Consciousness and Cognition 48 (2017) 149–160

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Perception of ensemble statistics requires attention

Molly Jackson-Nielsen [a], Michael A. Cohen [b], Michael A. Pitts [a,*]

[a] Department of Psychology, Reed College, United States
[b] Department of Brain and Cognitive Sciences, McGovern Institute for Brain Research, Massachusetts Institute of Technology, United States

ARTICLE INFO

Article history:
Received 7 September 2016
Revised 12 November 2016
Accepted 14 November 2016

Keywords:
Inattentional blindness
Visual ensembles
Summary statistics
Color diversity
Conscious perception
Dual-task

ABSTRACT

To overcome inherent limitations in perceptual bandwidth, many aspects of the visual world are represented as summary statistics (e.g., average size, orientation, or density of objects). Here, we investigated the relationship between summary (ensemble) statistics and visual attention. Recently, it was claimed that one ensemble statistic in particular, color diversity, can be perceived without focal attention. However, a broader debate exists over the attentional requirements of conscious perception, and it is possible that some form of attention is necessary for ensemble perception. To test this idea, we employed a modified inattentional blindness paradigm and found that multiple types of summary statistics (color and size) often go unnoticed without attention. In addition, we found attentional costs in dual-task situations, further implicating a role for attention in statistical perception. Overall, we conclude that while visual ensembles may be processed efficiently, some amount of attention is necessary for conscious perception of ensemble statistics.

1. Introduction

What is the relationship between attention and consciousness? Is our moment-to-moment awareness limited to the few things we can attend to, or are certain aspects of the world experienced outside of attention? These questions are central to a broader debate about whether visual perception is relatively "rich" or "sparse" (Block, 2011; Cohen & Dennett, 2011; Dehaene, 2014; Gross & Flombaum, in press; Lamme, 2010; Phillips, 2016; Ward, Bear, & Scholl, 2016).

Recently, a potential solution to this debate was proposed that attempts to reconcile our subjective impressions of a rich phenomenal world with empirical observations that suggest a severely limited perceptual bandwidth (Cohen, Dennett, & Kanwisher, 2016). According to this proposal, our perception is neither as rich as is intuitively believed, nor as sparse as dominant models of visual attention and working memory suggest. The key insight is that in addition to the few items that can be focally attended, the visual system represents large swaths of the visual world as "ensemble representations". These ensembles, or "summary statistics", can be represented across a wide range of visual dimensions including average orientation, motion direction, speed, size, position, density, facial expression, etc. (Whitney, Haberman, & Sweeny, 2014). Ensembles are encoded efficiently (Alvarez & Oliva, 2009), unconsciously (Moore & Egeth, 1997), and are likely processed by dedicated neural mechanisms in parallel to the detailed processing of individual items (Cant & Xu, 2012).

What remains unclear, however, is whether our *conscious experience* of visual ensembles requires attention. In a recent study, Bronfman, Brezis, Jacobson and Usher (2014) asked if one particular summary statistic — color diversity — could

* Corresponding author at: Department of Psychology, Reed College, 3203 SE Woodstock Blvd, Portland, OR 97202, United States.
E-mail address: mpitts@reed.edu (M.A. Pitts).

http://dx.doi.org/10.1016/j.concog.2016.11.007
1053-8100/

be perceived without focal attention. Using a variation of the Sperling paradigm (Sperling, 1960) in which the letters either had low color diversity (i.e., sampled from few adjacent regions on the color wheel) or high color diversity (i.e., sampled from the entire color wheel), they pre-cued one row of letters and measured how many letters could be reported during a single task (report letters only) versus a dual task (report letters and color diversity). Surprisingly, they found that there was no decrease in performance on the letter task when participants had to also report the color diversity of the non-cued letters, and that color diversity judgments for the non-cued letters were significantly above chance. These results have been interpreted as evidence for "cost free" awareness of ensemble statistics that does not rely on cognitive functions such as attention and working memory (Block, 2014).

Recently, Ward et al. (2016) replicated Bronfman et al.'s (2014) main results but challenged some of their interpretations. In a series of experiments, Ward et al. (2016) tested whether awareness of color diversity was possible despite a lack of awareness of the individual (differentiated) colors. Bronfman et al. posited that color diversity statistics could not be perceived without a "differentiated representation" of the individual colored elements. They further argued that this differentiated representation must have been briefly experienced. Ward et al.'s (2016) results, however, demonstrated that accurate color diversity judgments can persist even when participants fail to notice that the color of every single element changed halfway through the trial. By swapping the colors of individual letters while holding the color diversity statistic constant, Ward et al. (2016) found that performance on the statistical judgments was above chance despite robust change blindness for the individual colors. These results suggest that we can consciously perceive summary statistics that are derived from unconscious differentiated representations.

While Ward et al.'s (2016) study challenged some of the interpretations of Bronfman et al.'s (2014) results, participants in both studies were always partially attending to the colors in the non-cued rows because color diversity judgments were part of the task and were required on every trial. Thus, even though focal attention was allocated to the cued row of letters, it is likely that attention was divided, or perhaps that non-focal, diffuse attention was allocated to the stimuli in the non-cued rows, in order to carry-out the secondary color diversity judgment task. Therefore, the question of whether color diversity statistics, or any summary statistics, can be consciously perceived without (any type of) attention remains open. While it is obviously difficult to create a situation in which zero attention is allocated to a stimulus, one of the best methods to approximate such a situation is with inattentional blindness (Mack, 2003; Mack & Rock, 1998).

The current study asked if ensemble perception is possible without attention by modifying the Bronfman et al. task into an inattentional blindness paradigm. We hypothesized that Bronfman et al.'s and Ward et al.'s findings of above chance color diversity judgments may be explained by participants distributing their attention across the two tasks. With inattentional blindness paradigms, unlike well-practiced dual-tasks, participants do not have the motivation or opportunity to distribute their attention across tasks.

In a series of four experiments, we tested whether participants spontaneously experience ensemble statistics (color diversity, as well as size diversity and mean size) despite having no reason to attend to these ensembles. In essence, this set-up approximates many everyday situations in which we feel like we experience a rich phenomenal world without having a specific task or reason to attend to all aspects of the visual scene. If some form of attention is necessary for awareness of ensemble statistics, we should find robust inattentional blindness rates for color diversity and other summary statistics. Alternatively, if task-irrelevant ensemble statistics are consciously perceived "for free" (without attention), we should observe above-chance noticing rates for these statistics.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Participants were 50 Reed College students, all over the age of 18 with normal or corrected-to-normal vision. All were recruited volunteers and gave informed consent prior to beginning the experiment. Experimental procedures were approved by the Reed Institutional Review Board. Sample size was selected based on previous inattentional blindness experiments with similar designs (Cohen, Alvarez, & Nakayama, 2011; Mack & Clarke, 2012).

