Cognition 152 (2016) 78–86

Contents lists available at ScienceDirect

Cognition

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

Can you perceive ensembles without perceiving individuals?: The role of statistical perception in determining whether awareness overflows access

Emily J. Ward *, Adam Bear, Brian J. Scholl *

Yale University, United States

A R T I C L E   I N F O

Article history:
Received 16 June 2015
Revised 5 January 2016
Accepted 13 January 2016

Keywords:
Awareness
Iconic memory
Consciousness
Ensemble representation
Statistical perception

A B S T R A C T

Do we see more than we can report? Psychologists and philosophers have been hotly debating this question, in part because both possibilities are supported by suggestive evidence. On one hand, phenomena such as inattentional blindness and change blindness suggest that visual awareness is especially sparse. On the other hand, experiments relating to iconic memory suggest that our in-the-moment awareness of the world is much richer than can be reported. Recent research has attempted to resolve this debate by showing that observers can accurately report the color diversity of a quickly flashed group of letters, even for letters that are unattended. If this ability requires awareness of the individual letters' colors, then this may count as a clear case of conscious awareness overflowing cognitive access. Here we explored this requirement directly: can we perceive ensemble properties of scenes even without being aware of the relevant individual features? Across several experiments that combined aspects of iconic memory with measures of change blindness, we show that observers can accurately report the color diversity of unattended stimuli, even while their self-reported awareness of the individual elements is coarse or nonexistent—and even while they are completely blind to situations in which each individual element changes color mid-trial throughout the entire experiment. We conclude that awareness of statistical properties may occur in the absence of awareness of individual features, and that such results are fully consistent with sparse visual awareness.

1. Introduction

Two of the most central topics in visual cognition are conscious awareness and visual memory, yet how these capacities relate to each other is still not entirely clear. Do we see more than we can remember and report? One possibility is that we are aware of only that to which we attend and/or that which is encoded into memory. Another possibility, however, is that awareness "overflows" what is readily accessible in memory, such that in-the-moment percepts are richer than can be reported. The debate between these possibilities has engaged both psychologists and philosophers in recent years, in part because both possibilities seem to be supported by suggestive evidence.

1.1. Empirical measures of the richness of visual awareness?

On one hand, several stunning phenomena of visual awareness demonstrate that even highly salient events right in front of your eyes may often go unnoticed unless they are attended. For example, in demonstrations of inattentional blindness (e.g. Mack & Rock, 1998; Most, Scholl, Clifford, & Simons, 2005), many people fail to perceive stimuli such as a gorilla walking through a scene (Simons & Chabris, 1999) or a bright red cross traversing a display otherwise filled only with black and white shapes (Most et al., 2001), when attention is otherwise engaged. Such failures of awareness occur even when observers have instructions to immediately report unexpected events (in the moment, while they are occurring), confirming that this is a phenomenon of perception rather than memory (Ward & Scholl, 2015).

Similarly, in demonstrations of change blindness (e.g. Simons & Rensink, 2005), people fail to detect large changes made to scenes, when those changes do not draw attention. In one of the earliest and still most striking such demonstrations, viewers read text while having their eyes tracked, and failed to notice that every

---

* Corresponding authors at: Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, United States.

E-mail addresses: emily.ward@yale.edu (E.J. Ward), brian.scholl@yale.edu (B.J. Scholl).

http://dx.doi.org/10.1016/j.cognition.2016.01.010
0010-0277/© 2016 Elsevier B.V. All rights reserved.

letter in the display was an '*X*' except for the few near their fixation, as long as the changes were made during saccades (McConkie & Zola, 1979). Both sorts of phenomena seem readily explained by appeal to the sparse nature of visual awareness (though some philosophical work has challenged this assumption, e.g., Noë, Pessoa, & Thompson, 2000). In inattentional blindness, for example, attention may serve as a sort of gateway to awareness, such that we are not aware of unattended stimuli (such as the gorilla or the red cross) in the first place, even though they may be processed unconsciously. Some change blindness phenomena may be similarly explained, via the assumption that attention (and thus awareness) is often confined to the foveal region of a display. In cases such as McConkie & Zola's experiments, we may still *feel* like we see normal English text in the periphery, but in such cases that is clearly a mistaken inference, since there are no real words there (until you fixate on this region of the 'text').

On the other hand, experiments examining iconic memory suggest that our in-the-moment awareness of the world is much richer than can be reported. In the classic demonstration of such effects (Sperling, 1960), observers viewed a quickly flashed array of letters, and then were asked to report them. When asked about all of the letters, observers were only able to recall a few, demonstrating a stark limit on reportability. Those few letters that were recalled could be influenced by a cue, however: if prompted to report a specific row of letters, observers could do so. Critically, this was true even when the cue appeared *after* the offset of the letters. In such cases, observers were still reasonably accurate at reporting the letters in the postcued row, but not the others. These and related studies (e.g. Sligte, Scholte, & Lamme, 2008; Vandenbroucke, Fahrenfort, Sligte, & Lamme, 2014) have been taken to support the existence of rich visual awareness: if only some letters are reportable, but all letters are *potentially* reportable based on a postcue, then this may suggest that observers are initially phenomenally aware of all of the letters, but that only some are subsequently encoded into a memory durable enough to support subsequent report (Block, 2011; cf. Phillips, 2011). In this 'rich awareness' perspective, observers are thus aware of all the letters in the display, and the role of the postcue is simply to prompt the observers to encode a subset of them into a more durable (longer lasting, but lower capacity) memory store. (As explored in the General Discussion, this inference relies on the assumption that it is not possible for the postcue to, for the first time, pull into awareness cued letters that have been only unconsciously represented until that point; cf. Sergent et al., 2013.)

1.2. Resolving the debate by measuring statistical perception?

Ironically, though the debate between sparse vs. rich views of visual awareness was prompted in part by empirical evidence that seemed to favor both sides, the debate has proven difficult to resolve precisely because there doesn't seem to be any empirical way to directly measure the existence or nature of phenomenal awareness when there is no durable memory encoding. After all, at its core this view assumes that the contents of this form of awareness *are not reportable* (unless transferred into subsequent memory stores that also support 'access consciousness'; Block, 2011), and it is difficult to directly measure something that even in principle cannot be reported or accessed.

