Journal of Experimental Psychology: General

ISSN: 0096-3445

2021, Vol. 150, No. 9, 1695–1709
https://doi.org/10.1037/xge0000864

Characterizing a Snapshot of Perceptual Experience

Michael A. Cohen[1, 2], Caroline Ostrand[1], Nicole Frontero[1], and Phuong-Nghi Pham[3]
[1] Department of Psychology, Amherst College
[2] McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology
[3] Department of Psychology, Ryerson University

What can we perceive in a single glance of the visual world? Although this question appears rather simple, answering it has been remarkably difficult and controversial. Traditionally, researchers have tried to infer the nature of perceptual experience by examining how many objects and what types of objects are not fully encoded within a scene (e.g., failing to notice a bowl disappearing/changing). Here, we took a different approach and asked how much we could alter an entire scene before observers noticed those global alterations. Surprisingly, we found that observers could fixate on a scene for hundreds of milliseconds yet routinely fail to notice drastic changes to that scene (e.g., scrambling the periphery so no object can be identified, putting the center of 1 scene on the background of another scene). In addition, we also found that as observers allocate more attention to their periphery, their ability to notice these changes to a scene increases. Together, these results show that although a single snapshot of perceptual experience can be remarkably impoverished, it is also not a fixed constant and is likely to be continuously changing from moment to moment depending on attention.

Keywords: vision, perception, attention, scenes, awareness

Supplemental materials: https://doi.org/10.1037/xge0000864.supp

What is the resolution of perceptual awareness? What can observers perceive "in the blink of an eye?" To answer these questions, researchers have often used paradigms like change blindness and inattentional blindness to determine how much information observers fail to notice in a scene (Mack & Rock, 1998; Rensink et al., 1997; Simons & Chabris, 1998). In a typical experiment, individual items within a scene change or appear in some unexpected manner (i.e., a person's sweater will change color, a chair will disappear, a man in a gorilla suit will appear, etc.) and the frequencies with which observers notice those items are measured. Because so many observers routinely fail to notice these occurrences, many researchers have argued that observers are aware of surprisingly little of the world around them and perception is rather sparse (Baars, 2002; Cohen et al., 2012; Dehaene & Changeux, 2011; Dennett, 1991; Simons et al., 2005), though, of course, this view is extremely controversial (Block, 2007; Haun et al., 2017; Koch & Tsuchiya, 2007; Lamme, 2003).

Although these results are striking and informative, it is widely believed that there are firm limits as to how much information can go unnoticed when looking upon a scene. For example, in a typical change blindness or inattentional blindness experiment, the item that changes or appears is usually relatively simple and its change or appearance does not fundamentally alter the meaning (or "gist") of the scene. Indeed, it has been routinely claimed that if a change does affect the gist of the scene, that change will be easily noticed (Koch & Tsuchiya, 2007; Mack & Rock, 1998; Rensink, 2000; Sampanes et al., 2008; Simons & Levin, 1997; van Boxtel et al., 2010; but see Cohen et al., 2011 and Mack & Clarke, 2012). Intuitively, this idea makes quite a bit of sense: while it is easy to imagine a person failing to notice the keys on the dining room table, it is difficult to imagine how they could fail to notice that their dining room has been changed into a glacier (Block, 2007; Lamme, 2003).

The idea that perception of the scene as a whole is robust, even if individual items in that scene are often not encoded, is bolstered by numerous results that have shown how several dimensions of a scene can be extracted from just a single fixation. For example, with just a few hundred milliseconds of exposure, observers are able to perceive the gist of a scene, identify a few prominent objects within that scene, and describe a variety of properties such

---

Michael A. Cohen, Caroline Ostrand, Nicole Frontero, and Phuong-Nghi Pham designed research; Caroline Ostrand, Nicole Frontero, and Phuong-Nghi Pham performed research; all authors analyzed the data; Michael A. Cohen wrote the original manuscript; and all authors contributed to the revisions to the submitted manuscript and the published article.

This study was supported by a National Science Foundation Collaborative Research Award (1829470) to Michael A. Cohen. Thanks to Rachel Albert, Thomas Botch, Sarah Cormiea, Leyla Isik, Nancy Kanwisher, Julio Martinez, N. Ratan Murty, Ken Nakayama, and Caroline Robertson for extensive discussion and feedback on the project. Thanks to Sebastian Montesinos for help with data collection. Portions of the results reported here were presented at the Annual Meeting of the Vision Sciences Society, May 2019, Florida, United States and were published as a preprint on PsyArXiv. All stimuli and data are available for download from the Open Science Framework (https://osf.io/6c3xw/).

Correspondence concerning this article should be addressed to Michael A. Cohen, Department of Psychology, Amherst College, D208 Science Center, Amherst, MA 01002, United States. Email: michaelthecohen@gmail.com

as the navigability, depth, and temperature of a scene (Castelhano & Henderson, 2008; Fei-Fei et al., 2007; Greene & Oliva, 2009; Intraub, 1981; Potter, 1976; Torralbo et al., 2013; Walther et al., 2009). Together, these results raise a natural question: If observers routinely fail to perceive individual objects within a scene but are typically aware of the more global properties of the scene around them, what exactly are the limits of perceptual awareness in natural scenes? How much can observers truly encode in a single "snapshot" of perceptual experience?

To answer these questions, we took a novel approach. Rather than simply alter individual objects within a scene and measure if observers noticed those alterations, we asked how much of an entire scene we could alter before observers noticed. By examining when observers notice different types of alterations to images of natural scenes, we can directly examine the fidelity and resolution of visual cognition and perceptual experience.

The approach used in this set of studies was directly inspired by a series of pioneering experiments done by Rayner and colleagues that focused on the size of the perceptual span when reading (McConkie & Rayner, 1976; Rayner, 1998; Rayner, 2014). In these studies, the researchers used a gaze-continent technique such that only the part of the text that an observer was fixating on was presented normally, whereas the text just beyond the fovea as degraded (i.e., replaced by strings of *X*s, out of focus, etc.). In this case, rather than move the window of fixation through the visual field, we locked it in place in the middle of the display and kept our observers' eyes fixated on that location. While preserving the foveal region of each image, we sought to determine how much of a natural scene we could degrade beyond the fovea without observers noticing. Although the gaze-continent technique has since been widely used to separate foveal and peripheral processing in a variety of situations such as visual search (Pomplun et al., 2001; Reingold et al., 2001) and scene perception (Loschky & McConkie, 2002; Miellet et al., 2010; Võ & Henderson, 2011), the results reported here are the first to use this approach to directly probe the limits of perceptual experience.

Across three experiments, we used a modified inattentional blindness paradigm to determine how often observers will notice different types of alterations to images of natural scenes while performing a variety of perceptual tasks. In each experiment, we employed the following basic procedure. Observers performed a primary task for a few trials that involved looking at large images of scenes (i.e., 26° of visual angle) and doing a basic object detection or scene classification task. Then, on the last trial, without observers knowing anything different would occur, we presented a critical stimulus that had been altered in the periphery in one of several ways. For example, in some cases, the periphery was so scrambled that no object could be detected or identified. In another case, the center part of one image (e.g., a table in a dining room) was surrounded in the periphery by an inconsistent image (e.g., a glacier). In all of these cases, the only part of the images that was unaltered corresponded to the fovea/parafovea, which is a radius of approximately 4° of visual angle (Millodot, 2014; Polyak, 1941; Swanson & Fish, 1995). To put that number in perspective, the fovea/parafovea comprises less than 1% of the visual field in everyday life and comprises roughly 7.4% of the image presented in the experiment.[1] Thus, our alterations distorted an overwhelming majority of the stimuli (92.6%). How often would such drastic alterations be noticed?

Surprisingly, we found that many observers could fixate upon an image for several hundred milliseconds without noticing a wide variety of radical alterations made to the stimulus in the periphery. For example, a large portion of observers were completely oblivious that a scene they had just gazed upon was so degraded in the periphery that no single object could be recognized or identified. In addition, a majority of observers were totally unaware of the fact that they had looked directly at a scene in which the fovea/parafovea region came from one image (i.e., a dining room), while the periphery of that image came from an entirely different image (i.e., a glacier). Thus, contrary to prior claims that observers are always aware of the gist of the scene and can rapidly extract a variety of dimensions about a scene, we find several situations where this is simply not the case. Indeed, in our most drastic conditions, observers are not just unaware of individual items within a scene; they are unaware of the fundamental changes we have made to the global properties of the entire scene.

However, it should be stressed that even though observers were not aware of these extreme alterations, they were still able to process enough of the stimulus to pick it out of a lineup (e.g., "I saw that woman right there, not any of those other people"; see Figure 1). In other words, observers could fixate on a scene, process enough of that scene to identify the item they had just fixated upon, but still fail to realize that the vast majority of the scene was completely scrambled or replaced by another scene altogether.