2.1.2. Apparatus and stimuli

Stimuli were created in Presentation 17.0 (Neurobehavioral Systems, Berkeley, CA) and presented on a Dell 24″ LCD monitor with a resolution of 1920 × 1200 pixels and a screen refresh rate of 60 Hz. Screen viewing distance was approximately 41 cm. All stimuli were presented on a black background. All trials included a 6 × 4 array of capital letters, in Courier New Bold font, which were sampled randomly from the nine letters R, T, F, N, B, P, L, M, and K. Arrays were approximately 9 cm wide and 6 cm tall. In every trial, letters in one of the four rows were white with font size 50 pt. Colors and sizes of letters in the other three rows varied throughout the experiment (see procedure below). In "high color diversity" trials (i.e., high color variance), the colors of non-cued-row letters were sampled randomly from 19 colors selected from around the color wheel (see Bronfman et al., 2014, for RGB values). In "low color diversity" trials (i.e., low color variance), colors were sampled from a randomly selected set of six adjacent colors. In "high size diversity" trials, non-cued-row letter sizes were sampled randomly from font sizes 40–60 pt. In "low size diversity" trials, all letters were font size 50 pt (i.e., zero size variance). Example stimuli are depicted in Fig. 1. Before an array was presented, a white cue box appeared at the location of the upcoming white row of letters. Our rationale for rendering the cued row of letters white was based on (1) providing an additional attentional cue that would help participants focus and maintain their attention only on the cued row, and (2) improving the accuracy of the surprise questions on the critical trials (see procedure below), i.e. to better assess what participants perceived in the non-cued rows without contamination from the colors in the cued-row. Besides making the cued-row letters white and manipulating the size diversity of the letters in the non-cued rows, all other stimulus parameters were exactly the same as in Bronfman et al. (2014).

2.2. Procedure

Participants were told they would be performing a "visual working memory task". They were instructed to pay attention to the white letters in each trial, which would be cued by a white rectangle before the letters appeared, and to try to remember as many letters as they could. No performance feedback was provided. On every trial, a fixation cross was presented for 200 ms, followed by a cue box for 100 ms, then a letter array for 300 ms, then a blank screen for 900 ms, and then response prompts that varied throughout the experiment. On each trial, the position of the cued-row was selected randomly with three restrictions: (1) each letter row was cued an equivalent number of times, (2) on the 8th trial in each phase (the "critical trial"), the row second-from-top was always cued; and (3) in the seventh trial, the cued row was never second-from-top. The latter two restrictions ensured consistency across subjects in terms of where their attention was focused on the critical trial and the trial immediately preceding the critical trial.

In the first phase of the experiment, participants performed a single task (cued-row letter report) for four practice trials and seven experimental trials. For these 11 trials, letters in the non-cued rows always had high color diversity. At the end of each trial, a question mark appeared prompting participants to type all the letters from the cued row that they could remember. On the 8th experimental trial (the critical trial), the non-cued-row letters changed to low color diversity. After the 900 ms blank screen, instead of a question mark prompting a letter report, participants were immediately presented with a 3-alternative forced choice (3-AFC) recognition test. A screen appeared containing three side-by-side letter arrays. The arrays' left-to-right order on the screen was determined randomly. The arrays were numbered 1–3 from left to right and above them was the question, "Which picture looks most like the one you just saw? Type 1, 2, or 3." In all three options, the cued row was made of white X's to ensure that participants would make their choice based on the properties of the non-cued rows. In the correct array, the non-cued-row colors were sampled from the same 6 colors as those in the just-presented low color diversity array. In the incorrect arrays, the non-cued-rows had high color diversity. In one of the incorrect arrays (the foil), the non-cued row letters were flipped upside down. The purpose of the foil was to catch guesses of participants who did not perceive the color diversity change, but may have reasoned that something probably changed, as this was a psychology experiment with a surprise question.

It is important to note that the 3-AFC technique we used to test for inattentional blindness differs from previous studies (e.g., Cohen et al., 2011; Mack & Clarke, 2012). Typically, after the critical trial, participants are asked a series of open ended questions regarding what they noticed, and in some cases a recognition test is administered after the open questioning. Here, we wanted to give participants every possible chance to report what they had seen, assuming they had perceived the color diversity statistic without attention. This procedure inevitably overestimates noticing rates for two reasons. First, participants who did not see the color diversity change could still guess correctly due to the forced choice nature of the test and the limited number of options. Second, if the color diversity statistic was unconsciously processed, this unconscious trace could influence participants' decisions on the 3-AFC, even if the statistic was never consciously perceived. Our strategy

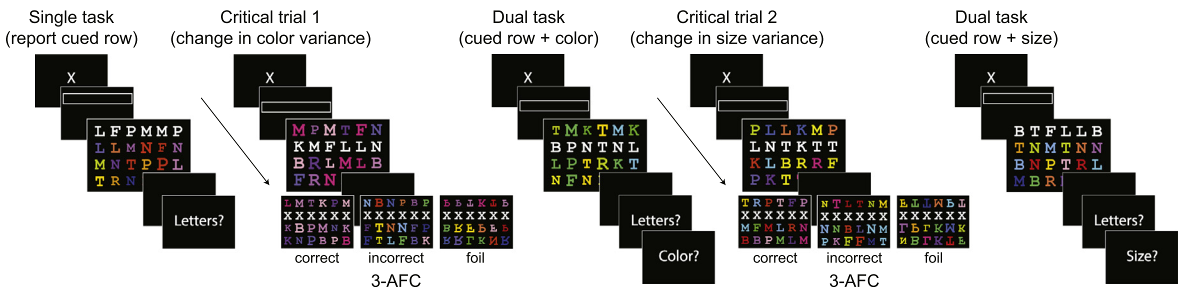

Fig. 1. Experiment 1: example stimuli and procedure. On the first 7 trials (single task: cued-row letter report), color diversity of the non-cued rows remained high. On the 8th, critical trial, color diversity changed to low and participants were immediately asked, "which display looks most like what you just saw?" from 3 alternatives (correct low color variance; incorrect high variance; incorrect high variance "foil" with upside-down letters). Participants who selected either of the incorrect options were considered inattentionally blind. Participants then performed a dual-task for 7 trials in which they reported the cued-row letters and the color diversity of the non-cued rows (high vs. low). On the 8th dual-task trial, size variance of the non-cued letters changed from high to low, and participants were immediately prompted with another 3-alternative forced choice. Finally, participants carried-out a second dual-task for 7 trials in which they reported the cued row letters and the size variance of the non-cued rows (high vs. low). See main text for additional details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

was thus to stack the cards *against* finding robust inattentional blindness rates by making the reporting procedure as inclusive as possible and simply accept the artificial inflation of noticing rates.

After the 3-AFC following the critical trial, participants were informed that they would enter the second phase of the experiment in which they would perform a dual-task. The dual-task involved the same cued row letter report as before, followed by a color diversity report (high or low). Participants were shown eight color diversity example arrays, four high and four low, and then given four practice trials on the dual-task, two each of high and low diversity. They then completed seven experimental trials, four with low color diversity and three with high color diversity, presented in random order.

There were three main purposes to this second phase. First, it was important to replicate Bronfman et al.'s (2014) and Ward et al.'s (2016) findings that participants could perform the color diversity task at above chance levels even while doing the primary letter report task. Second, we wanted to test whether color diversity is indeed "cost free" by comparing letter report accuracy between this dual-task and the single task of the first phase. Third, we used this dual-task to distract participants' attention from another ensemble statistic (size diversity) that we planned to unexpectedly change during the second critical trial.

Like all trials in the first phase, the non-cued-row letters in the second phase had high size diversity. On the 8th trial of the second phase (the second critical trial), the non-cued-row letters changed to low size diversity. Participants were again asked to perform a 3-AFC immediately after the critical trial. In the correct array, all letters were low size diversity (all 50 pt font) as they had been on the critical trial. In the incorrect arrays, the non-cued row letters were high size diversity, with the foil array containing upside down letters. Color diversity was always high on the critical trial as well as on the 3-AFC.