Recent research has attempted to resolve this debate by taking a somewhat different approach—supposing that even while the letters themselves in such situations aren't reportable, some other properties of the initial rich conscious experiences may still persist and so be measurable. As in previous studies of iconic memory, Bronfman, Brezis, Jacobson, and Usher (2014) presented observers with a brief array of (now colored) letters. Observers were precued to a specific row of letters, and then a postcue signaled the position of a single letter to be reported from the cued row. (In this design, the precue serves to orient attention to only a subset of the letters, with the others being entirely irrelevant to this task and thus presumably unreportable—though these researchers never actually directly measured the ability to report any letters from the uncued rows, and so were not directly assessing iconic memory as in Sperling, 1960.) Performance when reporting the postcued letter then serves as a measure of the degree to which other manipulations may or may not change the degree of attentional focus on the cued row, as described below.

Critically, observers also had a second task—to report a statistical property of the *colors* of the letters (in either the cued row or the uncued rows). The colors of the letters could be sampled from either a narrow region of a color wheel (low color diversity) or from the entire color wheel (high color diversity)—as in Fig. 1A—and observers were asked to report whether the specified group of letters (from either the cued row or the uncued rows) had high vs. low color diversity. As depicted in Fig. 1B, the displays were designed so that the diversity of the cued row vs. the uncued rows could vary independently. Observers in this experiment were above chance when reporting color diversity even for letters that were unattended, and color diversity judgments for unattended letters did not impair observers' ability to report the postcued letter (thus confirming that attention was still focused on the cued row).

These results led Bronfman et al. (2014) to conclude that the color diversity judgments were being made without attention, presumably on the basis of residual information from the observers' initial rich visual experience of all of the letters. Supporting this view—and purportedly ruling out an account based on unconscious visual color processing—observers in this experiment claimed to have seen the colors themselves: observers were asked to report on each trial whether they "did not see the colors", "partially saw the colors", or "saw the colors well", and the results indicated that when observers claimed to have not seen the colors, they could not accurately report the color diversity. (In other experiments, observers had an "escape" button that they could press whenever they failed to perceive the colors, but they never made use of this option.) These results were thus presented as a clear case of visual awareness overflowing access and reportability, based on the residual reportability of a statistical property of the letters: because observers could report color diversity even for the uncued rows, they must have visually experienced all of the letters. As such, this demonstration has impressed some researchers as a "dramatic advance" and an "astonishing result"—counting as a fairly decisive resolution to the debate over sparse vs. rich visual awareness (Block, 2014, p. 445).

1.3. Ensemble representation

Color diversity in these experiments is a type of statistical summary of a display, as may be stored in an *ensemble representation*. Ensemble representations are statistical summaries of features at an abstracted level that collapse across local details. In experiments on such representations, observers view an array of objects, and must report some summary statistic of the array—such as the *average* size of an array of discs (Ariely, 2001). The typical result from such experiments is that observers are impressively accurate at reporting the summary statistic, while also being generally terrible at reporting properties of any of the individual elements in the array (for reviews see Alvarez, 2011; Haberman & Whitney, 2012). The ability to form and use such 'statistical summary representations' appears to be highly general, as observers are readily able to report statistical summaries of properties ranging from size (Chong & Treisman, 2005), motion direction (Dakin & Watt, 1997), and location (Alvarez & Oliva, 2008) to facial identity (de Fockert &

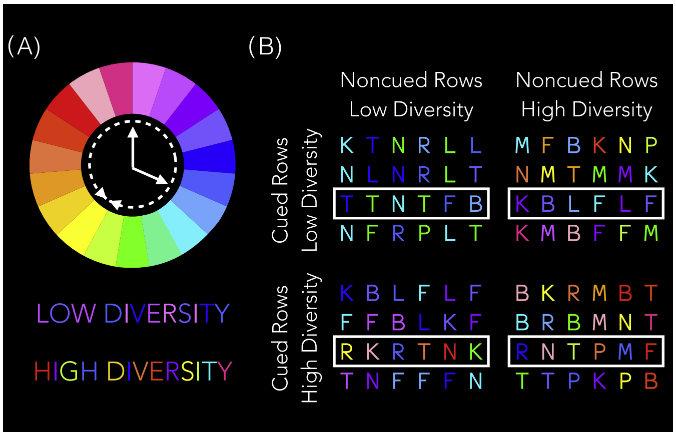


Fig. 1. Color diversity: experimental conditions. (A) The color wheel displaying the 19 color options. For high diversity conditions, the color of each letter was randomly selected from all 19 possibilities (indicated by the dashed circular arrow). For low diversity conditions, the colors were randomly sampled from 6 adjacent colors, the specific range of which was also randomly selected. A text example of high and low diversity appears underneath the color wheel. (B) Across all trials, there were four diversity conditions, wherein the cued row (which was also the row from which the postcued letter had to be reported) could be either low or high diversity, and the uncued rows could independently be either low or high diversity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Adapted from Bronfman et al., 2014.

Wolfenstein, 2009), emotion (Haberman & Whitney, 2007), and auditory frequency (Albrecht, Scholl, & Chun, 2012). Moreover, such representations are not even intrinsically spatial, insofar as they operate just as efficiently when extracting statistical properties from temporal sequences of items (Albrecht & Scholl, 2010; Haberman, Harp, & Whitney, 2009). Although such representations are often identified with perceptual *averaging* per se, forms of statistical diversity can be equally efficiently extracted (e.g. Albers, Correll, Gleicher, & Franconeri, 2014; Haberman, Lee, & Whiteny, 2015)—for example when observing not only the average density of a group of dots, but also its 'cluster', which is a measure of the *variance* in that density over space (Durgin, 1995). Indeed, with other features such as orientation, visual computations of statistical variance can even be *more* robust and precise than the corresponding computations of the means (e.g. Solomon, 2010).

Critically, the extraction of statistical summary properties appears to occur extremely efficiently and perhaps automatically. Measures of diversity such as dot cluster can be extracted at a glance (Durgin, 1995), and perceptual averaging can occur even for displays presented as briefly as 50 ms (Chong & Treisman, 2003), and even when the resulting accuracy necessitated processing of most if not all of the items in a display (Albrecht & Scholl, 2010; Chong, Joo, Emmmanouil, & Treisman, 2008; Haberman & Whitney, 2010). Perhaps most relevant in the current context, ensemble representations appear to be formed even outside the focus of attention. For example, in an attention-demanding multiple object tracking task in which observers' task was to track a subset of moving dots while ignoring other distractor dots, observers were able to localize the average position of the distractors, without being able to localize the individual distractors (Alvarez & Oliva, 2008). Moreover, ensemble representations can even be formed for stimuli that observers cannot see in the first place—as when simultanagnosic observers can accurately report average values despite only seeing one object at a time (Demeyere, Rzeskiewicz, Humphreys, & Humphreys, 2008), or when neglect patients include elements from the neglected visual field in their computations of statistical properties (Pavlovskaya, Soroker, Bonneh, & Hochstein, 2015). These results suggest that ensemble properties (that we may eventually become aware of, and so be able to *see* and report) may be formed even without visual awareness of the individual elements in the first place—a possibility that we directly address in the current experiments, in the context of visual awareness and color diversity judgments.

