Awareness doesn't come for free:
The attentional costs of gist perception

A Thesis

Presented to

The Division of Philosophy, Religion, Psychology, and Linguistics

Reed College

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Arts

Molly Louise Jackson-Nielsen

May 2015

Approved for the Division

(Psychology)

Michael Pitts

Acknowledgments

I extend many thanks to the Reed College SCALP Lab members who supported me in my first (and second!) semester as a new lab member: Chris, Phoebe, Orestis, Kat, Angelynn, Josie, and Nicole, I was really intimidated and you helped me feel like a part of the team. Kat, I don't know how I would have written my program without your help. Chris and Phoebe, you made me feel competent and welcome; thank you!

Enriqueta, you were always interested in and supportive of my research. You provided valuable methodological and theoretical critiques . . . thank you for being so engaged at every lab meeting. I appreciate your realistic yet optimistic eye.

Michael, thank you for being such a wonderful professor and advisor. You pushed me to gather more data, ask more questions, and think more deeply throughout the thesis year. I am grateful for your willingness to help me with figures one minute, and turn around and talk about consciousness the next. Your dedication to the field is inspiring, and your dedication to your students is even more so. I will contact you if I ever have questions about consciousness or wine.

I would also like to acknowledge Mac III: You were a comfortable, if smelly, fortress to come home to after a day of thesis work. You are all weird and I am so grateful to have been your HA this year.

Finally, thank you to my family and friends. Mom, Dad, Grandma, Grandpa, Sis, Mark, Oceano, Indigo, Fedora, Joseph, Tico, and Gypsy. Life is hard and exciting and you are my support system and co-conspirators. I love you.

Preface

This thesis is fundamentally about consciousness. Current theories of consciousness span disciplines from philosophy to neuroscience and differ in their claims concerning what consciousness is, what it is for, and how it should be studied (see Blackmore, 2012 for an overview). While I will focus on the contents of consciousness in this thesis, I am also motivated by questions regarding its nature and purpose. Truly, it is impossible to study the former without considering the latter. Any investigation into the mind's nitty-gritties inevitably leads to questions about its nature, function, and origins.

What, then, *is* consciousness? Intuitively, being conscious means (to us humans, who can reflect on it) having a *quality* that unconscious beings and other matter do not. Think, for a moment, how intractable it seems that matter can be organized in such a way that it *feels* things. At first glance, subjective experience cannot be explained just by our physical bodies. Many theories reflect this intuition to some degree, whether they posit that minds are functional, rather than physical, states (Putnam, 1975), or that consciousness can be explained with quantum physics (Schwartz, Stapp, & Beauregard, 2005). At the same time, such theories might argue that the "mystery" of consciousness is analogous to that of life, in that once consciousness is better explained, its élan vital-like quality will fade (Dennett, 1988). In other words, once we better understand its mechanism and function, we won't have so many questions about its "nature."

In this thesis, I use the richness of visual perception as a window into the contents of consciousness. Investigating the amount and variety of visual information that is truly "in consciousness", or, stated more accurately, that *comprises* visual awareness, provides insights into the structure and function of consciousness.

The following does not resolve any of the earlier quandaries, but it is guided by them.

Table of Contents

List of Figures

Abstract

The degree to which visual awareness exists outside focal attention is debated. This question can be addressed by presenting scenes from which ensemble statistics ("gist") can be extracted while attention is focused on a primary task; gist report is then used as a measure of visual richness. In a recent study, participants completed a focal task with a computer display and, in addition, reported the gist of the display (Bronfman et al., 2014). Gist performance was above chance and focal performance was unaffected by gist report, suggesting rich, cost-free phenomenal awareness outside focal attention. The present experiments tested whether gist would be perceived in the complete absence of attention by adding inattentional blindness manipulations to the Bronfman paradigm. Participants completed the Bronfman focal task, and then were unexpectedly asked to identify the gist of the computer display. They then performed the Bronfman dual-task (focal+gist). Next, they were unexpectedly asked about another aspect of the gist of the display. Finally, they completed a new dual-task (focal+new gist aspect). Performance in the surprise questions was at chance, suggesting no gist awareness in the absence of attention, even in the presence of task-irrelevant dispersed attention. Furthermore, contrary to the Bronfman et al. findings, focal performance in the present experiments dropped with the addition of gist report, suggesting an associated attentional cost. The present results suggest that conscious gist perception requires at least a small amount of attention; furthermore, no support was provided for the hypothesis that gist perception reflects rich phenomenal awareness. The performance in Bronfman et al. likely reflects divided attention rather than cost-free phenomenal awareness.

To the Rocky Mountains and the old-growth forest.

Introduction

The Brain – is wider than the Sky –
For – put them side by side –
The one the other will contain
With ease – and You – beside –

–Dickinson, *The Complete Poems of Emily Dickinson*

1.1 Background

Consciousness, like anything worthy of inquiry, is difficult to study. At first glance, we are all experts in the field: We all intimately know our own consciousness. However, consciousness becomes perplexing when we start asking questions about it. How can it come from a physical brain? Which animals are conscious, and which aren't? And, motivating this thesis, how much are we conscious of in a given moment? The last question is only one of the many that arise when we start asking questions about consciousness, but it is what I will focus on throughout these pages.

First, I want to clarify how the term "consciousness" will be used in this thesis. The word "conscious" is often used as a synonym for "awake" or "perceiving", denoting *level* of consciousness, as in conscious or unconscious. A different use of the word "consciousness", which I will employ here, is awareness *of* something. I will use "conscious(ness)" and "aware(ness)" interchangeably. Given that a person is conscious (awake, perceiving), what are they conscious *of*? I won't be focusing on level of consciousness here, though arguably, level of consciousness and contents of consciousness should be addressed together when possible (Bachmann & Hudetz, 2014).

I explore the contents of consciousness through the window of human visual awareness, a well-researched avenue of peoples' awareness of the external world (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Lamme, 2003). Visual perception itself (conscious or unconscious) is relatively well understood, and for good reason. Most of the human brain's occipital lobe is dedicated to visual processing; we pay a lot of attention to visual stimuli. If you ask what someone is conscious *of* in the external world in a given moment, you should expect vision to occupy a relatively large and complex portion of the answer.

This question regarding the informational capacity of consciousness was asked by William James (1890) and is still debated today. Introspectively, we have a sense of being aware of a lot, even more than what we can talk about or remember. In other words, we don't feel like our visual awareness is a spotlight, limited to what we are currently paying attention to. To confirm (or deny!) this, imagine walking down a tree-lined street on a nice day in autumn. (If visual imagery isn't your strength, the next paragraph may be helpful.) As you walk, golden sunlight streams through the upper branches, highlighting the crunchy leaves on the ground. Houses set back in their yards stand along each side of the street, casting long shadows in the low autumn sun. Sometimes during your walk you focus on the blue sky between the trees, sometimes on the rich reds, yellows and greens of their leaves, and sometimes you are lost in thought with a little bit of attention on everything around you. How much are you conscious of during your stroll? A lot? A little? Does it depend on what you're focusing on? To me, it seems I'm aware of most everything in my visual field. My vision is rich, full; there are no "holes" in my awareness. I can't describe to you everything I see, of course, but I'm seeing it nonetheless! In other words, it's not that I wasn't having the *experience* of that red mailbox over there before I paid direct attention to it; rather, it's that I wasn't paying attention to the red mailbox enough to think, "Oh, there's a red mailbox!" Now that I've paid attention to the mailbox, I can tell you about it. If I hadn't, I wouldn't be able to report it, but it still would have been in my awareness!

It may be useful to repeat the thought experiment above using your current surroundings. Without moving your eyes, focus your attention on something in your visual field that you weren't paying attention to before. What was the visual nature of the thing before you paid attention to it? Were you aware of it and simply unable to talk about it? Or, were you unaware of it entirely, even though light reflecting off the object entered your eyes and was processed by your visual system?

The above questions are contentious because of the aforementioned discrepancy between our subjective sense of informational richness and our ability to report (remember, talk about) the identity of said information. How can we possibly test whether something is in consciousness before someone pays attention to it? As soon as you asked about the mailbox, I paid attention to it, and its subjective nature changed. I can tell you I was always aware of it, but how can you be sure I'm right? The following paragraphs will detail two interpretations of the discrepancy. The first interpretation claims that the information that's in consciousness is greater than what our brains can process for the purposes of report or memory. The second argues that we are conscious of less than we think, perceptual richness is an illusion, and I wasn't really aware of the mailbox until I focused my attention on it.

Proponents of conscious richness, the first interpretation (Block, 2005, 2011, 2014; Lamme, 2003), agree with the intuition that we are aware of more than we can talk about. Block (2005) proposed a distinction between "phenomenal consciousness" (P-consciousness) and "access consciousness" (A-consciousness). Phenomenal consciousness is rich, high in capacity, not necessarily reportable, and focused in the visual cortex. Access consciousness is functional, closely tied with attention and working memory, limited in capacity, reportable, and instantiated only with the involvement of frontal and parietal cortices. A-consciousness is the subset of phenomenally conscious contents that are accessed.

Various experimental findings and their interpretations support the richness of consciousness. In the original Sperling experiment (Sperling, 1960), three rows of three or four letters each were briefly presented, and then one of the letter rows was retrospectively cued with a tone (high, middle, or low pitch). No matter which row was cued, participants could report about three or four of the letters, suggesting that all of the letters were stored in the brain in some capacity for a brief period of time. If one interprets the Sperling results through the P/A-consciousness distinction, all the letters were *phenomenally* conscious, but only three to four of them could be accessed and remembered. (The capacity of working memory is around three or four items, Luck & Vogel, 1997.) Regarding our intuition of richness, participants in Sperling-like experiments typically report a subjective impression of consciously seeing all of the letters presented, but only being able to remember a few of them.

Another study providing support for conscious richness is that of Chen and Wyble (2015). In their experiment, participants could identify the location of a letter among three numbers, but could not report the identity of the letter when they were unexpectedly asked about it immediately afterwards. Chen and Wyble interpreted their findings as reflecting rich consciousness: Participants had to be conscious of the letter's identity in order to report its location, they argued, but its identity was not preserved long enough for report. Their paradigm differs from that of Sperling in that participants were asked about a stimulus feature that they didn't expect to have to report. Participants in the Sperling experiment, on the other hand, knew they would be asked about one of the rows of letters, and therefore distributed their attention across all of the potentially relevant information.

A recent mechanistic proposal directly relevant to the concept of phenomenal consciousness is that of "fragile visual short term memory" (fVSTM) (Block, 2011; Vandenbroucke, Sligte, & Lamme, 2011). The proposal is that some phenomenal content may persist for a brief period of time, but that most experimental paradigms fail to measure it, capturing instead iconic (sensory) memory or visual working memory. fVSTM, if it exists, is smaller in capacity than retinal iconic memory but greater in capacity than visual working memory. It decays rapidly and only a portion of it can be transferred into working memory via attention.

Proponents of sparse consciousness, on the other hand, argue that phenomenal consciousness does not exist in the first place. According to the sparse view, the use of sensory information by the brain's consumer/output systems (attention, report) *is* consciousness (Dehaene et al., 2006). Inattentional blindness (Cohen, Alvarez, & Nakayama, 2011; Mack & Rock, 1998) and change blindness (Simons & Levin, 1997), wherein we fail to perceive significant objects or changes in a scene when we are not attending to them, provide support for the idea that we can be unaware of some details in our visual field if we are focused on other details.

The sparse theory argues that there is no need to extrapolate awareness from the results of Sperling (1960) and Chen and Wyble (2015), but rather that participants' visual systems were processing large amounts of information outside of consciousness. In the Sperling task, then, only about three to four letters were accessed by attention and held in working memory, and were therefore conscious. In Chen and Wyble, letter versus number discriminations were processed unconsciously and only letter location (not letter identity) was consciously accessed (see Dehaene et al., 2006 for proposed mechanisms).

A key difference between the phenomenal/access and access-only theories of consciousness is the credibility they give to the existence of awareness outside attention. While it is clear that access consciousness requires attention, most proponents of phenomenal consciousness assume that phenomenal experiences can occur without attention. Task-relevant attention, or attention to aspects of the environment that are relevant to current behavior, has been the focus of the recent debate between the rich and sparse views because it may help explain phenomena like inattentional blindness: When an aspect of the environment isn't relevant to the current task (e.g., a woman in a gorilla suit unexpectedly walking through a group of basketball players while observers try to count number of ball passes), the evidence suggests we are largely unaware of it (Neisser & Becklen, 1975; Simons & Chabris, 1999). Phenomenal consciousness, if it exists, does not depend on task-relevant attention, or perhaps on attention at all (Block, 2011; Lamme, 2003). Arguments in favor of phenomenal consciousness typically assume that inattentional blindness is actually inattentional amnesia, i.e., observers actually saw the gorilla but did not attend to it or remember it because it was task-irrelevant. Inattentional amnesia was the explanation that Chen and Wyble gave for ability to report letter location (task-relevant) but not letter identity (task-irrelevant). Note that, in the Sperling paradigm, all of the letters were task-relevant, but they exceeded the capacity of working memory.

Block (2005) argued, conversely, that there is no need to conclude the *absence* of awareness in cases of inattentional and change blindness. He proposed that the rich, phenomenal contents of visual awareness, independent of focal attention (or perhaps any attention), do not enter working memory and are therefore not reportable. These contents are not detected in most research paradigms in which awareness is operationalized as that which can be reported (Block, 2005; Lamme, 2003).

At the time of this writing, the rich/sparse debate continues. Cohen and Dennett (2011) argued that any theory of consciousness that separates awareness from function is scientifically unfalsifiable. That is, if there is some component of consciousness that causes no change in behavior (thought, report, movement), it cannot be examined using the scientific method, which relies on cause and effect (a function). Phenomenal consciousness, if it exists, does not necessarily have an effect, and therefore cannot be studied scientifically (but see a rebuttal by Fahrenfort & Lamme, 2012).

