Distinct Contributions of Attention and Working Memory to Visual Statistical Learning and Ensemble Processing.

Michelle G. Hall[1], Jason B. Mattingley[1,2], & Paul E. Dux[1]

[1]*School of Psychology, The University of Queensland*

[2]*Queensland Brain Institute, The University of Queensland*

Word Count: 7533

To whom correspondence should be addressed:
Michelle G. Hall
School of Psychology
The University of Queensland
McElwain Building, St Lucia
Queensland 4072
Australia
E-mail: michelle.hall@uqconnect.edu.au

Abstract

The brain exploits redundancies in the environment to efficiently represent the complexity of the visual world. One example of this is *ensemble processing*, which provides a statistical summary of elements within a set (e.g., mean size). Another is *statistical learning,* which involves the encoding of stable spatial or temporal relationships between objects. It has been suggested that ensemble processing over arrays of oriented lines disrupts statistical learning of structure within the arrays (Zhao, Ngo, McKendrick, & Turk-Browne, 2011). Here we asked whether ensemble processing and statistical learning are mutually incompatible, or whether this disruption might occur because ensemble processing encourages participants to process the stimulus arrays in a way that impedes statistical learning. In Experiment 1, we replicated Zhao and colleagues' finding that ensemble processing disrupts statistical learning. In Experiments 2 and 3, we found that statistical learning was unimpaired by ensemble processing when task demands necessitated: 1) focal attention to individual items within the stimulus arrays, and 2) the retention of individual items in working memory. Together, these results are consistent with an account suggesting that ensemble processing and statistical learning can operate over the same stimuli given appropriate stimulus processing demands during exposure to regularities.

Keywords: Statistical learning; ensemble processing; visual attention; working memory

The human visual system faces a fundamental challenge, which is to make sense of the vast amount of information it receives at any given moment. One key to solving this challenge is that many aspects of the visual world are stable and predictable; boats are usually found floating on water, a fridge is often close to an oven, and the appearance of particular objects (e.g., a block of cheese) often precedes the appearance of others (e.g., a bottle of wine). The visual system has evolved efficient strategies to exploit such redundancies in the environment, thereby reducing information processing demands. For example, informational complexity can be compressed into an efficient sensory representation by summarising over sets of similar items, rather than encoding each item separately. This has been referred to as *ensemble processing*. Thus, for example, the average height of the buildings making up a city skyline provides an informative ‘summary’ or *ensemble statistic* that is more computationally efficient than encoding the height of each building individually (Ariely, 2001). Redundancies in visual input can also be compressed via *statistical learning*, whereby reliable co-occurrences between objects in the environment are detected and encoded into long term memory (Fiser & Aslin, 2001). It has been hypothesized that ensemble processing and statistical learning draw on a common underlying resource (Zhao, Ngo, McKendrick, & Turk-Browne, 2011). Here we report three behavioural experiments that examined this claim. We find that ensemble processing and statistical learning can occur simultaneously over the same arrays of items when task demands are manipulated in a way that encourages focused attention to individual items within the arrays, as well as maintenance of individual items in working memory. Our results suggest that the allocation of attention and the engagement of working memory processes during exposure to regularities can account for the observation that ensemble processing disrupts statistical learning.

Ariely’s (2001) influential work on ensemble processing investigated how the visual system represents sets of similar items. Participants were presented with brief stimulus arrays of differently sized disks, and were asked to indicate whether a subsequent probe disk was larger or smaller than the average size of all disks within the array. Participants performed significantly above chance on this task, suggesting that the statistical properties of the array as a whole were represented with high fidelity. In contrast, participants were unable to identify individual items from within the array, either by indicating whether a single probe stimulus matched the size of an individual item from within the array, or by choosing which of two probe stimuli had appeared in the array. This result suggests that participants lacked explicit representations of individual items under these task conditions. On the basis of these results, Ariely proposed that the visual system might represent individual items and sets of items in fundamentally different ways. In addition to mean size, ensemble processing has since been demonstrated for a range of feature dimensions, including orientation (Dakin & Watt, 1997; Jacoby, Kamke, & Mattingley, 2013; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), location (Alvarez & Oliva, 2008), number (Dehaene, Dehaene-Lambertz, & Cohen, 1998; Odic, Libertus, Feigenson, & Halberda, 2013; Whalen, Gallistel, & Gelman, 1999), velocity (Atchley & Andersen, 1995), facial expression (Haberman, Harp, & Whitney, 2009), and biological motion (Sweeny, Haroz, & Whitney, 2013).

In addition to the visual redundancies in sets of similar items that are encoded through ensemble processing, redundancies also exist in the co-occurrences between objects over time. Visual learning mechanisms have evolved to detect and encode reliable spatial and temporal patterns between objects, in a process known as *visual statistical learning* (Fiser & Aslin, 2001, 2002b; Kirkham, Slemmer, & Johnson, 2002). The first study to demonstrate visual statistical learning of spatial contingencies between objects was conducted by Fiser and Aslin (2001), who presented participants with a series of stimulus arrays that were constructed from a number of pairs of novel shapes. The two members of each pair always appeared together in the same spatial configuration relative to one another, but the pair itself could appear anywhere within a 4 x 4 grid. Critically, participants were not informed of the repeating regularities embedded within the arrays. After several minutes' exposure, participants completed a two-alternative forced-choice recognition test. On each trial, they were presented with one of the pairs that had appeared repeatedly during the exposure phase of the experiment, alongside a foil pair that had never appeared in the given spatial configuration during exposure. Participants were reliably above chance in distinguishing between the true pairs and the foil pairs, reflecting incidental learning of the pairs that had been embedded in the stimulus arrays during exposure. Subsequent work has demonstrated that statistical learning generates long-term memory representations and results in persistent learning effects (Kim, Seitz, Feenstra, & Shams, 2009; Thiessen, Kronstein, & Hufnagle, 2013; Turk-Browne, Scholl, Chun, & Johnson, 2009).

A fundamental question is how the relational structure between objects is extracted from amongst inconsistent and unpredictable aspects of the environment. The most widely cited explanation for statistical learning invokes a mechanism that computes the conditional probabilities between items (Kirkham et al., 2002; Saffran, Aslin, & Newport, 1996). Recent modelling work has challenged this account, however, instead providing evidence in support of a "Bayesian chunk learner" (Orbán, Fiser, Aslin, & Lengyel, 2008). Rather than extracting and representing conditional probabilities between all possible combinations of stimuli, which would be computationally inefficient and would suffer from a potential combinatorial explosion, the Bayesian chunk learner model instead prioritises and encodes only those "chunks" of information necessary to represent repeating information within visual scenes, and provides a good account of behavioural performance.

Ensemble processing and statistical learning both involve extraction and compression of visual information, ultimately forming an efficient representation of complex information. However, whereas ensemble processing generates a representation of the statistical properties of a distribution of features within a set, statistical learning instead forms representations of probabilistic relationships between items over time and space. Given the apparent similarities and differences between these two processes, an open question is whether they rely on the same mechanisms for extracting statistical information from sensory signals, or whether they reflect distinct processes. As noted above, only the study by Zhao and colleagues (2011) has investigated whether ensemble processing and statistical learning can occur in parallel, over the same arrays of stimuli (see Figure 1). In that study, participants were presented with brief arrays of oriented line segments. As in previous studies on statistical learning, the arrays were constructed from pairs of stimuli that always appeared together in the same spatial configuration. Importantly, participants were unaware that over the course of the exposure phase they were presented with repeating structure in the form of the orientation pairs. While being exposed to these regularities embedded in the stimulus arrays, participants estimated the mean orientation (an ensemble statistic) of each array of lines. An active control group monitored for occasional duplicate orientations within each array. In the recognition test following the exposure phase, participants in the ensemble group were no better than chance in distinguishing consistent orientation pairs from foils, whereas the duplicate detection group were statistically above chance in identifying the consistent pairs, implying statistical learning.