1.4. The current study

The results of Bronfman et al. (2014) confirm theories of rich visual awareness only given the assumption that it is not possible to perceive color diversity without having experienced all of the individual colors themselves. Bronfman and colleagues make this assumption explicit in their discussion: "the availability of color diversity is best explained as resulting from the fleeting experience of the underlying individual colors... [This] follows from the fact that without a differentiated (albeit transient) representation of the colors, it is not possible to judge diversity" (p. 1395). And other commentators similarly argue that this must be the case, concluding that: "there must have been conscious awareness of specific colors ... because a trace of that conscious awareness in the form of a diversity judgment" survives (Block, 2014, p. 446). It seems to us that these conclusions are debatable, and that they must be empirical questions (though for incisive theoretical critiques, see Gross & Flombaum, in press; Phillips, in press). As such, the present experiments put these issues to the test. Why couldn't the observers have perceived color diversity without any visual experience of the individual colors?

Experiment 1 replicates Bronfman et al. (2014), but assesses observers' perception of the colors in a more fine-grained manner—explicitly contrasting the perception of color *in general* (as a statistical property) with the perception of color *of individual elements*. We reasoned that if asked directly, observers might simply be able to report the degree to which they experienced (the colors of) individual elements in the display. Experiments 2 and 3 then contrast awareness of diversity with awareness of individual elements via a change blindness manipulation: during the initial display, every single individual letter outside of the cued row *changed* its color, but in a way that preserved the same color diversity statistics. We reasoned that if color diversity could be perceived *without* perceiving individual colors, then such diversity reports could be accurate even when observers had no ability to detect that individual elements outside the focus of attention had changed.

2. Experiment 1: Detailed color perception of individual letters?

We first aimed to replicate the central results of Bronfman et al. (2014), as reported above. Critically, however, we assessed the perception of color in a more fine-grained way. Bronfman and colleagues had observers either (1) simply hit an 'escape key' when they failed to perceive any color, or (2) categorize their color perception in terms of whether they "did not see the colors", "partially saw the colors", or "saw the colors well". It seems to us that neither of these measures of color perception draws the necessary distinction between (1) experiencing colors largely or only as a statistical property, and (2) experiencing the colors of individual elements. Accordingly, we tried to draw this distinction directly when asking observers about their color experiences—having them judge for each trial which of the following options best matched their experience:

(1) I had no sense that any of the letters had any color at all.
(2) I had a vague sense that the letters were colored in general, but I didn't clearly perceive the individual colors of individual letters.

(3) I had a clear sense that the letters were colored in general, but I didn't clearly perceive the individual colors of individual letters.
(4) I had a clear sense that the letters were colored in general, and I could also clearly perceive the individual colors of individual letters.

In essence, Bronfman et al. inferred from the lack of "escape key" use in some of their experiments (which would correspond to our option #1) that the observers must have experienced the individual element colors in rich detail (corresponding to our option #4)—thus not allowing for the possibility that their observers' true percepts were better characterized by our options #2 or #3. Similarly, Bronfman et al. inferred from the fact in some of their other experiments that their observers "saw the colors well" that they must have experienced the *individual* colors well—thus not allowing for the possibility that such responses actually reflected an experience more akin to our option #3.[1] Here, then, we were primarily interested in whether observers would report their color experiences by always selecting our option #4 (as would be suggested by Bronfman et al.'s interpretation) or whether they would also select option #3 (which would certainly count as "seeing colors well", but would carry no necessary implications for the conscious perception of individual display elements).

2.1. Methods

2.1.1. Participants

Twelve members of the Yale community (mean age 24.1 years) participated for monetary compensation. This sample size was chosen to match that of Experiment 5 in Bronfman et al. (2014), and was identical for all of the experiments.

2.1.2. Apparatus

Stimuli were presented on an Acer monitor with a 60 Hz refresh rate, using custom software written in Python with the PsychoPy libraries (Peirce, 2007). Observers sat approximately 65 cm from the display, with all visual extents reported below computed based on this distance.

2.1.3. Stimuli and procedure

An array of 24 colored letters (4 rows × 6 columns, as in Fig. 1B) was centered on a black background (29.31° × 23.47°). Each letter was randomly sampled from nine consonants (R, T, F, N, B, P, L, M, K). The letters were presented in Arial font (up to 0.71° wide × 1.06° tall) and the entire array subtended 8.62° × 6.16°. The diversity of the letters' colors in the cued row and in the uncued rows was either low or high. High diversity was implemented by selecting a color for each letter sampled with replacement from 19 possible colors (the same used in Bronfman et al., 2014). Low diversity was implemented by limiting the sampling range to a randomly selected range of only six adjacent colors (see Fig. 1A). As depicted in Fig. 1B, there were four possible color-diversity combinations: high/low diversity in the cued row × high/low diversity in the uncued rows.

On each trial, a white (.71° × .71°) fixation cross appeared in the center of the display for 200 ms, after which a 200 ms visual spatial cue (a white 9.41° × 1.76° rectangle) appeared to indicate the task-relevant row (randomly selected for each trial). The cue was then replaced by the 24-letter array, appearing for 300 ms, followed by a 900 ms blank interval. A visual postcue (a 1.59° × 1.59° white square) then appeared at the location of one of the letters (also randomly selected) in the cued row. The postcue remained visible until observers pressed a letter key on a standard keyboard to indicate which letter had appeared at that location. After reporting the letter, observers were asked (via a prompt presented on the display; see Fig. 2) to press one of two keys to indicate the color diversity level (low or high) of either the cued row or of the three remaining uncued rows. (Following Bronfman et al., 2014, observers were always asked about the color diversity of the cued row during the first half of the experiment, and were always asked about the color diversity of the uncued rows during the second half.) Immediately afterwards, observers were also asked (via another visual prompt) to indicate which of 4 options (as listed above) best captured their experience of the colors of the letters in the same row(s) for which they had just reported the color diversity. Observers were told that there was no right answer to this question, that their primary task was still to recall the postcued letter, and they were assured that although it may seem odd to answer the question after each trial, they should just go with their first impression of the letters, regardless of whether they found themselves picking the same option frequently or picking different options from trial to trial. This trial sequence is summarized in cartoon form in Fig. 2.