1.2 "Gist"

At this point it may seem as though consciousness is stuck in a theoretical rut. How can phenomenal consciousness be demonstrated, if it is trapped within subjective experience? Conversely, how can its existence be disproven, if the same behavioral effects are predicted with or without it? One way that the opposing camps have moved forward constructively is to agree that we are often conscious of *something* outside of our focal attention. In other words, whether or not I was conscious of the mailbox or the gorilla per se, my awareness outside my focal attention wasn't just grey nothingness. In cases wherein some attention is dispersed, there is room to explore – via both rich and sparse frameworks – how much visual information is in consciousness.

Dispersed attention provides a window into the richness of subjective experience via peoples' ability to report the "gist" of their surroundings. If asking me about the red mailbox confounds phenomenal consciousness with task-relevant report, perhaps asking me about the overall colors of the trees is more fruitful and telling. The experiments in this thesis investigated awareness of gist via report of the gist of a computer display. The next section will discuss gist and its interpretations in more detail.

1.2.1 What is gist?

Gist is the visual information outside of focal attention that one *can* report, given whatever attentional resources may be necessary to do so. Returning to your afternoon stroll: Can you report the color of every leaf on the trees around you? Probably not. Can you, however, report the overall colors of the trees? Yes: They're part fading green, part brilliant red, and part mottled yellows and oranges. These *general visual aspects* of a scene are what is meant by gist. Gist can also be conceptualized as "ensemble statistics" (Haberman & Whitney, 2012), which are descriptive statistics about a scene (color, texture, etc.) that the visual system calculates in order to summarize that scene; these terms will be used more or less interchangeably throughout this thesis. The concept of gist is important to visual awareness research and theory because it concretizes this finding: While people can't report everything in their visual field (the red mailbox, the location of each individual yellow leaf) they can report a visual summary of the entire field. In some ways, such summary report reflects our subjective sense of visual perception: Our vision is filled-in, continuous. In other ways, we may feel that we *perceive* everything, but can't focus on all of it at once, and therefore can only *provide* a summary. We may feel, therefore, that there is a discrepancy between our perceptual richness and the more abbreviated summary information we are able to report. Here we return to our earlier discrepancy: Is gist report a cognitive distillation of rich phenomenal consciousness that underlies it (Block, 2011; Bronfman, Brezis, Jacobson, & Usher, 2014), or does gist report accurately mirror consciousness, meaning that awareness itself is a sparse, summary-like representation of what's in the visual field (Cohen & Dennett, 2011; Kouider, de Gardelle, Sackur, & Dupoux, 2010)? The study of gist report is a methodology that both sides of the richness-sparseness debate employ, but the interpretation of results differs based on underlying assumptions about the richness of consciousness.

1.2.2 Interpretations of gist awareness

There is general agreement on what constitutes gist and gist report. However, different theorists and researchers interpret peoples' ability to report gist in vastly different ways, starting with vastly different assumptions about the nature and structure of human visual awareness and of consciousness in general. This section details two interpretations of gist report and awareness that are motivated by rich and sparse theories of consciousness.

Gist awareness as reflecting underlying phenomenal consciousness

One interpretation of gist rests on the assumption that phenomenal consciousness exists (Block, 2011, 2014; Bronfman et al., 2014). This interpretation argues that we are aware of everything in our visual field, but that we cannot 'report' everything due to limits in cognitive systems like attention and working memory. To these theorists, it is most parsimonious to say that the visual information needed to ascertain the gist of a scene (e.g. the colors of individual leaves) is conscious, but that the gist itself (the overall colors of the leaves) is all that can be *accessed*.

Gist awareness as reflecting awareness of only ensemble statistics

A second interpretation of gist is that a report of general aspects of a visual scene reflects awareness of just that: Ensemble statistics. In this interpretation, conscious content does not include the visual information necessary to 'create' gist; rather, the summarizing occurs in primary and secondary visual cortices and only gist enters awareness (if it is attended) (Cohen & Dennett, 2011; Kouider et al., 2010; McConkie & Rayner, 1975).

The sparse interpretation of gist posits that perceptual and cognitive illusions (Cohen & Dennett, 2011; de Gardelle et al., 2010; Kouider et al., 2010; McConkie et al., 1975) are responsible for our subjective sense of seeing detail richer than that which we can remember and report. There is ample evidence that top-down influences on perception are wide-ranging and have effects on subjective experience and report, similar to the influence that schemas exert on memory for scenes (Pezdek, Whetstone, Reynolds, Askari, & Dougherty, 1989). It is therefore reasonable, according to the sparse framework, that bottom-up perceptual processing and top-down cognitive processes exert influence on subjective interpretations of gist awareness. In other words, perhaps unconscious perceptual and cognitive mechanisms cause us to interpret our awareness of ensemble statistics as evidence that we are phenomenally conscious (Cohen & Dennett, 2011).

A more experiential argument can be posed for the illusory nature of the sense of seeing everything in one's visual field: Every time we shift our visual attention (typically by moving our eyes, but not necessarily), we see what's there immediately. We have never tried to focus on something in our visual field without it immediately being there. For instance, as soon as I focus my attention on my teacup next to me, it's in my awareness. And my sense is that it was in my awareness all along. It's possible, then, that it is more parsimonious or adaptive to offer an internal explanation that awareness is greater than the information that is accessed. If the teacup was in my visual field all along, how could it not have also been in my awareness? How could I have been blind to it, when my eyes themselves could see it? It's analogous to the wall behind me. I can't see it, but every time I turn around, it's there. My assumption is that it stays there, regardless of whether I'm looking at it. The argument is that perhaps high-level visual processing without awareness is unintuitive to us. It is more comfortable to believe that consciousness is not a spotlight outside of which awareness does not exist.

1.3 Experimental motivation

The following experiments sought to elucidate the role that attention plays in awareness of gist; more specifically, whether gist awareness occurs when attention is completely focused elsewhere. We were motivated by the paradigm and findings of Bronfman et al. (2014) (Fig. 1.1). Bronfman et al. found that participants could report the gist of a display while completing a focal working memory task with no cost to focal task performance. Gist report consisted of determining whether uncued letters in a letter array had "high" or "low" color diversity (Fig. 1.1a, c), and the focal task was remembering letter identities from a cued row in the array (Fig. 1.1b, c). Bronfman et al. (2014) interpreted their results as evidence of rich phenomenal awareness outside focal attention that was independent of the working memory resources used to complete the focal task.

There are a number of alternate interpretations of the Bronfman results, which motivated the following experiments. First, due to the dual-task paradigm, both gist (color diversity level) and the focal task (letter report) were always task-relevant. Participants were likely 1) paying some attention to both informational sets, and 2) committing information from both sets to working memory. While this likelihood does not directly negate the claims of Bronfman et al. since they posited awareness outside of *focal* attention, not attention in general, their conclusions did suggest that awareness of color phenomenality is automatic. If it were automatic, one would predict the ability to report color diversity just after seeing it, even if it were not task-relevant.

Another important aspect of Bronfman et al.'s methods is that participants completed hundreds of trials and always received feedback for letter performance but never for color performance. Therefore, is likely that participants were allocating only a minimal amount of attention to the color task, given their extensive training and letter feedback. This interpretation could explain the consistent letter performance across the single and dual-tasks without claiming that color diversity awareness was cost-free.

Finally, Bronfman et al. claimed that participants' color diversity performance itself supported the existence of rich phenomenal experience. They argued that the capacity to judge diversity suggested a brief conscious experience of most or all of the individual colors on the screen. It is possible, however, that perception and statistical summary of individual colors could have occurred unconsciously, allowing participants to judge high vs. low color diversity at above-chance levels. This final interpretation wasn't directly explored in the present experiments, but it is worth addressing in the future.

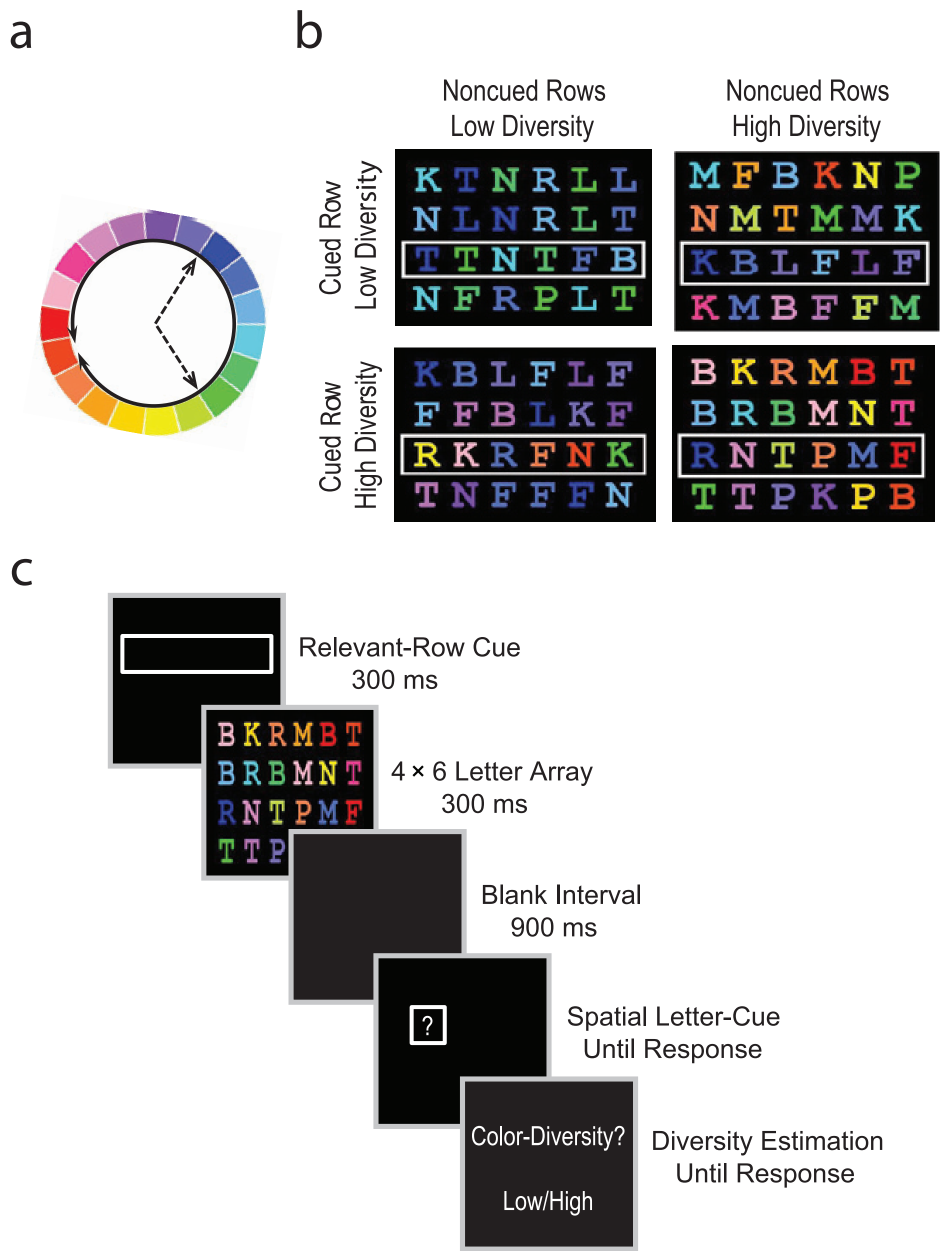


Figure 1.1. Methods from the Bronfman et al. (2014) letter-and-color dual-task a) High color diversity of uncued-row letters consisted of all 19 colors shown. Low color diversity consisted of a random subset of six adjacent colors. b) Ability to report uncued-row color diversity was tested under all four high/low combinations shown. c) Stimuli and timing of trials. Reprinted from "We see more than we can report: 'Cost free' color phenomenality outside focal attention," by Z. Z. Bronfman, N. Brezis, H. Jacobson, and M. Usher, 2014, *Psychological Science, 25*, p. 1397. Copyright 2015 by Association for Psychological Science.

In the present series of experiments, we added inattentional blindness to the Bronfman et al. paradigm by surprising participants who were completing only the letter task with a question about the color diversity of the just-viewed letter array. The surprise question was designed to probe ability to report gist when it had not been allocated any attention and had not been committed to working memory.

The present experiments and the Bronfman paradigm differed in small but significant ways. In their dual-task, Bronfman et al. were examining the cost of task-relevant gist perception. They were probing the amount of information that *can* be attended during a difficult task. The present experiments, on the other hand, examined the amount of task-irrelevant information that is in consciousness during a difficult task. To clarify this difference, I'll use the example of viewing a landscape. Bronfman et al. wanted to know how much information about a specific aspect of the landscape one *can* perceive while one completes a difficult task: How much can you report about the colors in the landscape when you're paying attention to them but also counting telephone poles? The current experiment sought to test how much of the landscape one perceives when one is only completing a difficult task. How much about the colors in the landscape do you perceive when you're only counting telephone poles? If awareness is rich, one ought to be conscious of the colors in the landscape in both cases. One ought to be conscious of color diversity in Bronfman et al. and in the following experiments.

Experiments 1, 2, and 3

We tested the hypothesis that ensemble statistics (gist) are only consciously perceived when they are allocated at least some attention. Participants performed a visual working memory task with a screen of letters; then, on a critical trial (CT1) when a change in color diversity of uncued letters occurred, they were unexpectedly asked about color diversity. This question was designed to probe awareness of color diversity in the absence of attention. Then, while participants performed the Bronfman et al. (2014) dual-task of letter report and color diversity level report, a change in either size diversity or mean size of uncued letters occurred (CT2) and they were unexpectedly asked about size diversity/mean size. The size manipulation was designed to probe awareness of task-irrelevant ensemble statistics when attention was spatially dispersed across the stimuli comprising those statistics.