Finally, it should also be stressed that the resolution of perception we are describing here does not hold in all situations at all times. In another experiment, we found that the size of these effects varied significantly as a function of the way participants allocated their attention across the scene. In our first two experiments, participants performed a simple object detection task at fixation (i.e., Is a human face present or absent?). In the third experiment, participants performed a task that drew their attention more toward the periphery (i.e., Is the image of an indoor or outdoor scene?). In this case, we found that when attending more toward the periphery, observers were more likely to notice the alterations to the critical stimuli.

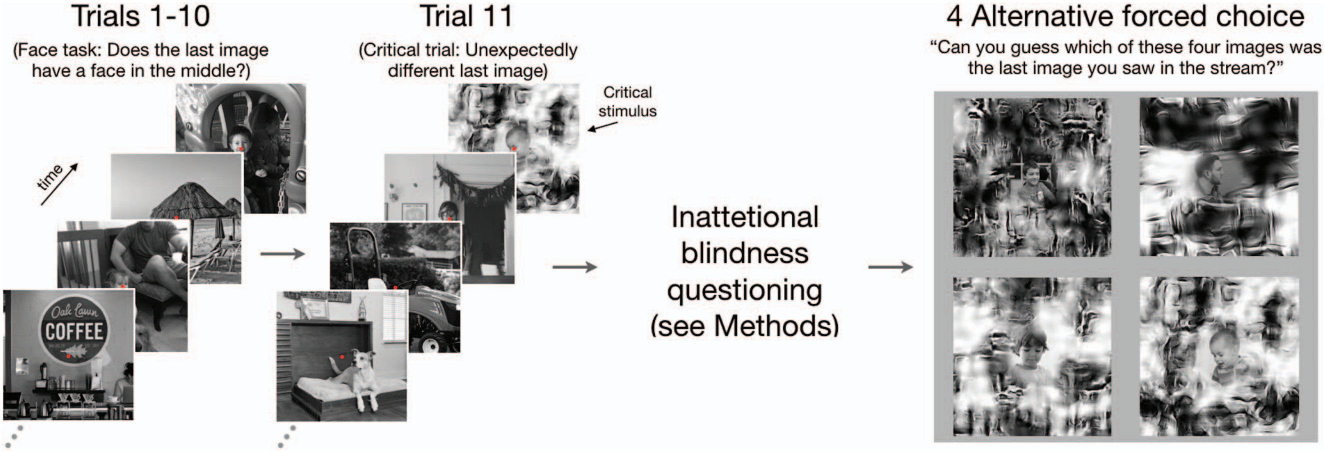
Taken together, these results highlight how a single snapshot of perceptual experience may be far more impoverished than is widely believed. Although it has previously been shown that individual objects or items can be changed or replaced without observers noticing, these are the first results showing how an entire scene can be radically altered without observers noticing. Moreover, the fact that these effects varied as a function of attention also highlights how the resolution of perceptual experience is not a static constant. These results have serious implications for many theories of visual cognition and perceptual awareness.

---

[1] Because the full visual field extends roughly 110° to the left and right of fixation (i.e., temporally) and roughly 60° above fixation and 70° to 75° below fixation, our preserved circular aperture made up an incredibly small portion of the visual field (Dagnelie, 2011; Frisén, 1990; Rönne, 1915; Savino & Danesh-Meyer, 2012).

Figure 1

Procedures for Experiments 1 and 2



Note. Participants performed 10 trials in which they determined whether the last image in the stream did or did not contain a human face in the middle. Then, on Trial 11, an unexpected critical stimulus was presented at the end of the trial and participants were immediately probed to determine whether they noticed the critical stimulus. Finally, participants were shown four examples of different critical stimuli and were asked to choose which image they saw at the end of the last trial (in most cases, participants reported picking the image based on the preserved/middle part of the image). Images with faces shown here are in the public domain and freely available (https://commons.wikimedia.org); images in the original stimulus set shown to participants are available via the Open Science Framework (https://osf.io/6c3xw/). See the online article for the color version of this figure.

Experiment 1: Scrambled Periphery

Method

Here, we asked how much of the periphery we could scramble before observers would notice. In this case, we systematically altered the extent to which the periphery of an image was scrambled while preserving certain basic natural scene statistics. At its most extreme level, images were so severely scrambled that individual objects in the periphery were completely unidentifiable. How often will observers notice such drastic alterations?

Participants

Forty participants were run in each of the three experimental conditions, resulting in a total of 120 participants. All participants were recruited from the Amherst College and Massachusetts Institute of Technology communities. All participants gave informed consent and were compensated with either $5.00 or course credit. All experimental procedures were approved by the Committee on the Use of Human Subjects in Research under the Institutional Review Board of Amherst College.

Procedure

Participants performed a simple face detection task at fixation. On each trial, they were shown between seven and 30 images of natural scenes and reported whether the last image they saw in the stream contained a face in the middle of it (see Figure 1). The number of images shown on every trial was randomly sampled from between seven and 30. Each image was shown for 288 ms, with a 100-ms gap in between. We chose a presentation rate of 288 ms per item because an analysis of several eye-tracking studies revealed that this is approximately the duration of a single fixation period in naturalistic viewing conditions (Castelhano & Heaven, 2011; Hayhoe et al., 2003; Henderson, 2003; Nuthmann, 2017; Pelz & Canosa, 2001; Rayner, 1998; Rayner et al., 2008; Unema et al., 2005; see the Appendix). Regardless of whether the trial had a face at the end of it, there were anywhere between two and five images that had faces in the middle of them throughout the stream. Participants were explicitly told that this would happen and that they were only to say whether the last image in a given trial had a human face in the middle. All images were randomly selected from trial to trial from the different stimulus folders (i.e., "face targets" and "no face distractors"). No image was ever seen twice within a single trial, although images could be seen multiple times across trials. For the first 10 trials, half of the trials had a face target present at the end and half did not. At the end of each trial, a screen appeared that prompted the participant to say whether the last image had a human face in the middle. Participants pressed one key to report that a target face was present and another key to report that no target face was present.

On the critical trials, the selection procedure for the number of stimuli used in the entire trial (i.e., seven to 30) and the number of face target stimuli used throughout the trial (i.e., two to five) was identical to all previous trials. The main difference was primarily that the very last image was a critical stimulus. Half of the participants were shown a critical stimulus with a human face in the middle and the other half were shown a critical stimulus that did not have a human face in the middle. As soon as the critical stimulus disappeared, the display went blank. No text was presented to prompt the participants to give any response. Instead, the moment the trial ended, the experimenter immediately began asking the following series of questions to the participant:

1. "Did you notice anything strange or different about that last trial?"
2. "If I were to tell you there was something different about that last trial, could you tell me what it was?"
3. "If I were to tell you there was something different about the very last image you saw on that last trial, could you tell me what it was?"

If participants could not correctly identify the critical alteration after any of those questions, the trial ended with a four alternative forced choice task that asked them to identify the specific image they saw. In this case, observers would be shown four stimuli, one of which was the exact image that had been presented on the critical trial. All four images had been altered in the same way (i.e., four images that were medium scramble, four images that had an abstract periphery, etc.). In addition, all four images would either all contain a face target in the middle or none of them would contain a face target in the middle depending on the specific stimulus participants saw on the last trial. Upon being shown those four images, participants were told what had happened on the last trial and were asked to confirm whether they had noticed. Only those participants who claimed to not notice the critical manipulation when seeing the four images were classified as being completely inattentionally blind and were reported in the main figures.[2] Finally, participants were asked to select the image they saw on the last trial by pressing one of four different keys on the keyboard. Showing four example images like this served a few purposes. First, it allowed us to show participants the alterations we made to the images and explicitly verify that they did not notice those modifications. Second, it is possible that many of our participants were unaware of the critical manipulations because they did not notice the last image at all for one reason or another (e.g., sneezing, blinking, or a phone ringing at an inopportune time). Thus, we wanted to determine how often participants did not notice the alterations but did process and remember the critical stimulus.

Stimuli

All stimuli were shown on a 27-in. Apple iMac Model A1419 computer with a refresh rate of 60 Hz and 2560 pixels × 1440 pixels. A chin-and-forehead rest was used to minimize head movements and maintain a constant viewing distance of 57 cm. The experiments were created and controlled using MATLAB 2018a (MathWorks, Inc., Natick, MA) and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997). The background color of the screen was gray, with a luminance of 68.5 cd/m^2. Each individual stimulus subtended 26° × 26° of visual angle. For the critical stimuli, the center part of the image was preserved by a flat top Gaussian that subtended 4° of visual angle (radius). The flat top portion of the Gaussian subtended 2.5° of visual angle (radius), which approximately corresponds to the fovea and was entirely unaltered. Meanwhile, the remaining 1.5° of visual angle were the slopes of the Gaussian that served as a gradual fade from the untouched center of the image to the altered periphery of the image. This sloping/faded region comprised the remaining portion of the preserved region and would land approximately on the parafoveal part of the retina. All initial stimuli were acquired through online resources (i.e., Google) and personal photograph collections of the authors and were then converted to grayscale (See Figure 1 in the online supplemental material for information on stimulus generation procedures).