The experiment began with a supervised 70-trial practice block in which observers' only task was to report the postcued letter. Observers were then shown an example of a row of letters with high color diversity, and another with low color diversity. Observers then completed 272 experimental trials, receiving a short, self-terminated break every 96 trials and a 1.5-min mandatory break every 192 trials. After their session, each observer completed a funneled debriefing procedure during which they were asked about their experiences and about any particular strategies that they had employed.

2.2. Results

The first step in our analyses was to average six key measurements across all participants: Letter Recall Accuracy, Letter Recall Accuracy by Cue Type (whether observers were asked about the color diversity of cued or uncued rows), Color Diversity Accuracy, Color Diversity Accuracy by Cue Type (cued or uncued rows), Color Diversity Accuracy by Diversity Type (high or low), and Color Diversity Accuracy by Cue Type and Diversity Type (interaction). These measurements are included in Table 1, along with the relevant statistics that highlight significant performance. In general, letter recall accuracy was well above chance (11.11%), and (as depicted in Fig. 3A) observers were also able to correctly report color diversity above chance (50.00%).

When asked to rate their subjective impression of the colors of individual letters in the uncued rows, observers gave a variety of responses,[2] as depicted in Fig. 4. Inspection of this figure suggests two key patterns: (a) observers chose option #1 only rarely, and (b) they chose the other options at approximately equal rates. These impressions were verified by the following statistical tests. There was no difference in judgment rates when measured by an omnibus

---

[1] Note that Bronfman et al.'s option #2 ("partially saw the colors") would not necessarily be an appropriate choice for observers whose experience was best captured by our option #3: such an observer might still have a fully rich (and far from "partial") experience of the colors *in general*, as a statistical property. Thus it seems to us that these possibilities can never be differentiated in a unidimensional set of ratings (as Bronfman et al. used), without explicitly distinguishing experiences of individual colors vs. experiences of colors in general.

[2] There was a marginally significant effect of Cue Type on the judgment rates ($F(1,11) = 4.61$, $p = 0.055$, partial $\eta^2 = 0.3$). However, when considering subjective impressions of the colors of individual letters in only the cued row, all critical comparisons were the same (i.e. observers were just as likely to choose option #3 as option #4, and just as likely to choose options #2 and #4). The only difference when looking at the ratings for cued letters was that participants were less likely to choose option #1 than any of the other options (including option #4).

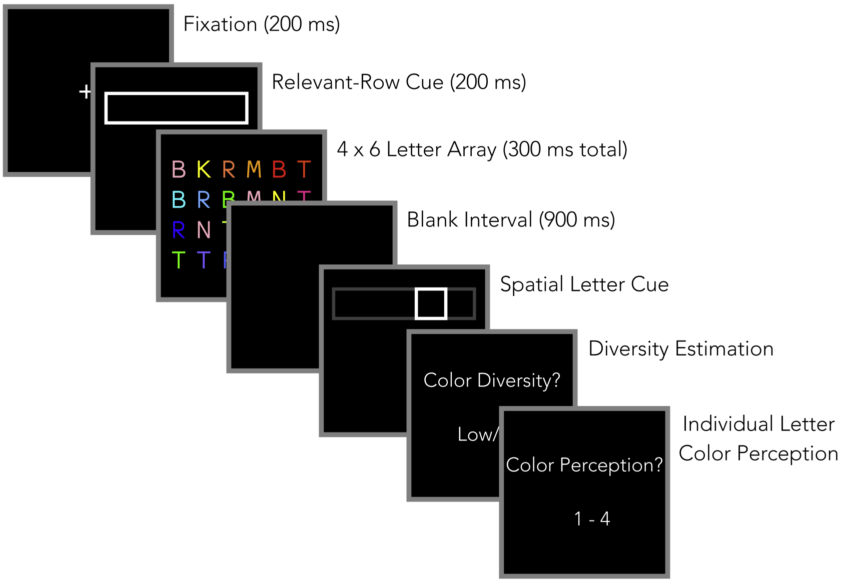


Fig. 2. The experimental procedure from Experiment 1. On each trial, a fixation cross appeared in the center of the display for 200 ms, after which a 200 ms visual spatial cue (a white rectangle) appeared alone to cue the (randomly selected) task-relevant row. The cue was then replaced by the 24-letter array, appearing for 300 ms, followed by a 900 ms blank interval. A visual postcue (a letter-sized white square) then appeared at the location of one of the letters (also randomly selected) in the cued row. The postcue remained on the screen until observers pressed a letter key on a standard keyboard to indicate which letter had appeared at that location. After reporting the letter, observers were asked to estimate the color diversity level (low or high) of either the cued row or of the uncued rows. Finally, observers reported their visual experiences of the letters' colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Overall letter recall and color diversity performance across all experiments.

	Expt. 1 Detail	Expt. 2 Switch	Expt. 3 Digital
Letter recall			
Accuracy (%)	57.32 (17.78)	57.18 (14.28)	47.46 (14.46)
p	**<0.001**	**<0.001**	**<0.001**
t	9.0	11.18	8.71
Cohen's *d*	2.6	3.23	2.51
Letter recall – cue type			
M_{cued} (%)	59.10 (19.51)	60.81 (13.94)	48.05 (13.84)
M_{uncued} (%)	55.19 (17.23)	53.56 (15.94)	46.88 (15.39)
p	0.17	**0.002**	0.38
F	2.11	16.42	0.84
η_p^2	0.16	0.6	0.07
Color diversity			
Accuracy (%)	64.07 (7.81)	67.88 (6.11)	67.32 (7.67)
p	**<0.001**	**<0.001**	**<0.001**
t	10.10	10.13	7.82
Cohen's *d*	2.92	2.92	2.26
Color diversity – cue type			
M_{cued} (%)	63.66 (7.78)	69.53 (6.37)	66.75 (6.93)
M_{uncued} (%)	64.74 (5.30)	66.23 (8.01)	67.88 (9.53)
p	0.67	0.17	0.56
F	0.19	2.18	0.36
η_p^2	0.02	0.17	0.03
Color diversity – diversity type			
M_{high} (%)	62.98 (9.84)	67.49 (7.39)	65.58 (10.58)
M_{low} (%)	65.16 (5.32)	68.27 (6.79)	69.05 (9.50)
p	0.77	0.71	0.38
F	0.09	0.14	0.86
η_p^2	0.01	0.01	0.07
Color diversity – cue × diversity type interaction			
p	0.86	0.59	0.70
F	0.03	0.31	0.16
η_p^2	<0.01	0.03	0.01

Note: Degrees of freedom for all one-sample *t*-tests shown is 11 and for all ANOVAs is (1, 11). Boldface entries highlight statistically significant results ($p < 0.05$).