These results reveal interference at the behavioural level between ensemble processing and statistical learning of the line pairs presented during exposure, but do not provide evidence as to the source of the interference. One possibility is that ensemble processing and statistical learning tap a common mechanism involved in extracting statistical information from a visual scene. According to this hypothesis, statistical learning is disrupted when participants engage in ensemble processing because processing resources are preferentially allocated to the ensemble task, in accordance with its top-down priority. Alternatively, differences in stimulus processing, encouraged by task demands, might account for the differences in learning between the experimental groups (e.g., Aru, Bachmann, Singer, & Melloni, 2012; de Graaf, Hsieh, & Sack, 2012). As an example, selective attention is broadly conceptualised as a filter that allows certain stimuli to undergo further processing at the expense of others, and therefore might constrain statistical learning (Broadbent, 1958; Deutsch & Deutsch, 1963; Treisman & Gelade, 1980). Similarly, working memory has been hypothesised to serve as a selection mechanism, determining which information is available for further processing (Cowan, 2001). In addition, it features as a precursor to long-term memory in several prominent theories

(Baddeley, 1992; Baddeley & Hitch, 1974; Cowan, 2001), suggesting that it may also play a role in statistical learning.

Accumulating evidence suggests that attention does indeed modulate statistical learning. For example, under conditions in which interleaved sequences of items were presented in different colours, with structure repeating independently within each subset of stimuli, Turk-Browne, Jungé, and Scholl (2005) observed statistical learning for only the attended stimuli (but see Musz, Weber, & Thompson-Schill, 2014). In addition, Zhao, Al-Aidroos, and Turk-Browne (2013) reported that regularities introduced covertly into streams of visual stimuli captured attention spatially, suggesting that regularities are detected and attended automatically. Given this evidence, and the fact that statistical learning involves encoding contingencies between items, it seems plausible that processing of each item might be a necessary minimum to encode relationships between them. However, when ensemble statistics are extracted from an array, little information is available about individual items (Alvarez & Oliva, 2008; Ariely, 2001; Corbett & Oriet, 2011; Parkes et al., 2001). Thus, the lack of representation of individual item identities when engaging in ensemble processing could explain the interference with statistical learning reported by Zhao and colleagues (2011). On the other hand, ensemble processing appears to be flexible – whereas evidence suggests that ensemble processing is facilitated by distributed attention, only a modest decrement in performance is observed when ensemble statistics are estimated under focused attention (Chong & Treisman, 2005a). This raises the possibility that if participants focus on individual elements in an array while estimating ensemble statistics, statistical learning of regularities within the arrays might proceed without disruption.

To examine the source of behavioural interference between ensemble processing and statistical learning, we conducted three experiments in which we manipulated how participants processed the stimulus arrays during the exposure phase. After replicating Zhao and colleagues' (2011) results, in Experiment 2 we evaluated the hypothesis that statistical learning might be facilitated when processing is directed toward individual items within each array, relative to when processing is distributed over the entire array. In Experiment 3, we attempted to dissociate the spatial scale of attention (local vs. global) from the requirement to maintain item-specific information in working memory, to further characterise when statistical learning does and does not occur.

Experiment 1

Experiment 1 was a partial replication of the design of Zhao and colleagues (2011; Experiment 1). These authors reported that estimating ensemble statistics of arrays of oriented lines impaired statistical learning of orientation pairs that appeared together consistently within the arrays. Zhao and colleagues observed that monitoring for occasional duplicate orientations during the exposure phase did not impair statistical learning, and actually facilitated learning relative to that of a passive viewing group. We therefore had our control group perform duplicate detection, and excluded the passive viewing group from our design.

Method

Participants. Thirty-eight participants were recruited from The University of Queensland community and were paid $10 for participation. All participants had normal or corrected-to-normal vision and provided written informed consent in accordance with a protocol approved by The University of Queensland ethics committee. Data from two participants were excluded due to technical difficulties, leaving a final sample of 36 participants (18 per group; mean age = 24.9 years, 12 male). The groups were similar in age (mean age of 25.4 years in the duplicate group, and 24.3 years in the ensemble group) and gender (7 males in the duplicate group, 5 males in the ensemble group).

Apparatus and stimuli. Participants were seated approximately 57cm from a 21 inch Sony-Dell Trinitron CRT monitor. Stimuli were presented via a Mac Mini running MATLAB and the Psychophysics Toolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). The stimuli were eight line segments, each tilted at a unique orientation (0° - horizontal, 15°, 45°, 75°, 90° - vertical, 105°, 135°, 165°), presented in white against a grey background. At the beginning of the experiment, the eight orientations were pseudo-randomly assigned without replacement to four pairs, each defined by the unique spatial arrangement of the two line segments making up the pair. This pair assignment had the constraint that the mean orientation of the line segments for all permutations of three pairs did not equal 90° (see Procedure below). One pair of lines was arranged horizontally (side by side), one vertically (one above the other), and two pairs were arranged diagonally (the second line of the pair centred at 45° upward and to the left of the first line, for one diagonal configuration, and upward and to the right of the first line, for the other; see Figure 1 for an example diagonal pair configuration).

Exposure phase. On each trial during the exposure phase, a stimulus array was created by pseudo-randomly placing three pairs into 6 of the cells of an invisible 3 x 3 grid (approximately 13.5° x 13.5° of visual angle). On 20% of exposure trials, one orientation was selected randomly from the array and its orientation was changed to match the other member of that pair. The direction of the mean orientation of the array (leftward or rightward) was counterbalanced across trials, and the order of presentation of arrays was randomised.

Test phase. On each trial of the recognition test, participants were presented with one of the consistent pairs from the exposure phase, alongside a foil pair. Foils were generated for each pair by replacing the first orientation of each pair with an orientation selected pseudo-randomly from the other three pairs. This procedure was repeated to generate a second foil for each pair, this time replacing the second orientation from the pair. Each of the orientations from the exposure phase appeared equally often as a pair member and as a member of a foil pair. Each pair-foil combination appeared twice during the recognition test, for a total of 16 trials. Whether the consistent pair or foil pair appeared on the left or right of the screen was counterbalanced across trials.

Procedure. Participants were assigned to one of two experimental groups (ensemble group, duplicate group). The groups differed only in the task that they completed during the exposure phase. Participants in the ensemble group were instructed to indicate whether the average orientation of the line segments in each array was rightward or leftward (i.e., pointing toward the upper right or upper left, respectively), whereas the duplicate group were instructed to indicate whether all of the orientations in the array were unique, or whether one orientation had been repeated. All responses were made by pressing one of two keys on the keyboard. Participants were instructed to respond as accurately as possible, and speed was not emphasised, except that responses needed to be entered before the stimuli offset from the screen. No information was provided about the pair structure embedded within the exposure arrays.

To ensure participants in both groups understood the task instructions, they first observed 5 example trials demonstrating the task for their condition, followed by an additional 10 untimed practice trials, with feedback. The experimenter was available to answer participants' questions. The arrays presented during the example and practice trials were identical for all participants, and the line segments were placed randomly into the grid so that there was no embedded structure. This was followed by the exposure phase, in which the pair structure was introduced to the arrays. This consisted of 400 experimental trials with no feedback (example trials are shown in Figure 2, top left). Each trial began with a fixation point at the centre of the screen for 950ms, a blank screen for 50ms, and finally the stimulus array for 2000ms. Responses made after the stimulus array offset from the screen were recorded as incorrect. Short breaks were offered every 100 trials.

After the exposure phase, participants completed 16 trials of a two-alternative forced-choice recognition test, in which one of the pairs from the exposure phase was presented on the screen alongside a foil pair (see Figure 2, bottom left). Recognition test trials began with a fixation point on screen for 1000ms followed by the two test pairs (one consistent pair and one foil) for 2000ms. After the stimuli offset, the screen remained blank until a response was entered.