Face Task Stimuli. For the face detection task, there were 30 face target images and 30 distractor images (Figure 2). Each face target had a human face that was unambiguously in the middle of the image, viewed either head on or from the side. Individual faces came from a variety of genders, ages, and ethnicities. Each no-face distractor was defined by not having a human face in the middle. In other words, there could be animal faces in the middle and humans in the periphery, but without a human face in the middle, those images were considered distractors. Before the experiment began, observers were shown multiple examples of both target and distractor images to understand what differentiated targets and distractors. For both face and no-face stimuli, the images comprised a wide variety of settings and locations (e.g., a man in a kayak, a construction worker, a chef in a kitchen, a dog on a dog bed, an empty office, an empty beach).

Critical Stimuli. All critical stimuli were generated using texture synthesis algorithms that preserve a variety of low-level image statistics within prespecified pooling windows (Rosenholtz, 2016). In this case, we used three different pooling window sizes for our three experimental conditions, which were as follows: (a) *small scramble*—stimuli in this condition were generated using the Freeman and Simoncelli (2011) algorithm, with a scaling parameter of 0.5, which corresponds to the ratio of the diameter of a pooling window relative to the eccentricity of that pooling window in the periphery (see Figure 3); (b) *medium scramble*—stimuli in this condition were also generated using the Freeman and Simoncelli algorithm, except this time they were created with a scaling parameter of 3.0, resulting in images that appeared far more scrambled than in the small scramble condition; (c) *large scramble*—stimuli in this condition were generated using the Portilla and Simoncelli (2000) algorithm, in which one large pooling window spans the entire image. Stimuli from the large scramble condition are so thoroughly scrambled that no individual object can be identified. Finally, in each of the three conditions, after generating the stimuli with these different algorithms, we also replaced the central parts of the scrambled images with a preserved section of the original image. This preserved section subtended a circular aperture with a radius of 4° of visual angle, which roughly accounts for the entire foveal/parafoveal regions of the retina (Millodot, 2014; Polyak, 1941; Swanson & Fish, 1995; see Figure 1 in the online supplemental material). Thus, for each critical stimulus, the region of the image that landed on the fovea/parafovea was normal, whereas the information on the periphery was scrambled to varying.

Results

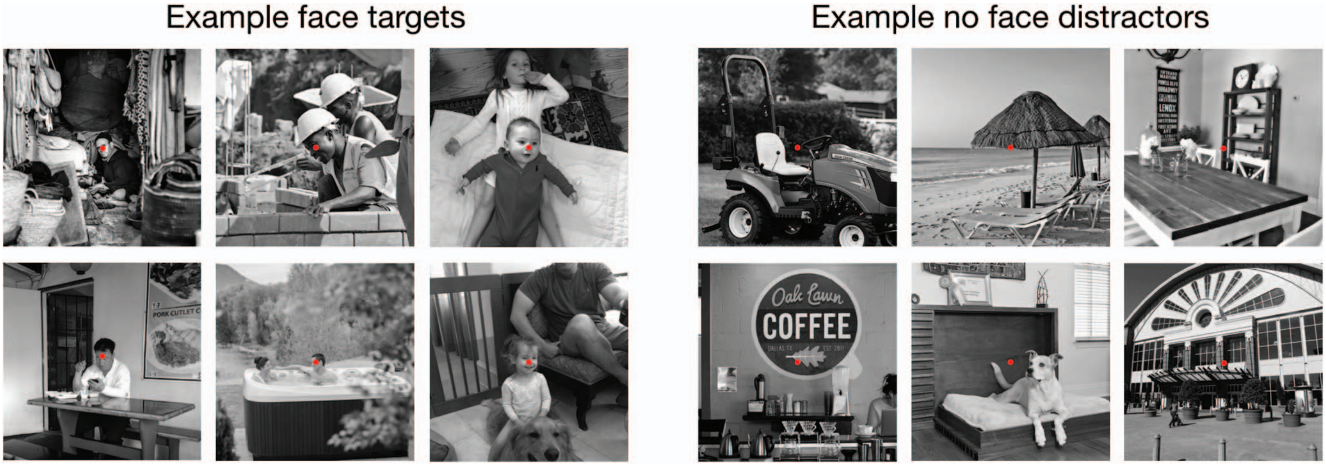
What percentage of participants will be unaware of the scrambling done to these images when they are attending at fixation? In the small scramble condition, 100% of participants (40/40) were unaware of the stimulus alterations (see Figure 4; see also the online supplemental material for more detailed analyses of the data). This result is unsurprising since the scrambling is small relative to that in the other conditions, and thus is not always

---

[2] It should be noted that we neither made audio recordings of participants' verbal responses nor did we have multiple raters code those responses to quantify interrater reliability. However, in any situation in which a participant's response was ambiguous, the experimenters wrote down verbatim what participants said and then met to discuss those responses to ensure consistency in coding across experiments.

Figure 2

Target and Distractor Stimuli From Experiments 1 and 2



Note. Representative examples of stimuli used in Experiments 1 and 2. Images with faces shown here are in the public domain and freely available (https://commons.wikimedia.org); images in the original stimulus set shown to participants are available via the Open Science Framework (https://osf.io/6c3xw/). See the online article for the color version of this figure.

immediately apparent. However, this amount of scrambling is indeed detectable when observers attend directly to the periphery (Wallis et al., 2019). Somewhat surprising, however, is the fact that 50% of participants (20/40) were unaware of the stimulus alterations in the medium scramble condition. In fact, an even more surprising result is the fact that in the large scramble condition, where the periphery of the stimuli are scrambled so much that no individual objects can be identified, 48% of participants (19/40) were completely unaware of the stimulus alterations. In this case, there was a significant difference in noticing rates between the small and medium scramble conditions, $\chi^2(1) = 24.07, p < .001$, as well as between the small and large scramble conditions, $\chi^2(1) = 25.83, p < .001$. Meanwhile, there was no significant difference between the medium and large scramble conditions, $\chi^2(1) = 0$, $p = 1$. Finally, performance on the primary face task was nearly perfect for each of the three experimental conditions (small scramble: 97.3% correct; medium scramble: 98.5% correct; large scramble: 97% correct), with no significant difference in performance among any of the three conditions: $t(39) > 1.22$, $p > .22$, in all cases.

As mentioned earlier, the final step in our experimental procedure was showing participants four examples of different critical stimuli in order to (a) verify that they did not notice the stimulus alterations and (b) ask if they could identify the last image even if they were unaware of the critical manipulations. In this case, the vast majority of participants were able to identify the last image they saw in each experimental condition (small: 90%, 36/40 participants; medium: 85%, 34/40 participants; large: 90%, 36/40 participants). Even though these rates are quite high, it is still possible that they underestimate the percentage of participants who were aware of the last image. Because a good amount of time would sometimes elapse between the presentation of the critical stimulus and the presentation of the four options, it is possible that some participants were aware of the critical stimulus but simply confused it with another stimulus after the questioning period. Regardless, our next analysis focused on determining the inattentional blindness rates when only examining the data from participants who could identify the last image. In this case, we found that the inattentional blindness rates were virtually identical to when all participants' data was analyzed (small scramble: 100%, 36/36 participants; medium scramble: 41%, 14/34 participants; large scramble: 50%, 18/36 participants). In other words, for each condition, more than 40% of participants were completely unaware of the significant alterations that had been made to the stimulus in the periphery even though they processed enough of the critical stimulus to pick it out of a lineup of four options.