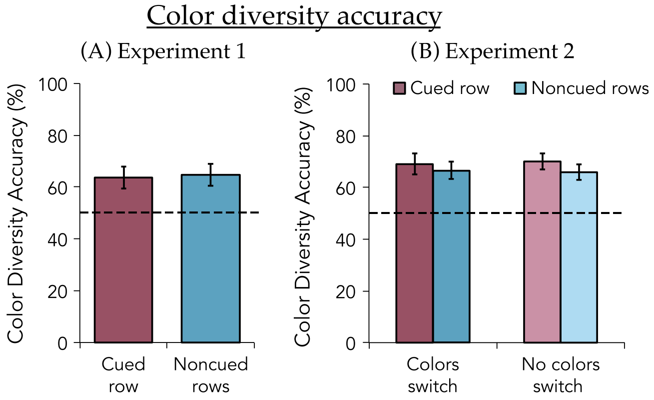


Fig. 3. Color diversity judgment accuracy. Observers were able to judge the color diversity of both the cued and uncued rows above chance. (A) Color diversity accuracy in Experiment 1. (B) Color diversity accuracy in Experiment 2. There was no impact on performance when the colors of the letters switched. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

test ($M_1 = 8.63 \pm 9.96\%$, $M_2 = 38.41 \pm 24.84\%$, $M_3 = 32.89 \pm 21.95\%$, $M_4 = 20.07 \pm 28.05\%$), $F(1,11) = 0.72$, $p = 0.41$, $\eta_p^2 = 0.06$. Critically, observers were just as likely to choose option #3 (indicating that while they did have a clear sense of color in general, they did not clearly perceive the individual colors of individual letters) as they were to choose option #4 (indicating that they could clearly perceive the individual colors of individual letters) (3 vs. 4: $t(11) = 1.09$, $p = 0.30$, $d = 0.31$). Indeed, they were even just as likely to choose option #2 (indicating that they only had a vague sense that the letters were colored in general) as they were to choose option #4 (2 vs. 4: $t(11) = 1.29$, $p = 0.22$, $d = 0.37$). Finally, observers were much less likely to choose option #1 (indicating that they had no sense of color) than option #2 or #3 ($ps < 0.02$), but just as likely to choose option #1 as #4 (1 vs. 4: $t(11) = 1.23$, $p = 0.25$, $d = 0.35$). Three

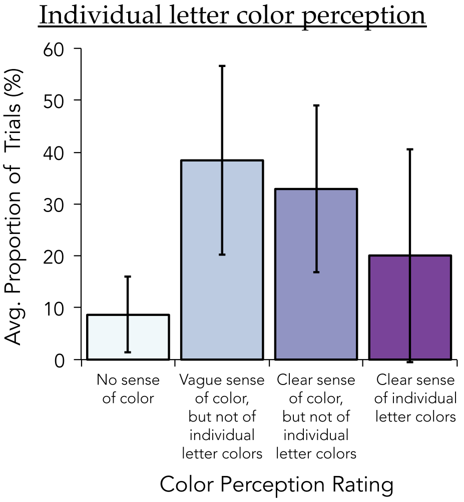


Fig. 4. Individual letter color perception. After each trial in Experiment 1, observers rated the quality of their visual experiences of the colors of individual items. Plotted are the average percentage of trials corresponding to each of the four ratings. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

observers never chose option #1, but when the other 9 did choose that option, they were at chance at reporting color diversity ($t(8) = 0.69$, $p = 0.51$, $d = 0.23$), replicating Bronfman et al. (2014). In contrast, color diversity reports were above chance for trials in which observers chose each of the other three options ($ps < 0.03$). Most critically, color diversity reports were above chance (63.54 ± 10.69%) even for those trials in which observers chose option #2 (i.e. when they reported being unable to see individual letters' colors; $t(10) = 4.20$, $p = .002$, $d = 1.27$)—and, among the 10 subjects who chose both options #2 and #4 at some point during the experiment, color diversity report accuracy on those trials when observers chose option #2 did not differ significantly from the color diversity report accuracy on those trials when they chose option #4 (i.e. when they did report being able to see individual letters' colors; 63.29 ± 10.66% vs. 73.53 ± 27.77%, $t(9) = 1.05$, $p = .32$, $d = 0.35$). Nonparametrically, one observer never chose option #2 to describe their impression of the uncued rows, but of the remaining 11, 10 reported the uncued rows' diversity above chance. All twelve observers chose option #3 to describe their impression of the uncued rows and 11 reported the uncued rows' diversity above chance.

2.3. Discussion

The results of this experiment replicated the color diversity results of Bronfman et al. (2014), but fuel a different conclusion about the observers' visual experiences of color. Instead of always perceiving the individual colors in rich detail, our more nuanced color report options indicated that observers often perceived color only in a general sense, without perceiving individual letters' colors. Thus observers can accurately report a statistical property (color diversity) of unattended stimuli in a display, even while their self-reported awareness of the individual elements is coarse or nonexistent. As explored in the General Discussion, these results are consistent with accounts of sparse visual awareness, and with the possibility that observers can experience ensemble properties without experiencing individual elements.

3. Experiment 2: Combining color diversity and change blindness

Another way of testing whether accurate color diversity judgments entail a rich experience of the colors of individual items is to assess performance on a task that can only be completed with information about (at least some) individual colors. Per our discussion in the Introduction of studies such as those of McConkie and Zola (1979), change blindness seems like the perfect tool for this job.