Results and Discussion

Results for Experiment 1 are shown in Figure 2. During the exposure phase, mean accuracy for the duplicate group (86.2%, SD = 3.6%) was significantly better than chance (50%), $t(17) = 42.66$, $p < .001$, $d = 10.06$. Mean accuracy for the ensemble group (61.8%, SD = 8.3%) was also significantly above chance, $t(17) = 6.02$, $p < .001$, $d = 1.42$. Although this level of performance is somewhat poorer than has previously been reported in the literature, we note that the orientations of the stimuli used here were highly variable within each stimulus array, and heterogeneity of stimuli has been shown to reduce ensemble performance (Morgan, Chubb, & Solomon, 2008). Mean reaction time during the exposure phase did not differ for the duplicate group (1477ms) and the ensemble group (1444ms), $t(34) = 0.56$, $p = .579$, $d = 0.19$, although it should be noted that participants were not encouraged to respond quickly beyond the requirement to respond before the array offset from the screen (2000ms).

Accuracy on the recognition test was taken as a measure of statistical learning of the pairs presented during the exposure phase. Mean accuracy for the duplicate group was 62.2% (SD = 14.9%), which was significantly above chance performance (50%), $t(17) = 3.45$, $p = .003$, $d = 0.81$. This result demonstrates that the duplicate group could distinguish between the pairs and the foils, even while engaged in a task during exposure to statistical regularities. In contrast, accuracy for the ensemble group was 46.2% (SD = 13.1%), which did not differ significantly from chance, $t(17) = 1.24$, $p = .232$, $d = 0.29$. Recognition test accuracy was significantly better for the duplicate group than the ensemble group, $t(34) = 3.42$, $p = .002$, $d = 1.14$. These results replicate those reported by Zhao and colleagues (2011), revealing disrupted statistical learning when participants estimated ensemble statistics over arrays containing regularities.

The lack of statistical learning observed for the ensemble group might have resulted from competition for a domain-general statistical processing mechanism. Alternatively, differences in general task demands of the ensemble estimation and duplicate detection tasks might account for the differences in learning. Given that statistical learning involves encoding spatial relationships between individual items (Orbán et al., 2008), this learning mechanism might rely on serial attention to each item within the arrays, in order to detect and encode contingencies between them. However, previous work has shown that when participants are able to accurately estimate ensemble statistics over arrays of items, they are no longer able to report featural information about individual items (Ariely, 2001;

Corbett & Oriet, 2011). In addition, ensemble estimation proceeds automatically when attention is distributed over an array of items (e.g., Chong & Treisman, 2005a, 2005b; Parkes et al., 2001), implying that ensemble processing may have an inherent bias toward distributed processing. Thus, participants in the ensemble group might have been inclined to distribute attention over the entire array, whereas the duplicate detection task might have favoured focusing attention serially over individual line segments. This differential allocation of attention over the stimulus arrays during the exposure phase may have disrupted learning for the ensemble group.

The maintenance of item-specific information in working memory may also play an important role in gating statistical learning, and may have differed systematically for participants in each group. For the ensemble group, the binary (leftward/rightward mean orientation) ensemble estimation task is unlikely to have required precise featural processing (Alvarez, 2011; Alvarez & Oliva, 2008), whereas the duplicate detection task encouraged more precise processing because it required comparison of features between items in the array. Within the visual working memory literature, it has been argued that the precision of representations maintained in memory is variable (Bays & Husain, 2008; Fougnie, Suchow, & Alvarez, 2012; van den Berg, Shin, Chou, George, & Ma, 2012; Zhang & Luck, 2008). Evidence from human neuroimaging research corroborates this claim by showing that the reproducibility of patterns of brain activity in sensory regions - thought to reflect the fidelity of the representation of the stimulus - is strongly related to behavioural performance on memory recall tasks, and is modulated by attention (Emrich, Riggall, LaRocque, & Postle, 2013; Ester, Anderson, Serences, & Awh, 2013). Analogously, precision of processing and encoding may be important for statistical learning. This might be particularly important in the current experiments, as the orientation pairs used were defined solely by co-occurrences between orientations, with some orientations differing from one another by as little as 15°. Noisy, imprecise representations of the orientations might have prevented differentiation between them, and therefore prevented learning.

Experiment 2

The aim of Experiment 2 was to investigate whether the disrupted statistical learning in the ensemble group relative to the duplicate group in Experiment 1 was driven by task-induced differences in stimulus processing, rather than competition for a shared statistical processing mechanism. Therefore, Experiment 2 was designed to encourage participants to prioritise the number of items processed and encoded, as well as the fidelity of those representations. To achieve this, we designed a dual-task version of Experiment 1, in which participants not only estimated the average orientation (ensemble estimation group) or detected duplicate orientations (duplicate detection group) within each array, but also judged whether the orientation of a single probe item, appearing after each stimulus array, had changed relative to the stimulus that was previously in that location. Neither a global allocation of attention nor imprecise processing and encoding into working memory would be effective for this task because participants did not know in advance which item from the stimulus array would be probed. If precise featural information about individual items is important for statistical learning of contingencies among items, the detailed encoding of each item necessitated by the change detection task might overcome the previously observed interference in the ensemble group.

Method

The method for Experiment 2 was identical to that for Experiment 1, except where specified.

Participants. Forty-three new participants were recruited from The University of Queensland community, and received $10 for participation. Results from six participants were excluded for having missing data on more than 75% of trials during the exposure phase, due to slow or otherwise missing responses (on the primary task, see below). One additional participant was excluded for low accuracy during the exposure phase (7.25% for the primary task). This left a final sample of 36 participants (18 per condition; mean age = 22.1 years, 11 male). The groups were similar in age (mean age of 22.6 years in the duplicate group, and 21.6 years in the ensemble group) and gender (4 males in the duplicate group, 7 males in the ensemble group).

Apparatus, stimuli and procedure. Each trial began as in Experiment 1, with participants responding to the primary task – the ensemble estimation or duplicate detection task – before the initial stimulus array offset from the screen. The stimulus array was now followed by a masking stimulus for 100ms, after which a probe stimulus appeared in one of the locations that had been occupied in the stimulus array (example trials are shown in Figure 3, top left). The probe stimulus could either be identical to the line segment previously appearing in that location, or could be rotated by 45° in either direction. Whether the orientation was a match or a mismatch was counterbalanced across trials, as was the direction of offset for mismatch stimuli. Participants made their second response, to indicate whether the orientation of the probe was a match or a mismatch, by pressing one of two keys on the keyboard. Responses for the change detection task were untimed, and the probe stimulus remained on the screen until a response was entered. As this was now a challenging dual task, we instructed participants to prioritise the primary task (duplicate detection or ensemble estimation) over the change detection task, in order to maintain above chance performance on the primary task.

Results and Discussion

Results for Experiment 2 are shown in Figure 3. During the exposure phase, mean accuracy for the duplicate group was 76.2% (SD = 12.8%), which was significantly better than chance (50%), $t(17) = 8.69$, $p < .001$, $d = 2.05$, but significantly worse than performance for this group in Experiment 1, $t(34) = 3.19$, $p = .003$, $d = 1.06$. Mean accuracy for the ensemble group was also significantly better than chance, $t(17) = 4.77$, $p < .001$, $d = 1.21$ (57.6%, SD = 6.8%), and did not significantly differ to that observed for the ensemble group in Experiment 1, $t(34) = 1.66$, $p = .105$, $d = 0.55$. During the exposure phase, mean reaction time did not differ significantly between the duplicate group (1528ms) and the ensemble group (1477ms), $t(34) = 0.83$, $p = .414$, $d = 0.28$.

Accuracy on the change detection task was significantly better than chance (50%) for both the duplicate (mean accuracy 53.7%, SD = 5.8%, $t(17) = 2.68$, $p = .016$, $d = 0.63$), and ensemble groups (52.2%, SD = 4.4%, $t(17) = 2.12$, $p = .049$, $d = 0.50$). Although these accuracy results for the change detection task are numerically low, this is not surprising given that they were obtained during a challenging dual task, and we instructed participants to prioritise the primary task (ensemble or duplicate judgement) over the change detection task. In addition, working memory is widely accepted to be a capacity limited process (Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Todd & Marois, 2004), and performance on a change detection task with a set size of 6 items would be expected to be well below ceiling.