Experiment 2: Inconsistent, Abstract, and No Periphery

Method

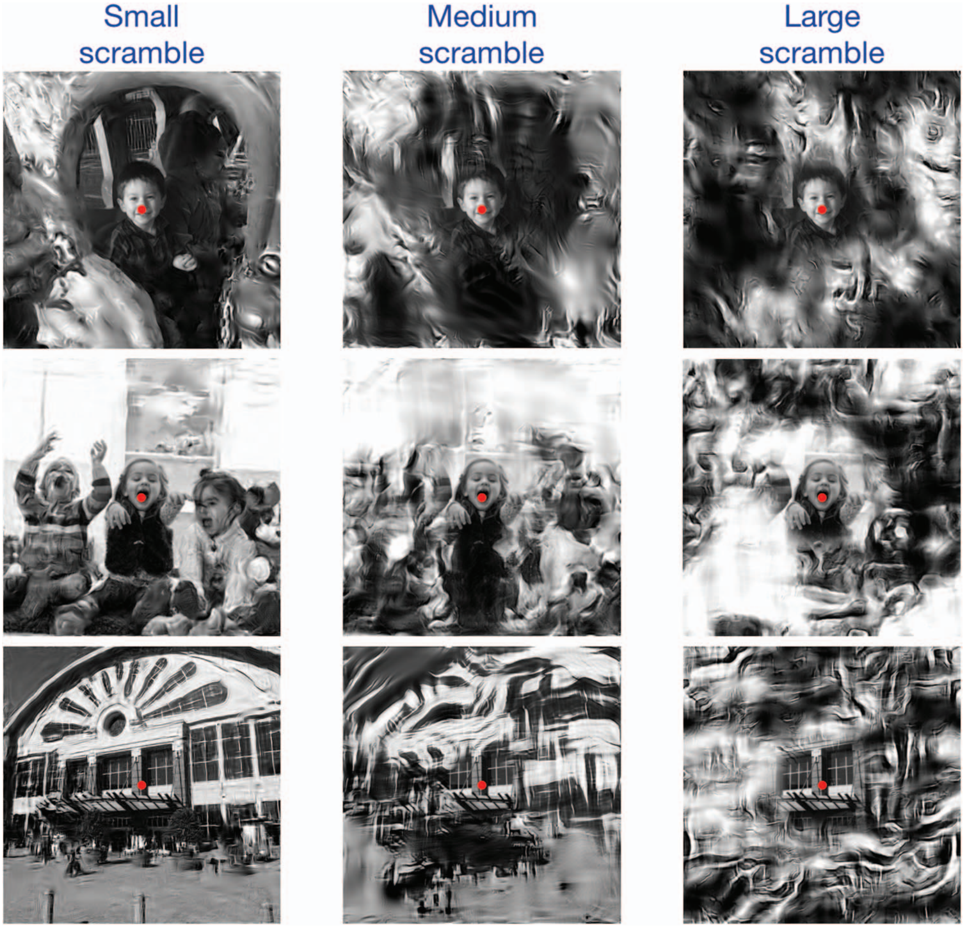
Perhaps the most interesting result in Experiment 1 is the fact that so many observers failed to notice the alterations to the stimuli in the large scramble condition. If several of our observers are unaware of those alterations, what does it take to get them to routinely notice stimulus alterations? In Experiment 2, we tried to answer this question by repeating the procedures of Experiment 1 with a new set of critical stimuli that altered the periphery in a variety of ways.

Participants

Forty participants, none of whom participated in Experiment 1, were run in each of the three experimental conditions, resulting in a total of 120 participants. All participants were recruited from the Amherst College and Massachusetts Institute of Technology communities. All participants gave informed consent and were compensated with either $5.00 or course credit. All experimental procedures were approved by the Committee on the Use of Human

Figure 3

Example Target Stimuli From Experiment 1



Note. Each image subtended 26° of visual angle in the experiment, with only the center 4° of visual angle (radius) being preserved. Written permission for publication in scientific journals has been obtained from each photographed individual.

Subjects in Research under the Institutional Review Board of Amherst College.

Procedure

The only procedural difference between Experiments 1 and 2 was regarding the critical stimuli. Besides changing the critical stimuli, the procedures between the two experiments were identical.

Stimuli

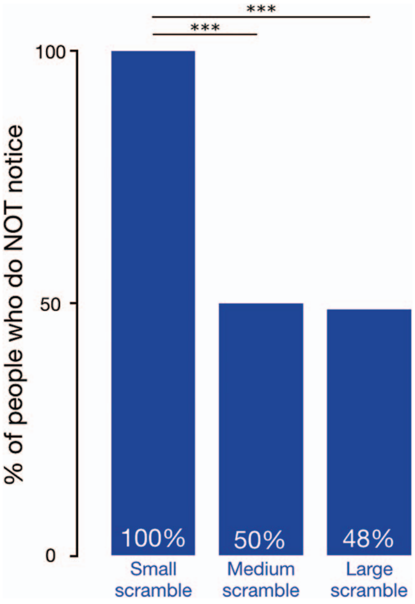
There were three types of critical stimuli in this experiment: (a) *inconsistent periphery*—the foveal/parafoveal part of one image is placed on top of a semantically inconsistent peripheral image (i.e., a part of a coffee shop placed on top of a desert scene, etc.; see Figure 5); (b) *abstract periphery*—the foveal/parafoveal part of one image is placed on top of high-contrast abstract shapes and patterns; and (c) *no periphery*—the foveal/parafoveal part of one image is placed on top of a uniform gray background. (See Figure 2 in the online supplemental material for information on stimulus generation procedures.)

Results

How often will participants be aware of these critical stimuli when performing the same face detection task as in Experiment 1? The results from Experiment 2 are plotted in Figure 6. In the inconsistent periphery condition, 73% of participants (29/40) were unaware of the stimulus alterations on the critical trial. Put more simply, most participants were unaware of the fact that they had seen stimuli comprised of parts from two entirely different scenes. It was not until the abstract periphery and no periphery conditions that most participants noticed the stimulus alterations (abstract periphery: 18%, 7/40; no periphery: 13%, 5/40). In this experiment, there was a significant difference in noticing rates between the inconsistent and abstract periphery conditions, $\chi^2(1) = 22.27$, $p < .001$, as well as between the inconsistent and no periphery conditions, $\chi^2(1) = 27.06$, $p < .001$. However, there was no significant difference between the abstract and no periphery conditions, $\chi^2(1) = 0.10$, $p = .75$. As in Experiment 1, the majority of participants were able to identify the last image they saw in each experimental condition (inconsistent periphery: 87.5%, 35/40 par-

Figure 4

Inattentional Blindness Rates for Experiment 1



Note. The percentage of participants who failed to notice the critical stimulus is plotted on the vertical axis. Each bar corresponds to a different experimental condition.
*** $p < .001$.

ticipants; abstract periphery: 92.5%, 37/40 participants; no periphery: 92.5%, 37/40 participants). Furthermore, we again found that the inattentional blindness rates were virtually identical when only participants who could identify the critical image were analyzed (inconsistent periphery: 71%, 25/35 participants; abstract periphery: 16%, 6/37 participants; no periphery: 11%, 4/37 participants). As was the case in Experiment 1, performance on the primary face task was nearly perfect for each of the three experimental conditions (inconsistent periphery: 98.5% correct; abstract periphery: 96.8% correct; no periphery: 98% correct) with no significant difference in performance among any of the three conditions: $t(39) = 1.30$, $p > .20$, in all cases.

Experiment 3

Method

Taken together, the results from Experiments 1 and 2 suggest that a single snapshot of perceptual experience is remarkably impoverished when observers are attending at fixation. Of course, in everyday life, observers can deploy their attention in a variety of ways. For example, the scale of attention will be radically different when searching for an individual in a crowded room compared to when staring at an open landscape and trying to take in its distributed layout. For the first two experiments, we chose a simple task that required participants to attend at fixation since it is generally the case that people attend to whatever they are fixating on and wherever they will fixate next (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Kowler et al., 1995; Posner, 1980). However, because attention can easily be directed more toward the periphery, we sought to determine how the scale of attention affects a snapshot of perception. Thus, in Experiment 3, we replaced the face detection task with an indoor/outdoor task that was designed to naturally get participants to attend more to the periphery. In this case, rather than say whether the last image had a human face in the middle, observers simply said whether the last image was of an indoor scene or an outdoor scene. The logic behind this experiment is that by asking participants to make a judgment about the entire scene, rather than just about an object at fixation, their attention would naturally be distributed across the scene in a way that is somewhat ecologically valid. In addition, by distributing attention more toward the periphery, we are also bolstering the processing of peripheral vision, which has been shown to be more important for gist perception in a single fixation (Larson & Loschky, 2009). Thus, we predicted that the rates of inattentional blindness would decrease in Experiment 3 relative to those in Experiments 1 and 2.

Participants

Forty participants were run on the indoor/outdoor task. All participants were recruited from the Amherst College and Massachusetts Institute of Technology communities and none of whom participated in either Experiments 1 or 2. All participants gave informed consent and were compensated with either $5.00 or course credit. All experimental procedures were approved by the Committee on the Use of Human Subjects in Research under the Institutional Review Board of Amherst College.

