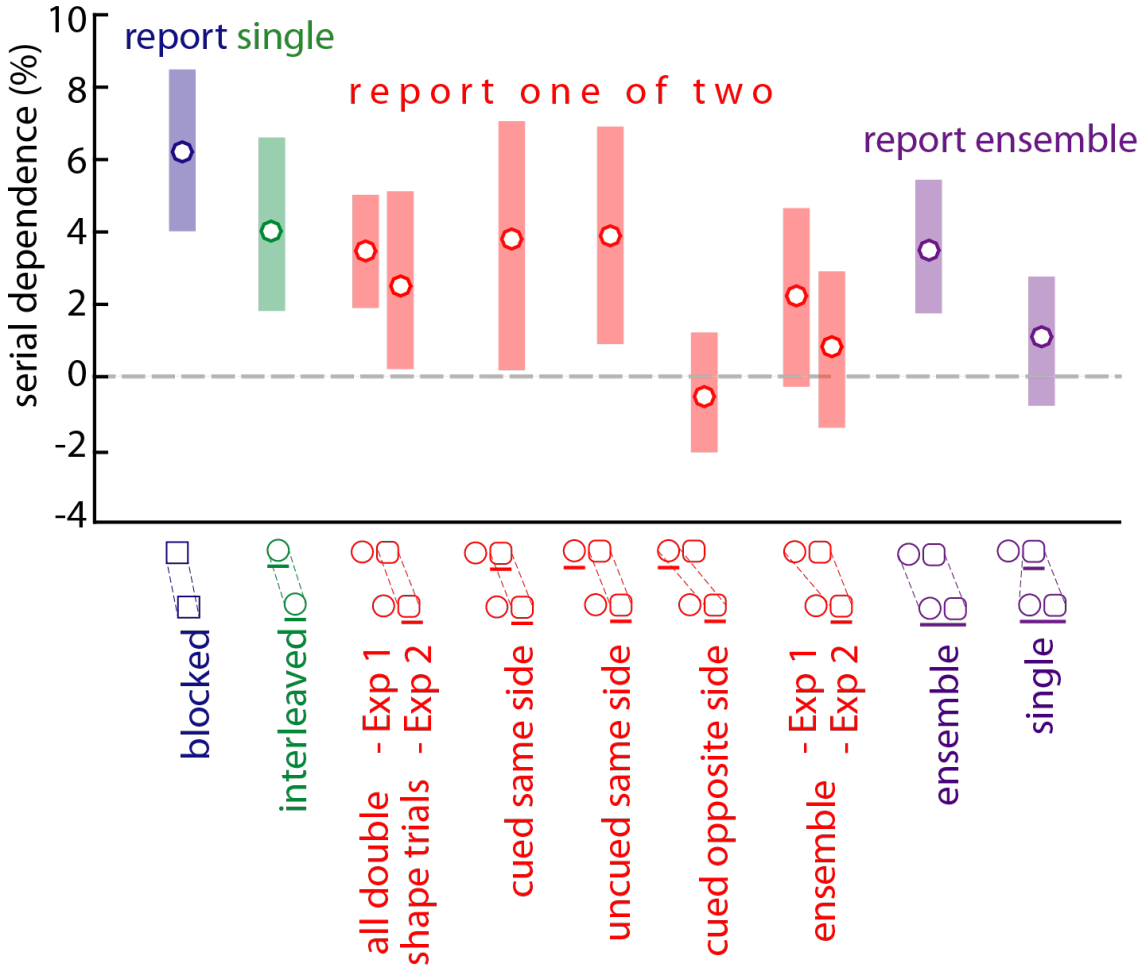

Figure 4. Percent serial dependence on the 1-back trial as a function of report (one, two or average), in Experiments 1 and 2. Blue and green: single shapes in blocked and interleaved designs, respectively. Red: double shape trials in which subjects reported only one of the two. In order, the conditions are the following: serial dependence on the previous shape on the same side for all double-shape trials (data from Exp. 1 & 2); SD on previous shape on the same side when it was cued (Exp. 1); SD on previous shape on the same side when it was uncued (Exp. 1); SD on previous cued shape on the opposite side (Exp. 1); SD on the average of the two previous shapes (Exp. 1 & 2). Purple (Exp. 2): double-shape trials in which subjects reported the average of the two shapes: SD on the average of the two previous shapes; SD on the previously cued single shape.

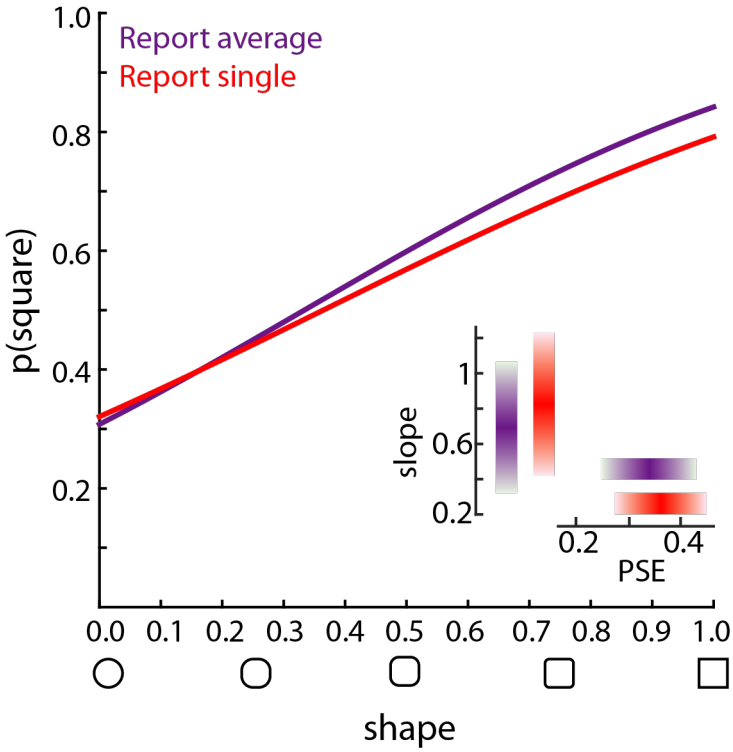

Figure 5. Psychometric functions (probability of responding "square" as a function of physical shape value), in the ensemble experiment (Exp. 2). Purple: double-shape trials in which subjects reported the average shape. Red: double-shape trials in which subjects reported one of the two shapes. Inset: Average slope and PSE (± bootstrapped 95% confidence intervals).