For the recognition test, mean accuracy for the duplicate group was 59.0% (SD = 16.4%), which was statistically above chance (50%), $t(17) = 2.34$, $p = .032$, $d = 0.55$, despite the poorer performance during the exposure phase for this group relative to that observed in Experiment 1. Thus, as in Experiment 1, the duplicate group demonstrated learning of the regularities embedded in the stimulus arrays. Importantly, however, accuracy for the ensemble group was now 58.7% (SD = 13.2%), significantly better than chance, $t(17) = 2.78$, $p = .013$, $d = 0.66$, and significantly greater than that seen for this condition in Experiment 1, $t(34) = 2.85$, $p = .007$, $d = 0.95$. This provides evidence that the ensemble group also learned the regularities embedded in the stimulus arrays, despite estimating ensemble statistics over the arrays during exposure. In addition, recognition test accuracy was not significantly different between the two groups in Experiment 2, $t(34) = 0.07$, $p = .945$, $d = 0.02$. Thus, ensemble processing and statistical learning can co-occur when task demands encourage stimulus processing that favours both operations.

Together, the results of Experiments 1 and 2 reveal that statistical learning can occur during ensemble encoding, provided that item-specific processing is encouraged, as was the case with the addition of the change detection component of Experiment 2. However, the results do not reveal whether attention to individual items within the stimulus arrays is sufficient to induce statistical learning during ensemble processing, in an obligatory manner. An additional constraint may be that item representations also need to be maintained in working memory, as this was also encouraged by the change detection task.

Experiment 3

The aim of Experiment 3 was to determine whether the retention of items in working memory is necessary for statistical learning under ensemble processing conditions. To address this, we kept constant the requirement for focal attention to each item in the array, but now limited the requirement that individual items be encoded into working memory, to reduce the demands placed on this resource relative to Experiment 2. To encourage participants to process each line segment equally during the exposure phase, and to equate the allocation of attention to the stimuli across the two conditions, in Experiment 3 we presented the line segments in each array sequentially. Statistical learning has previously been demonstrated for regularities embedded within temporal sequences of items (e.g., Fiser & Aslin, 2002a; Turk-Browne et al., 2005; Turk-Browne & Scholl, 2009; Turk-Browne et al., 2009). Similarly, ensemble statistics can be estimated for sequentially presented stimuli as well as simultaneously presented arrays, with only a modest decrement in accuracy for sequential relative to simultaneously presented arrays (Chong & Treisman, 2005a; Corbett & Oriet, 2011).

The task requirements of Experiment 3 necessitated serial processing of each item, but did not place the same demands on retention in working memory as the change detection paradigm of Experiment 2. Specifically, rather than maintaining each item in memory (as in Experiment 2, in service of the change detection task), and integrating them into an average ensemble representation immediately before making a response at the end of the trial, participants approach this task by continually updating their ensemble representation as each new stimulus appears (Chong & Treisman, 2005a; Corbett & Oriet, 2011; Haberman et al., 2009). In other words, information is maintained in working memory, but in the form of an ensemble representation rather than representations of individual items.

Method

The method for Experiment 3 was identical to that for Experiment 1 except where specified.

Participants. Thirty-six new participants were recruited from The University of Queensland community, in exchange for $10 (18 per condition; mean age = 23.4 years, 10 male). The groups were similar in age (mean age of 23.1 years in the duplicate group, and 23.8 years in the ensemble group) and gender (6 males in the duplicate group, 4 males in the ensemble group).

Apparatus, stimuli and procedure. During the exposure phase, each of the 6 line segments in the stimulus array appeared sequentially on the screen, while maintaining their spatial locations within the array (example trials are shown in Figure 4, top left). This encouraged participants to attend equally to each item, as the onset of each stimulus likely captured attention (Yantis & Jonides, 1990). Each line segment appeared on the screen for 333ms (2000ms for all 6 line segments in an array), thereby matching the total stimulus exposure duration used during the exposure phase in Experiment 1. The sixth and final line segment of each array was followed by a blank screen for 2000ms, during which participants made their response. The order in which the pairs appeared on a given trial was randomised, but the members of any given pair always appeared on successive frames, and always in the same temporal order.

Each recognition test trial began with a fixation point flashing briefly at the centre of the screen, as an alerting signal, by alternately displaying the fixation point for 200ms, and a blank screen for 200ms, for two cycles. The fixation point then reappeared and remained on screen for 1500ms, before a brief blank interval of 500ms preceded the stimuli. The members of the first of the two test pairs (the consistent pair or the foil pair, counterbalanced) then appeared sequentially on the left of screen for 333ms each, followed by a blank interval of 1500ms, before the members of the second test pair appeared sequentially for 333ms each on the right of screen (see Figure 4, bottom left). As in Experiment 1, the screen remained blank until a response was entered. We matched the stimulus presentation parameters for the recognition test to those of the exposure phase to increase the sensitivity of the recognition test to detect learning of the pairs.

Results and Discussion

Results for Experiment 3 are shown in Figure 4. During the exposure phase, mean accuracy for the duplicate group was 71.7% (SD = 9.1%), which was significantly above chance performance (50%), $t(17) = 10.13$, $p < .001$, $d = 2.39$. Mean accuracy for the ensemble group was 67.2% (SD = 14.2%), which was also significantly better than chance, $t(17) = 5.12$, $p < .001$, $d = 1.21$.

For the recognition test, mean accuracy for the duplicate group was 59.0% (SD = 16.4%), which was statistically better than chance (50%), $t(17) = 2.34$, $p = .032$, $d = 0.55$. This result was not statistically different to the recognition test performance for the duplicate group in Experiment 1, $t(34) = 0.60$, $p = .553$, $d = 0.20$, and was identical to that in Experiment 2. Again, this demonstrates that the duplicate group could distinguish between the pairs and the foils, despite having been engaged in a task while being exposed to the statistical regularities. In contrast, and consistent with the pattern observed in the Experiment 1, accuracy for the ensemble group was 49.3% (SD = 13.9%), which was not reliably different from chance, $t(17) = 0.21$, $p = .834$, $d = 0.05$. Here, test accuracy was not significantly different to Experiment 1, $t(34) = 0.70$, $p = .491$, $d = 0.23$, but was significantly reduced compared to Experiment 2, $t(34) = 2.07$, $p = .046$, $d = 0.69$. Statistically comparing test accuracy for the two groups in Experiment 3 revealed a trend toward better recognition in the duplicate group than in the ensemble group, $t(34) = 1.92$, $p = .063$, $d = 0.64$.

Thus, despite presenting the line segments sequentially to equate the serial allocation of attention in the duplicate and ensemble groups, only the former displayed statistical learning. We interpret these results, together with those of Experiments 1 and 2, as evidence that not only focal attention to individual items, but also encoding of those individual items into working memory, are necessary conditions for statistical learning under ensemble processing conditions.

General Discussion

In three experiments, we investigated the extent to which ensemble processing and visual statistical learning interact. It has been reported that estimating ensemble statistics over arrays of stimuli disrupts statistical learning (Zhao et al., 2011), which has previously been shown to proceed automatically for attended inputs, even while participants engage in a cover task (Turk-Browne et al., 2005). We assessed whether this interaction between statistical learning and ensemble processing results from a difference in stimulus processing required for each. After replicating Zhao and colleagues' result, we equated the stimulus processing for both groups during the exposure phase in Experiment 2, by introducing an additional change detection task to encourage precise processing of the features of each item. Here, both the duplicate group and the ensemble group were sensitive to the pairs that had been embedded in the stimulus arrays during the exposure phase, relative to foil pairs - our measure of statistical learning. In Experiment 3, we attempted to dissociate the effects of attention and working memory on statistical learning. Specifically, we held constant the requirement for serial attention to each item, by presenting individual items within each stimulus array sequentially, but reduced the requirement for individual items to be maintained in working memory, as there was no change detection task. Under these conditions we observed statistical learning only for participants in the duplicate group, for which maintenance of items in working memory was necessitated by the duplicate detection component of their task. On the basis of these findings, we suggest that ensemble processing and statistical learning can occur over the same visual displays when attention is focused on individual items, and when processing and encoding of their features is encouraged, even under challenging dual-task conditions.

