

Natural-Scene Perception Requires Attention

Michael A. Cohen, George A. Alvarez, and Ken Nakayama

Harvard University

Psychological Science
22(9) 1165–1172
© The Author(s) 2011
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797611419168
<http://pss.sagepub.com>



Abstract

Is visual attention required for visual consciousness? In the past decade, many researchers have claimed that awareness can arise in the absence of attention. This claim is largely based on the notion that natural scene (or “gist”) perception occurs without attention. This article presents evidence against this idea. We show that when observers perform a variety of demanding, sustained-attention tasks, inattentional blindness occurs for natural scenes. In addition, scene perception is impaired under dual-task conditions, but only when the primary task is sufficiently demanding. This finding suggests that previous studies that have been interpreted as demonstrating scene perception without attention failed to fully engage attention and that natural-scene perception does indeed require attention. Thus, natural-scene perception is not a preattentive process and cannot be used to support the idea of awareness without attention.

Keywords

inattentional blindness, dual tasks, attention, conscious awareness, scenes, gist

Received 11/12/10; Revision accepted 4/7/11

A central debate in consciousness studies revolves around whether visual attention is required for visual awareness (Block, 2005; Koch & Tsuchiya, 2007; Lamme, 2003; O'Regan & Noë, 2001; Posner, 1994). Classic models of vision posit that conscious perception of feature conjunctions and whole objects requires attention (Treisman & Gelade, 1980). Studies using a variety of psychophysical paradigms involving inattentional blindness, change blindness, and the attentional blink have supported this view by showing that significant events and changes will not reach awareness without attention (Joseph, Chun, & Nakayama, 1997; Mack & Rock, 1998; Rensink, O'Regan, & Clark, 1997).

Amid all of this research, however, findings for one stimulus category stand out. Natural scenes appear to be perceived without attention. This idea was initially based on everyday visual phenomenology: The rest of a scene does not disappear when the observer focuses attention on one item (Block, 1995; Wolfe, 1999). Indeed, a variety of empirical studies have demonstrated the robustness of scene perception. In one powerful demonstration, Mack and Rock (1998) showed that natural scenes are immune to inattentional blindness (i.e., the inability to perceive an otherwise salient stimulus because attention is directed elsewhere). Similarly, other studies have shown that the ability to detect and classify items in a scene is not impaired under dual-task conditions involving a demanding primary task (Li, VanRullen, Koch, & Perona, 2002; Rousset, Fabre-Thorpe, & Thorpe, 2002). Furthermore, in change-blindness

studies, although certain objects embedded in a scene can appear and disappear for more than a minute without being noticed, participants instantly notice if the meaning, or “gist,” of the scene is altered (Rensink et al., 1997; Simons & Levin, 1997). Finally, for several decades, it has been known that certain types of semantic information can be extracted from a scene in under 150 ms, which is thought to be too fast for focused attention to play any substantive role (Biederman, 1972; Potter, 1975; Thorpe, Fize, & Marlot, 1996). The apparent ability to consciously perceive scenes without focusing attention on them has been repeatedly cited as support for the notion of awareness in the absence of attention (Koch & Tsuchiya, 2007; Lamme, 2004; Tononi & Koch, 2008).

Is it the case that conscious perception of natural scenes does not require attention? Alternatively, perhaps natural-scene perception has been classified incorrectly as a preattentive process because it is so efficient that it requires very little attention, and thus is relatively impervious to dual-task interference and inattentional blindness. This second possibility has been hinted at in studies using the attentional blink (Marois, Yi, & Chun, 2004; Slagter, Johnstone, Beets, &

Corresponding Author:

Michael A. Cohen, Harvard University, Department of Psychology, Vision Sciences Laboratory, William James Hall, 33 Kirkland St., Cambridge, MA 02138
E-mail: michaelthecohen@gmail.com

Davidson, 2010), dual-task situations (Rousselet, Thorpe, & Fabre-Thorpe, 2004; Walker, Stafford, & Davis, 2008), rapid presentation (Evans & Treisman, 2005), and perception of two superimposed scenes (Neisser & Becklen, 1975) to show that natural-scene perception can be compromised under specific experimental conditions.

In the experiments reported here, we used multiple-object tracking (MOT; Pylyshyn & Storm, 1988) and rapid serial visual presentation (RSVP) as our primary tasks and found that natural-scene perception was not immune to inattentional blindness and was impaired under dual-task conditions. We also found that dual-task interference occurred only when the primary task was sufficiently demanding. Taken together, these results demonstrate that attention is required to consciously perceive natural scenes, and they suggest that the tasks used in previous studies were not demanding enough, allowing excess attention to be allocated to the natural-scene stimuli. Furthermore, our findings pose a major challenge for models of consciousness in which awareness can arise without attention. Indeed, if previous studies do not provide evidence of preattentive scene perception, there is no longer any empirical evidence of consciousness in the absence of attention.

General Method

Participants

In total, 71 participants (ages 18–37) were tested. All reported normal or corrected-to-normal visual acuity. All participants gave informed consent to participate and were paid \$10 per hour. All procedures were approved by the Harvard University Human Subjects Institutional Review Board.

Apparatus

Stimuli were presented on a 24-in. color iMac running OS X 10.5.6 and were programmed in MATLAB 7.5 (The MathWorks, Natick, MA) using the Psychophysics Toolbox routines (Brainard, 1997; Pelli, 1997). The monitor was set to a spatial resolution of 1,920 × 1,200 pixels and a refresh rate of 60 Hz. Observers were seated approximately 57 cm from the monitor, so 1 cm on the screen subtended 1° of visual angle.