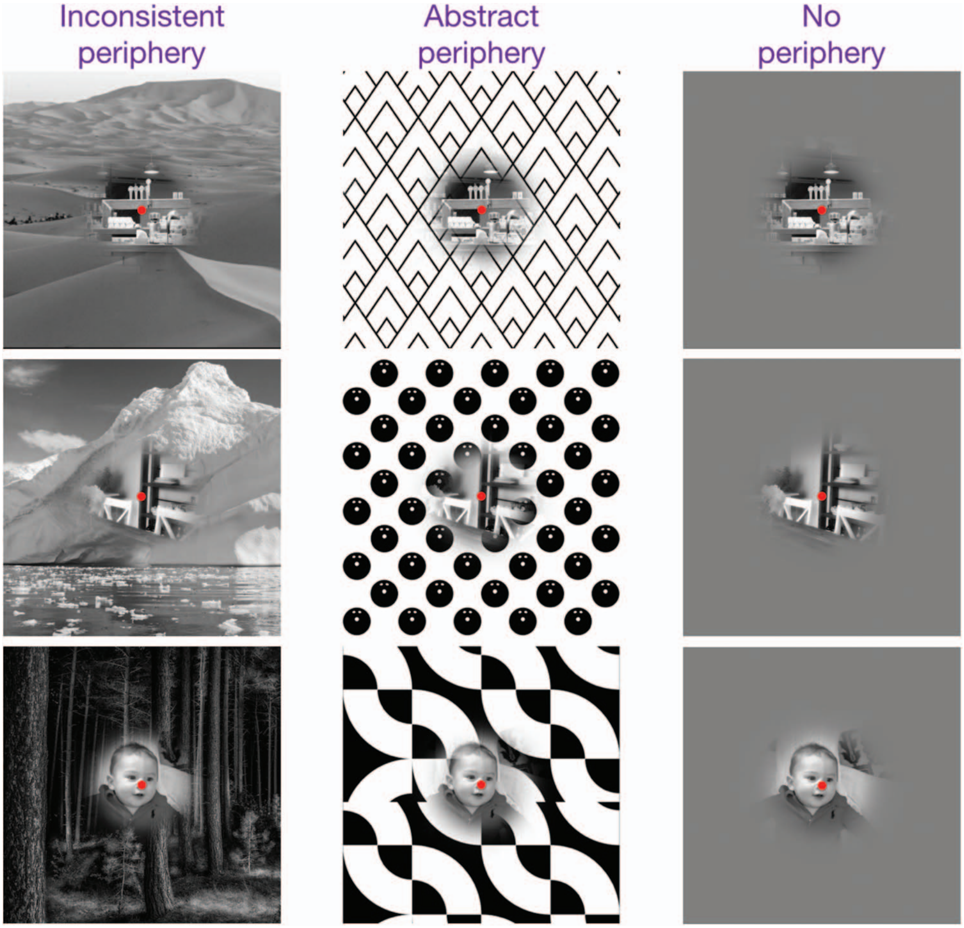
Procedure

In Experiment 3, the procedures were identical to those in Experiments 1 and 2 with a few minor exceptions. First, an eye-tracker (i.e., a Gazepoint GP3 eye tracker [https://www.gazept.com/product/gazepoint-gp3-eye-tracker/]) was used to ensure that participants maintained fixation throughout the trial. Given the task demands of Experiments 1 and 2, it seemed safe to assume that participants maintained fixation throughout. In fact, if participants did not, that would only decrease our inattentional blindness rates since participants are far more likely to notice critical stimuli if they are foveating on parts of the image we did not preserve. For Experiment 3, meanwhile, participants performed a five-point calibration sequence at the beginning of the experiment that required 0.5° of visual angle to be accepted. Once a trial began, if participants broke fixation (i.e., more than 2.5° of visual angle away from fixation), the trial would start over. This was infrequent and occurred in less than 4% of trials.

For the first 10 trials of Experiment 3, the primary difference between this experiment and the previous two was that participants reported whether the last image was of an indoor or outdoor scene (see Figure 7a). Again, before the experiment began, participants were shown multiple examples of indoor and outdoor stimuli so they could familiarize themselves with the task. In addition, for the last image of the trial, half of the trials had an indoor scene at the end and half had an outdoor scene at the end. Beyond those differences, the remaining procedures of Experiment 3 were iden-

Figure 5

Example Target Stimuli From Experiment 2



Note. Each image subtended 26° of visual angle in the experiment, with only the center 4° of visual angle being preserved. Written permission for publication in scientific journals has been obtained from each photographed individual.

tical to Experiments 1 and 2 up to and including the procedures of the critical trial.

Stimuli

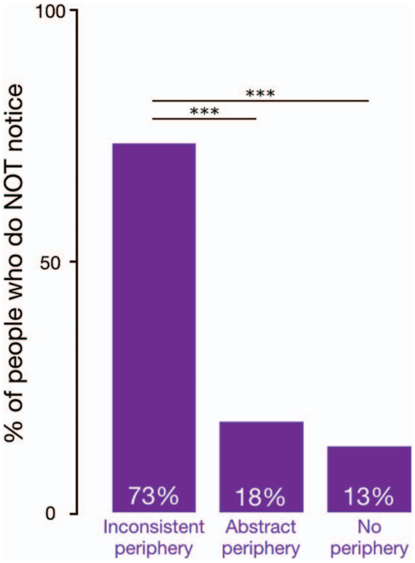
For this experiment, we used two different types of stimuli. Specifically, we used stimuli from the medium scramble condition (Experiment 1) and the inconsistent periphery condition (Experiment 2) as our critical stimuli. For the indoor/outdoor task, the same stimuli were used as in the face task and the procedures used to select the stimuli from trial to trial were identical as the face task. Thus, the stimuli did not differ between the two tasks; only the tasks being performed on those stimulus streams were different. The lone exception to this rule came in selecting the stimuli for the last image in a stream. In certain cases, we decided that some of the images were somewhat ambiguous in terms of whether they were indoors or outdoors (e.g., an image of a woman working in a greenhouse). To avoid any confusion, those ambiguous images were never shown as the last image in the stream (although they could be shown during a trial in any position that was not the last image in a stream). However, on the critical trial, no stimuli were excluded from consideration. The same exact stimulus batch and the same selection procedures used in the face task were used in the indoor/outdoor task.

Results

Overall, we found that the scale of attention had a dramatic effect on the resolution of perception. As a reminder, in Experiment 1, we found that 50% of participants (20/40) did not notice the medium scramble stimuli when performing the face task. By comparison, in Experiment 3, the inattentional blindness rates were significantly lower with only 15% of participants (6/40) failing to notice the critical stimuli when performing the scene task, $\chi^2(1) = 9.63$, $p < .01$ (see Figure 8a). When only examining the data from participants who could identify the last image (92.5%, 37/40 participants), we found that the inattentional blindness rates dropped to 8% (3/37 participants). Meanwhile, a similar pattern of results was found when using the inconsistent periphery

Figure 6

Inattentional Blindness Rates for Experiment 2



Note. The percentage of participants who failed to notice the critical stimulus is plotted on the vertical axis. Each bar corresponds to a different experimental condition.
*** $p < .001$.

stimuli. Again, as a reminder, in Experiment 2, we found that 73% of participants (29/40) did not notice the critical stimuli when performing the face task. By comparison, in Experiment 3, the inattentional blindness rates were significantly lower with 45% of participants (18/40) failing to notice the critical stimuli when performing the scene task, $\chi^2(1) = 4.22$, $p < .05$ (see Figure 8b). When only examining the data from participants who could identify the last image (85%, 34/40 participants), we found that the inattentional blindness rates dropped to 38% (13/34 participants). These results highlight how the resolution of perceptual experience is not a fixed constant but is an ever-changing concept that varies as a function of attention.

When comparing inattentional blindness rates between primary tasks (i.e., face vs. scene tasks), it is important to also compare performance on those very primary tasks since differences in primary task performance can lead to differences in inattentional blindness rates. However, in this case we found no such evidence of a difference in primary task performance for the medium scramble condition (face task performance: 98.5%; scene task performance: 96.8%; $t[39] = 0.39$, $p = .70$) or the inconsistent periphery condition (face task performance: 98.5%; scene task performance: 97.5%; $t[39] = 0.81$, $p = .42$). Of course, these results should be interpreted with some caution because there is likely a ceiling effect in these cases with performance being above 96.8% in all cases.

General Discussion

When discussing the limits of visual cognition, it has been often claimed that even though observers may fail to encode individual objects, they will effectively always process the general scene around them (Koch & Tsuchiya, 2007; Mack & Rock, 1998; Rensink, 2000; Sampanes et al., 2008; Simons & Levin, 1997; van Boxtel et al., 2010; but see Cohen et al., 2011 and Mack & Clarke, 2012). This idea is consistent with the fact that observers cannot only recognize the gist of a scene very briefly (Fei-Fei et al., 2007; Intraub, 1981; Potter, 1976) but can also extract a variety of other higher level dimensions about a scene as well (Castelhano & Henderson, 2008; Greene & Oliva, 2009; Tatler et al., 2003; Torralbo et al., 2013; Walther et al., 2009). Across multiple experiments, however, we showed that there are many cases in which observers can fixate upon a scene for several hundred milliseconds, process and remember the items they were fixating upon, and yet still fail to notice shockingly drastic alterations that alter the entire meaning and interpretation of the scene. These findings suggest that awareness of the contents of peripheral vision in a single snapshot of perceptual experience can be surprisingly impoverished. Indeed, we found that a large portion of observers failed to realize that their periphery was so degraded that no object could be identified or that it was comprised of an entirely different scene. Of course, it should be emphasized that it is not always the case that observers will routinely fail to notice such drastic alterations. In Experiment 3, we found that significantly more observers noticed the scrambled or inconsistent periphery when their attention was deployed away from fixation to perform the indoor/outdoor task. Together, this collection of findings highlights how the resolution of perception can vary tremendously based on the deployment of attention. This is a crucial idea for researchers to keep in mind when trying to understand the richness of perceptual experience since it appears that there cannot be one definitive answer to the question, how rich is perceptual experience?