2.1 Experiment 1

In Experiment 1, color diversity changed from high to low in CT1, and size diversity changed from high to low in CT2. We began with a high-to-low order because we predicted that changes in the other direction – from low to high color/size diversity – would facilitate bottom-up attention capture, creating a potential confound: If participants showed above-chance performance in the critical trials, it might have been due to bottom-up attention capture rather than to awareness outside of attention.

2.1.1 Methods

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 experiments, while being naive to the surprise manipulations. Each participant received one department-sponsored lottery ticket for the chance to win $50. Experimental approval was obtained from the Reed Institutional Review Board before data collection began.

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 x 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 6x4 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. Colors and sizes of letters in the other three rows varied throughout the experiment. In some trials, the colors of uncued-row letters were sampled randomly from 19 colors selected from around the color wheel (Fig. 1.1a; see Bronfman et al. (2014) for RGB values), and in other trials they were sampled randomly from a randomly selected six adjacent colors out of the 19. In some trials, uncued-row letter sizes were sampled randomly from sizes 40-60, and in others, all letters were size 50. Before an array was presented, a white cue box appeared where the white letter row would be. Participants made responses on a standard keyboard.

Procedure

Before a participant began the experiment, they were told they would be performing a "visual working memory" task. They were given instructions to pay attention to the white letters in each trial, which would be cued by a white rectangle, and to try to remember as many as they could.

The experiment was composed of three phases, each phase made up of four practice trials and seven or eight experimental trials (Figure 2.1). The eighth experimental trial in the first phase was Critical Trial 1 CT1), and the eighth trial in the second phase was Critical Trial 2 (CT2). In 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 (interstimulus interval/ISI), and finally a prompt that varied throughout the experiment. A 100-ms delay preceded the following trial. In each trial, the position of the cue box was selected randomly with three restrictions: 1) each letter row was cued twice (except that in the third phase, which only had seven trials, the second row was only cued once); 2) on the eighth trial, the row second-from-top was cued; and 3) in the seventh trial, the cued row was not second-from-top in Phases 1 and 2. Performance feedback was not given for any task in the experiment.

Phase 1

Phase 1 tested whether color diversity like that in Bronfman et al. (2014) would be perceived when participants were not paying attention to it. We hypothesized that color diversity performance would be at chance, suggesting that the above-chance color performance in Bronfman et al. (2014) was due to some attention being purposefully allocated to color diversity, rather than being due to "'cost free' color phenomenality" (p. 1394). Phase 1 also provided a working memory (WM) capacity for each participant, to be compared to WM during the dual-tasks in later phases.

In all trials in Phase 1, uncued-row letters had "high size diversity", meaning their sizes were sampled randomly from sizes 40-60.

Participants completed only the letter task from Bronfman et al. for four practice trials and seven experimental trials. For these 11 trials, letters in the uncued rows had "high color diversity", meaning that each letter's color was sampled randomly from 19 colors (Fig. 1.1a). After the letter array and 900-ms ISI, a question mark appeared, at which point participants typed all the letters from the cued row that they could remember (Fig. 2.1a).

After the first 11 trials, participants were presented with the first critical trial (CT1; Fig. 2.1a). CT1 tested awareness of color diversity when it was not task-relevant. In CT1, uncued-row letters had "low color diversity," i.e. their colors were sampled randomly from a randomly selected set of six adjacent colors out of the original 19. Up through the 900-ms ISI, everything else was the same in CT1 as in the previous 11 trials.

After the ISI, instead of a question mark, participants were presented with a three-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 arrays, the cued row was made of white X's, which served to direct participants' attention away from cued rows.

In one array in the 3-AFC, uncued-row letters had high color diversity, like the arrays of all trials prior to CT1. In another, uncued-row letter colors were sampled randomly from the same 6 colors as those in the just-presented array; this was the low color diversity option and the correct choice. In the final array, uncued-row letters had high color diversity, and they were upside-down. This was a "foil" array; its purpose was to "catch" guesses of sly participants who didn't perceive the color diversity change, but may have reasoned that something probably changed, as this was a psychology experiment with a surprise question. The upside-down array was drawn randomly from 10 pre-made upside-down arrays. We predicted that participants would choose at chance, indicating that they had not perceived the color diversity of the just-presented array since their attention had been occupied with the letter task.

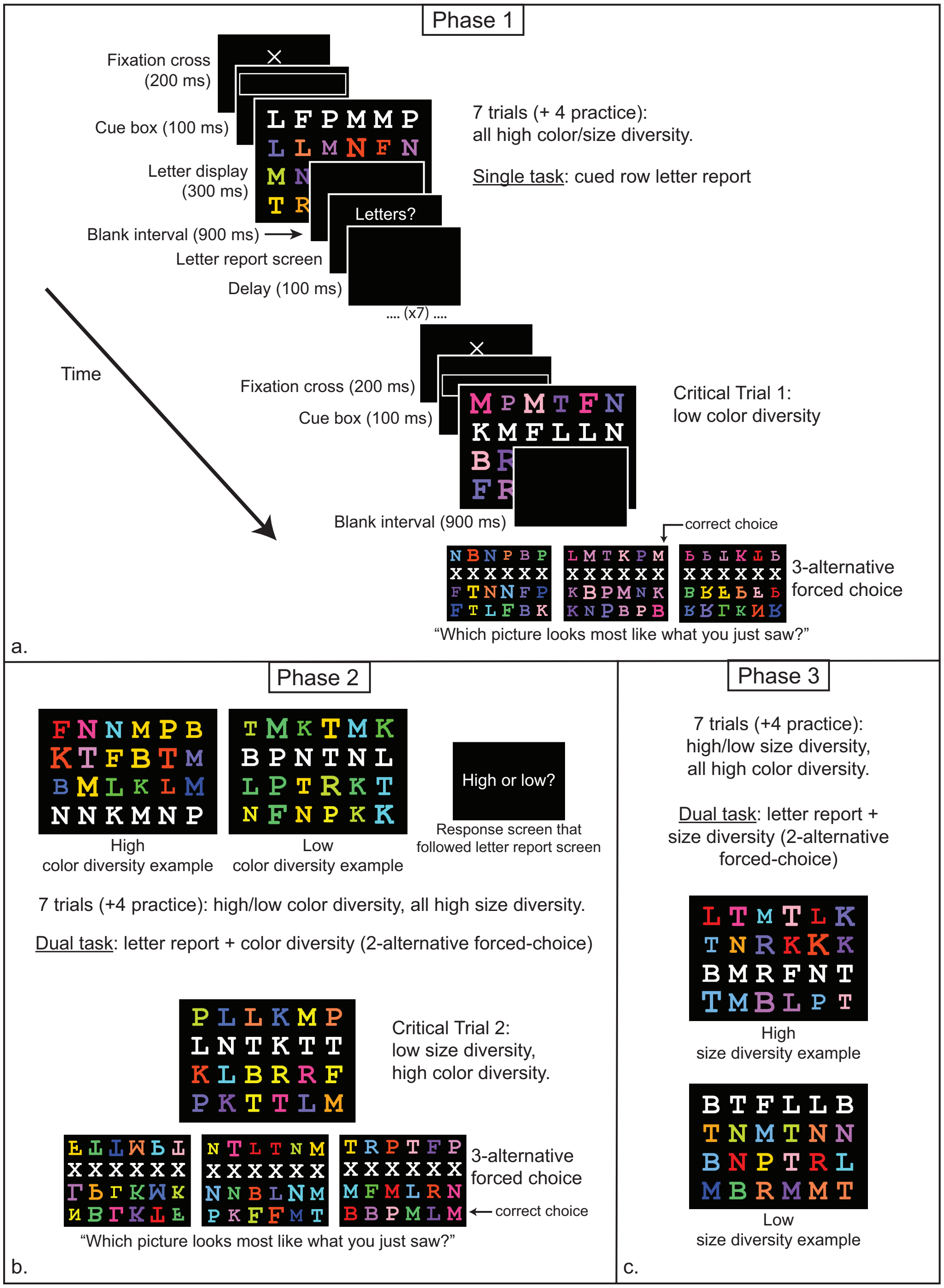


Figure 2.1. Diagram of Phases 1, 2, & 3 in Experiment 1

2.1a. Phase 1 (letter task) and Critical Trial 1 (change in, and surprise question about, color diversity) showing timing of individual trials. 2.1b. Phase 2 (dual-task: letter report and identification of color diversity level) and Critical Trial 2 (change in, and surprise question about, size diversity). 2.1c. Phase 3 (dual-task: letter report and identification of size diversity level).

Phase 2

Immediately after CT1, participants entered Phase 2 (Fig. 2.1b). This phase's purpose was to replicate the Bronfman et al. (2014) results wherein participants were able to complete the WM task and color diversity tasks simultaneously. A replication would show that participants could complete the dual-task, but only when some attention was allocated to both tasks. Phase 2 also included CT2, which probed awareness of size diversity when attention was spatially dispersed for the color diversity task.

In Phase 2, color diversity of uncued-row letters could be either high or low, and following the same question mark screen as that of Phase 1, a screen appeared asking "High or low?" Participants indicated high color diversity by pressing the up arrow on the keyboard and low color diversity by pressing the down arrow.

Before the phase began, participants were informed that they would be doing the new letter-and-color dual-task. They were shown eight color diversity example arrays, four each of high and low, and then given four practice trials of the dual-task, two each of high and low. They then completed seven experimental trials, four with low color diversity and three with high color diversity, presented in random order. Uncued-row letters in these 11 trials had high size diversity, like all trials in Phase 1.

After the first seven trials of Phase 2, participants were presented with the second critical trial (CT2; Fig. 2.1b). Uncued-row letters in this array had "low size diversity" (all size 50), and color diversity was high. Like CT1, everything else was the same in CT2 as it had been in the previous seven trials up through the 900-ms ISI. After the ISI, instead of a question mark, participants were presented with another three-alternative forced choice, similar to that in CT1.

In one array in the 3-AFC, letter sizes were sampled randomly from sizes 40-60, as in all trials prior to CT2. In another, all letters were size 50; this was the correct choice. The third array, the foil, was sampled from the same 10 foils as in CT1: letters were sizes 40-60 and upside-down. Uncued-row letters in all three arrays had high color diversity.

Phase 3

Immediately after CT2, participants entered Phase 3, the last phase of the experiment. Trials in this phase were similar to those in Phase 2, except that the second part of the dual-task (the first being letter report) consisted of reporting level of size diversity. The purpose of Phase 3 was to confirm that participants could perform the letter report and size diversity dual-task if attention was allocated to both tasks. Size diversity of uncued-row letters in this phase could be either high or low. Participants indicated high size diversity by pressing the up arrow on the keyboard and low size diversity by pressing the down arrow.

Before the phase began, participants were informed that they would be doing the new dual-task. They 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 the new dual-task. There were four low size diversity trials and three high size diversity trials, presented in random order. Color diversity of uncued-row letters was always high.

When participants finished the experiment, they were thanked and debriefed.

2.1.2 Results

Focal working memory task

Performance in the letter task decreased with the addition of a secondary task (Fig. 2.2). A one-way repeated measures ANOVA showed an effect of phase on letter performance, $F(2,98)=19.03$, $p<.01$. Dependent means t-tests revealed that letter performance in Phase 1 ($M=3.60$, $SD=0.79$) was significantly higher than that in Phase 2 ($M=3.08$, $SD=0.77$), $t(49)=6.12$, $p<.01$, and Phase 3 ($M=3.17$, $SD=0.85$), $t(49)=4.25$, $p<.01$. No difference was found between Phases 2 and 3, $t(49)=1.04$. A post hoc Tukey test confirmed the results of all three t-tests, $p<.01$.

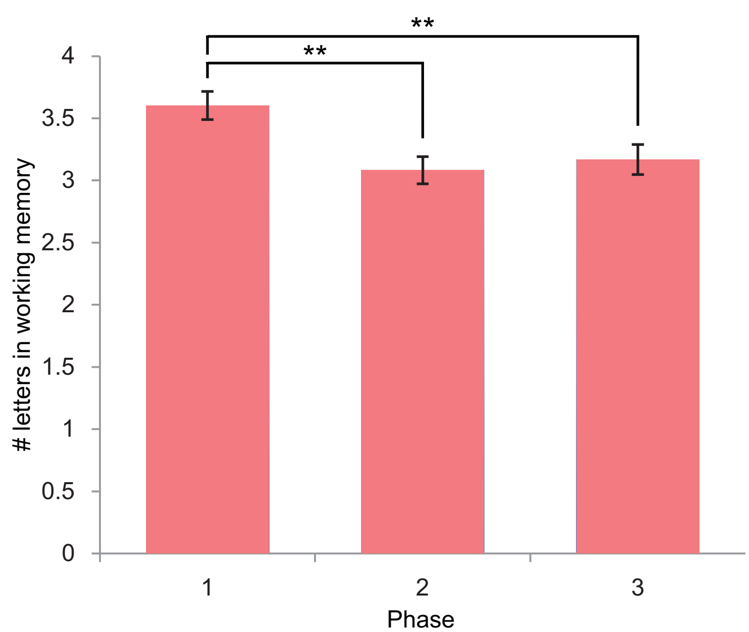


Figure 2.2. Mean letter accuracy (# correct) across phases, Exp. 1

Letter accuracy across phases. **Significant at $p<.01$. Error bars denote ±1 *SEM*.

Color and size diversity tasks

Color diversity performance in the Phase 2 dual-task (*M*=83.7%, *SD*=15.0) was significantly above chance (Fig. 2.3), $t(49)=15.89$, $p<.01$. Size diversity performance in the Phase 3 dual-task was also significantly above chance (Fig. 2.3) (*M*=67.4%, *SD*=18.3), $t(49)=6.75$, $p<.01$. Color diversity performance was significantly higher than size diversity performance, $t(49)=4.99$, $p<.01$.