After the second critical trial, participants entered the third and final phase of the experiment. In this third phase, participants performed another dual-task, but in this case the secondary task was to report size diversity instead of color diversity. The purposes of this third phase were to confirm that participants could perceive the size diversity statistic at above chance levels (when attended), and to test for an attentional cost on the primary letter report task. Before the third phase began, participants were shown eight examples of high and low size diversity arrays, four of each, and then given four practice trials, two each of high and low. They then completed seven trials of this dual-task. There were four low size diversity trials and three high size diversity trials, presented in random order. Color diversity of non-cued-row letters was always high. Fig. 1 provides an overview of the procedure.

2.3. Results and discussion

Inattentional blindness rates are plotted in Fig. 2a. On the 3-AFC following the first critical trial, 78% of participants were deemed "inattentionally blind" to the color diversity statistic, i.e. these participants selected one of the incorrect options (58% chose high color diversity, 20% chose the foil). Only 22% of the participants "noticed" the color diversity change, i.e. they chose the correct low color diversity display. On the 3-AFC following the second critical trial, 58% of participants were inattentionally blind to the size diversity statistic (40% chose high size diversity, 18% chose the foil), while 42% noticed the size diversity change (indicated by their choosing the correct low size diversity display).

Because only one of the three options was correct on the surprise 3-AFC recognition tests, we also compared participant's responses against chance (33%). However, it is important to note that the three options in the 3-AFC were not balanced. Two of the options contained the same (incorrect) summary statistic, one of which served as the "lure" (upside-down letters) intended to catch participants who might be thinking too much about possible "tricks" in the experimental design. Therefore, slightly above-chance recognition may not be the best indicator of successful "noticing" of the ensemble statistics. It might be more appropriate to lump together the two incorrect options and consider chance performance to be 50%. Nevertheless, for completeness, we report comparisons against a chance level of 33%. In this first experiment, binomial tests revealed that the proportion of participants correctly reporting color diversity was significantly *below* chance (22%, $p = 0.03$), while the proportion of participants correctly reporting size diversity was slightly above chance (42%, $p = 0.05$).

Performance on the letter report task decreased in the dual-tasks compared to the single task (Fig. 2b). A one-way repeated measures ANOVA showed an effect of phase on letter accuracy, $F(2, 98) = 19.03$, $p < 0.01$, $\eta\rho^2 = 0.28$. Dependent

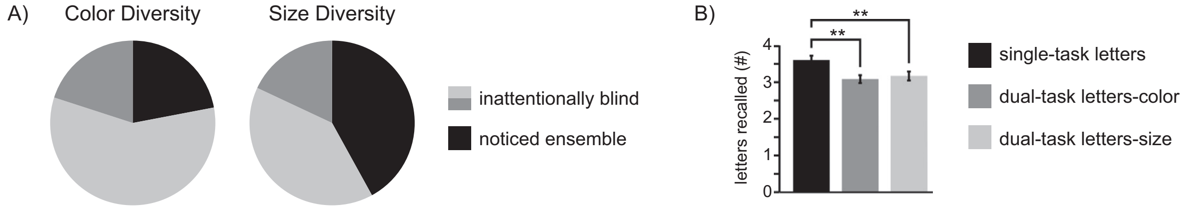

Fig. 2. Experiment 1 results. (A) Inattentional blindness rates from critical trial 1 (color diversity) and critical trial 2 (size diversity). Black regions represent the percent of participants who selected the correct (low diversity) ensemble display on the 3-AFC that immediately followed the critical trials. Participants who selected one of the incorrect (high diversity) options were considered inattentionally blind (light gray = normal letters; dark gray = upside-down foil). A majority of participants (>50%) were inattentionally blind to the color and size ensemble statistics. (B) Mean number of cued-row letters correctly reported as a function of task. An attentional cost to letter recall performance was observed for the dual-tasks (dark gray = letter-color task; light gray = letter-size task) compared to the single-task (black = letter task), $^{**}p < 0.01$.

means *t*-tests revealed that letter performance in the single task ($M = 3.60$, $SD = 0.79$) was significantly higher than that in dual-task with color ($M = 3.08$, $SD = 0.77$), $t(49) = 6.12$, $p < 0.01$, $\eta\rho^2 = 0.44$, and the dual-task with size ($M = 3.17$, $SD = 0.85$), $t(49) = 4.25$, $p < 0.01$, $\eta\rho^2 = 0.27$. No difference was found between the two dual-tasks, $t(49) = 1.04$, $p > 0.05$, $\eta\rho^2 = 0.02$.

Color diversity performance in the phase 2 dual-task ($M = 83.7\%$, $SD = 15.0$) was significantly above chance, $t(49) = 15.89$, $p < 0.01$, *Cohen's d* = 2.25. Size diversity performance in the phase 3 dual-task was also significantly above chance ($M = 67.4\%$, $SD = 18.3$), $t(49) = 6.75$, $p < 0.01$ *Cohen's d* = 0.95.

To summarize, experiment 1 revealed robust inattentional blindness for color and size ensemble statistics. More than 50% of participants failed to spontaneously notice these statistics on the critical trials, while they did so with ease during the subsequent partial attention (dual-task) conditions. In addition, attentional costs to letter recall were evident when participants switched from the single-task (letters-only) to the dual-tasks (letters-color and letters-size). These results suggest that ensemble perception requires attention.

3. Experiment 2

In our first experiment, the color and size ensemble statistics were always high-variance on the lead-up trials and low-variance on the critical trials. It is possible that this type of change (i.e., highly diverse colors and sizes becoming more uniform) is easier to miss than the inverse. In the next experiment, we swapped these statistics such that color and size were low in variance on the lead-up trials and high in variance on the critical trials.

3.1. Methods

3.1.1. Participants

A new set of 25 Reed College students was recruited for experiment 2. One participant was excluded due to a failure to follow instructions (final N = 24).

3.1.2. Procedure

Apparatus, stimuli, and procedure were almost identical to experiment 1. The only differences were that the color and size diversity changes were inverted relative to experiment 1. In other words, color and size diversities were low during the lead-up trials, and switched to high on the critical trials. Color was still tested first and size second. On the 3-AFC following the first critical trial, the two incorrect arrays (low color diversity) were sampled from two sets of 6 adjacent colors at opposite ends of the color wheel. Size diversity was low for the entire experiment up until the second critical trial in which it switched to high size diversity.

3.2. Results and discussion

Inattentional blindness rates are shown in Fig. 3a. On the 3-AFC following the first critical trial, 54% of subjects were inattentionally blind to the color diversity ensemble (33% chose low color diversity, 21% chose the foil), while 46% selected the correct high color diversity display. On the 3-AFC following the second critical trial, 79% of subjects selected one of the incorrect options (58% chose low size diversity, 21% chose the foil), while only 21% chose the correct high size diversity display. Again, keeping in mind the caveats mentioned above about whether 33% should be considered "chance" in the current paradigm, binomial tests suggested that the proportion of correct responses for the color and size diversity statistics were not statistically different than chance (color: 46%, p = 0.07; size: 21%, p = 0.07).