Which aspects of the design of Experiment 2 allowed ensemble processing and statistical learning to occur without disruption? It has been argued that ensemble statistics are extracted pre-attentively and in parallel over arrays of items (Chong & Treisman, 2003; Corbett & Oriet, 2011; Utochkin & Tiurina, 2014), and that ensemble representations come at the expense of information about individual items within the array (Alvarez & Oliva, 2008; Ariely, 2001; Corbett & Oriet, 2011; Parkes et al., 2001). The inclusion of the additional change detection task in Experiment 2 required participants not only to estimate ensemble statistics (or monitor for duplicate orientations), but also to attempt to encode as many items as possible into working memory, for later comparison to the probe item. Thus, this additional task requirement served to attenuate any differences in the allocation of attention and working memory resources between the experimental groups. Under these conditions, statistical learning proceeded without disruption for both groups.

The results of Experiment 3 help to clarify the source of interference between statistical learning and ensemble processing observed in Experiment 1, and in Zhao and colleagues' (2011) work. Here, participants in both groups attended each item, necessitated by the sequential presentation of the stimuli within each array, but only the duplicate group showed statistical learning. Whereas the duplicate detection task required participants to maintain precise item representations in working memory, in order to compare each item with the next and identify duplicate orientations, the ensemble task required participants to maintain an ensemble representation in memory, rather than individual item identities, and update this representation upon the appearance of each subsequent item (Chong & Treisman, 2005a; Corbett & Oriet, 2011; Haberman et al., 2009). This result points to working memory as a candidate precursor for statistical learning, suggesting that representations of each item must be encoded into working memory in order to detect spatial co-occurrences between them.

In addition to the duplicate detection and ensemble estimation tasks included in the current work, statistical learning has also been demonstrated under passive viewing conditions, in which participants are typically instructed to attend to the stimulus displays because they will be asked some simple questions about them later (Fiser & Aslin, 2001; Kirkham et al., 2002; Saffran et al., 1996; Turk-Browne, Isola, Scholl, & Treat, 2008). Although there are no explicit task demands under these conditions to serially attend to items, or encode them into working memory, participants are likely to inspect and explore the display visually. Working memory processes may also be engaged under passive viewing conditions, if participants attempt to memorise aspects of the stimuli that they anticipate will become relevant later in the experiment. Furthermore, recent evidence suggests that visual information can be encoded into working memory automatically (Hassin, Bargh, Engell, & McCulloch, 2009; Marshall & Bays, 2013; Umemoto, Scolari, Vogel, & Awh, 2010), and that this can occur without conscious awareness (Dutta, Shah, Silvanto, & Soto, 2014; Soto, Mantyla, & Silvanto, 2011).

The precision of encoding of individual line segments might be additionally important in modulating statistical learning. A wealth of behavioural and neuroimaging work has demonstrated that the fidelity of representations maintained in working memory is variable, and this variability interacts with attention (Bays & Husain, 2008; Emrich et al., 2013; Ester et al., 2013; Fougnie et al., 2012; van den Berg et al., 2012; Zhang & Luck, 2008). If precision is important for statistical learning, as it is for goal-directed working memory performance, the change detection component of Experiment 2 might have facilitated learning in the ensemble group because it encouraged participants to encode the stimuli with high precision. Similarly, the duplicate detection task might have required more precise encoding of individual item features relative to the ensemble estimation task, in both experiments. Whereas statistical learning might depend on the precision of featural processing, ensemble estimation can be quite accurate even when individual representations are imprecise (e.g., Alvarez & Oliva, 2008). Hence, in Experiment 1 imprecise encoding of items might have been sufficient for the ensemble task, which required only a binary decision as to whether the average orientation of the lines within the array was leftward (less than 90°) or rightward (greater than 90°), but insufficient to induce statistical learning.

Precise encoding may have been particularly important in the current experiments because the line segments employed in the present study varied on a single feature dimension – orientation – rendering the differences between the stimuli quite subtle. Previous work has typically used more complex stimuli, such as novel shapes (Fiser & Aslin, 2001; Turk-Browne & Scholl, 2009) or alphabetic glyphs (Turk-Browne et al., 2009), which vary on multiple features and are thus more distinguishable from one another. These studies have typically reported stronger learning effects than those reported here, and by Zhao and colleagues (2011). Although we cannot directly address this precision hypothesis with the current design and results, the role of precision in statistical learning poses an interesting question for future investigations.

Although not statistically reliable, we observed a numerical reduction in ensemble performance from Experiment 1 to Experiment 2. This could be due to the introduction of the secondary change detection task in Experiment 2, but there are two alternative possibilities. First, ensemble processing may indeed interfere with statistical learning at a mechanistic level, which is nonetheless attenuated by focal attention and encoding of individual items into working memory. Second, participants may have switched tasks throughout the exposure phase, at times prioritising the ensemble task, and at other times prioritising the change detection task. However, participants were clearly instructed to prioritise the primary task over the change detection task, and were given practice trials to familiarise themselves with the tasks prior to the experiment.

Somewhat surprisingly, we found very similar recognition test accuracy for the duplicate group across all three experiments. This was observed despite clear differences in the difficulty of the task during encoding across the experiments, as revealed by the significant drop in performance on the duplicate detection task from Experiment 1 to Experiment 2. This raises the intriguing possibility that statistical learning might occur independently of task difficulty or general cognitive load during exposure to regularities.

Our findings converge with other recent evidence for a modulatory role of attention on statistical learning (Turk-Browne et al., 2005; Zhao, Al-Aidroos, & Turk-Browne, 2013), and provide an informative constraint for mechanistic accounts of this phenomenon. Recent modelling work suggests that statistical learning is highly efficient, representing only those contingencies that are uniquely informative for recurring visual inputs (Orbán et al., 2008), and does not rely on computing conditional probabilities between all possible combinations of objects (e.g., Fiser & Aslin, 2001). An interesting question, then, is how an unsupervised learning mechanism can be efficient if it is restricted to only those inputs that are attended and encoded into working memory. A possible reconciliation of this apparent paradox comes from recent evidence that attention is captured toward locations where regularities are present. That is, Zhao and colleagues (2013) showed that regularities, introduced covertly, had a biasing effect on spatial attention. Similarly, infants show a preference to attend to the location of a sequence of stimuli with regularities that are moderately complex, compared with sequences with regularities that are either more or less complex (Kidd, Piantadosi, & Aslin, 2012). As noted above, there is also an emerging body of work showing that encoding of visual information into working memory can occur automatically (Hassin et al., 2009; Marshall & Bays, 2013; Umemoto et al., 2010), and implicitly (Dutta et al., 2014; Soto et al., 2011).

In addition to showing disrupted statistical learning when estimating ensemble statistics, Zhao and colleagues (2011) reported a complementary interference effect of statistical regularities on ensemble processing, showing that the mere presence of regularities within stimulus arrays impaired estimation of ensemble statistics over those arrays. Their results revealed superior ensemble estimation for unstructured arrays relative to structured arrays (Experiment 2), and for participants who were pre-exposed to the regularities embedded within the arrays prior to the ensemble estimation task (Experiment 3). These results were taken as further evidence for interference between ensemble processing and statistical learning. However, these results are also consistent with an alternative account, that the allocation of attention differed when regularities were present versus absent (Experiment 2), and when regularities were learned versus novel (Experiment 3). Chong and Treisman (2005a) found that attention can modulate ensemble estimation performance. They reported a modest decrement in the accuracy of ensemble estimation under focused attention compared with distributed attention. Furthermore, visual attention is captured to locations at which regularities are present, both in adults (Umemoto et al., 2010; Zhao et al., 2013) and in infants (Kidd et al., 2012), suggesting that the allocation of attention may indeed be contingent on the presence of regularities.