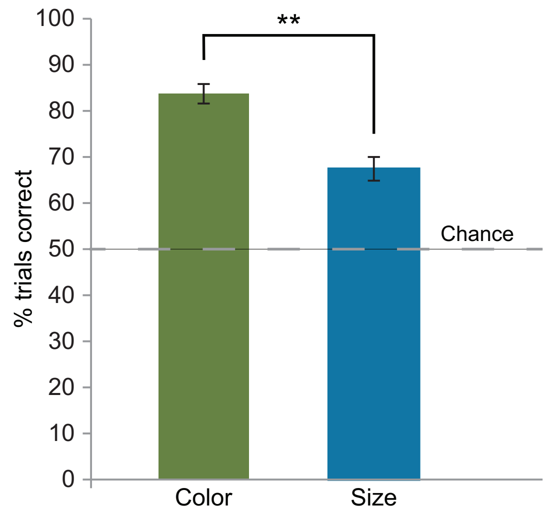


Figure 2.3. Percentage of trials correct for color diversity and size diversity, Exp. 1 Percent trials correct in the color diversity task (Phase 2) and size diversity task (Phase 3). **Significant at $p<.01$. Error bars denote ±1 *SEM*.

Critical trials

Proportion of choices in Critical Trial 1 (Fig. 2.4) deviated significantly from that expected by chance, X^2 (2, N=50) = 13.72, p<.01. Choices for color diversity of the critical trial (22%; low color diversity; correct) and the foil display (20%) were fewer than those for color diversity of Trials 1-7 and the four practice trials (58%; high color diversity; incorrect). Proportion of choices in Critical Trial 2 deviated marginally from that expected by chance, X^2 (2, N=50) = 5.32, p=0.07. Choices for size diversity of the critical trial (42%; low size diversity; correct) and for all trials before that (40%; high size diversity; incorrect) were greater in number than those for the foil display (18%).

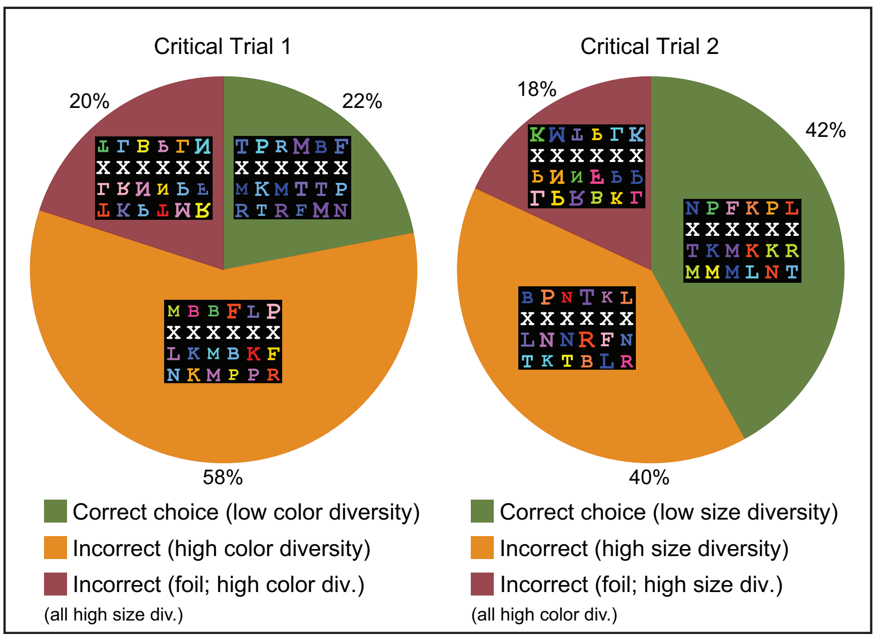


Figure 2.4. Choice percentages in Critical Trials 1 and 2, Exp. 1

Percentage chosen of each of the 3 options in the forced choices in Critical Trial 1 and Critical Trial 2. "Correct choice" was shown in the critical trial, "Incorrect" in orange is what was shown in Trials 1-7, and "Incorrect" in red is the upside-down foil array that was never shown. Examples of each array are shown in the slices.

2.1.3 Discussion

Performance in the letter task decreased from Phase 1 to Phase 2, suggesting an attentional cost of the color diversity task. Performance was constant from Phase 2 to Phase 3, suggesting that the color diversity and size diversity tasks were allocated similar amounts of attention. Performance in both the color and size diversity tasks was above chance, demonstrating that participants were able to complete both dual-tasks when the ensemble statistics were allocated some attention. The size diversity task was probably more difficult than the color diversity task, since size performance was significantly lower than color performance. Finally, performance in both critical trials was at or below chance, suggesting that conscious perception of color and size diversity did not occur in the critical trials.

In the 3-AFC in CT1, the high color diversity option was chosen more frequently than the other two. The disproportionate number of choices for high color diversity could be due to three possible causes: First, high color diversity could be more salient than low color diversity due to color contrast, leading participants to *guess* the high option more often than the low option. Second, high color diversity could have elicited bottom-up attention capture, also due to color contrast, at some point prior to CT1. Finally, participants could have perceived high color diversity at some point in the 11 trials leading up to CT1 for a reason unrelated to high color diversity per se. This perception could have occurred in practice trials before attention was focused, or by chance on one out of the 11 high color diversity trials, or unconsciously in a cumulative fashion.

Experiment 2 was motivated by the high frequency of high color diversity choices in CT1 in Experiment 1. We sought to test whether the disproportion was due to stimulus salience or bottom-up attention capture, or whether it was simply due to participants being more likely to consciously perceive color diversity in the first 11 trials than in the critical trial.

2.2 Experiment 2

In Experiment 2, color diversity changed from low to high in CT1, and size diversity changed from low to high in CT2. If we found a high frequency of high color diversity choices in CT1, it would suggest bottom-up attention capture or salience of high color diversity. If, on the other hand, we found a disproportionate number of low color diversity choices in CT1, it would suggest that participants were more likely to perceive color diversity in the 11 trials leading up to CT1 than in CT1 itself. Still, we did not predict awareness of gist *outside of attention*, as there was no evidence for it in Experiment 1. We predicted similar results for CT2 as we had found in Experiment 1.

2.2.1 Methods

Participants

Participants were 25 Reed College students, none of whom had participated in Experiment 1; see Section 2.1.1 for other details. One participant was excluded from analyses due to failure to follow instructions (final $N = 24$).

Apparatus and stimuli

See Section 2.1.2.

Procedure

Details of overall procedural structure, participant instructions, performance feedback, trial timing, and cueing can be found at the beginning of Section 2.1.3. Otherwise, the procedure of Experiment 2 is detailed below.

Phase 1

The four practice trials and first seven experimental trials in Phase 1 were identical to those of Phase 1 in Experiment 1, except for two details. First, size diversity was low (all letters were size 50) in all trials. Second, in the four practice trials and first seven experimental trials, letters in uncued rows had “low color diversity”, i.e. their colors were sampled randomly from a randomly selected set of six adjacent colors from the original 19. The six-color set was selected anew for each trial.

After the first 11 trials in Phase 1, participants were presented with the first critical trial (CT1). The structure of CT1 was the same as it had been in Experiment 1. Color diversity in this trial was high, unlike that of all previous trials. Size diversity was low.

In one array in the three-alternative forced-choice (3-AFC) following the CT1 display, uncued-row letters had low color diversity, like the arrays of all trials prior to CT1. This array was drawn randomly from 10 pre-made arrays with different six-color ranges. In another array, the foil array, uncued-row letters were upside-down. Uncued-row letters in the foil had low color diversity, but their color range was opposite the other low-color array on the color wheel. This array was drawn from 10 pre-made arrays and its selection was yoked to the selection of the first array. The final array had high color diversity; this was the correct choice. The reason for setting the two low color arrays’ ranges opposite each other was to prevent a “pop-out” effect wherein participants chose the high color diversity array more often than either of the two low color diversity arrays simply because the low color arrays, by chance, had similar color ranges. Size diversity was low in all three arrays.

Phase 2

Immediately after CT1, participants entered Phase 2. The four practice trials and first seven experimental trials in Phase 2 were similar to those of Phase 2 in Experiment 1, except for one detail: size diversity was low in these trials. As in Experiment 1, there were 4 low color diversity and 3 high color diversity trials in Phase 2, presented in random order.

After the first 11 trials of Phase 2, participants were presented with the second critical trial (CT2). The structure of CT2 was the same as it had been in Experiment 1. Unlike all previous trials, uncued-row letters in this array had high size diversity (all size 50); color diversity was high.

In one array in the 3-AFC, letters were all size 50, as in all trials prior to CT2. In another, letter sizes were sampled randomly from sizes 40-60; this was the correct choice. The third array, the foil, contained letters all of size 50, but uncued-row letters were upside-down. This array was sampled randomly from 10 pre-made arrays. Uncued-row letters in all three arrays had high color diversity.

Phase 3

Phase 3 of Experiment 2 was identical to that of Experiment 1.

When participants finished the experiment, they were thanked and debriefed.

2.2.2 Results

Focal working memory task

As in Experiment 1, a one-way repeated measures ANOVA showed an effect of phase on letter performance, $F(2, 46)=5.68$, $p<.01$ (Fig. 2.5). Dependent means t-tests revealed that letter performance in Phase 1 ($M=3.71$, $SD=0.92$) was significantly higher than that in Phase 2 ($M=3.34$, $SD=0.88$), $t(23)=3.28$, $p<.01$, and Phase 3 ($M=3.43$, $SD=0.87$), $t(23)=2.88$, $p<.05$. No difference was found between Phases 2 and 3, $t(23)=0.67$. A post hoc Tukey test confirmed the results of all three t-tests ($p<.01$ between Phases 1 and 2 and $p<.05$ between Phases 1 and 3).

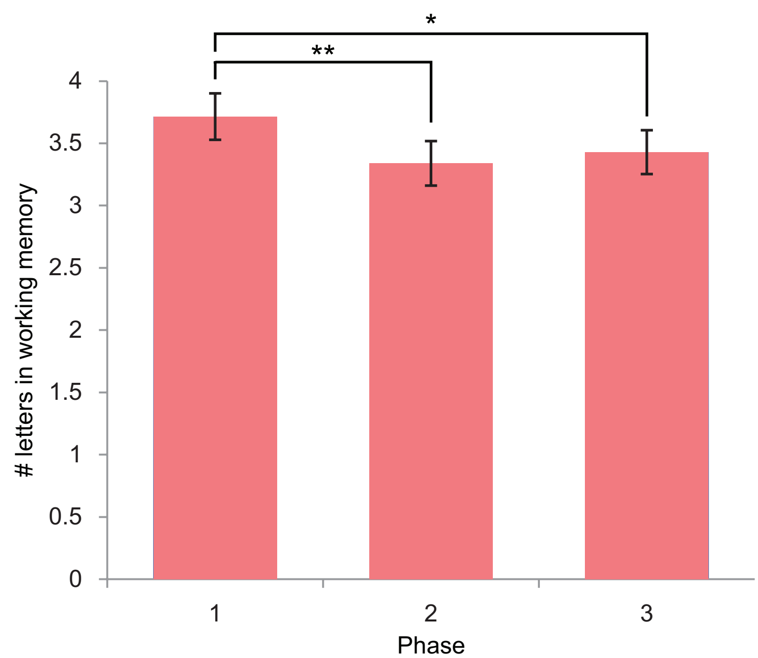


Figure 2.5. Mean letter accuracy (# correct) across phases, Exp. 2
Letter accuracy across phases. **Significant at $p<.01$. *Significant at $p<.05$. Error bars denote ±1 *SEM*.

Color and size diversity tasks

As in Experiment 1, color diversity performance (*M*=83.3%, *SD*=16.7) was above chance, $t(23)=9.79$, $p<.01$, as was size diversity performance (*M*=70.8%, *SD*=19.5), $t(23)=5.23$, $p<.01$ (Fig. 2.6). Color diversity performance was significantly higher than size diversity performance, $t(23)=2.52$, $p<.05$.

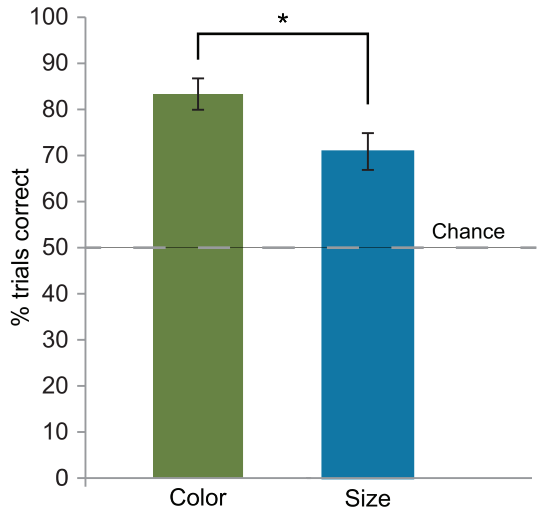


Figure 2.6. Percentage of trials correct for color diversity and size diversity, Exp. 2 Percent trials correct in the color diversity task (Phase 1) and size diversity task (Phase 2). *Significant at $p<.01$. Error bars denote ±1 *SEM*.

Critical trials

Proportion of choices in Critical Trial 1 (Fig. 2.6) did not deviate significantly from that expected by chance, X^2 (2, N=24) = 2.25. However, participants chose color diversity of the critical trial (46%; high color diversity; correct) more often than they chose color diversity of Trials 1-7 and the four practice trials (33%; low color diversity; incorrect) or the foil display (21%). Proportion of choices in Critical Trial 2 did deviate significantly from that expected by chance, X^2 (2, N=24) = 6.75, p<.05. Choices for size diversity of the critical trial (21%; high size diversity; correct) and the foil display (21%) were fewer in number than those for size diversity of all trials previous to the critical trial (58%; low size diversity; incorrect).