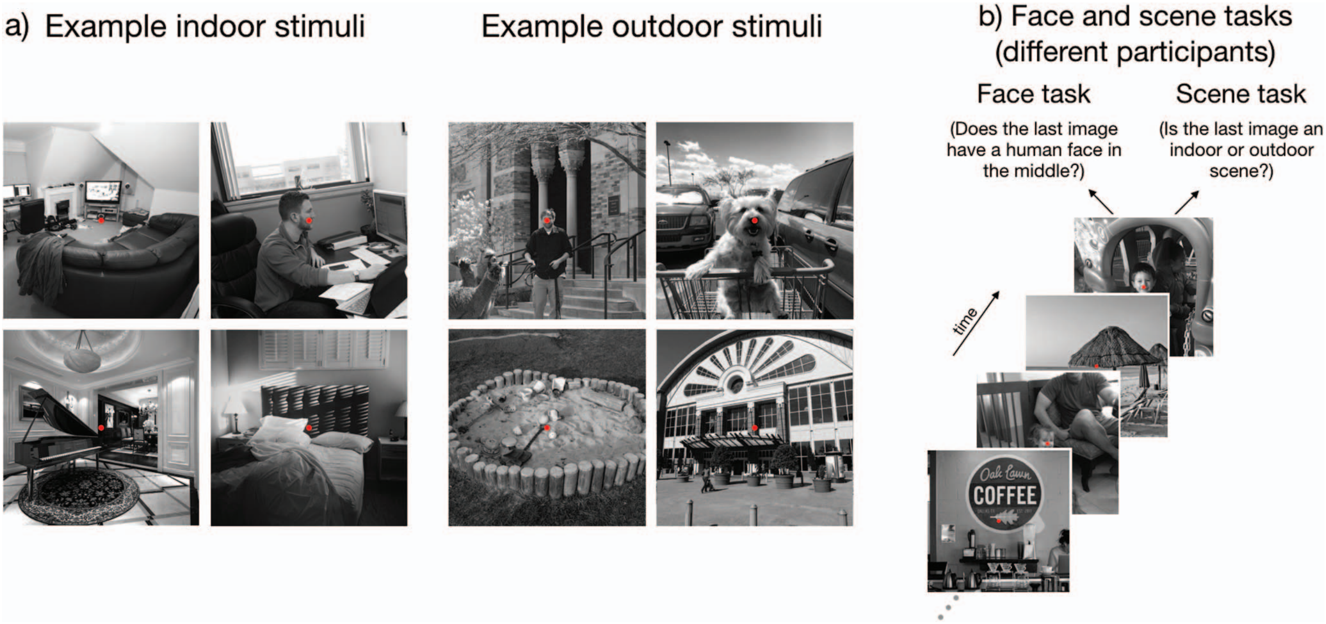
Relation to the Useful Field of View

The idea of the resolution of perception is closely related to a similar idea in the literature referred to as the useful field of view (UFOV; Mackworth, 1965). The UFOV is the amount of visual information that can be processed within a single fixation without any eye or head movements. Critically, however, is the fact that UFOV appears to vary as a function of attentional load and task complexity (Atchley & Dressel, 2004; Chan & Courtney, 1998; Ringer et al., 2016), much in the same way that we suggest the resolution of perception varies as a function of attentional allocation. In other words, as the attentional load increases or gets more focused around the fovea, the UFOV shrinks and the amount of information observers can usefully process decreases. One interesting idea for future research raised by research on the UFOV is the notion that as people age, the resolution of perception varies. It has been systematically shown that the UFOV varies as a function of age (Ball et al., 1988, 1993; Sekuler et al., 2000). These findings raise the interesting possibility that older and younger adults may have different rates of inattentional blindness and may actually differ in terms of how much they can perceive "in the blink of an eye" (Pringle et al., 2001; Pringle et al., 2004).

What Are the Important Elements of Natural Scenes?

When examining the results from Experiments 1 and 2, a natural question emerged: Why are some alterations to the periphery

Figure 7
Indoor and Outdoor Stimuli From Experiment 3



Note. Visualization of the trial procedures for the face and scene tasks. Note that two groups of participants performed the face and scene tasks, respectively. Images with faces shown here are in the public domain and freely available (https://commons.wikimedia.org); images in the original stimulus set shown to participants are available via the Open Science Framework (https://osf.io/6c3xw/). See the online article for the color version of this figure.

noticed whereas others are not? Although we cannot definitively answer this question, we believe a few natural possibilities emerge from the results described here.

For example, one potential critical factor may be whether or not the periphery of the critical stimulus contains enough information such that the gist of the scene can be extracted (i.e., “office,” “gym,” “beach,” etc.; Castelhano & Henderson, 2008; Oliva, 2005; Oliva & Schyns, 1997; Potter, 1976). Here, the two highest rates of inattentional blindness were with the small scramble and inconsistent periphery conditions. In both cases, the gist of the scenes could be easily extracted from the peripheral information shown in the critical stimuli (see the present Figures 3 and 5 and Figures 2 and 6 in the online supplemental material). However, one particularly surprising finding is the fact that in the inconsistent periphery condition, the gist of the peripheral scene was inconsistent with the center of the image (e.g., a computer monitor on top of a jungle, a face on top of mountain). Is it the case that so long as a gist—any gist—can be extracted from a stimulus, observers will routinely fail to notice the critical alterations to the image? This idea is somewhat supported by the fact that when the gist of the scene could not be extracted, as was likely the case in the large scramble condition (see the present Figure 3 and Figure 4 in the online supplemental material), the inattentional rates dropped significantly. Together, these results raise a series of questions that future work could easily address: Is it really the case that so long as any discernable gist in the periphery is presented, observers will often fail to notice the alterations to the critical stimuli? What if, for example, the periphery of an image is negated or rotated 180° but observers can still successfully extract the gist of the scene (Loschky et al., 2015; Torralbo & Oliva, 2003)? Does the extent to which gist can be extracted from a set of synthesized textures (Experiment 1) predict the rates of inattentional blindness? Understanding the role that gist perception plays in this paradigm could shed important light on the boundary limits of perceptual awareness with natural scenes.

However, it is certainly possible that gist is not the critical dimension driving our results at all and the key features are other aspects of natural scenes. For example, in addition to changing the presence/absence of gist information, a wide variety of other lower-level properties of natural scenes vary between our experimental conditions. Consider the large scramble condition from Experiment 1 and the abstract periphery condition in Experiment 2. In both cases, there are numerous differences between the two stimulus conditions even though neither contains gist information in the periphery (i.e., differences in nonstationarity, 1/f power spectra, etc.). Thus, it might be the case that it is alterations to these lower-level statistical properties of the periphery, not alterations to the high-level meaning of the scene (i.e., gist), which determines whether or not inattentional blindness will occur.

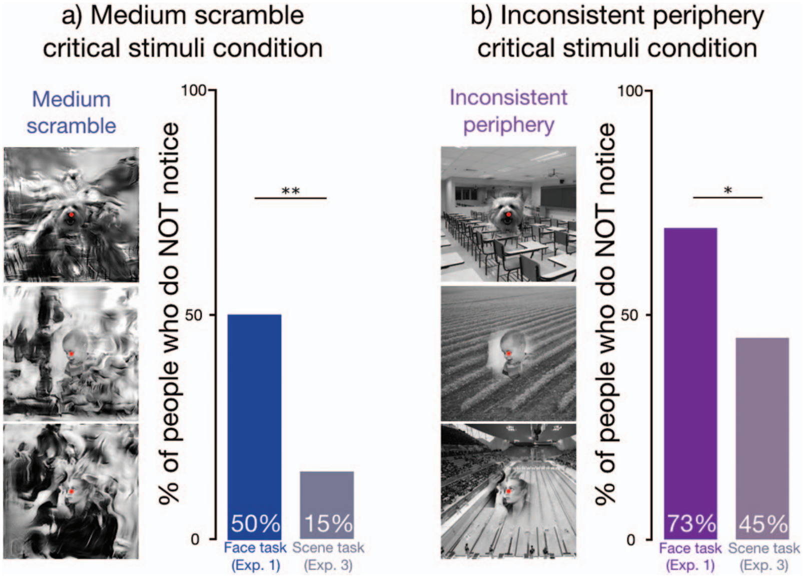
Of course, it is also quite reasonable to suspect that there is not one simple answer to the question of what causes some alterations to be noticed and others not. Indeed, identifying the factors that determine the effects reported here will require a more complete and thorough understanding of the bandwidth of perception in natural scenes.