In summary, the results of the current study provide an alternative account for the finding that ensemble processing disrupts statistical learning. Rather than reflecting competition for a shared mechanism for statistical computations, our results suggest that ensemble processing and statistical learning can occur over the same arrays of stimuli, given appropriate task demands. Statistical learning is unimpaired when participants estimate ensemble statistics if attention is allocated to individual items within the stimulus arrays, and if individual items are encoded into working memory. Our findings add to existing evidence for interactions between attention and statistical learning (Turk-Browne et al., 2005; Zhao et al., 2013), suggesting that focal attention to individual items is necessary for statistical learning to occur. In addition, we suggest that maintenance of individual items in working memory is also important for statistical learning, in contrast to maintenance of global representations such as ensemble statistics. An interesting avenue for future research concerns characterising the role of precision of item processing and encoding for statistical learning.

References

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

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. doi: 10.1111/j.0963-7214.2004.01502006.x

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*(4), 392-398. doi: 10.1111/j.1467-9280.2008.02098.x

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157-162. doi: 10.1111/1467-9280.00327

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews, 36*(2), 737-746. doi: 10.1016/j.neubiorev.2011.12.003

Atchley, P., & Andersen, G. J. (1995). Discrimination of speed distributions: Sensitivity to statistical properties. *Vision Research, 35*(22), 3131-3144. doi: 10.1016/0042-6989(95)00057-7

Baddeley, A. (1992). Working memory. Science, 255(5044), 556-559. doi: 10.1126/science.1736359

Baddeley, A., & Hitch, G. (1974). Working memory. In H. B. Gordon (Ed.), Psychology of Learning and Motivation (Vol. Volume 8, pp. 47-89): Academic Press.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*(5890), 851-854. doi: 10.1126/science.1158023

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436. doi: 10.1163/156856897x00357

Broadbent, D. (1958). *Perception and Communication*: Pergamon Press.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*(4), 393-404. doi: 10.1016/s0042-6989(02)00596-5

Chong, S. C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics, 67*(1), 1-13. doi: 10.3758/bf03195009

Chong, S. C., & Treisman, A. (2005b). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*(7), 891-900. doi: 10.1016/j.visres.2004.10.004

Corbett, J. E., & Oriet, C. (2011). The whole is indeed more than the sum of its parts: Perceptual averaging in the absence of individual item representation. *Acta Psychologica, 138*(2), 289-301. doi: 10.1016/j.actpsy.2011.08.002

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-+. doi: 10.1017/s0140525x01003922

Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research, 37*(22), 3181-3192. doi: 10.1016/s0042-6989(97)00133-8

de Graaf, T. A., Hsieh, P.-J., & Sack, A. T. (2012). The 'correlates' in neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews, 36*(1), 191-197. doi: 10.1016/j.neubiorev.2011.05.012

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences, 21*(8), 355-361. doi: 10.1016/s0166-2236(98)01263-6

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review, 70*(1), 80.

Dutta, A., Shah, K., Silvanto, J., & Soto, D. (2014). Neural basis of non-conscious visual working memory. *NeuroImage, 91*, 336-343. doi: 10.1016/j.neuroimage.2014.01.016

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. *Journal of Neuroscience, 33*(15), 6516-6523. doi: 10.1523/jneurosci.5732-12.2013

Ester, E. F., Anderson, D. E., Serences, J. T., & Awh, E. (2013). A neural measure of precision in visual working memory. *Journal of Cognitive Neuroscience, 25*(5), 754-761. doi: 10.1162/jocn_a_00357

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*(6), 499-504. doi: 10.1111/1467-9280.00392

Fiser, J., & Aslin, R. N. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology-Learning Memory and Cognition, 28*(3), 458-467. doi: 10.1037//0278-7393.28.3.458

Fiser, J., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America, 99*(24), 15822-15826. doi: 10.1073/pnas.232472899

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications, 3*. doi: 10.1038/ncomms2237

Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision, 9*(11). doi: 10.1167/9.11.1

Hassin, R. R., Bargh, J. A., Engell, A. D., & McCulloch, K. C. (2009). Implicit working memory. *Consciousness and Cognition, 18*(3), 665-678. doi: 10.1016/j.concog.2009.04.003

Jacoby, O., Kamke, M. R., & Mattingley, J. B. (2013). Is the whole really more than the sum of its parts? Estimates of average size and orientation are susceptible to object substitution masking. *Journal of Experimental Psychology-Human Perception and Performance, 39*(1), 233-244. doi: 10.1037/a0028762

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *Plos One, 7*(5). doi: 10.1371/journal.pone.0036399

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters, 461*(2), 145-149. doi: 10.1016/j.neulet.2009.06.030

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35-B42. doi: 10.1016/s0010-0277(02)00004-5

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36*, 14-14.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281. doi: 10.1038/36846

Marshall, L., & Bays, P. M. (2013). Obligatory encoding of task-irrelevant features depletes working memory resources. *Journal of Vision, 13*(2). doi: 10.1167/13.2.21

Morgan, M., Chubb, C., & Solomon, J. A. (2008). A 'dipper' function for texture discrimination based on orientation variance. *Journal of Vision, 8*(11). doi: 10.1167/8.11.9

Musz, E., Weber, M. J., & Thompson-Schill, S. L. (2014). Visual statistical learning is not reliably modulated by selective attention to isolated events. *Attention, Perception, & Psychophysics*, 1-19. doi: 10.3758/s13414-014-0757-5

Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology, 49*(6), 1103-1112. doi: 10.1037/a0029472

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences of the United States of America, 105*(7), 2745-2750. doi: 10.1073/pnas.0708424105

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*(7), 739-744. doi: 10.1038/89532

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437-442. doi: 10.1163/156856897x00366

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Soto, D., Mantyla, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology, 21*(22), R912-R913.

Sweeny, T. D., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology-Human Perception and Performance, 39*(2), 329-337. doi: 10.1037/a0028712

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*(4), 792-814. doi: 10.1037/a0030801

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*(6984), 751-754. doi: 10.1038/nature02466

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology, 12*(1), 97-136.

Turk-Browne, N. B., Isola, P. J., Scholl, B. J., & Treat, T. A. (2008). Multidimensional visual statistical learning. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*(2), 399-407. doi: 10.1037/0278-7393.34.2.399

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology-General, 134*(4), 552-564. doi: 10.1037/0096-3445.134.4.552

Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: Transfer across space and time. *Journal of Experimental Psychology-Human Perception and Performance, 35*(1), 195-202. doi: 10.1037/0096-1523.35.1.195

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*(10), 1934-1945. doi: 10.1162/jocn.2009.21131

Umemoto, A., Scolari, M., Vogel, E. K., & Awh, E. (2010). Statistical learning induces discrete shifts in the allocation of working memory resources. *J Exp Psychol Hum Percept Perform, 36*(6), 1419-1429. doi: 10.1037/a0019324

Utochkin, I. S., & Tiurina, N. A. (2014). Parallel averaging of size is possible but range-limited: A reply to Marchant, Simons, and De Fockert. *Acta Psychologica, 146*, 7-18. doi: 10.1016/j.actpsy.2013.11.012

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 of the United States of America, 109*(22), 8780-8785. doi: 10.1073/pnas.1117465109

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10*(2), 130-137. doi: 10.1111/1467-9280.00120

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention - Voluntary versus automatic allocation. *Journal of Experimental Psychology-Human Perception and Performance, 16*(1), 121-134. doi: 10.1037//0096-1523.16.1.121

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233-U213. doi: 10.1038/nature06860

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological Science, 24*(5), 667-677. doi: 10.1177/0956797612460407

Zhao, J., Ngo, N., McKendrick, R., & Turk-Browne, N. B. (2011). Mutual interference between statistical summary perception and statistical learning. *Psychological Science, 22*(9), 1212-1219. doi: 10.1177/0956797611419304

Acknowledgements

This work was supported by the Australian Research Council Special Research Initiative - Science of Learning Research Centre (SR120300015) to PED and JBM. In addition, MGH was supported by an Australian Postgraduate Award, PED by an ARC Future Fellowship (FT120100033) and JBM by an ARC Australian Laureate Fellowship (FL110100103) and the ARC Centre of Excellence for Integrative Brain Function (ARC Centre Grant CE140100007). The authors thank Amy Taylor for assistance with data collection.