This experiment was thus qualitatively identical to Experiment 1, except that instead of asking directly about color experience, we simply introduced a massive change into the display during the initial presentation on half of the trials: all of the unattended letters' colors switched, but since this was implemented by colors actually just being reshuffled among the letters, the color diversity of those letters was held constant. In addition to assessing color diversity judgments, we then also simply asked whether observers ever noticed this massive change. We reasoned that if observers are experiencing only the statistical diversity of the colors without experiencing individual element colors at all, then they might fail to notice such changes—not just occasionally, but perhaps throughout the entire experiment, and even though every unattended letter is dramatically changing. In contrast, if observers do (at least occasionally) notice such changes while making accurate color diversity judgments, this would falsify our interpretation, suggesting that conscious perception of the colors goes beyond statistical properties and includes at least some information about individual colors.

3.1. Methods

This experiment was identical to Experiment 1 except as follows. Twelve additional observers participated (mean age 21.3 years). Following the initial cue to one of the rows, the array of letters appeared for 150 ms, followed by a 17 ms blank screen, and then by the array that appeared again for 150 ms. (As in Experiment 1, this array then disappeared and was followed by the appearance of a postcue at the location of one of the letters from the cued row.) On half of the trials (randomly chosen, differently for each observer), the array was identical for both 150 ms presentations. On the other half, all of the colors of the letters in the uncued rows were randomly reshuffled during the blank screen. Color experience was not assessed. Because the trials were shorter, there were 384 trials in total—192 with color switches, and 192 without color switches. (As in Experiment 1, observers were always asked about the color diversity of the cued row during the first half of the experiment, and were always asked about the color diversity of the uncued rows during the second half.) After the experiment, two additional questions were added to the debriefing questionnaire to determine observers' awareness of the color switches: (a) "Did you notice if the colors of the uncued letters ever changed mid-trial? If so, how did they change?" and (b) "Did you suspect during the course of the experiment that that flashing had anything to do with the purpose of the experiment?". Only observers who answered the first question negatively were counted as 'unaware' observers.

3.2. Results

The data from this experiment, as analyzed via the same six key measurements as used in Experiment 1, and as summarized in detail in Table 1, replicated all of the primary results from Experiment 1—primarily the above-chance color diversity performance for uncued rows, as in Bronfman et al. (2014). The only result that was different was that when the color diversity of the uncued rows was queried, letter recall performance decreased. However, accuracy was still well above chance for both conditions (cued: $t(11) = 12.35$, $p < 0.001$, $d = 3.56$; uncued: $t(11) = 9.64$, $p < 0.001$ $d = 2.78$).

Critically, none of our 12 subjects noticed that the colors switched. The switch also had no behavioral consequences, on either color diversity judgments ($M_{switch} = 67.80 \pm 6.31\%$,

$M_{no\ switch} = 67.97 \pm 6.44\%$, $F(1,11) = 0.03$, $p = 0.87$, $\eta_p^2 < 0.01$; see Fig. 3B; no interaction with Cue Type $p = 0.30$) or letter recall ($M_{switch} = 56.99 \pm 14.19\%$, $M_{no\ switch} = 57.38 \pm 14.70\%$, $F(1,11) = 0.09$, $p = 0.76$, $\eta_p^2 = 0.01$; no interaction with Cue Type $p = 0.68$).

3.3. Discussion

Observers in this experiment performed above chance (as in Bronfman et al., 2014) on color diversity judgments for unattended letters, yet at the same time they failed to notice massive changes in the colors to the individual elements, when those colors did not change the diversity. As explored in the General Discussion, and as with Experiment 1, these results are consistent with accounts of sparse visual awareness, and with the possibility that observers can experience ensemble properties without experiencing individual elements.

4. Experiment 3: Color diversity and change blindness with longer exposures

In this experiment we replicated Experiment 2 with an even starker change blindness manipulation: whereas the pre-change letter array colors were only visible for 150 ms in Experiment 1, here they were fully visible and unchanging for a full 650 ms. (To implement this change without changing the timing of the letters themselves—which of course would dramatically change the letter identification performance—we used colored placeholders, as depicted in Fig. 5). This manipulation made it incredibly easy to see the massive color changes when you knew to look for them, as in the demonstration presented online at http://www.yale.edu/perception/ColorDiversity/.

4.1. Method

This experiment was identical to Experiment 2 except as follows. Twelve additional observers participated (mean age 23.8 years). The array first appeared as identical digital placeholders (as in Fig. 5) for 500 ms. These placeholders then instantaneously 'dropped' segments to form real letters (with the same colors and in the same font), which then stayed visible for an additional 150 ms, after which the change could occur as in Experiment 2.

4.2. Results and discussion

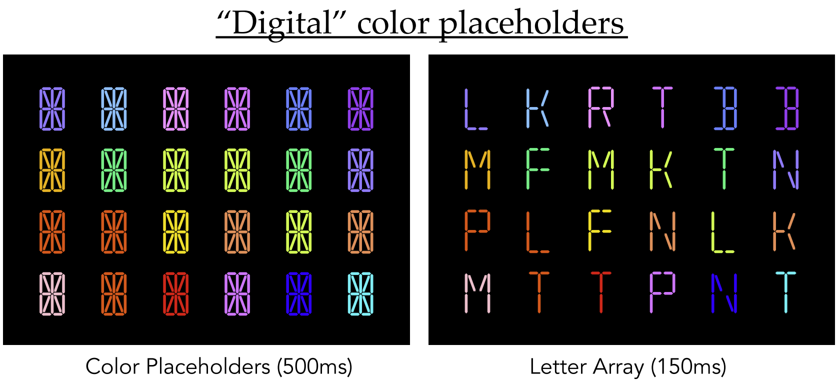
The data from this experiment, as analyzed via the same six key measurements as used in Experiment 1, and as summarized in detail in Table 1, replicated all of the primary results from Experiment 1—primarily the above-chance color diversity performance for uncued rows, as in Bronfman et al. (2014). In addition, replicating Experiment 2, not a single one of the observers noticed the color changes. Again, the switch also had no behavioral consequences, on either color diversity judgments ($M_{switch} = 66.84 \pm 7.05\%$, $M_{no\ switch} = 67.80 \pm 8.63\%$, $F(1,11) = 0.86$, $p = 0.37$, $\eta_p^2 = 0.07$; no interaction with Cue Type $p = 0.18$) or letter recall ($M_{switch} = 46.79 \pm 13.28\%$, $M_{no\ switch} = 48.13 \pm 15.90\%$, $F(1,11) = 1.02$, $p = 0.34$, $\eta_p^2 = 0.08$; no interaction with Cue Type $p = 0.97$). This again seems consistent with the possibility that observers do not have rich visual experiences of the individual elements, despite being able to judge color diversity.