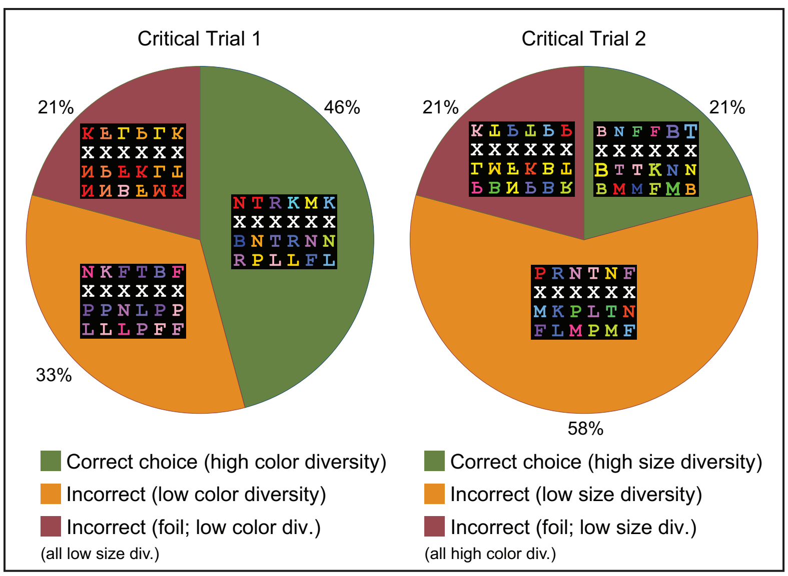


Figure 2.7. Choice percentages in Critical Trials 1 and 2, Exp. 2

Percentage chosen of each of the 3 options in the forced choices in CT1 and CT2. "Correct choice" is what was shown in the critical trial, "Incorrect" in orange is what was shown in Trials 1-7, and "Incorrect" in red is the upside-down letter array that was never shown. Examples of each array are shown in the slices.

2.2.3 Discussion

Experiment 2

Results from the letter task and dual-tasks in Experiment 1 were replicated in Experiment 2; therefore, the same conclusions can be made about the attentional costs of the ensemble statistic tasks. As in Experiment 1, performance in both critical trials was at or below chance, again suggesting that conscious perception of color and size diversity did not occur in the critical trials.

The patterns of CT choices in Experiment 2 differed from those of Experiment 1. In the 3-AFC in CT1, choices were at chance, but the high color diversity option was chosen more frequently (46%) than the two low-color options, meaning that choices trended toward correct. (Because sample size in Exp. 2 was half that of Exp. 1 (N=24 vs. N=50), it is possible that more participants would have driven CT1 performance above chance.) The high proportion of high color diversity choices prompt one of three interpretations: First, it is again possible that participants guessed high color diversity in the 3-AFC due to salience caused by color contrast. Second, they may have consciously perceived color diversity in CT1 due to bottom-up attention capture elicited by high color diversity. Third, participants may have perceived color diversity outside of attention in the critical trial. The third interpretation seems less likely if one considers the low proportion of low color diversity choices in CT1 in Experiment 1; there is not a clear reason why awareness outside attention would occur for high color diversity but not for low color diversity. Furthermore, because less than half of participants in Experiment 2 chose the correct display in CT1, it is evident that there was little conscious perception of color diversity on critical trials, with or without attention.

In the 3-AFC in CT2, participants chose low size diversity, that of the first 11 trials, more than they chose high size diversity, the correct choice. Independently, this result suggests that participants perceived size diversity at some point before CT2, but combined with Experiment 1 results, this interpretation seems less likely. More likely is that participants did not consciously perceive size diversity in lead-up trials or in the critical trial.

The following section is a brief discussion further synthesizing critical trial results from Experiments 1 and 2.

Experiments 1 and 2

The combined critical trial results from Experiments 1 and 2 allow for more conclusive interpretations of what participants consciously perceived in critical trials and in the preceding trials. Across CT1 in Experiments 1 and 2, participants chose high color diversity more often than low color diversity. In other words, the higher proportion of high color diversity choices was independent of whether it had been shown in the critical trial or in the preceding 11 trials. Therefore, either participants were more likely to notice high color diversity due to bottom-up attention capture, or they were more likely to choose it due to salience of the color contrast. The combined proportion of choices result strongly suggests that participants did not perceive color diversity outside of attention: choices only trended toward correct when CT1 color diversity was high. If participants perceived color diversity independently of attention, we would expect equal numbers of correct choices in CT1 in Experiment 1 and Experiment 2.

If we strictly consider correct CT1 choices across both experiments, the average is 34%. The most generous interpretation of this result is that only 34% of participants perceived color diversity on the critical trial, which provides further support that awareness of gist largely did not occur when it was task-irrelevant.

The combined results of CT2 in Experiments 1 and 2 are less conclusive but suggest that participants did not consciously perceive size diversity in the critical trials: High and low size diversity were chosen at equal frequencies in Experiment 1, and low size diversity (the incorrect choice) was chosen at higher frequencies than high size diversity (the correct choice) in Experiment 2. In other words, the correct display was chosen at or below chance.

The following experiment (Experiment 3) was conducted to confirm that high color diversity was chosen more often than low color diversity because of bottom-up attention capture or salience, not because it had been shown repeatedly (Exp. 1) or had just been shown (Exp. 2). Experiment 3 also tested a new size-related ensemble statistic, mean size.

2.3 Experiment 3

Experiments 1 and 2 provided strong support for the lack of conscious awareness of gist outside of attention. However, the pattern of choices in the 3-AFC in CT1 left questions about whether choices (specifically high color diversity choices) were, in fact, due to awareness of color diversity at some point: Participants chose high color diversity more often both when it had been shown repeatedly (Exp. 1), and when it had just been shown (Exp. 2). Were these biases to choose the high color diversity option due to repeated exposure to it in the first 11 trials in Experiment 1 and to bottom-up attention capture in CT1 in Experiment 2? Or, did the bias occur in the 3-AFC itself due to greater salience of high color diversity than low color diversity?

Experiment 3 sought to address the high color diversity bias by balancing high and low color diversity in lead-up trials, and presented either high or low color diversity in the critical trial. By presenting high and low an equal number of times during Phase 1, we could examine the effect of what was presented on the critical trial without the confound of having the opposite diversity level presented in all previous trials.

Experiment 3 also tested a new size ensemble statistic, mean size. It has been suggested that people are incapable of perceiving size diversity as an ensemble statistic (Haberman & Whitney, 2012), but that mean size can be consciously perceived and reported with relatively little cost. Because it was clear that participants were not aware of size diversity on CT2 in both Experiments 1 and 2, we decided to switch to mean size in order to test cost free gist perception with greater construct validity. Two levels of mean size were tested in CT2 and were balanced across participants.

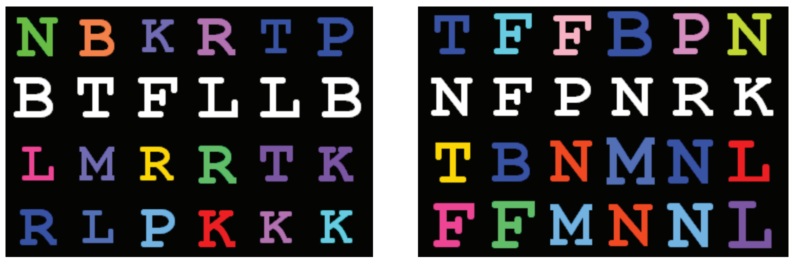
2.3.1 Methods

Participants

Participants were 30 Reed College students, none of whom had participated in Experiment 1 or 2; see Section 2.1.1 for other details.

Apparatus and stimuli

Mean size had two levels, low and high. Low mean size was made by sampling randomly from the size range 40-50 pt., (Fig. 2.8a), and high mean size was made by sampling randomly from the size range 50-60 pt. (Fig. 2.8b). “Medium” mean size (range 45-55 pt.) was used in the 3-AFC in CT1. See Section 2.1.2 for other details about apparatus and stimuli.



a. b.

Figure 2.8. Examples of high and low mean size arrays.
a. Low mean size (sampled randomly from sizes 40-50 pt.). b. High mean size (sampled randomly from sizes 50-60 pt.).

Procedure

There were two conditions in Experiment 3, which will be called Condition L and Condition H, with 15 participants in each condition. Participants in Condition L were presented with low color diversity in CT1 and low mean size in CT2. Participants in Condition H were presented with high color diversity in CT1 and high mean size in CT2.

Details of overall procedural structure, participant instructions, performance feedback, trial timing, and cueing can be found at the beginning of Section 2.1.3. Otherwise, the procedure of Experiment 3 is detailed below.

Phase 1

Unlike in the previous two experiments, both low and high color diversity were presented in trials leading up to CT1. Each color diversity level was presented twice in the four practice trials in both conditions, in random order. In Condition L, high color diversity was then presented four times and low color diversity three times in the first seven experimental trials, in random order. In Condition H, high color diversity was presented three times and low color diversity four times in the first seven experimental trials, in random order.

Low and high mean size were each presented twice in practice trials in both conditions, and four times in the eight experimental trials in both conditions (the eighth trial being the CT1 display), always in random order.

After the first 11 trials in Phase 1, participants were presented with the first critical trial (CT1). The structure of CT1 was the same as in the previous experiments.

Mean size in the 3-AFC was medium in both conditions so participants would have equal exposure to high and low mean size at the time of CT2. Because the 3-AFC in CT1 was presented until response, participants would have been exposed to either high or low mean size for a relatively long period of time. This may have biased them toward choosing the size mean from the 3-AFC in CT1 during the 3-AFC in CT2.

Condition L

In CT1, uncued-row letters had low color diversity. Mean size was randomly either high or low. In one array in the 3-AFC following the CT1 display, uncued-row letters had low color diversity using the same six colors as the CT1 display; this was the correct choice. In another array, uncued-row letters had high color diversity, an incorrect choice. In the final array, the foil array, uncued-row letters had high color diversity and were upside-down. The foil was drawn randomly from 10 pre-made arrays. Note that the foil and the incorrect choice had the same color diversity.

Condition H

In CT1, uncued-row letters had high color diversity. Mean size was randomly either high or low. In one array in the 3-AFC, uncued-row letters had high color diversity, like that in CT1; this was the correct choice. In another array, uncued-row letters had low color diversity, an incorrect choice. This array was drawn randomly from 10 pre-made arrays with different six-color ranges. In the final array, the foil array, uncued-row letters were upside-down. Uncued-row letters in the foil had low color diversity, but their color range was opposite the other low-color array on the color wheel. This array was drawn from 10 pre-made arrays and its selection was yoked to the selection of the first array. Note that the foil and the incorrect choice had the same color diversity.

Phase 2

Immediately after CT1, participants entered Phase 2, which employed the same letter-and-color dual-task as in Experiments 1 and 2. Each color diversity level and mean size level was presented twice in the four practice trials in both conditions, in random order and independent of each other. Both low and high mean size were presented in the seven experimental trials leading up to CT2. In both conditions, high color diversity was presented three times and low color diversity was presented four times in the seven experimental trials leading up to CT2.

After the first 11 trials in Phase 2, participants were presented with the second critical trial (CT2). CT2 color diversity and color diversity of the arrays in the 3-AFC were high in both conditions.

CT2 was otherwise different for participants in the L and H conditions.

Condition L

In CT2, uncued-row letters had low mean size. In one array in the 3-AFC, uncued-row letters had low mean size; this was the correct choice. In another array, uncued-row letters had high mean size, an incorrect choice. In the foil array, uncued-row letters had high mean size and were upside-down. The foil was drawn randomly from 10 pre-made arrays.

Condition H

In CT2, uncued-row letters had high mean size. In one array in the 3-AFC, uncued-row letters had high mean size; this was the correct choice. In another array, uncued-row letters had low mean size, an incorrect choice. In the foil array, uncued-row letters had low mean size and were upside-down. The foil was drawn randomly from 10 pre-made arrays.

Phase 3

Phase 3 was similar to that in Experiments 1 and 2, except that the second part of the dual-task was reporting level of mean size. High mean size was presented four times and low mean size three times for Condition L participants in the seven practice trials; the opposite was true for Condition H participants. Color diversity was always high.

When participants finished the experiment, they were thanked and debriefed.

2.3.2 Results

Focal working memory task

A one-way repeated measures ANOVA showed a significant effect of phase on letter performance, $F(2, 58)=7.18$, $p<.01$. Dependent means t-tests revealed that letter performance in Phase 1 ($M=3.64$, $SD=1.07$) was significantly higher than that in Phase 2 ($M=3.17$, $SD=1.12$), $t(29)=2.47$, $p<.01$. Letter performance in Phase 3 ($M=3.28$, $SD=1.36$) was marginally higher than that in Phase 2, $t(29)=2.09$, $p=.05$. No difference was found between Phases 1 and 3, $t(29)=1.71$.

Color diversity and mean size tasks

Color diversity performance ($M=81.9\%$, $SD=14.5$) was significantly above chance, $t(29)=12.05$, $p<.01$, as was mean size performance ($M=62.9\%$, $SD=17.8$), $t(29)=3.95$, $p<.01$. Color diversity performance was significantly higher than mean size performance, $t(29)=4.38$, $p<.01$.

Critical trials

To assess awareness of color diversity and mean size in critical trials independently of whether high or low color diversity/high or low mean size was presented, the results of critical trials in conditions L and H were combined into correct, incorrect, and foil choices. Then, to better assess what was perceived on CT1, and to assess response biases on the 3-AFC while controlling for exposure during the lead-up trials (equal exposure to high and low diversity), we analyzed the choices for Conditions L and H separately (N=15 in each case).

Total

Proportion of choices in Critical Trial 1 (Fig. 2.9) did not deviate significantly from that expected by chance, X^2 (2, N=30) = 4.20, though there were more correct (47%) than incorrect (37%) choices. Proportion of choices in Critical Trial 2 (Fig. 2.9) deviated significantly from that expected by chance, X^2 (2, N=30) = 7.20, p<.03. Incorrect choices (53%) were more frequent than correct choices (33%).