Questions for Future Research

How can the findings reported here be reconciled with personal introspection, which strongly suggests we see far more than these

Figure 8

Target Stimuli and Results From Experiment 3



Note. a) Examples of the target stimuli used on the critical trial of Experiment 3 with the medium scramble, which were also used in the medium scramble condition of Experiment 1, and the inattentional blindness rates for the face and scenes tasks. b) Representative examples of the target stimuli used on the critical trial of Experiment 3 with the inconsistent periphery stimuli (and used in the inconsistent periphery condition of Experiment 2) and inattentional blindness rates for the face and scene tasks. In both cases, the percentage of participants who failed to notice the critical stimulus is plotted on the vertical axis. Each bar corresponds to a different experimental condition. Images with faces shown here are in the public domain and freely available (https://commons.wikimedia.org); images in the original stimulus set shown to participants are available via the Open Science Framework (https://osf.io/6c3xw/).
$^{*}\ p < .05.$ $^{**}\ p < .01.$

findings imply? We believe there are several possible answers to this question. For starters, in our experiments, observers are only presented with a single presentation of images that approximates the length of an average fixation period (288 ms per item). In everyday life, however, observers do not see a single snapshot of a scene in isolation and then another snapshot of an entirely new scene. Instead, observers are usually embedded in an environment that they can look upon and explore for seconds, minutes, or even hours. One intriguing possibility is that observers build up a detailed representation of the scene around them over time, and the knowledge/memory of that representation provides the foundation for an apparently rich perceptual experience (Hollingworth, 2009). Alternatively, some researchers would argue that these results are consistent with the notion of *subjective inflation* (Knotts et al., 2019). Under this view, observers systematically overestimate the richness of their perceptual experience when attention is primarily focused elsewhere. Support for this idea comes from several studies showing that when objective performance is matched between two experimental conditions (e.g., cued vs. uncued stimuli), participants are more likely to mistakenly claim an item is present, when in reality it is not, whenever attention is reduced (i.e., in the uncued condition; Rahnev et al., 2011; Solovey et al., 2015). In the context of our results, observers might overestimate the richness of their peripheral experience and be unaware of the degree to which an image can be altered without their realizing it when attention is focused on the fovea during the face detection task. Of course, reconciling the results reported here with the subjective feeling of a richer perceptual experience is not a trivial endeavor, especially given the longstanding controversy surrounding this issue (Block, 2011; Boly et al., 2017; Cohen et al., 2012, 2016; Dehaene & Changeux, 2011; Koch & Tsuchiya, 2007; Kouider et al., 2010; Lamme, 2010; Mack & Rock, 1998; Rensink et al., 1997; Simons et al., 2005).

Fortunately, the results reported here raise a variety of questions that are far more tractable and could easily be addressed by future research. For example, earlier we discussed how one potential shortcoming of focusing on individual snapshots of perception is that it discounts the possible role of built up representations over time. One approach to address this issue would be to compare inattentional blindness rates with familiar versus unfamiliar scenes. It is possible that if observers are already familiar with the critical stimulus, which was not the case in the present experiments, their preexisting knowledge of that scene will enable them to spot alterations more easily. However, it should be noted that an alternative pattern of results could also emerge. If observers are familiar with a scene, their visual system may be able to “fill-in” certain details from memory and the inattentional blindness rates could actually increase (Komatsu, 2006). Another factor not addressed in this study, but which could be easily incorporated into future experiments, focuses on the richness of color perception. Another long running debate focuses on how much color observers perceive in their periphery (Haun et al., 2017; Tyler, 2014). The paradigm described here could be adapted to contribute to this controversy by making all of the images in color, rather than in black and white, and making the critical stimulus black and white in the periphery, and colored in the fovea. The extent to which observers notice such color alterations could shed important light on the amount of color perceived in a snapshot of experience.

Another important question raised by these results is related to the relationship between our findings and the well-established limits of peripheral vision. Specifically, can our findings simply be accounted for by factors such as visual crowding or the drop off in retinal acuity (Loschky et al., 2005; Pelli & Tillman, 2008; Rosenholtz, 2016; Whitney & Levi, 2011)? We believe that while those bottlenecks on perception are thematically related to the results reported here, they do not constrain perception nearly enough to explain our results. Perhaps the clearest bit of data to explain this idea comes from Wallis et al. (2019) who directly explored the boundary limits of perception in the periphery using the same texture synthesis algorithm as we did in Experiments 1 and 3 (Freeman & Simoncelli, 2011). In their study, Wallis and colleagues explicitly asked participants to determine which image in a set had been synthesized (i.e., an odd-ball detection task). In this case, the critical image had been scrambled in a manner that corresponds to our small scramble condition in Experiment 1. In that case, observers were able to reliably identify the synthesized image among natural images with little difficulty. As a reminder, in our Experiment 1, not a single person noticed the critical stimuli in the small scramble condition. Thus, it seems as if our results do not stem from a fundamental limit on peripheral vision imposed by crowding or the drop off in retinal acuity; instead, we argue that the primary bottleneck is likely the way in which observers are deploying their attention across these scenes. An interesting line of future research would be to explore exactly how attention is generally deployed in natural viewing conditions and what the exact resolution of perceptual awareness is in those situations.

Conclusion

Overall, our results reveal that observers can gaze upon an image, process enough of that image to identify it out of multiple options, but completely fail to notice drastic changes to the periphery of that image. These results raise a variety of experimentally tractable questions to further understand the richness of visual experience in terms of both an individual snapshot as well as when naturally embedded in the real world.

References

Atchley, P., & Dressel, J. (2004). Conversation limits the functional field of view. *Human Factors*, *46*(4), 664–673. https://doi.org/10.1518/hfes.46.4.664.56808

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*(1), 47–52. https://doi.org/10.1016/S1364-6613(00)01819-2

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America*, *5*(12), 2210–2219. https://doi.org/10.1364/JOSAA.5.002210

Ball, K. K., Owsley, C., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1993). Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology & Visual Science*, *34*(11), 3110–3123. https://iovs.arvojournals.org/article.aspx?articleid=2160828

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30*(5–6), 481–5478. https://doi.org/10.1017/S0140525X07002786

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

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(40), 9603–9613. https://doi.org/10.1523/JNEUROSCI.3218-16.2017

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review*, *18*(5), 890–896. https://doi.org/10.3758/s13423-011-0107-8

Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on scene gist. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 660–675. https://doi.org/10.1037/0096-1523.34.3.660

Chan, H. S., & Courtney, A. J. (1998). Stimulus size scaling and foveal load as determinants of peripheral target detection. *Ergonomics*, *41*(10), 1433–1452. https://doi.org/10.1080/001401398186199

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

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

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

Dagnelie, G. (2011). *Visual prosthetics: Physiology, bioengineering, & rehabilitation*. Springer. https://doi.org/10.1007/978-1-4419-0754-7

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 220–227. https://doi.org/10.1016/j.neuron.2011.03.018

Dennett, D. C. (1991). *Consciousness explained*. Little, Brown.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837. https://doi.org/10.1016/0042-6989(95)00294-4

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision*, *7*(1), 10. https://doi.org/10.1167/7.1.10

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*(9), 1195–1201. https://doi.org/10.1038/nn.2889

Frisén, L. (1990). *Clinical tests of vision*. Raven Books.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*(4), 464–472. https://doi.org/10.1111/j.1467-9280.2009.02316.x

Haun, A. M., Tononi, G., Koch, C., & Tsuchiya, N. (2017). Are we underestimating the richness of visual experience? *Neuroscience of Consciousness*. Advance online publication. https://doi.org/10.1093/nc/niw023

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *10*(3), 15. https://doi.org/10.1167/3.1.6

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504. https://doi.org/10.1016/j.tics.2003.09.006

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787–795. https://doi.org/10.3758/BF03206794

Hollingworth, A. (2009). Memory for real-world scenes. In J. R. Brockmole (Ed.), *The visual world in memory* (pp. 89–119). Psychology Press.

Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 604–610. https://doi.org/10.1037/0096-1523.7.3.604

Knotts, J. D., Odegaard, B., Lau, H., & Rosenthal, D. (2019). Subjective inflation: Phenomenology's get-rich-quick scheme. *Current Opinion in Psychology*, *29*, 49–55. https://doi.org/10.1016/j.copsyc.2018.11.006

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. https://doi.org/10.1016/j.tics.2006.10.012

Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience*, *7*(3), 220–231. https://doi.org/10.1038/nrn1869

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, *14*(7), 301–307. https://doi.org/10.1016/j.tics.2010.04.006

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897–1916. https://doi.org/10.1016/0042-6989(94)00279-U

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7*(1), 12–18. https://doi.org/10.1016/S1364-6613(02)00013-X

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

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*(10), 6.

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology: Applied*, *8*(2), 99–117. https://doi.org/10.1037/1076-898X.8.2.99

Loschky, L. C., McConkie, G. W., Yang, J., & Miller, M. E. (2005). The limits of visual resolution in natural scene viewing. *Visual Cognition*, *12*(6), 1057–1092. https://doi.org/10.1080/13506280444000652

Loschky, L. C., Ringer, R. V., Ellis, K., & Hansen, B. C. (2015). Comparing rapid scene categorization of aerial and terrestrial views: A new perspective on scene gist. *Journal of Vision*, *15*(6), 11. https://doi.org/10.1167/15.6.11

Mack, A., & Clarke, J. (2012). Gist perception requires attention. *Visual Cognition*, *20*, 300–327.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. https://doi.org/10.7551/mitpress/3707.001.0001

Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic Science*, *3*, 67–68. https://doi.org/10.3758/BF03343023

McKonkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, *8*, 365–368. https://doi.org/10.3758/BF03335168

Miellet, S., Zhou, X., He, L., Rodger, H., & Caldara, R. (2010). Investigating cultural diversity for extrafoveal information use in visual scenes. *Journal of Vision*, *10*(6), 21. https://doi.org/10.1167/10.6.21

Millodot, M. (2014). *Dictionary of optometry and visual science*. Elsevier Health Sciences. https://www.elsevier.com/books/dictionary-of-optometry-and-vision-science/unknown/978-0-7020-7223-9

Nuthmann, A. (2017). Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review*, *24*(2), 370–392. https://doi.org/10.3758/s13423-016-1124-4

Oliva, A. (2005). Gist of a scene. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–256). Elsevier. https://doi.org/10.1016/B978-012375731-9/50045-8

Oliva, A., & Schyns, P. (1997). Course blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*(1), 72–107. https://doi.org/10.1006/cogp.1997.0667

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*(10), 1129–1135. https://doi.org/10.1038/nn.2187

Pelz, J. G., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, *41*(25–26), 3587–3596. https://doi.org/10.1016/S0042-6989(01)00245-0

Polyak, S. L. (1941). *The retina*. University of Chicago Press.

Pomplun, M., Reingold, E. M., & Shen, J. (2001). Peripheral and parafoveal cueing and masking effects on saccadic selectivity in a gaze-contingent window paradigm. *Vision Research*, *41*(21), 2757–2769. https://doi.org/10.1016/S0042-6989(01)00145-6

Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, *40*(1), 49–71. https://doi.org/10.1023/A:1026553619983

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*, 3–25. https://doi.org/10.1080/00335558008248231

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(5), 509–522. https://doi.org/10.1037/0278-7393.2.5.509

Pringle, H. L., Irwin, D. E., Kramer, A. F., & Atchley, P. (2001). The role of attentional breadth in perceptual change detection. *Psychonomic Bulletin & Review*, *8*(1), 89–95. https://doi.org/10.3758/BF03196143

Pringle, H. L., Kramer, A. F., & Irwin, D. E. (2004). Individual differences in the visual representation of scenes. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 165–185). MIT Press.

Rahnev, D., Maniscalco, B., Graves, T., Huang, E., de Lange, F. P., & Lau, H. (2011). Attention induces conservative subjective biases in visual perception. *Nature Neuroscience*, *14*(12), 1513–1515. https://doi.org/10.1038/nn.2948

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2014). The gaze-contingent moving window in reading: Development and review. *Visual Cognition*, *22*(3–4), 242–258. https://doi.org/10.1080/13506285.2013.879084

Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2008). Eye movements and visual encoding during scene perception. *Psychological Science*, *20*(1), 6–10. https://doi.org/10.1111/j.1467-9280.2008.02243.x

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements.

Psychological Science, *12*(1), 48–55. https://doi.org/10.1111/1467-9280.00309

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*(1–3), 17–42. https://doi.org/10.1080/135062800394667

Rensink, R., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368–373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Ringer, R. V., Throneburg, Z., Johnson, A. P., Kramer, A. F., & Loschky, L. C. (2016). Impairing the useful field of view in natural scenes: Tunnel vision versus general interference. *Journal of Vision*, *16*(2), 7. https://doi.org/10.1167/16.2.7

Rönne, H. (1915). Zur Theorie und Technik der Bjerrumschen Gesichtsfelduntersuchung. *Archiv fur Augenheilkunde*, *78*(4), 284–301.

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, *2*, 437–457. https://doi.org/10.1146/annurev-vision-082114-035733

Sampanes, A. C., Tseng, P., & Bridgeman, B. (2008). The role of gist in scene recognition. *Vision Research*, *48*(21), 2275–2283. https://doi.org/10.1016/j.visres.2008.07.011

Savino, P. J., & Danesh-Meyer, H. V. (2012). *Color atlas and synopsis of clinical ophthalmology—Wills Eye Institute—neuro-ophthalmology*. Lippincott Williams & Wilkins.

Sekuler, A. B., Bennett, P. J., & Mamelak, M. (2000). Effects of aging on the useful field of view. *Experimental Aging Research*, *26*(2), 103–120. https://doi.org/10.1080/036107300243588

Simons, D. J., & Chabris, C. F. (1998). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*, 1059–1074.

Simons, D., & Levin, D. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267. https://doi.org/10.1016/S1364-6613(97)01080-2

Simons, D. J., & Rensink, R. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. https://doi.org/10.1016/j.tics.2004.11.006

Solovey, G., Graney, G. G., & Lau, H. (2015). A decisional account of subjective inflation of visual perception at the periphery. *Attention, Perception & Psychophysics*, *77*(1), 258–271. https://doi.org/10.3758/s13414-014-0769-1

Swanson, W. H., & Fish, G. E. (1995). Color matches in diseased eyes with good acuity: Detection of deficits in cone optical density and in chromatic discrimination. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *12*(10), 2230–2236. https://doi.org/10.1364/JOSAA.12.002230

Tatler, B., Gilchrist, I., & Risted, J. (2003). The time course of abstract visual representation. *Perception*, *32*(5), 579–593. https://doi.org/10.1068/p3396

Torralbo, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computational and Neural Systems*, *14*, 391–412. https://doi.org/10.1088/0954-898X_14_3_302

Torralbo, A., Walther, D. B., Chai, B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2013). Good exemplars of natural scene categories elicit clearer patterns than bad exemplars but not greater BOLD activity. *PLoS ONE*, *8*(3), e58594. https://doi.org/10.1371/journal.pone.0058594

Tyler, C. W. (2014). Peripheral color demo. *Perception*, *6*, 1–5.

Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*(3), 473–494. https://doi.org/10.1080/13506280444000409

Van Boxtel, J. A., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: On sufficiency and necessity. *Frontiers in Psychology*, *1*, 1–13. https://doi.org/10.3389/fpsyg.2010.00217

Võ, M. L.-H., & Henderson, J. M. (2011). Object–scene inconsistencies do not capture gaze: Evidence from the flash-preview moving-window paradigm. *Attention, Perception & Psychophysics*, *73*(6), 1742–1753. https://doi.org/10.3758/s13414-011-0150-6

Wallis, T. S. A., Funke, C. M., Ekcer, A. S., Gatys, L. A., Wichmann, F. A., & Bethge, M. (2019). Image content is more important than Bouma's Law for scene metamers. *eLife*, *8*, 42512. https://doi.org/10.7554/eLife.42512

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(34), 10573–10581. https://doi.org/10.1523/JNEUROSCI.0559-09.2009

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168. https://doi.org/10.1016/j.tics.2011.02.005

Appendix

Fixation Duration Estimates

Table A1

References to Prior Studies That Measured the Duration of Fixations When Gazing Upon a Variety of Stimuli and Performing a Variety of Tasks

Source	Fixation duration measurement
Rayner, 1998	275 ms
Pelz & Canosa, 2001	327 ms
Hayhoe et al., 2003	225 ms
Henderson, 2003	330 ms
Unema et al., 2005	260 ms
Rayner et al., 2008	256 ms
Castelhano & Heaven, 2011	260 ms
Nuthmann, 2017	275 ms

Note. The average of these measurements were used to determine how long stimuli were presented across all experiments.

Received July 25, 2019
Revision received May 1, 2020
Accepted May 1, 2020 ■