Figure 1. Example trials from the exposure phase and recognition test. During the exposure phase (top), participants viewed arrays of oriented lines and made either a duplicate judgement (‘Was an orientation repeated in the array?’) or ensemble judgement (‘Was the average orientation of the lines in the array leftward or rightward?’) over each array. The arrays were constructed from pairs of orientations. The two orientations making up a given pair always appeared together in the same spatial configuration. An example of a pair within an array is highlighted to the right (dark grey line segments in a diagonal spatial configuration). During the recognition test (bottom), participants indicated which of two pairs (the pair on the left or the pair on the right) was more familiar. One pair had appeared reliably during the exposure phase, and the other was a foil pair that had not appeared reliably. The correct pair for each example trial is highlighted to the right.

Figure 2. Trial sequences and results for the exposure phase and recognition test in Experiment 1. During the exposure phase (top), line segments within the stimulus arrays appeared simultaneously. Participants in both the duplicate group and the ensemble group performed above chance. During the recognition test (bottom), a correct pair and a foil pair appeared simultaneously on each trial, one on the left of the screen and the other on the right of the screen. Participants in the duplicate group performed significantly above chance on the recognition test, demonstrating statistical learning, whereas participants in the ensemble group performed at chance.

Figure 3. Trial sequences and results for the exposure phase and recognition test for Experiment 2. During the exposure phase (top), the line segments of each stimulus array appeared simultaneously, followed by a brief mask, and finally a probe item. In addition to their primary task (duplicate or ensemble judgement), all participants completed a secondary change detection task. Participants reported whether the probe item was the same orientation or a different orientation than the line appearing in the same location in the initial stimulus array of that trial. Participants in both groups performed above chance on the primary task and on the change detection task. On each trial of the recognition test (bottom), the correct pair and foil pair appeared simultaneously, one on either side of fixation. Participants in both the duplicate group and the ensemble group performed significantly above chance on the recognition test, demonstrating statistical learning.

Figure 4. Trial sequences and results for the exposure phase and recognition test for Experiment 3. During the exposure phase (top), the line segments of each stimulus array appeared sequentially. Participants in both the duplicate group and the ensemble group performed above chance. During the recognition test (bottom), the line segments constituting each pair also appeared sequentially - On each

trial, one of the pairs (correct or foil pair) appeared sequentially, followed by a pause, before the second pair. Participants in the duplicate group performed significantly above chance on the recognition test, demonstrating statistical learning, whereas participants in the ensemble group performed at chance.

Figure 1

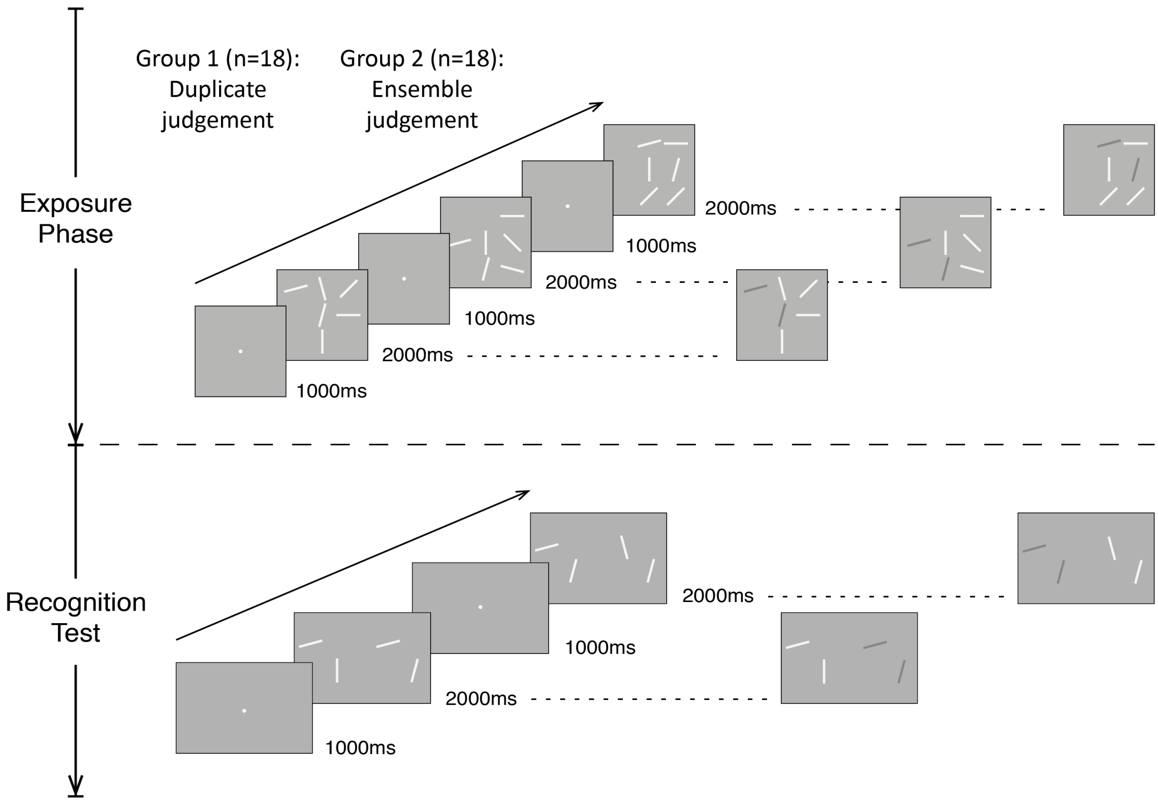

Group 1 (n=18):
Duplicate
judgement
Group 2 (n=18):
Ensemble
judgement
Exposure
Phase
1000ms
2000ms
1000ms
2000ms
1000ms
2000ms
Recognition
Test
1000ms
2000ms
1000ms
2000ms