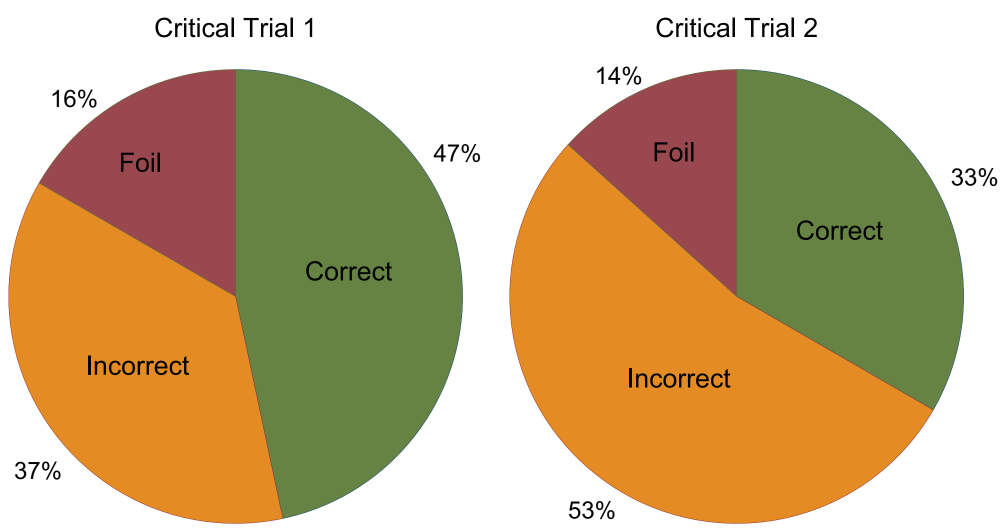


Figure 2.9. Choices in Critical Trials 1 and 2, Exp. 3, conditions combined
Percentage chosen of the correct option, the incorrect option, and the foil in CT1 and CT2. "Correct choice" is what was shown on the critical trial (and previous trials), "Incorrect" in orange is what was shown only in previous trials, and "Incorrect" in red is the upside-down letter array that was never shown.

Condition L

Proportion of choices in Critical Trial 1 (Fig. 2.10) did not deviate significantly from that expected by chance, X^2 (2, N=15) = 3.60. However, participants chose low color diversity (14%, correct) less often than they chose high color diversity (53%, incorrect) or the foil display (33%). Proportion of choices in Critical Trial 2 (Fig. 2.10) also did not deviate significantly from that expected by chance, X^2 (2, N=15) = 3.60. However, choices for low mean size (33%, correct) and the foil display (14%) were fewer in number than those for high mean size (58%, incorrect).

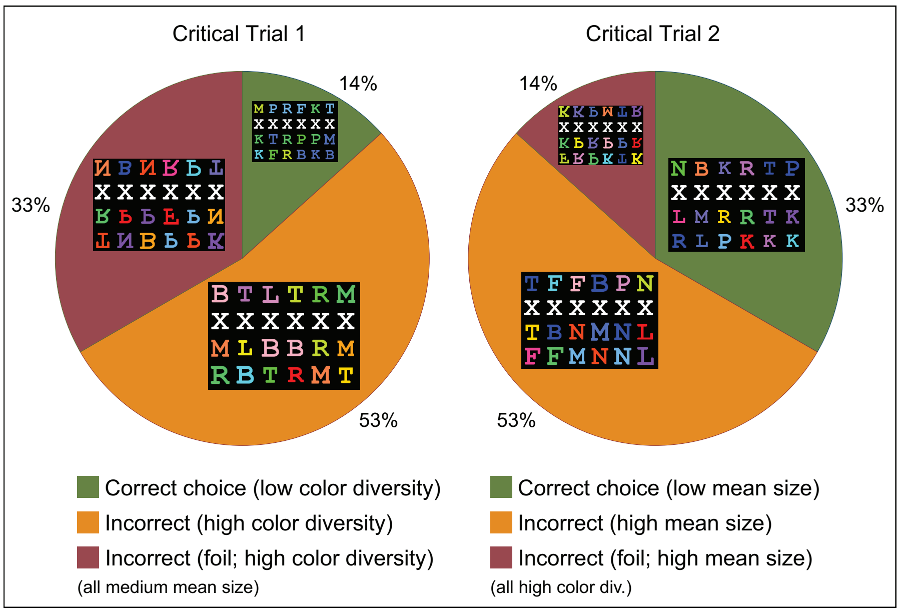


Figure 2.10. Choice percentages in Critical Trials 1 and 2, Exp. 3, Condition L Percentage chosen of each of the 3 options in the forced choices in CT1 and CT2. “Correct choice” is what was shown on the critical trial (and previous trials), “Incorrect” in orange is what was shown only in previous trials, and “Incorrect” in red is the upside-down letter array that was never shown.

Condition H

Proportion of choices in Critical Trial 1 (Fig. 2.11) deviated significantly from that expected by chance, X^2 (2, N=15) = 15.60, p<.01. A greater proportion of participants chose high color diversity (80%, correct) than the incorrect low color diversity (20%) No participant chose the foil display. Proportion of choices in Critical Trial 2 (Fig. 2.11) did not deviate significantly from that expected by chance, X^2 (2, N=15) = 3.60. However, choices for high mean size (33%, correct) and the foil display (14%) were fewer in number than those for low mean size (58%, incorrect).

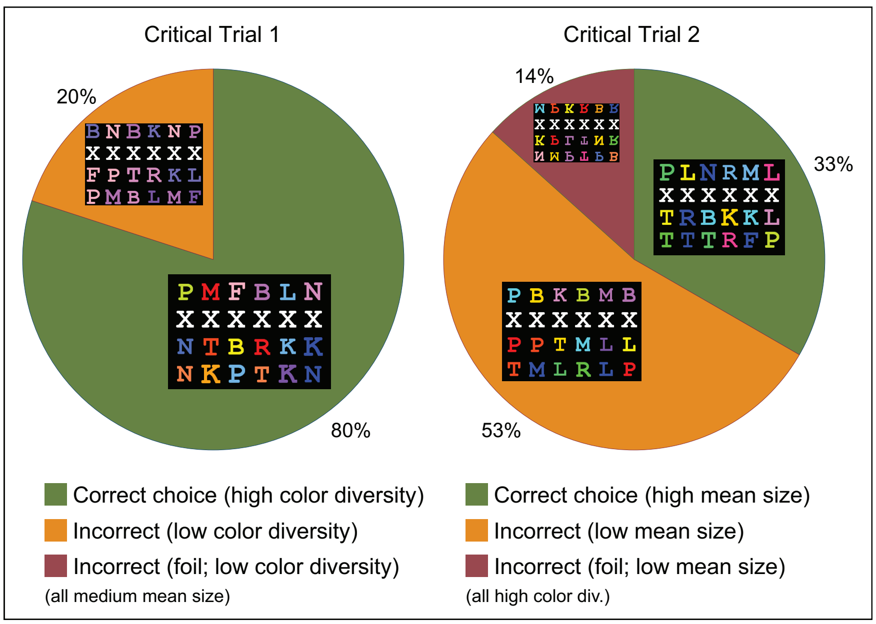


Figure 2.11. Choice percentages in Critical Trials 1 and 2, Exp. 3, Condition H Percentage chosen of each of the 3 options in the forced choices in CT1 and CT2. "Correct choice" is what was shown on the critical trial (and previous trials), "Incorrect" in orange is what was shown only in previous trials, and "Incorrect" in red is the upside-down letter array that was never shown.

2.3.3 Discussion

Results from the Experiment 3 letter task were not an exact replication of results from Experiments 1 and 2. However, the pattern of results was the same: performance was highest in Phase 1, dropped in Phase 2, and then rose slightly in Phase 3, but remained lower than that in Phase 1.

Above-chance performance in the color and size dual-tasks was replicated in Experiment 3; therefore, similar conclusions can be made about the attentional costs of the ensemble statistic tasks.

Performance in the mean size task (62.9%) was lower than performance in the size diversity task in Experiments 1 (67.4%) and 2 (70.8%), suggesting that it was a more difficult ensemble statistic to perceive and report in conjunction with the focal task. This increased difficulty of the size task was not predicted, since it has been suggested that individuals can perceive mean size but not size diversity (Haberman & Whitney, 2012). One possible interpretation of lower mean size performance is that participants did not perceive size diversity per se in the dual-task in Experiments 1 and 2; rather, perhaps they looked for the presence of a size difference in a small number of letters. This strategy would have worked because “low size diversity” was comprised of letters of all the same size (see Supplemental Discussion for an explanation of this manipulation). Therefore, if a size difference was not perceived in a few letters, participants could have chosen low size diversity and shown above-chance performance.

When CT1 results from both conditions were pooled into a correct/incorrect/foil sample, choices did not deviate significantly from chance, supporting lack of awareness of color diversity on a trial-by-trial basis when it was not task-relevant.

The pattern of choices in CT1 in both conditions provided support for the hypothesis that high color diversity was chosen more often than low color diversity because of a combination of bottom-up attention capture and stimulus salience, rather than repeated exposure or phenomenal awareness outside attention. When high and low color diversity had been presented an equal number of times before the critical trial, participants were more likely to choose high color diversity, regardless of what was actually presented on the critical trial. Stimulus salience of high color diversity in the 3-AFC may also have played a role in the proportion of high color choices; see Section 3.4 for a future manipulation that would isolate the effect of salience.

Experiment 3 did provide some evidence that participants consciously perceived high color diversity on CT1: When high color diversity was presented on CT1, 80% of participants chose it. When it was not presented in CT1, only 53% chose it. However, keep in mind that the foil was always made of the incorrect color diversity. Therefore, if we interpret CT1 results as about 80% of participants choosing high color diversity in both conditions (53% high + 33% foil in Condition L, and 80% high + 0% foil in Condition H), it suggests that participants chose high color diversity at equal rates independently of whether it was presented in CT1. Again, however, these results do not conclusively show that high color diversity was perceived in the absence of attention, as seeing the high color diversity (whether in the practice, lead-up trials, or CT1) was probably due to bottom-up attention capture.

In the 3-AFC in CT2, Condition H participants chose low mean size more often (incorrect), and Condition L participants chose high mean size more often (incorrect). This result suggests that participants did not consciously perceive mean size, even with attention dispersed for the color task, and were choosing at chance. This chance performance in a task-irrelevant size statistic in the presence of dispersed attention builds on our previous findings and suggests two ideas: 1) Perception is not rich, and 2) size ensemble statistics are unlikely to be perceived during a difficult task, even via bottom-up attention capture. When CT2 results were pooled, choices deviated significantly from chance but in the direction of incorrect, demonstrating no awareness of mean size when attention was spatially dispersed but mean size was task-irrelevant.

General Discussion

3.1 Results

Overall, no evidence was found for cost-free phenomenal awareness outside of focal or dispersed attention in Experiments 1, 2, and 3. Participants were not aware of color diversity, size diversity or mean size when these ensemble statistics were not task-relevant. Furthermore, performance in the focal task dropped with the addition of both gist tasks, demonstrating that there was an associated attentional cost of the secondary tasks.

3.1.1 Attentional costs of the secondary task

Unlike the results in Bronfman et al. (2014), the present experiments suggest small attentional trade-offs between the letter and color/size tasks. Letter performance in the present experiments decreased with the addition of the color diversity task: Accuracy dropped from 3.60 to 3.08 in Exp. 1, from 3.71 to 3.34 in Exp. 2, and from 3.64 to 3.17 in Experiment 3, all significant decreases. In Bronfman et al.'s Experiment 2, letter performance was 2.94 in the single task and 3.08 in the dual-task, a nonsignificant difference. This apparent discrepancy in the cost of the secondary task may be explained by differences in trial quantity and performance feedback. Participants in Bronfman et al.'s Experiment 2 completed 70 single-task practice trials, 130 single-task experimental trials, and 200 dual-task experimental trials, compared to the present experiments' 11 trials total in each of the single and dual-tasks. It is likely that, over the course of 200 dual-task trials, Bronfman et al. participants learned to allocate a minimal amount of attention to the color task, such that no decrease in overall letter performance occurred. It is reasonable to assume that participants in the present experiments allocated more attention to the secondary task, since they had less experience with it, which adversely affected letter report accuracy in the primary task. It is possible that in Bronfman et al., a greater attentional allocation may have also occurred during the first 10-20 dual-task trials, but that over the course of the remaining 180-190 dual-task trials, participants learned to divide their attention such that there was no measurable decrease in letter performance.

Further supporting the minimal-attention hypothesis is the fact that participants in Bronfman et al. received feedback for letter performance, while the current experiments provided no such feedback. In addition, Bronfman et al. included an escape button in the color task for participants to use if they were unsure of color diversity. These procedural differences mean that participants in Bronfman et al. were likely more motivated to only allocate a minimal amount of attention to the color task, since they experienced consequences for their letter responses (performance feedback), and were given reason to believe that their color performance was less consequential (via the escape button).

Finally, color diversity judgments were more accurate in the dual-task in the present experiments (about 83%) than in Bronfman et al.'s experiment (66%). This difference further suggests that more attention was allocated to color diversity in the present experiments. Again, it is likely that, due to the limited number of trials in the present experiments and lack of performance feedback on the letter task, participants did not learn to or were not motivated to allocate only a minimal amount of attention to color diversity.

These results support the hypothesis of small attentional trade-offs rather than cost-free phenomenality in both Bronfman et al. and in the present experiments.

3.1.2 Inattentional blindness to color diversity

The present experiments found that, during a difficult task, high rates of inattentional blindness to color diversity occur when color diversity is task-irrelevant. In CT1 in Experiment 1, choices deviated from chance but in the opposite direction of what would be expected if participants had perceived color diversity on the critical trial: Participants chose high color diversity (what had been presented in earlier trials) significantly more often than low color diversity (the correct choice). This suggests that the color diversity of the CT1 array was not perceived, at least by a majority of participants. In CT1 in Experiment 2, choices did not deviate significantly from chance. In CT1 in Experiment 3, choices were only above chance in Condition H. The proportion of correct choices in CT1 with all experiments and conditions combined was about 41%, which is comparable to other inattentional blindness findings (Cohen et al., 2011; Mack & Clarke, 2012).