As in experiment 1, a one-way repeated measures ANOVA showed an effect of phase on letter performance, $F(2, 46) = 5.68$, $p < 0.01$, $\eta\rho^2 = 0.20$. Dependent means *t*-tests revealed that letter performance in the single task ($M = 3.71$, $SD = 0.92$) was significantly higher than that in the dual-task with color ($M = 3.34$, $SD = 0.88$), $t(23) = 3.28$, $p < 0.01$,

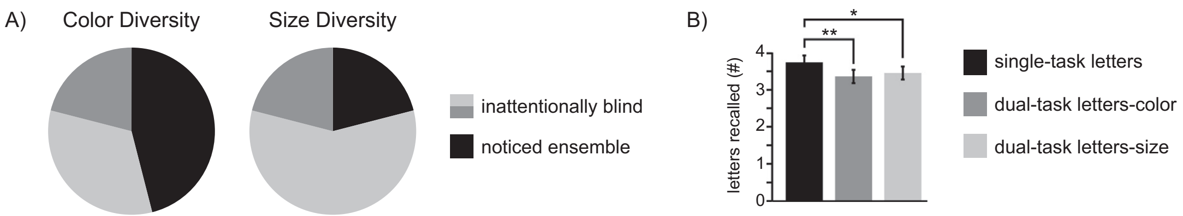

Fig. 3. Experiment 2 results. (A) Inattentional blindness rates from critical trial 1 (color diversity) and critical trial 2 (size diversity). Black regions represent the percent of participants who selected the correct (high diversity) ensemble display on the 3-AFC. Participants who selected an incorrect (low diversity) option were considered inattentionally blind (light gray = normal letters; dark gray = upside-down foil). As in experiment 1, a majority of participants (>50%) were inattentionally blind to the color and size ensemble statistics. (B) Mean number of cued-row letters correctly reported as a function of task. Once again, an attentional cost to letter recall performance was observed for the dual-tasks (dark gray = letter-color task; light gray = letter-size task) compared to the single-task (black = letter task), $^{*}p < 0.05$, $^{**}p < 0.01$.

$\eta\rho^2 = 0.32$, and the dual-task with size ($M = 3.43$, $SD = 0.87$), $t(23) = 2.88$, $p < 0.05$, $\eta\rho^2 = 0.27$. No difference was found between dual-tasks, $t(23) = 0.67$, $p > 0.05$, $\eta\rho^2 = 0.02$. Fig. 3b summarizes these results.

Also consistent with experiment 1, in the dual-task conditions color diversity judgments ($M = 83.3\%$, $SD = 16.7$) were well above chance, $t(23) = 9.79$, $p < 0.01$, *Cohen's d* = 2.00, as were size diversity judgments ($M = 70.8\%$, $SD = 19.5$), $t(23) = 5.23$, $p < 0.01$, *Cohen's d* = 1.07.

While spontaneous noticing rates of the color diversity statistic increased in experiment 2 compared to experiment 1, robust inattentional blindness was still observed. This was the case even though the colors changed from being relatively uniform in the lead-up trials to highly diverse on the critical trial. Interestingly, the reverse trend was observed for the size statistic, i.e., noticing rates decreased in experiment 2 even though the size of the letters changed from uniform on the lead-up trials to diverse on the critical trial. Attentional costs to the letter recall task were also evident in experiment 2. Thus, both of the main findings from experiment 1 replicated in experiment 2.

4. Experiment 3

In the first two experiments, the color and size ensemble statistics were held constant on the lead-up trials and changed on the critical trials. When the surprise question was posed, it is possible that participants responded based on accumulated information from the 7 lead-up trials instead of the critical trial itself. To control for this, in the next experiment, we varied the ensemble statistics on the lead-up trials (high and low were presented on an equal number of trials, including the critical trial). Additionally, because mean size is a well-established summary statistic (compared to size diversity, e.g. see Haberman & Whitney, 2012), we tested for awareness of this statistic instead of size diversity in experiment 3.

4.1. Methods

4.1.1. Participants

A new set of 30 Reed College students was recruited for experiment 3.

4.1.2. Procedure

Apparatus, stimuli, and procedure were nearly identical to experiment 1 & 2, except for the following modifications. Instead of holding constant the color and size ensemble statistics and switching them on the critical trial, the statistics varied between high and low on the lead-up trials and the particular statistic (high or low) presented on the critical trial was counterbalanced across participants. For half of the participants (N = 15), the non-cued letters were of high color diversity on the critical trial, with four of the lead-up trials having low color diversity and three having high color diversity. For the other half of participants (N = 15), the non-cued letters were low in color diversity on the critical trial, with four of the lead-up trials having high color diversity and three having low color diversity.

In high mean size displays, the letters of the non-cued rows were pseudo-randomly drawn at font sizes between 50 and 60 pts (ensuring an average of 55 pts). In low mean size displays, non-cued row letters varied between 40 and 50 pt font (averaging 45 pts). As in experiments 1 & 2, the cued-row letters were always sized at 50 pt font throughout all phases of the experiment. Just as we randomized high and low color diversity on the lead-up trials, mean size varied during the first two phases (such that an equal number of trials contained high and low mean size). On the second critical trial, half of the participants were presented a low mean size display and half were presented a high mean size display.

Consistent with experiments 1 & 2, on the 3-AFC for the critical trials, the incorrect and foil displays matched in terms of summary statistics. On the first 3-AFC, in order to avoid subjects selecting a display based on mean size instead of color diversity, all non-cued row letters were drawn in "medium" mean size (range 45–55 pt.). On the second critical trial and the second 3-AFC, all displays were always of high color diversity to ensure testing of the mean size (rather than color diversity) statistic.

4.2. Results and discussion

Fig. 4a presents the inattentional blindness rates for experiment 3. On the 3-AFC of the first critical trial, 53% of participants were inattentionally blind to the color diversity statistic (37% chose the incorrect color diversity display; 16% chose the foil), while 47% chose the correct (high or low) color diversity option. This rate of noticing was slightly above chance ($p = 0.046$). The participants who were presented the high color diversity display on the critical trial were more likely to select the correct display on the 3-AFC compared to those who received the low color diversity display (80% vs. 14%). For the second critical trial, 67% of participants were inattentionally blind to the mean size statistic (53% selected the incorrect mean size; 14% selected the foil), while 33% chose the correct (high or low) mean size display (not different than chance, $p = 0.15$). Here, there was no difference based on the type of display (high or low) presented on the critical trial (33% chose the correct mean size option in both cases).

A one-way repeated measures ANOVA showed a significant effect of phase on letter performance, $F(2, 58) = 7.18$, $p < 0.01$. Dependent means *t*-tests revealed that letter performance in the single-task ($M = 3.64$, $SD = 1.07$) was significantly higher than that in the dual-task with color ($M = 3.17$, $SD = 1.12$), $t(29) = 2.47$, $p < 0.01$. Letter performance in the dual-task with size

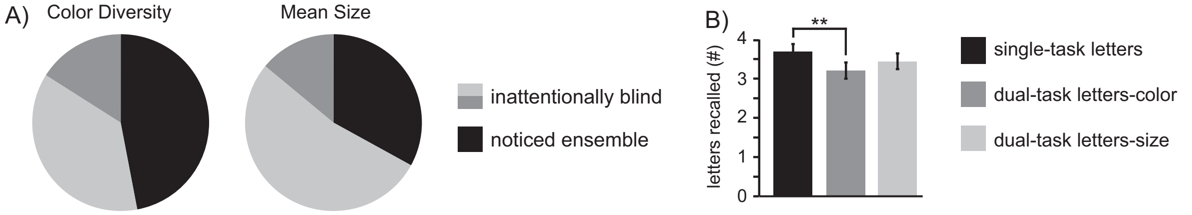

Fig. 4. Experiment 3 results. (A) Inattentional blindness rates from critical trial 1 (color diversity) and critical trial 2 (mean size). Black regions represent the percent of participants who selected the correct ensemble display on the 3-AFC. Participants who selected an incorrect option were considered inattentionally blind (light gray = normal letters; dark gray = upside-down foil). As in experiments 1–2, a majority of participants (>50%) were inattentionally blind to the color and size ensemble statistics. (B) Mean number of cued-row letters correctly reported as a function of task. An attentional cost to letter recall was observed for the letter-color dual-task (dark gray) but not the letter-size dual-task (light gray), $^{**}p < 0.01$.