5. General discussion

The three experiments presented here replicate the primary results of Bronfman et al. (2014) but suggest very different conclusions. We showed that observers were able to report the statistical ensemble property of color diversity for an array of letters even when those letters were unattended. Nevertheless, our key manipulations suggested that this ability may be present *without* robust visual experience of the individual letters' colors themselves. First, in Experiment 1, we showed that judging color diversity was possible even during trials in which observers explicitly reported that their experience of *individual* letters' colors was coarse or nonexistent. Second, in Experiments 2 and 3, we showed that judging color diversity was possible even when observers failed to notice changes to the letters' colors. The extent of this change blindness was striking. The key manipulation in Experiments 2 and 3 was a massive color-change that involved shuffling the color of every single unattended letter in the display. This occurred 192 times for each of the 12 observers in each experiment (totaling 3456 changes), yet these changes were never noticed by even a single observer—despite the changes in Experiment 3 being easily visible when looking for them, as in our online demonstration. These results suggest to us that the ability to judge color diversity may not involve (much less require) rich color experiences of the individual letters.

5.1. Revisiting sparse vs. rich awareness

Our results suggest that it may be possible to experience ensemble properties without necessarily experiencing the individual elements and features that make up those ensembles. We suggest that this empirical observation casts doubt on the underlying inferences that have been based on the results of Bronfman et al.



Color Placeholders (500ms)

Letter Array (150ms)

Fig. 5. Placeholders manipulation. A depiction of the placeholders manipulation from Experiment 3, which allowed the initial colors to be presented for 650 ms, while the letters were only visible with their initial colors for 150 ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(2014). Recall that Bronfman and colleagues argued that "the availability of color diversity is best explained as resulting from the fleeting experience of the underlying individual colors... [This] follows from the fact that without a differentiated (albeit transient) representation of the colors, it is not possible to judge diversity" (p. 1395). But upon close examination, this appears to be a non-sequitur. We enthusiastically agree that color diversity is not possible without a differentiated *representation* of the colors, but that does not require that this representation itself be conscious. Instead, the (conscious) percept of the ensemble property could be based on an unconscious 'differentiated' visual representation of the colors. At any rate, it is an empirical question whether this is possible, and our results are consistent with that possibility. And so when Bronfman et al. (2014) cautiously license the possibility of "generic (undetailed) or fragmentary information about objects at unattended locations" (p. 1395), we would suggest that statistical summary information may also properly belong in this list. It may very well be that a "trace" of the initial encounter with the objects survives in the form of a diversity judgment (see Block, 2014); but if, as suggested by our results, observers were never aware of the individual specific colors, then such a trace may not have been conscious after all.

It is important to note here that our change blindness results do not demonstrate that this novel statistical 'sparse' interpretation of the results of Bronfman et al. (2014) is necessarily correct. In particular, it is still possible that observers had a rich phenomenological awareness of *both* the pre-change and post-change individual colors, on every single trial. This is possible for a specific reason, which is that change blindness can also occur due to a failure to *compare* the pre- and post-change displays, even when both are reliably encoded (e.g. Mitroff, Simons, & Levin, 2004; Simons, Chabris, Schnur, & Levin, 2002). But it is also possible for a general reason, which is that such interpretations are *always* possible—given that phenomenally conscious information can be, by definition, entirely inaccessible and unreportable. (As such, we don't see any way that this possibility could ever be scientifically disconfirmed.) The question at hand, though—both here and in Bronfman et al. (2014)—is whether there is evidence that a 'rich awareness' view is not only *possible*, but is also *correct*. Bronfman et al. (2014) take their results (as does Block, 2014) to have decisively demonstrated this. But we think that this inference is mistaken (see also Phillips, in press), because of a fascinating and previously unrecognized possibility—that we may be able to consciously perceive summary statistics without perceiving individual features.

Despite the ever-present possibility of a 'rich awareness' interpretation, our experiments still constitute a strong test of these possibilities, insofar as they could have easily disconfirmed the statistically-based 'sparse' interpretation. In Experiment 1, for example, it could have been that our observers always indicated that they clearly saw the individual colors (which is what Bronfman et al., 2014, assumed, but did not directly test with their dichotomous measure). And in Experiments 2 and 3, it could have been that whenever observers were successful at reporting the color diversity, then they always would have detected at least *one* of the 18 elements that had changed on that trial—at least once during the course of the entire experiment (encompassing more than 3000 changes). Either of those patterns of results would have decisively ruled out our interpretation. That they did not—and that the extent of the change blindness in Experiments 2 and 3 was so extreme—thus lends support to the statistically-based 'sparse' interpretation.

5.2. Related evidence

If color diversity judgments can be made without awareness of the individual elements, that at least casts doubt on the supposedly decisive evidence of Bronfman et al. (2014)—which may not end up speaking to the sparse-vs.-rich awareness debate at all. Of course, our results do not themselves speak to the seemingly rich awareness that derives from iconic memory results themselves. However, other recent studies have pointed out a key empirical assumption in those inferences too: such results (e.g. Sligte et al., 2008; Sperling, 1960; Vandenbroucke et al., 2014), to speak to the issue of sparse vs. rich awareness, must assume that postcues cannot bring into awareness a stimulus that was never before consciously perceived. But this too is an empirical question, and in fact, several early studies hinted that this was possible. For example, during motion-induced blindness, the sudden offset of a visual stimulus—in an orientation that was never before seen (since the item was rotating while rendered invisible)—can cause that stimulus to suddenly appear in awareness in its new and never-before-perceived orientation (Mitroff & Scholl, 2004).

A recent study demonstrates a related phenomenon, but much more impressively—suggesting that a stimulus can be called into awareness for the first time, even up to 400 ms after it had disappeared (Sergent et al., 2013). Observers were shown a display wherein a single at-threshold Gabor grating appeared either on the right or left side of the screen, and observers had to indicate its orientation. On some trials, a pre-cue appeared before the onset of the Gabor, which improved orientation judgments. Amazingly, when the same cue appeared between 100 and 400 ms *after* the onset of the Gabor, these postcues also improved orientation judgments. Follow-up experiments showed that this result was driven by subjectively changing observers' impressions in two ways. First, observers were more likely to report not seeing a Gabor when the cue was absent, and second, they were more likely to report increased visibility when there was one present. This combination suggests that the postcue can actually elicit the conscious perception of a visual stimulus that was previously unconscious (Sergent et al., 2013). These results, like those of the present paper, are thus completely consistent with the possibility of sparse visual awareness—and with the possibility that awareness does not overflow access (see also de Gardelle, Sackur, & Kouider, 2009; Kouider, de Gardelle, Sackur, & Dupoux, 2010; Phillips, 2011). Moreover, the present results suggest a specific mechanism for how such iconic memory effects might arise.