It is important to note that, because our critical trials were forced-choice, it is highly unlikely that all correct choices reflected awareness. Our foil displays may provide a rough estimate of percentage of guesses. The portion of foil choices was consistently about 20%. Therefore, it is possible that the percentage of participants who consciously perceived color diversity on the critical trials is close to 20% (41% - 20% = 21%). However, as was suggested in the Experiment 3 discussion, perhaps participants only chose based on color diversity, meaning that the foil display actually acted as one of two high color or low color diversity choices.

The present finding of lack of awareness for task-irrelevant color diversity is incongruent with the Bronfman et al. (2014) claim of cost-free phenomenal awareness outside of focal attention. While Bronfman et al. did not posit awareness outside of attention altogether, they did suggest that rich phenomenal color awareness occurs without a cost to focal tasks. If it is true that color phenomenality occurs with no cost to focal attention, one might predict greater accuracy in the 3-AFC following critical trials, given that it was a recognition test and occurred only 900 ms after the offset of the display. The rates of inattentional blindness found do not support Bronfman et al.'s claims. Rather, they suggest that task-relevance is crucial for the awareness of color diversity outside of focal attention when a difficult task is being performed.

3.1.3 Disproportionate number of choices of the high color diversity display

In CT1 across all three experiments, high color diversity was chosen more frequently than low color diversity, independently of whether it was the correct choice. One possible reason for the disproportionate frequency of high color diversity choices is that high color diversity led to higher rates of awareness due to bottom-up attention capture. This interpretation could hold for all three experiments, given that participants chose high color diversity more often than low color diversity in all three. Another possibility for the disproportionate frequency of high color diversity choices in CT1 trials is that participants were guessing, but that they were more likely to pick the high color diversity display than the low color diversity display due to the higher stimulus salience engendered by greater diversity.

3.1.4 Inattentional blindness in the presence of dispersed attention

The at-or-below-chance performance in CT2 in all three experiments suggests that, even when attention was spatially dispersed across the letter array, only the ensemble statistic that was task-relevant (color diversity) was in awareness. If dispersed attention were sufficient for awareness of gist to occur, as was suggested by Bronfman et al. (2014), one would predict above-chance performance in CT2, demonstrating awareness of size diversity/mean size when it was not task-relevant. A sparse consciousness interpretation of the CT2 results is that, because of the high attentional load posed by the letter-and-color dual-task, there was little attention left for awareness of size diversity/mean size to occur. Although size performance was lower than color diversity performance overall, suggesting that the size tasks were more difficult, letter performance did not differ between Phases 2 and 3 in either experiment, suggesting that a comparable amount of attention was given to the secondary task in both cases. In other words, given that participants could complete the size task with a comparable amount of attention to the color task, the rich phenomenal awareness of Bronfman et al. would predict size diversity to be in awareness in Critical Trial 2, given that dispersed attention was present. The fact that it was not provides further support for sparse consciousness in terms of gist.

3.1.5 Summary

The present results contrast with the conclusions of Bronfman et al. (2014). There appears to be a small attentional cost of reporting color diversity or size diversity/mean size while completing a focal working memory task, such that higher performance in one leads to lower performance in the other. Furthermore, awareness of color diversity or size diversity/mean size was absent when participants completed a difficult focal task or a difficult focal+dispersed attention dual-task. It appears that, when a difficult task is being completed, conscious perception of ensemble statistics does not occur when they are not task-relevant, even when they are spatially attended.

3.2 Methodological relevance to inattentional blindness data

The methods detailed above deviate from similar inattentional blindness (IB) paradigms in terms of stimulus duration, masking, and IB probing procedure. This section summarizes the methods of those paradigms and discusses the ways in which the present methodology differs and thereby contributes to the existing IB literature.

Cohen et al. (2011) had participants complete a focal multiple object tracking task (MOT) while the background, consisting of pattern masks, changed every 67 ms. On a critical trial, the second-to-last mask was replaced by a natural scene. After the final mask, awareness of the natural scene was immediately probed with a line of questioning. Cohen et al. found 64% inattentional blindness for the gist of the scene. Using a similar paradigm in which participants counted the occurrences of an alphanumeric character during rapid serial visual presentation (RSVP), while the background changed every 100 ms, they found 50% IB for the gist of a natural scene replacing the second-to-last background mask. When natural scene classification was added as a secondary task to the MOT and RSVP tasks, focal task performance diminished significantly, while gist performance was almost perfect (96% in MOT and 93% in RSVP).

Mack and Clarke (2012) asked participants to complete a focal task of identifying the long arm of a cross; the cross plus a mosaic pattern were shown for 100 ms and backward masked for 500 ms. On a critical trial, the mosaic was replaced with a scene and, following the 500-ms mask, awareness of the gist of the scene was probed with questioning. When the scene was presented in the center of the screen, Mack and Clarke found 75% IB, and when it was presented peripherally, they found 85% IB. When scene identification was part of a dual-task (scene and cross arm identification), performance was significantly above chance across experiments (82%, 50%, 57%, 80%). Furthermore, cross performance was inversely related to scene performance in the dual-task. Finally, Mack et al. found that presenting the critical scene twice (once in the critical trial and once in an earlier trial) did not decrease the incidence of IB.

The present experiments are similar to the Cohen et al. (2011) and Mack and Clarke (2012) studies in terms of investigating inattentional blindness to gist, but several methodological differences are worth consideration. First, we did not employ masking. In addition, our stimuli were presented for a full 300 ms, which is 3 times longer than the 100-ms presentations of Cohen et al. and Mack and Clarke. We found rates of inattentional blindness comparable to those found in both Cohen et al. and Mack and Clarke, along with high gist performance when gist became task-relevant. The comparable IB rates and dual-task performance provide support for the validity of our paradigm in terms of testing inattentional blindness, which strengthens our findings that people can be inattentionally blind to the color and size diversity/mean size of unmasked stimuli presented for a full 300 ms under conditions of high working memory load.

Another methodological difference between the present experiments and those discussed above is our use of immediate gist recognition, rather than gist report followed by a recognition test. Both Cohen et al. and Mack and Clarke presented questions immediately after critical trials, only giving a forced-choice recognition test if questioning did not produce accurate gist report. We argue that failure to recognize gist is stronger evidence for inattentional blindness than is failure to report gist, as a weak memory or familiarity could potentially aid recognition without aiding free recall. The present experiments contribute to current findings on IB for gist by lowering the threshold of awareness and/or memory that is necessary for accurate performance.

Also meriting discussion is that, in our forced choice following critical trials, one option always resembled the gist of noncritical trials. Neither the Cohen et al. nor the Mack and Clarke paradigm included a noncritical display (pattern mask/mosaic) as one of the options. Providing participants with the option of choosing the gist of previous displays allowed us to probe whether gist awareness had occurred in any manner prior to the critical trials. It is possible that if Mack and Clarke and Cohen et al. had included the noncritical displays (mosaic noise patterns) as options in the critical trial recognition test, they would have found relatively high percentages of participants choosing those displays. In both previous studies as well as in the present study, inattentional blindness to a *change* in gist was tested rather than inattentional blindness to the *presence* of gist. When natural scenes or colorful displays of letters are used as stimuli, these salient stimuli tend to automatically capture attention, preventing inattentional blindness to the presence of gist per se. To circumvent bottom-up attention capture, the two previous studies as well as the present study tested inattentional blindness for *changes* in gist.

Our Experiment 3 methods are similar to those of Experiment 4 in Mack and Clarke, in which the critical scene was presented on a noncritical trial in addition to on the critical trial. In our experiment, half of the trials leading up to each critical trial had the same gist as the critical trial (color diversity for CT1 and size diversity for CT2). Our manipulation, like that of Mack et al., allowed us to explore whether seeing one gist formulation multiple times would contribute to its being in awareness.

Finally, our operationalization of gist was modeled after that of Bronfman et al. (2014): we measured gist awareness by manipulating specific ensemble statistics, while keeping other aspects of the display constant. Cohen et al. and Mack and Clarke, on the other hand, tested awareness of a "setting" of an image when it replaced a mask/mosaic. At this point it is unclear what functional differences exist between gist awareness for individual ensemble statistics and for gist awareness natural scenes, but it is a difference worth exploring in future studies.

The differences discussed above highlight the contributions that our experiments make to an understanding of the role of attention in gist awareness.

3.3 Alternate interpretations

3.3.1 Awareness without attention

In order to explore alternate interpretations of the present results, it is necessary to return to the theories discussed earlier involving the richness of consciousness, specifically whether it overflows that which can be attended, stored in working memory, and reported (Block, 2005, 2011; Cohen & Dennett, 2011). Recall that Block (2005) argued for a distinction between phenomenal consciousness, which is rich and largely independent of working memory and attention, and access consciousness, the phenomenal content that is attended, transferred to working memory, and therefore reportable. Our paradigm for probing the richness of consciousness relied on report, which, in the context of the phenomenal/access distinction, may have revealed reportable conscious contents rather than phenomenally conscious contents per se. It is possible that participants in the present experiments were always phenomenally conscious of gist (the ensemble statistics of color diversity, size diversity, and mean size), but could not access it for report when it was not attended and held in working memory. Through a rich consciousness interpretation, a lack of *accessibility* of gist may be what was revealed during the critical trials: When participants were unexpectedly asked about color diversity and size diversity/mean size, perhaps they had consciously perceived these statistics when the display was presented, but could not access them for report.

Rich consciousness is also compatible with the decrease in letter performance in the dual-tasks in the present experiments. It is possible that decreased performance reflects the resources involved in storing and reporting gist, rather than perceiving it per se.

3.3.2 Rich awareness with dispersed attention

In a recent paper (2011) and in his 2014 review of Bronfman et al. (2014), Block extended the argument for rich phenomenal awareness by addressing the role it plays in task-relevant gist report. Gist, he argued, is a summary that occurs *after* rich detail enters consciousness, not before. In other words, because of the limitations of working memory, only gist can be reported when a scene is attended, but rich detail (everything that comprises gist) is consciously perceived. This claim relates to the claim of rich awareness outside attention because it argues that consciousness overflows what can be reported. It differs in that it addresses what occurs when more information is allocated task-relevant attention than can be transferred to working memory.

Block (2011) reviewed fragile visual short-term memory (fVSTM) as a potential mechanism by which report can be used to probe phenomenal consciousness underlying task-relevant gist. He cited Vandenbroucke, Sligte, and Lamme (2011), who employed a paradigm similar to Sperling's but with lines of different orientations. They found the capacity to report orientation with a post-cue occurring from one to four seconds after array offset. Vandenbroucke et al. and Block interpreted this capacity as being due to fVSTM, a conscious form of memory. Therefore, Block argued, rich phenomenal content can be probed with fVSTM. The sparse view of consciousness would argue that, after stimulus offset, the line orientations were all being processed unconsciously until one of them was fully attended.

It is difficult, however, to use fVSTM as an alternate explanation for the present results, and for inattentional blindness results in general, since fVSTM seems to require attention. According to Block (2011), fVSTM persists for up to a few seconds. If fVSTM occurs independently of attention, one would predict little to no evidence of inattentional blindness for up to a few seconds after stimulus offset. In terms of the present experiments, if gist had persisted in fVSTM during critical trials, one would predict above-chance performance on the forced-choices; this was not the case. It seems reasonable in light of these results and other inattentional blindness research to posit that fVSTM, if it exists, requires attention to the stimuli – or relevant aspects of the stimuli – about which questions are asked. If fVSTM requires attention, then it is not a mechanism by which phenomenal consciousness outside of attention can be probed with report. Furthermore, recent evidence suggests that there is a cost to storing visual information even for a brief amount of time, such that fVSTM or its likeness does not completely overflow access (Mack, Erol, & Clarke, 2015).

3.4 Future manipulations

There are a number of possible manipulations that would clarify and expand upon the present results. First, in order to tease out the effect of salience on high color diversity choices in CT1, it would be fruitful to run a Phase 1 in which only low color diversity was displayed, but to present the same options in the CT1 3-AFC as in the present experiments: a high color diversity display, a low color diversity display (with the same six colors as in CT1), and a high color diversity display with upside-down letters. This manipulation would eliminate high color choices due to bottom-up attention capture in lead-up trials, thereby allowing analysis of the effect of salience on the disproportionate number of high color choices.

Another helpful manipulation would be to shorten the 900-ms ISI in critical trials in order to have a more sensitive probe of awareness. It is possible that the length of the ISI was longer than the duration of conscious memory in the present paradigm. By shortening the ISI, any persisting visual memory may be more likely to be picked up by the 3-AFC. A limitation associated with this manipulation is that retinal iconic memory, which can occur unconsciously, lasts up to a few hundred milliseconds.

A third manipulation might test the limits of inattentional blindness by making cued letters color diverse as well as uncued-row letters (like the displays in Bronfman et al., 2014), and testing participants' awareness of a change in color diversity of cued letters. This manipulation is analogous to CT2 in the present experiments, in that a spatially attended, task-irrelevant ensemble statistic is tested.

Finally, two manipulations might be carried out to show whether accurate color diversity/size diversity/mean size report when these statistics are task-relevant actually reflects rich awareness, as was suggested by Bronfman et al. First, one could manipulate colors or sizes of just a few individual letters to test whether gist judgments are made from many letters or from just a few. Second, one could manipulate colors or sizes of letters closest to the cued row to see if there is a bias toward judging gist from those letters, which would suggest that perceiving gist farther from focal attention is more costly than perceiving it close to focal attention.