($M = 3.28$, $SD = 1.36$) was marginally higher than that in the dual-task with color, $t(29) = 2.09$, $p = 0.05$, $\eta\rho^2 = 0.21$. No difference was found between the single-task and the dual-task with size, $t(29) = 1.71$, $\eta\rho^2 = 0.09$. These results are summarized in Fig. 4b.

In the dual-tasks, color diversity judgments were significantly above chance ($M = 81.9\%$, $SD = 14.5$), $t(29) = 12.05$, $p < 0.01$, *Cohen's d* = 2.20, as were mean size judgments ($M = 62.9\%$, $SD = 17.8$), $t(29) = 3.95$, $p < 0.01$, *Cohen's d* = 0.72.

Once again, substantial inattentional blindness rates (>50%) were found for both color and size ensemble statistics. By varying the statistics on the lead-up trials such that high and low variance (or means) were presented an equal number of times, this experiment allowed us to rule out an explanation based on accumulated evidence biasing responses on the 3-AFC. For the color diversity statistic, we did observe a bias to select the high-variance option on the 3-AFC (overall, 83% of participants chose high diversity and 17% chose low diversity). We revisit this result in the general discussion. The finding of attentional costs on the letter task replicated for one of the dual-tasks (letters-color), while no significant costs were observed for the other dual-task (letters-size).

5. Experiment 4

Experiments 1–3 showed robust inattentional blindness to color and size ensemble statistics, however, one could argue that these results are due to a lack of training or knowledge. In other words, at the time we asked the surprise questions, subjects had not yet learned what the statistics were or how to condense this information into binary judgments (high versus low diversity). Indeed, Bronfman et al. (2014) emphasize that reportability of the color diversity statistic depends on one's ability to condense complex color information into a simple binary statistic that can more easily persist in memory alongside the letter identities from the focal task. The 3-AFC method we used to assess inattentional blindness in experiments 1–3 was designed to alleviate this concern, i.e. this recognition test only required selection of the display that "looked most similar" to the display on the critical trial, without the need to convert this percept into a "high" versus "low" diversity judgment.

Nevertheless, a more direct way to address this concern is to first train participants to discriminate and report the summary statistics, and then on subsequent trials, divert their attention with the letter recall task and ask the surprise questions about the statistics. If robust inattentional blindness for the color and size ensembles is still observed even after subjects are trained and tested on the very same statistics a dozen or so trials earlier, the results from experiments 1–3 cannot be explained by a lack of knowledge or training in translating the color or size information into reports about the summary statistics.

5.1. Methods

5.1.1. Participants

A new set of 30 Reed College students was recruited for experiment 4.

5.1.2. Procedure

Apparatus and stimuli were identical to experiment 3, but the procedure was modified to address the concern outlined above. Two phases were added to the beginning of the procedure in which participants carried out single-tasks on color diversity and mean size, respectively. The very first displays presented to participants were the eight examples of high and low color diversity stimuli. They then performed a single-task for 7 trials in which they reported the color diversity of the non-cued letters (which varied randomly between high and low). After the single-task on color diversity, participants were shown eight examples of high and low mean size and performed 7 trials of a single-task in which they reported the mean size of the non-cued row letters. After this second single-task, the same exact procedure from experiment 3 commenced, i.e. participants were instructed to attend to the cued-row and report as many letters as they could (single-task letters).

After 7 trials of the letter task, the first critical trial was presented and participants were immediately given the 3-AFC recognition test for color diversity. On the critical trial, half of the participants were presented a high color diversity stimulus and half were presented a low color diversity stimulus. Note that in this experiment, subjects were trained to make the high/low color diversity discrimination from the outset, and had just been performing the color diversity judgment task 15 trials before the surprise question about color diversity.

As in experiments 1–3, the first critical trial was followed by a dual-task on the cued-row letters and the color diversity of the non-cued rows. On the eighth trial of this dual-task, a second critical trial probed for awareness of the mean size statistic. The mean size (high or low) on the critical trial was counterbalanced across participants. Following the 3-AFC, participants then completed a dual-task on the cued-row letters and mean size of the non-cued row letters for 7 trials.

Finally, to fully test the limits of asking "unexpected" questions and to examine the recently reported phenomenon of "repeated inattentional blindness" (Ward & Scholl, 2015), we added a third critical trial at the very end of the experiment in which we again asked participants about color diversity. By the time this third critical trial occurred, not only had participants been trained to perceive and report color diversity, but they had already been asked a surprise question about it and were currently focusing attention on the non-cued letters in order to carry-out the dual-task on mean size. Therefore, we expected that noticing rates would increase, but the question was whether any subjects at all would still fail to report the correct color diversity statistic on this repeated critical trial.

5.2. Results and discussion

Inattentional blindness rates are shown in Fig. 5a. On the first critical trial, 67% of participants were inattentionally blind to the color diversity statistic (44% chose the incorrect color diversity display; 23% chose the foil), while 33% chose the correct (high or low) color diversity option (not different than chance, $p = 0.15$). Similar to experiment 3, participants who were presented the high color diversity display on the critical trial were more likely to select the correct display on the 3-AFC compared to those who received the low color diversity display (53% vs. 13%). For the second critical trial, 60% of participants were inattentionally blind to the mean size statistic (30% selected the incorrect mean size; 30% selected the foil), while 40% chose the correct (high or low) mean size display (not different than chance, $p = 0.11$). Participants who received the high mean size display on the critical trial were slightly more likely to respond correctly on the 3-AFC than those who received the low mean size display (47% vs. 33%).

On the third critical trial, in which we again asked about color diversity, 33% of participants were still inattentionally blind (20% chose the incorrect option; 13% chose the foil), while 67% noticed the color diversity of this display (significantly above chance, $p < 0.001$). All participants who responded correctly to the 3-AFC on the first critical trial (33%) also did so on the third critical trial. Of the 67% who were inattentionally blind on the first critical trial, half responded correctly on the third critical trial (33%), while half remained inattentionally blind (33%).

ANOVA showed a significant effect of phase on letter performance, $F(2, 58) = 11.54$, $p < 0.01$. Dependent means t-tests revealed that letter performance in the single-task ($M = 3.90$, $SD = 0.77$) was significantly higher than that in the dual-task with color ($M = 3.40$, $SD = 0.78$), $t(29) = 4.02$, $p < 0.01$, $\eta\rho^2 = 0.20$, and the dual-task with size ($M = 3.50$, $SD = 0.85$), $t(29) = 3.99$, $p < 0.01$, $\eta\rho^2 = 0.18$. No difference was found between the dual-tasks, $t(29) = 0.43$, $p > 0.05$, $\eta\rho^2 = 0.02$. These results are summarized in Fig. 5b.