5.3. Seeing ensembles

Leaving debates about sparse vs. rich awareness aside, the present results are also interesting from the perspective of research on statistical summary representations, in two ways. First, though ensemble processing is very much a hot topic at present, no previous studies to our knowledge have directly explored the subjective quality of observers' experiences of individual elements when (e.g.) reporting perceptual averages—nor have they tested change detection performance during such tasks. Our results suggest that ensemble representations can be formed and reported even with very coarse experiences of the individual elements (at best) and even when observers utterly fail to notice changes in individual elements during such judgments. This is all consistent with the observations that ensemble representations are formed quickly (Chong & Treisman, 2003), without focused attention (Alvarez & Oliva, 2008; Chong & Treisman, 2005), early in development (Sweeny, Wurnitsch, Gopnik, & Whitney, 2015; Zosh, Halberda, & Feigenson, 2011), and (in patient populations) without any awareness at all of the individual elements themselves (much less their features; Demeyere et al., 2008; Pavlovskaya et al., 2015).

Second, whereas past studies of ensemble representations have mostly involved perceptual averaging, others have stressed that such abilities also apply to other statistical measures, especially those of variance or diversity (Albers et al., 2014; Durgin, 1995;

Haberman et al., 2015; Solomon, 2010). Nevertheless, to our knowledge, no previous studies have explored color diversity judgments. Along with the results of Bronfman et al. (2014), the current results suggest that such representations can be robust, even when you do not notice changes to the colors of the individual elements.

From both the perspectives of ensemble representations and sparse-vs.-rich awareness, our results thus add to a growing recognition that the unconscious mind is capable of surprisingly sophisticated processing.

Author note

For helpful conversation and/or comments on earlier drafts, we thank Ned Block, Zohar Bronfman, Steve Gross, Ian Phillips, Louis Thibault, Marius Usher, and the members of the Yale Perception and Cognition Lab.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.01.010.

References

Albers, D., Correll, M., Gleicher, M., & Franconeri, S. (2014). Ensemble processing of color and shape: Beyond mean judgments. *Journal of Vision, 14*, 1056.

Albrecht, A. R., & Scholl, B. J. (2010). Perceptually averaging in a continuous visual world: Extracting statistical summary representations over time. *Psychological Science, 21*, 560–567.

Albrecht, A. R., Scholl, B. J., & Chun, M. M. (2012). Perceptual averaging by eye and ear: Computing summary statistics from multimodal stimuli. *Attention, Perception, & Psychophysics, 74*, 810–815.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*, 122–131.

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

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*, 157–162.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences, 15*, 567–575.

Block, N. (2014). Rich conscious perception outside focal attention. *Trends in Cognitive Sciences, 18*, 445–447.

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*, 1394–1403.

Chong, S. C., Joo, S. J., Emmmanouil, T. A., & Treisman, A. (2008). Statistical processing: Not so implausible after all. *Perception & Psychophysics, 70*, 1327–1334.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*, 393–404.

Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics, 67*, 1–13.

Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research, 37*, 3181–3192.

de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *The Quarterly Journal of Experimental Psychology, 62*, 1716–1722.

de Gardelle, V., Sackur, J., & Kouider, S. (2009). Perceptual illusions in brief visual presentations. *Consciousness & Cognition, 18*, 569–577.

Demeyere, N., Rzeskiewicz, A., Humphreys, K. A., & Humphreys, G. W. (2008). Automatic statistical processing of visual properties in simultanagnosia. *Neuropsychologia, 46*, 2861–2864.

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception & Performance, 21*, 149–169.

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

Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision, 9*, 1–13.

Haberman, J., Lee, T., & Whitney, D. (2015). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *Journal of Vision, 15*(4), 16. http://dx.doi.org/10.1167/15.4.16. 1–11.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17*, R751–R753.

Haberman, J., & Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Attention, Perception, & Psychophysics, 72*, 1825–1838.

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* (pp. 339–349). Oxford University Press.

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences, 14*, 301–307.

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

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics, 25*, 221–224.

Mitroff, S. R., & Scholl, B. J. (2004). Seeing the disappearance of unseen objects. *Perception, 33*, 1267–1273.

Mitroff, S. R., Simons, D. J., & Levin, D. T. (2004). Nothing compares 2 views: Change blindness can occur despite preserved access to the changed information. *Perception & Psychophysics, 66*, 1268–1281.

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

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science, 12*, 9–17.

Noë, A., Pessoa, L., & Thompson, E. (2000). Beyond the grand illusion: What change blindness really teaches us about vision. *Visual Cognition, 7*, 93–106.

Pavlovskaya, M., Soroker, N., Bonneh, Y., & Hochstein, S. (2015). Computing an average when part of the population is not perceived. *Journal of Cognitive Neuroscience, 27*, 1397–1411.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*, 8–13.

Phillips, I. B. (2011). Perception and iconic memory: What Sperling doesn't show. *Mind & Language, 26*, 381–411.

Phillips, I. B. (in press). No watershed for overflow: Recent work on the richness of consciousness. Philosophical Psychology.

Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L., & Tallon-Baudry, C. (2013). Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. *Current Biology, 23*, 150–155.

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

Simons, D. J., Chabris, C. F., Schnur, T., & Levin, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness and Cognition, 11*, 78–97.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*, 16–20.

Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLoS ONE, 3*, e1699.

Solomon, J. (2010). Visual discrimination of orientation statistics in crowded and uncrowded displays. *Journal of Vision, 10*(14), 19, 1–19, 16.

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

Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2015). Ensemble perception of size in 4–5-year-old children. *Developmental Science, 18*, 556–568.

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

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*, 722–727.

Zosh, J. M., Halberda, J., & Feigenson, L. (2011). Memory for multiple visual ensembles in infancy. *Journal of Experimental Psychology: General, 140*, 141–158.