Conclusions

Here I return to the theory and issues discussed in the introduction, and place the present experimental findings within the context of current literature. Recall that there is controversy not only over whether consciousness is rich or sparse, but (relatedly) over whether it merits an explanation that gives great weight to subjective experience, or whether it should come to be defined solely in terms of neural mechanisms (Lamme, 2006) or cognitive and behavioral function (Cohen & Dennett, 2011). Our human intuition is that subjectivity is a mysterious beast that science has yet to (or cannot) successfully contain. However, perhaps we will soon find that, like evolution, reproduction, and the dying of a star, all that needs explanation is the mechanism or function. Perhaps attempts to find evidence of subjectivity that overflows report are analogous to a scientific search for God, in that the impression of rich consciousness can be explained scientifically, but rich consciousness itself may or may not exist and cannot be studied using the scientific method. In addition, perhaps attempting to explain subjectivity distracts from the important task of establishing a scientifically tractable definition of consciousness, and using that as a tool to evolutionarily and functionally explain consciousness.

Despite the aforementioned uncertainties, the scientific search for awareness that overflows report is directly relevant to the present experiments. We continued this search through the window of visual ensemble statistics, or gist. We used a form of report, although, again, there could exist conscious contents that are completely unreportable and have no behavioral effects, making report an ineffective method for probing consciousness (Cohen & Dennett, 2011). The present paradigm employed immediate gist recognition as a proxy for awareness, with the assumption that an immediate recognition task would be a relatively sensitive measure of awareness persisting in some form of memory outside of attention (e.g. in fragile visual short-term memory, Block, 2005). Our paradigm was developed from that of Bronfman et al. (2014). They found that, in a focal-task+gist report dual-task, there was no performance cost associated with above-chance gist report. We changed their paradigm by unexpectedly asking about gist, in order to more realistically simulate day-to-day life in which the gist of one's surroundings is not necessarily task-relevant. Not only did we find at-chance gist recognition performance when gist was outside of attention, but we also failed to replicate Bronfman et al.'s finding of cost-free gist awareness/report when it was allocated some attention.

It is important to note that starting assumptions of the Bronfman et al. paradigm and of the present experiments differed, and likely contributed to hypothesis-formation and methodology, thereby affecting results. It may be through these differences in assumptions, not necessarily differences in results per se, that consciousness theory and research becomes polarized. Bronfman et al. probably began with the assumption that consciousness is rich. Gist was always task-relevant in their paradigm, which eliminated the potential for inattentional blindness/amnesia. They gave each participant hundreds of trials, probably so that participants would learn to allocate only a minimal amount of attention to color diversity. By repeating the task hundreds of times, participants were able to keep their performance on the letter task steady while still judging color diversity at above-chance accuracy. These are all reasonable manipulations that led Bronfman et al. to the conclusion that rich, cost-free phenomenal awareness exists outside focal attention.

The present experiments began with the assumption that consciousness is sparse and depends upon attention. To test this assumption, we looked for evidence of gist awareness outside of attention altogether. We did not view inattentional blindness as a confound, as we viewed our surprise recognition task as a sufficiently sensitive measure of awareness. In addition, each participant completed only a few trials in which gist was task-relevant, which simulated awareness of gist in everyday life. Excessive experience with the color and size dual-tasks may have led to perceptual or strategic learning aiding performance, thereby causing an incorrect conclusion of cost-free awareness. It is clear that differences in theoretical beliefs behind the two paradigms played a role in their hypotheses, methods, results, and conclusions.

Again, perhaps the discrepancy between intuitions about consciousness and the evidence is an intractable problem, at least via the scientific method. This intractability may be due to circularity, as was suggested in the preceding paragraphs: Results can often be interpreted to demonstrate both rich and sparse consciousness, and the interpretation depends upon starting assumptions about what evidence does or does not constitute consciousness. To prevent this circularity, the field may need to develop a more coherent definition of consciousness and implement new methodologies that do not prioritize subjective experience to such an extent.

For the time being, the present results demand an interpretation within the context of existing literature. Our results do not support the existence of rich phenomenal awareness. We presented participants with a 3-alternative forced-choice gist recognition test only 900 ms after they viewed a 300-ms, unmasked display, and performance was at chance. One of the most vocal proponents of rich consciousness, Ned Block, proposed fragile visual short-term memory (fVSTM) as a mechanism by which rich awareness can be preserved long enough to test it (2011). Because our 900-ms delay was well within the proposed duration of fVSTM, participants should have had above-chance performance in critical trials. They did not. This null result suggests that fVSTM, if it exists, depends upon attention, which brings us back to the earlier claim that awareness requires attention.

Of course, the possibility remains that accurate gist report in the presence of attention reflects rich underlying phenomenal consciousness, as proposed by Block. The problem with this interpretation, again, is that the same behavioral results are predicted if accurate gist report is due to unconscious processing of ensemble statistics, rather than to rich phenomenal awareness. Because of the demonstrated capacity of unconscious processing (Dehaene et al., 2006; Kouider et al., 2010), this sparse interpretation is more precedented than a rich one.

The inattentional blindness to gist demonstrated in the present experiments, in conjunction with the cost associated with gist report, suggests that awareness is inextricably linked to attention and is therefore costly. It doesn’t come for free.

Supplemental discussion

This section covers important methodological and theoretical considerations that do not directly relate to the main hypotheses, methods, or findings of the present experiments.

5.1 Procedural divergences from Bronfman et al. (2014)

A brief discussion is provided here of differences between the methods of the present experiments and those of Bronfman et al. (2014). First, our cued-row letters were white rather than color-diverse like those in Bronfman et al. The change was made in order to shorten experiment duration without introducing confounds: Bronfman et al. found an influence of cued-row color diversity on uncued-row diversity judgment (methods in Fig. 1.1b; see Bronfman et al. (2014) for results). This influence would have likely been present in our experiments, requiring more trials and conditions per participant in order to isolate its effects. We therefore made cued rows white, which eliminated any potential influence of focal color diversity on nonfocal color diversity judgment. We believe the switch to white cued rows was justified because Bronfman et al. found above-chance performance on uncued-row color diversity judgment independent of cued-row color diversity, suggesting that awareness of color diversity outside focal attention occurred independently of focal color. One might argue, however, that our white cued row letters facilitated inattentional blindness (IB) in Critical Trial 1 in all three experiments. If IB was indeed facilitated, lower rates of IB may have been found with the use of color-diverse cued rows. Future studies should explore the effects of focal/nonfocal color contrast on awareness of color outside focal attention.

In order to balance out the attention capture induced by white cued-row letters, we shortened the duration of the cue box from 300 ms (Bronfman et al., 2014) to 100 ms. A 300-400 ms stimulus onset asynchrony (SOA; length of time from cue to stimulus) results in a shorter reaction time (RT) in an RT task than does a 100-ms SOA (Posner, 1980), suggesting that a “sweet spot” of attention capture exists at about a 350-ms SOA. We moved the SOA away from the sweet spot, which likely hindered attention capture, thereby providing participants with more of a “chance” to be aware of color diversity. Our inattentional blindness results are arguably made stronger given the brief amount of time participants had to focus their attention on cued-row location before the letter array appeared.

The briefness of our experiments required two more procedural changes. First, participants used a standard keyboard rather than a 9-key response box; we assumed that preexisting keyboard proficiency was more valuable than the specificity of a response box, given the small number of trials participants would have had to become comfortable with it. Second, participants typed all the cued letters they remembered rather than being asked to recall a specific cued letter (the Bronfman et al. method; see Fig. 1.1c). This change may have led to a less accurate assay of working memory (WM) since letter repeats often occurred in a single cued row, but there were not enough trials in our experiments to calculate WM capacity based on accurate recall of a single letter per trial.

Finally, letter report performance was not given in the present experiments, again due to experiment briefness. The lack of feedback is important considering differences between our results and the results of Bronfman et al. (2014); see main discussion.

5.2 Methodological concerns

5.2.1 Low color diversity

Because of the way color is perceived, color diversity in the Bronfman et al. (2014) methods, and therefore in the present experiments, may have been confounded with average color. High color diversity was sampled from 19 colors from around the color wheel, while low color diversity was sampled from 6 adjacent colors from within those 19. The 6 low-diversity colors may have created a perceptible average color, while the average color of high diversity displays would have likely been brown or imperceptible to the human eye. Bronfman et al. found that people could unconsciously report average color, but could not unconsciously report color diversity, suggesting that average color perception is a lower-level process than is color diversity perception. Therefore, in the letter+color dual-task, participants may have used the rule “can compute average color/cannot compute average color” to decide on high or low color diversity, which may have made the task easier and could have potentially occurred unconsciously.

5.2.2 Low size diversity

In Experiments 1 and 2, “low” size diversity was actually zero size diversity; all letters were size 50. Piloting of the program revealed limitations that necessitated this metric: Keeping the mean size the same (size 50) and keeping the overall size of the letter array the same (about 9 cm x 6 cm) imposed limitations on the maximum range of letter sizes for high size diversity; letters of sizes greater than 60 significantly occluded each other. So, the greatest size diversity possible was the range 40-60. Then, there was a minimum range necessary for any size diversity at all to be reasonably detectable. When the threshold range was reached that the presence of any size diversity would be detected above chance, discrimination of this range versus the range 40-60 was tested, and performance was not reliably above chance. So, the decision was made to make low size diversity actually zero size diversity.

One might argue that discrimination between zero and any size diversity does not actually demonstrate perception of ensemble statistics, since one could detect only two letters of different sizes and rule out zero size diversity. However, in Experiment 1, size diversity switched from high to low (zero) in Critical Trial 2. In order to consciously perceive a switch to low size diversity, one would arguably have had to have perceived the “gist” of the zero size diversity array. This argument is conceptually similar to the data for pop-out stimuli: it takes people longer to say that something isn’t there than to say that it is. The argument is that participants had to perceive more information in order to identify a change to no size diversity than to identify high size diversity. At any rate, we still found high rates of IB in Critical Trial 2 in both Experiments 1 and 2.

Whether we were actually testing gist awareness or not (in other words, whether or not performance in Phase 3 in Experiments 1 and 2 was based on gist awareness or simply on pop-out stimuli – this could explain the data which was above chance but still not high), high rates of IB in Critical Trial 2 demonstrated that people weren't perceiving anything about size, either pop-out stimuli or the size diversity ensemble statistic.

Bibliography

Bachmann, T. & Hudetz, A. G. (2014). It is time to combine the two main traditions in the research on the neural correlates of consciousness: C = L x D. *Frontiers in Psychology, 5*, 1-13. doi: 10.3389/fpsyg.2014.00940

Baddeley, A. (1992). Working memory. *Science, 31*, 556-559. doi: 10.1126/science.1736359

Blackmore, S. (2012). *Consciousness: An introduction* (2nd ed.). New York, NY: Oxford University Press.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences, 9*, 46-52. doi: 10.1016/j.tics.2004.12.006

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

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

Bronfmann, Z. Z., Brezis, N., Jacobson, H., & Usher, M. We see more than we can report: "Cost free" color phenomenality outside focal attention [Supplemental Material]. *Psychological Science, 25*, 1394-1403. doi: 10.1177/0956797614532656

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*, 200-219. url: http://ingentaconnect.com/content/imp/jcs/1995/00000002/00000003/653

Chen, H. & Wyble, B. (2015). Amnesia for object attributes: Failure to report attended information that had just reached conscious awareness. *Psychological Science, 26*, 203-210. doi: 10.1177/0956797614560648

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

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

de Gardelle, V., Sackur, J., & Kouider, S. (2009). Perceptual illusions in brief visual presentations. *Consciousness and Cognition, 18*, 569-577. doi: 10.1016/j.concog.2009.03.002

Dehaene, S., Changeux, J.P., Naccache, L, Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*, 204-211.

Dennett, D. (1988). Quining qualia. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Contemporary Science* (381-414). Oxford, United Kingdom: Oxford University Press.

Dennett, D. (1996). Facing backwards on the problem of consciousness. *Journal of Consciousness Studies, 3*, 4-6. http://www.ingentaconnect.com/content/imp/jcs/1996/00000003/00000001/674

Fahrenfort, J. J. & Lamme, V. A. F. (2012). A true science of consciousness explains phenomenology: Comment on Cohen and Dennett. *Trends in Cognitive Sciences, 16*, 138-139. doi: 10.1016/j.tics.2012.01.004

Haberman, J. & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. M. Wolfe & L. C. Robertson (Eds.), *From Perception to Consciousness: Searching with Anne Treisman* (339-349). New York: Oxford University Press.

James, W. *Principles of Psychology*. New York, NY: Henry Holt and Company.

Kouider, S, de Gardelle, V., Sackur, J., and Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences, 14*, 301-307. doi: 10.1016/j.tics.2010.04.006

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7*, 12-18. doi: 10.1016/S1364-6613(02)00013-X

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

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

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

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

McConkie, G. W. & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. Perception & Psychophysics, 17, 578-586. doi: 10.3758/BF03203972

Neisser, U. & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology, 7*, 480-494. doi: 10.1016/0010-0285(75)90019-5

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 587-595. doi: 10.1037/0278-7393.15.4.587

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25. url: http://www.ncbi.nlm.nih.gov/pubmed/7367577

Putnam, H. (1975). The nature of mental states. *Mind, Language and Reality: Philosophical Papers, Volume 2* (429-440). Cambridge, United Kingdom: Cambridge University Press.

Schwartz, J. M., Stapp, H. P., & Beauregard, M. (2005). Quantum physics in neuroscience and psychology: A neurophysical model of mind-brain interaction.

Philosophical Transactions of the Royal Society B-Biological Sciences, 360, 1309-1327. doi: 10.1098/rstb.2004.1598

Simons, D. J. & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 9*, 1059-1074. PMID: 10694957

Simons, D. J. & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*, 261-267. doi: 10.1016/S1364-6613(97)01080-2

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*, 1-29. doi: 10.1037/h0093759

Vandenbroucke, A. R. E., Sligte, I. G., & Lamme, V. A. F. (2011). Manipulations of attention dissociate fragile visual short-term memory from visual working memory. *Neuropsychologia, 49*, 1559-1568. doi: 10.1016/j.neuropsychologia.2010.12.044