On the single-tasks for color diversity and mean size, performance was very strong (color: $M = 97\%$, $SD = 6.9$; size: $M = 83\%$, $SD = 15.5$). In the dual-tasks, color diversity performance decreased but remained above chance ($M = 62\%$,

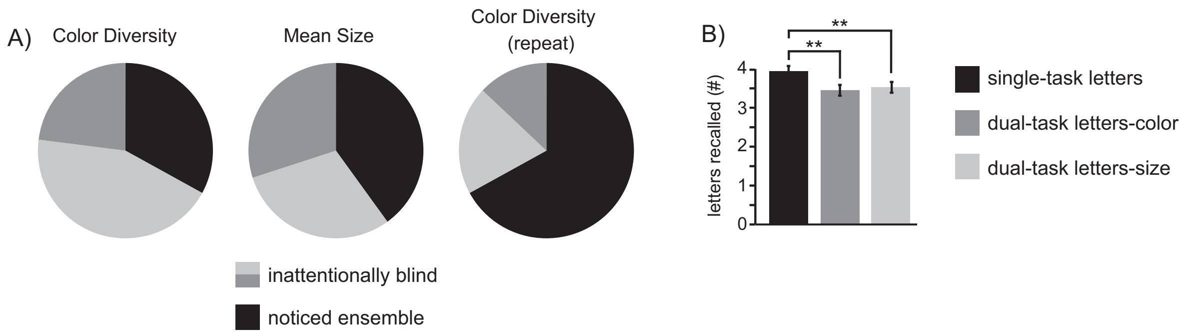

Fig. 5. Experiment 4 results. (A) Inattentional blindness rates from critical trial 1 (color diversity), critical trial 2 (mean size), and critical trial 3 (color diversity, again). In this experiment, prior to any of the critical trials, participants were trained to perceive and report the color and size ensembles in single-task conditions. Black regions represent the percent of participants who selected the correct ensemble display on the critical trial. Participants who selected an incorrect option were considered inattentionally blind (light gray = normal letters; dark gray = upside-down foil). Once again, a majority of participants (>50%) were inattentionally blind to the color and size ensemble statistics, except for critical trial 3, in which we asked the "surprise" question about color diversity a second time. Here, noticing rates increased because participants probably started to expect such questions and therefore devoted some attention to the color statistic. (B) Mean number of cued-row letters correctly reported as a function of task. An attentional cost to letter recall performance was observed for the dual-tasks (dark gray = letter-color task; light gray = letter-size task) compared to the single-task (black = letter task), $^{*}p < 0.05$, $^{**}p < 0.01$.

SD = 24.7), $t(29) = 2.74$, $p < 0.01$, *Cohen's d* = 0.50, as did mean size performance (*M* = 60%, *SD* = 19.4), $t(29) = 2.96$, $p < 0.01$, *Cohen's d* = 0.54.

6. General discussion

In the current series of experiments, we found robust inattentional blindness to color and size ensemble statistics (inattentional blindness rates ranged from 53 to 79%, with chance-level performance on the 3-AFC recognition tests in almost all cases). Experiment 4 confirmed that this inability to accurately report these summary statistics was not due to a lack of knowledge or training, as 67% of participants were inattentionally blind to color diversity despite having carried-out a task on this statistic a mere 15 trials earlier. These findings are consistent with two previous studies that investigated inattentional blindness to the "gist" of natural scenes (Cohen et al., 2011; Mack & Clarke, 2012).

In addition, we found substantial rates of "repeated inattentional blindness" in experiment 4. When we asked the surprise question about color diversity a second time, 33% of subjects remained inattentionally blind. This finding is consistent with a recent study in which participants were either instructed to "be on the lookout for anything unexpected", or were asked surprise questions about unexpected stimuli on multiple occasions (Ward & Scholl, 2015).

In all four experiments, we also observed dual-task costs. The number of cued-row letters participants could recall decreased during dual-task conditions (letters-color and letters-size) compared to single-task conditions (letters-only). During the dual-tasks, participants reported the color and size statistics at well-above chance levels, despite minimal instruction and very little practice. This ability to accurately judge color diversity in non-cued rows during a demanding central task is consistent with Bronfman et al.'s (2014) and Ward et al.'s (2016) results. However, the finding of dual-task costs on letter recall differs from these previous studies and is discussed in more detail below.

Overall, the current results suggest that conscious perception of certain visual ensembles, specifically color diversity, size diversity, and mean size, requires some attentional resources. While there were methodological differences between our study and Bronfman et al.'s (discussed below), these results demonstrate that there are clearly situations in which ensemble statistics are not perceived automatically. While it may be true that ensembles can be perceived without focal attention, there appear to be certain attentional requirements (e.g., via diffuse, distributed attention) that are necessary for ensemble perception.

6.1. A broader view of the role of attention in ensemble perception

A natural question to ask is, why did we find high rates of inattentional blindness for color diversity while Bronfman et al. (2014) and Ward et al. (2016) found above-chance performance on the color diversity task coupled with a lack of dual-task interference? We suspect that even though Bronfman et al. and Ward et al. deemphasized the color task in their studies to encourage participants to devote only a minimal amount of attention to this ensemble statistic, participants still allocated *some* attention to the color stimuli. Since ensembles are efficiently formed and can be represented with limited attention (Alvarez & Oliva, 2008, 2009), only a small amount of attention was needed to perform the task adequately. Under this view, the amount of attention was so small that it had no measureable effect on the letter-reporting task. If this idea is correct, then the amount of attention allocated to the color task should systematically alter performance on making color diversity judgments. Across experiments, we created conditions that ranged from full attention to the color diversity statistic (single-task in experiment 4), to near-zero attention to this statistic (critical trial 1 in all experiments). If we take a broader view of the relationship between attention and ensemble perception, Bronfman et al.'s (2014) and Ward et al.'s (2016) results might only reflect a small snapshot of a larger function. Fig. 6 combines some of the current results with those of Bronfman et al. (2014), and shows a linear relationship between the amount of attention allocated to the color statistics in the non-cued rows and participants' ability to perceive and judge these statistics (see also Huang, 2015). We argue that Bronfman et al. (2014) and Ward et al. (2016) measured color diversity perception under conditions of *partial* attention. Here, we measured color diversity perception during *in*attention, partial attention, and full attention, and found that this summary statistic, along with two others (size variance and mean size), are not consciously perceived during inattention by a majority of participants.