Figure 2

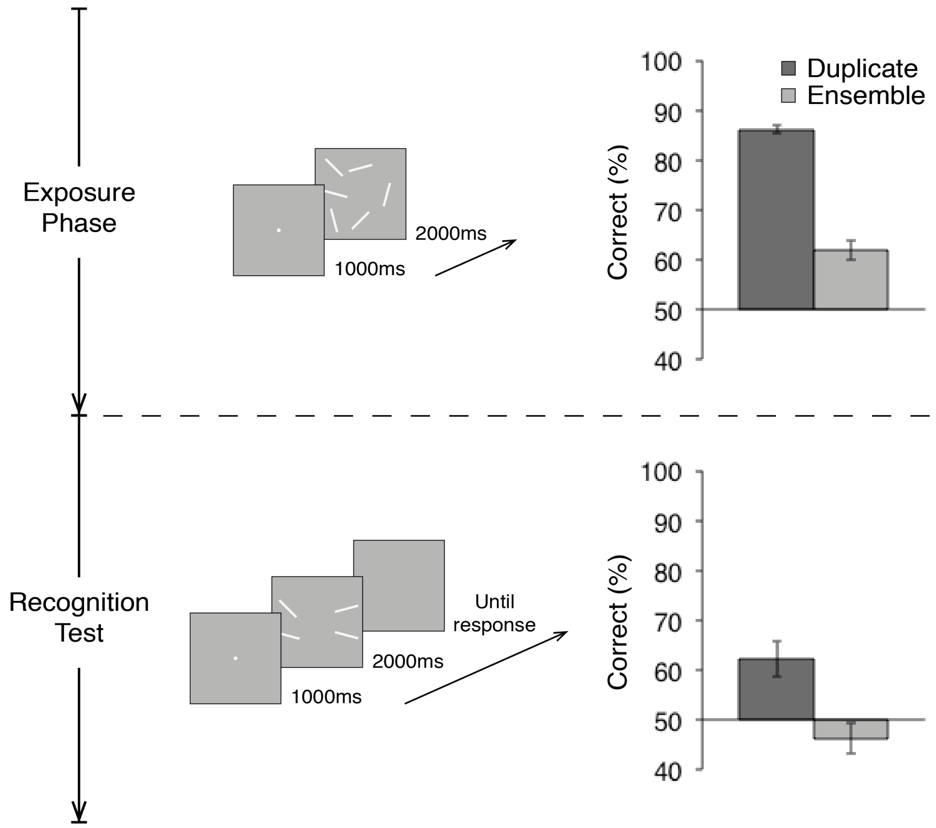

Exposure
Phase
1000ms
2000ms
100
90
80
70
60
50
40
Correct (%)
Duplicate
Ensemble
Recognition
Test
1000ms
2000ms
Until
response
100
90
80
70
60
50
40
Correct (%)

Figure 3

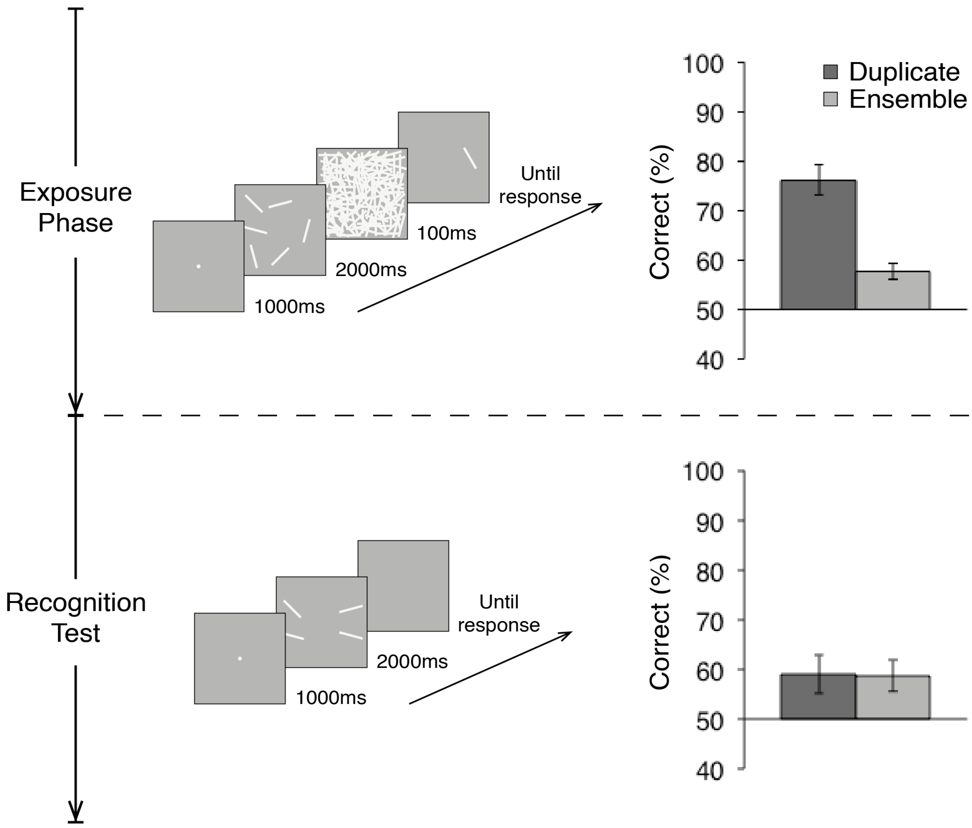
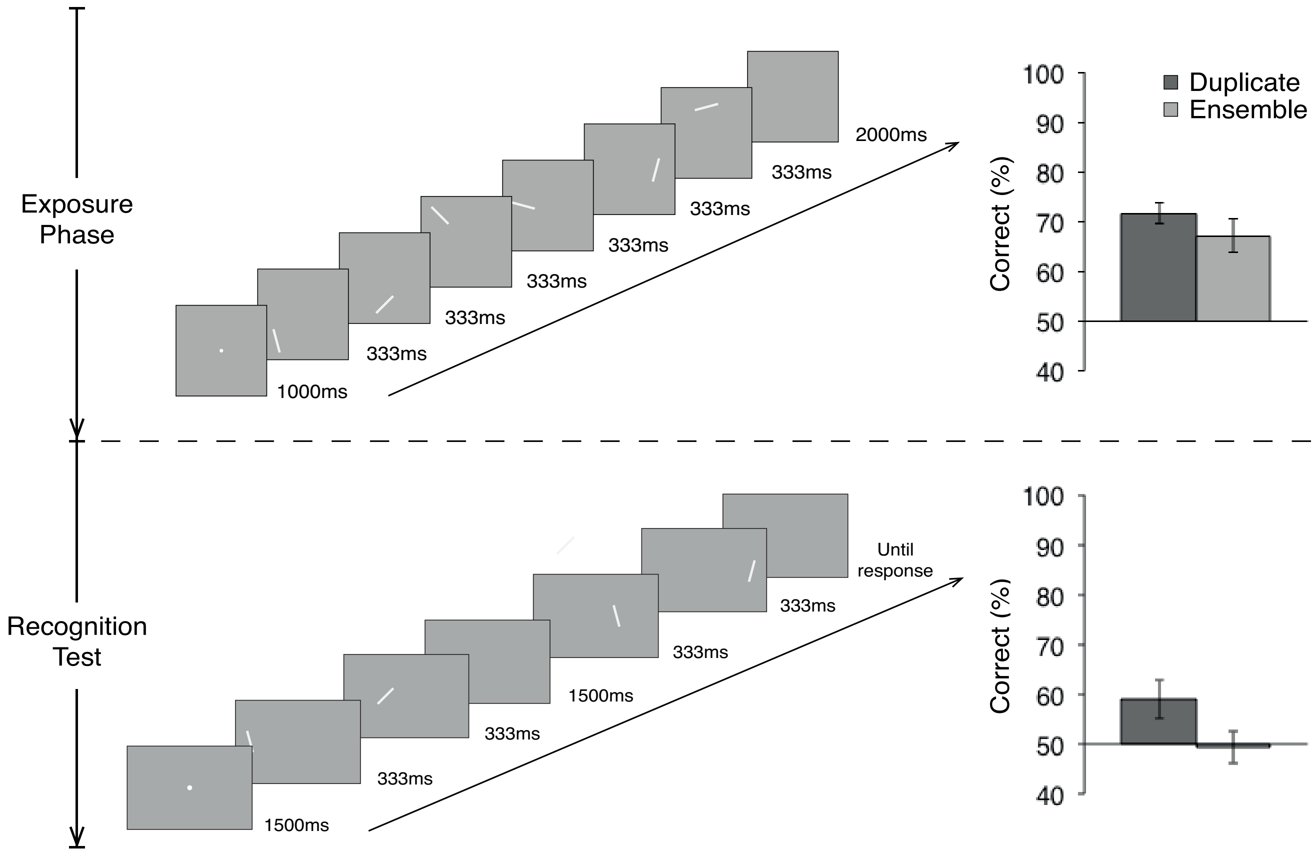

Exposure Phase
1000ms
2000ms
100ms
Until response
Recognition Test
1000ms
2000ms
Until response
Correct (%)
100
90
80
70
60
50
40
Duplicate
Ensemble

Figure 4


Exposure
Phase
1000ms
333ms
333ms
333ms
333ms
333ms
333ms
2000ms
Recognition
Test
1500ms
333ms
333ms
1500ms
333ms
333ms
Until
response
Correct (%)
100
90
80
70
60
50
40
Duplicate
Ensemble