Another way to interpret the attentional-dependence of ensemble perception observed here is with the "zoom lens" metaphor of attention (Eriksen & St James, 1986; Eriksen & Yeh, 1985; Muller, Bartelt, Donner, Villringer, & Brandt, 2003). Instead of conceiving of visual attention as a unitary "spot light" that can be directed to different locations or as a limited cognitive resource with fixed "amounts" that can be allocated to various sensory stimuli, the zoom lens model posits a continuum between focal and diffuse attention. The zoom lens account would predict that attention would be maximally diffuse in the single-task on color diversity statistics, minimally diffuse (most focal) in the single-task on reporting the cued-row letters, and intermediately diffuse on the dual-tasks (ensemble judgments + letter reports). In other words, the scale plotted in Fig. 6 that ranges from full to zero attention could be interpreted to reflect the amount of diffuse attention allocated to the visual ensembles. With maximally focal attention (no diffuse attention), the ensemble statistics were not consciously perceived (as observed in our critical trials). With slightly more diffuse attention (dual-tasks), the ensembles were perceived above chance, and with maximally diffuse attention (single-task), the ensembles were perceived with almost perfect preci-

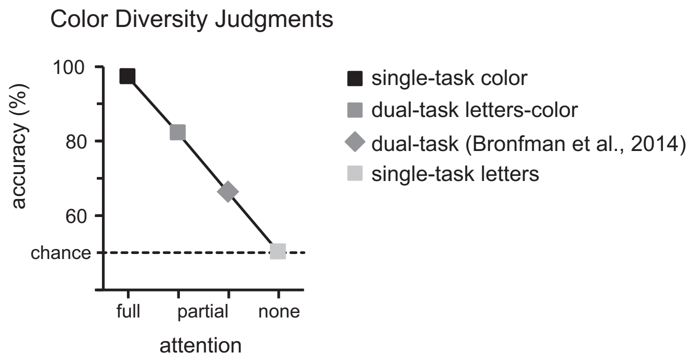

Fig. 6. Accuracy for judging high versus low color diversity in the non-cued rows as a function of attention. Square symbols = data from current study; diamond symbol = data from Bronfman et al. (2014). The black square shows data from experiment 4 in which participants first performed a single-task on color diversity in the non-cued rows (full attention). The dark gray square shows composite data from experiments 1–3, in which effort on the dual-task was balanced (partial attention). The diamond symbol shows data from the dual-task in Bronfman et al.'s experiment 1 in which the letter-task was primary and the color-task was secondary (partial attention, color-task deemphasized, 66% = mean accuracy for judging color diversity in their congruent and incongruent conditions which was not manipulated here as our cued-row letters were always white). The light gray square depicts results from the critical trials (approximately zero attention to the color statistic) across all four of the current experiments (chance performance for color diversity judgments is plotted here as 50%; see main text for analyses of the 3-AFC recognition tests). Note that under the "zoom lens" model of attention (Eriksen & St James, 1986; Eriksen & Yeh, 1985; Muller et al., 2003), the x-axis can be interpreted as the amount of "diffuse attention" allocated to the color diversity ensembles (full = maximally diffuse; partial = intermediate diffusion; none = minimally diffuse).

sion. Regardless of the attentional metaphor employed, the current results suggest that at least a minimal amount of attention is required for conscious perception of visual ensembles.

The current findings of clear dual-task costs were likely due to methodological differences between our study and previous studies. Bronfman et al. (2014) and Ward et al. (2016) deemphasized the color diversity task by explicitly instructing participants to consider this task "subjective" and "secondary" to the letter report task, and by providing feedback on the letter task but not the color task. They also gave participants substantial practice and their main experiments included 100's of trials. Thus, participants likely developed strategies for devoting only a small amount of attention to the color diversity task. In the current study, the color diversity task was not deemphasized and the dual-task (letters-color) only consisted of 7 trials (preceded by 4 practice trials). Thus, participants in the current study likely over-estimated the amount of attention (or the degree of attentional diffusion) needed for this task, thereby drawing enough attention away from the cued-row letters to result in observable performance decrements on the letter task. This interpretation is further supported by our finding of overall higher performance on the color diversity judgments during the dual-task compared to Bronfman et al. (2014) and Ward et al. (2016), e.g.~82% in the current experiments versus ~66% in Bronfman et al. and ~65% in Ward et al.

6.2. Differences in noticing high versus low variance

The overall finding reported here is >50% inattentional blindness rates (if one lumps together the 2 incorrect options in the 3-AFC tests) for color and size ensemble statistics across the four experiments, which is very much in line with rates considered "substantial" in the literature (Mack & Rock, 1998; Most, Scholl, Clifford, & Simons, 2005; Simons & Chabris, 1999; Ward & Scholl, 2015). Measured differently, participants exhibited chance-level performance (or worse) in selecting the correct color diversity display in the 3-AFC tests for 3 out of the 4 experiments. That being said, why did we observe differences in inattentional blindness rates for color diversity across the 4 experiments?

For example, a larger number of participants (78%) were inattentionally blind in experiment 1 compared to experiment 2 (54%). In experiment 1, color diversity was high on the lead-up trials and switched to low on the critical trial. In experiment 2, color diversity was low on the lead-up trials and switched to high on the critical trial. Similarly, in experiments 3 and 4 in which color diversity varied randomly during the lead-up trials and diversity on the critical trial was counterbalanced (high or low for different participants), those who received the high diversity display on the critical trial were more likely to respond correctly on the 3-AFC compared to those who received the low diversity display (80% vs. 14% in experiment 3; 53% vs. 13% in experiment 4).

One potential explanation for these differences in inattentional blindness rates is that the high color diversity displays were more likely to attract attention. If so, this difference in salience between the high and low diversity displays could have influenced attention, perception, memory, or decision-making. For example, the high color diversity stimuli may have been more likely to capture attention on the critical trial itself (compared to the low diversity stimuli), leading to increased noticing and therefore more accurate reporting of this ensemble. Alternatively, the high color diversity displays may have attracted attention during the 3-AFC, thereby introducing a response bias despite no differences in perception on the critical trial. Finally, at least in experiments 3 and 4, the high diversity stimuli may have captured attention more often during the lead-up trials, and regardless of what participants saw on the critical trial, their memory of the color in the non-cued rows was biased towards high diversity, thus they selected this option more frequently in the 3-AFCs.

Another potential factor that may have contributed to these differences is the minimal number of items that must be compared in order to determine whether a display is "high" or "low" in color diversity. On a given trial, if attention is allocated (or spills over) to only two individual non-cued letters, it is possible to determine that the display is high in color diversity (e.g. if one letter is green and the other is red), but it is not possible to determine that the display is low in color diversity (e.g. if one letter is green and the other is bluish-green, this could be consistent with either type of stimulus, and additional items must be apprehended to make the distinction). Note that while Ward et al.'s (2016) results appear to argue against this explanation (since color diversity judgments were steady despite change blindness for individual items), it is possible that participants in Ward et al.'s experiments perceived and compared the colors of two items during *either* the first *or* the second half of the trial in order to determine if the display was high in color diversity.

6.3. Inattentional blindness or amnesia?

Advocates of the "rich" view of visual awareness are likely to interpret the present results as stemming from memory failures rather than perceptual failures. In other words, the entire display including the color and size ensembles may have been briefly experienced, but due to the fragility of phenomenal awareness, this rich experience did not persist long enough to enable reports of what was just perceived. In this view, the role of attention is to select and maintain information in a more durable form of memory (working memory) rather than to enable conscious perception in the first place (Lamme, 2003, 2006; Sligte, Scholte, & Lamme, 2008; Vandenbroucke, Fahrenfort, Sligte, & Lamme, 2014; Wolfe, 1999).

Determining whether a failure to report seeing something is due to having a brief conscious experience followed by rapid-decay/memory-interference or to a lack of conscious perception in the first place, has been previously argued to be difficult, if not impossible, to address scientifically (Cohen & Dennett, 2011). Nevertheless, several recent studies appear to favor the inattentional blindness (versus amnesia) account. Ward and Scholl (2015) instructed participants to "be on the lookout" for anything unexpected and to immediately report any unexpected items on the display. In spite of these instructions, inattentional blindness for unexpected stimuli was still observed. The fact that participants still failed to notice unexpected items when they were told to be ready for such items suggests that rather than perceiving items and quickly forgetting them, people do not perceive them in the first place. Consistent with these findings, Mack, Erol, Clarke, and Bert (2016) presented Sperling-like letter arrays in the center of the display along with colored circles in the periphery. Participants were initially trained on each task separately (letter-recall task or color-circle task) but in the main experiment, retro-cues at different probabilities directed participants to perform one task or the other. In the key condition, participants were retro-cued on 90% of trials to perform the color-circle task, and on the critical trial the entire Sperling letter array was removed (replaced by a blank white screen). Strikingly, 50% of the participants failed to notice that the entire letter array was missing (Mack et al., 2016). In this case, it is difficult to argue that participants had a fleeting conscious experience of a blank screen but this fragile memory decayed or was disrupted by the surprise question leading to reports of *seeing letters that were not there*. Instead, we argue that a more parsimonious explanation is that half of the participants did not perceive the missing letter array in the first place because they allocated no attention to that region of the screen.

In the present study, several steps were taken to help participants report the ensembles they may have seen, even if their perceptual experience was brief and fragile. First, the stimuli were unmasked, lasting 300 ms in duration, and the surprise question was presented 900 ms following stimulus offset. Of course, even with only 900 ms between the critical stimuli and the surprise question, we acknowledge that a form of immediate visual memory is necessary to allow for accurate reporting. Second, the 3-AFC used to assess perception on the critical trial only required recognition based on familiarity, rather than explicit recall, stimulus categorization, or verbalization. Finally, as described above, the 3-AFC method is very liberal in terms of estimating noticing rates. This method over-estimates noticing because participants could guess correctly based on chance alone or based on unconscious processing, despite not having consciously perceived the stimulus on the critical trial. Given these considerations, the most parsimonious explanation for the current results is that ~50–80% of participants failed to *perceive* the ensemble statistics on the critical trial, and were thus inattentionally *blind* (versus amnesic).

7. Conclusion

The pattern of results observed here suggests that awareness of ensemble statistics requires attention. This view is consistent with previous studies that demonstrated inattentional blindness to the gist of natural scenes (Cohen et al., 2011; Mack & Clarke, 2012), attentional modulations of statistical perception (Huang, 2015), and attentional requirements for iconic memory (Mack, Erol, & Clarke, 2015; Mack et al., 2016). It appears that attention is necessary for conscious perception (Cohen, Cavanagh, Chun, & Nakayama, 2012), even for basic ensemble percepts such as color and size.

Acknowledgements

Thanks to Chris Graulty and Caleb Kalisher for assisting with various aspects of data collection and analysis. Thanks to Ed Awh, Zohar Bronfman, and Marius Usher for useful comments on earlier versions of this manuscript. Thanks to Enriqueta Canseco-Gonzalez for helpful discussions throughout this project.

References

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*(4), 392–398.

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences USA, 106*(18), 7345–7350. http://dx.doi.org/10.1073/pnas.0808981106.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences, 15*(12), 567–575. http://dx.doi.org/10.1016/j.tics.2011.11.001.

Block, N. (2014). Rich conscious perception outside focal attention. *Trends in Cognitive Sciences, 18*(9), 445–447. http://dx.doi.org/10.1016/j.tics.2014.05.007.

Bronfman, Z. Z., Brezis, N., Jacobson, H., & Usher, M. (2014). We see more than we can report: "Cost free" color phenomenality outside focal attention. *Psychological Science, 25*(7), 1394–1403. http://dx.doi.org/10.1177/0956797614532656.

Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *Journal of Neuroscience, 32*(22), 7685–7700. http://dx.doi.org/10.1523/JNEUROSCI.3325-11.2012.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science, 22*(9), 1165–1172. http://dx.doi.org/10.1177/0956797611419168.

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences, 16*(8), 411–417. http://dx.doi.org/10.1016/j.tics.2012.06.013.

Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences, 15*(8), 358–364. http://dx.doi.org/10.1016/j.tics.2011.06.008.

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences, 20*(5), 324–335. http://dx.doi.org/10.1016/j.tics.2016.03.006.

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. New York, NY: Penguin Books.

Eriksen, C. W., & St James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40*(4), 225–240.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*(5), 583–597.

Gross, S., & Flombaum, J. (in press). Does perceptual consciousness overflow cognitive access? The challenge from probabilistic, hierarchical processes. Mind and Language.

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. Wolfe & L. Robertson (Eds.), *From perception to consciousness: Searching with anne treisman*. New York, NY: Oxford University Press.

Huang, L. (2015). Statistical properties demand as much attention as object features. *PLoS ONE, 10*(8), e0131191. http://dx.doi.org/10.1371/journal.pone.0131191.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7*(1), 12–18.

Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10*(11), 494–501.

Lamme, V. A. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience, 1*(3), 204–220. http://dx.doi.org/10.1080/17588921003731586.

Mack, A. (2003). Inattentional blindness: Looking without seeing. *Current Directions in Psychological Science, 12*(5), 180–184.

Mack, A., & Clarke, J. (2012). Gist perception requires attention. *Visual Cognition, 20*(3), 300–327. http://dx.doi.org/10.1080/13506285.2012.666578.

Mack, A., Erol, M., & Clarke, J. (2015). Iconic memory is not a case of attention-free awareness. *Consciousness and Cognition, 33*, 291–299. http://dx.doi.org/10.1016/j.concog.2014.12.016.

Mack, A., Erol, M., Clarke, J., & Bert, J. (2016). No iconic memory without attention. *Consciousness and Cognition, 40*, 1–8. http://dx.doi.org/10.1016/j.concog.2015.12.006.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 339–352.

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review, 112*(1), 217–242.

Muller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the "Zoom Lens" of visual attention. *Journal of Neuroscience, 23*(9), 3561–3565.

Phillips, I. (2016). No watershed for overflow: Recent work on the richness of consciousness. *Philosophical Psychology, 29*(2), 236–249. http://dx.doi.org/10.1080/09515089.2015.1079604.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*(9), 1059–1074.

Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores? *PLoS ONE, 3*(2), e1699. http://dx.doi.org/10.1371/journal.pone.0001699.

Sperling, G. (1960). The information available in brief visual presentation. *Psychological Monographs: General and Applied, 74*, 1–29.

Vandenbroucke, A. R., Fahrenfort, J. J., Sligte, I. G., & Lamme, V. A. (2014). Seeing without knowing: Neural signatures of perceptual inference in the absence of report. *Journal of Cognitive Neuroscience, 26*(5), 955–969. http://dx.doi.org/10.1162/jocn_a_00530.

Ward, E. J., Bear, A., & Scholl, B. J. (2016). Can you perceive ensembles without perceiving individuals?: The role of statistical perception in determining whether awareness overflows access. *Cognition, 152*, 78–86. http://dx.doi.org/10.1016/j.cognition.2016.01.010.

Ward, E. J., & Scholl, B. J. (2015). Inattentional blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review, 22*(3), 722–727. http://dx.doi.org/10.3758/s13423-014-0745-8.

Whitney, D., Haberman, J., & Sweeny, T. D. (2014). From textures to crowds: Multiple levels of summary statistical perception. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences*. Cambridge: MIT Press.

Wolfe, J. (1999). Inattentional amnesia. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief visual stimuli* (pp. 71–94). Cambridge, MA: MIT Press.