Stimuli

In the MOT task, displays consisted of identical discs that moved while images in the background changed rapidly. In the RSVP task, letters and digits were presented against changing background images. The background images were masks and natural scenes. The masks were colored checkerboards consisting of a 4 × 4 grid of equally sized boxes, an 8 × 8 grid of equally sized boxes, or a grid combining boxes of both sizes (see the Supplemental Material available online for examples); colors were randomly assigned to the boxes. In all experiments, scenes were always the second-to-last background

image presented on the critical trials and were followed by one mask. Masks of the three types were presented in a randomized order. In every experiment, the background images took up 21° of visual angle, and the MOT discs and RSVP streams were drawn in a 20.5° × 20.5° imaginary window.

Experiment 1: Inattentional Blindness for Natural Scenes

It has previously been claimed that inattentional blindness does not occur for natural scenes (Mack & Rock, 1998). Reasoning that the primary tasks used in previous research might not have fully engaged attention, in Experiment 1 we employed paradigms that require continuous, sustained attention and therefore are likely to be more attentionally demanding than the tasks employed in previous work. In Experiment 1a, we used an MOT task in which participants tracked four of eight identical black discs moving at 10.5° per second (see Fig. 1a). This task has two advantages over the tasks used previously: Tracking requires continuous, sustained attention, because even a momentary lapse in attention can cause observers to lose track of targets without any possibility of recovery. In Experiment 1b, we used an RSVP task (see Fig. 1c), which also required sustained attention, to ensure that the effects we obtained in Experiment 1a were due to attentional limitations and were not a specific artifact of motion processing. Twenty-five participants were recruited for Experiment 1a, and 30 different participants were recruited for Experiment 1b.

Method

Inattentional blindness procedure. In Experiment 1a (MOT task), each trial began with eight identical discs randomly placed on a single checkerboard background. The four discs that were the targets to be tracked turned green for 3 s. All eight discs then moved for a duration randomly selected from a uniform distribution ranging from 3 to 6 s in 100-ms steps (e.g., 3,000 ms, 3,100 ms, 3,200 ms). (See the Supplemental Material for information on the motion algorithm.) The background checkerboard changed every 67 ms (Fig. 1a). On the first four, noncritical trials, all the discs stopped moving at the end of the trial, and participants clicked on each of the targets with the mouse. Feedback was given immediately. On the fifth, critical trial, a scene unexpectedly replaced the second-to-last mask, and no MOT response was given at the end of the trial; instead, participants were immediately probed about the scene (see the Supplemental Material for the script that was used).

The scenes used in Experiment 1a came from five categories (beach, building, highway, indoor, and mountain; see the Supplemental Material for the images used). Each scene was presented to 5 participants.

In Experiment 1b (RSVP task), participants' task was to count the number of times a digit was presented in a stream of letters. At points where the characters overlapped with the background, they were reduced to 35% transparency so that

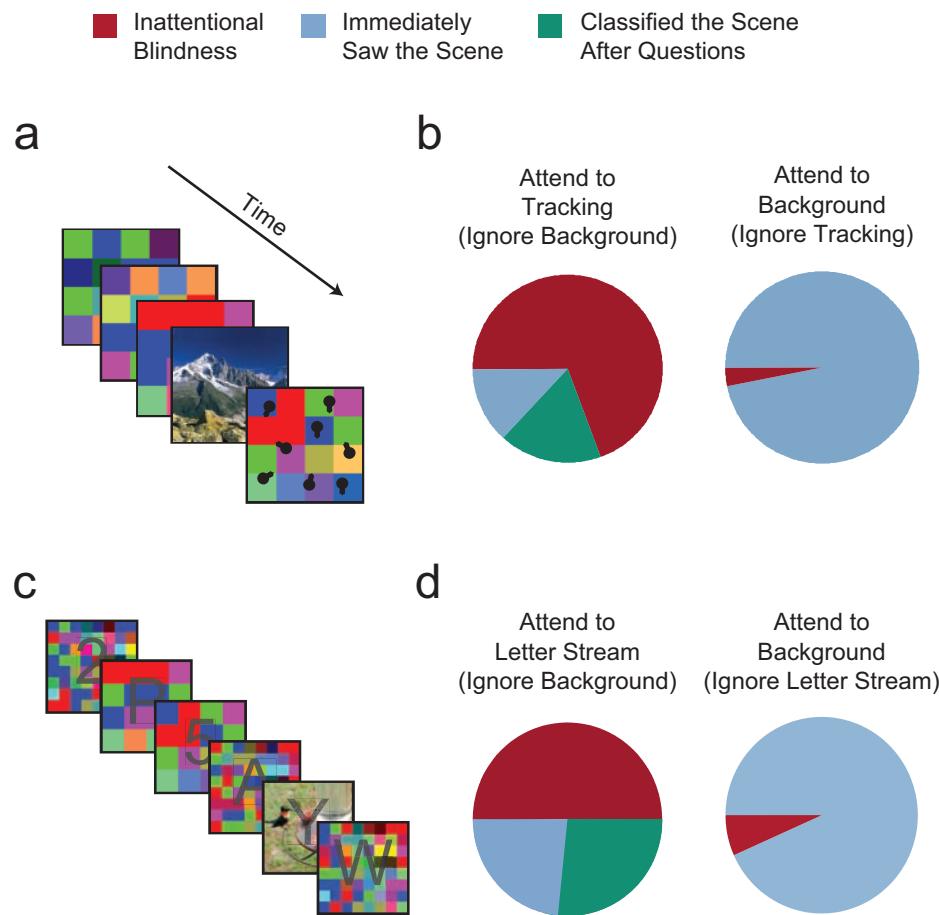


Fig. 1. Illustration of the inattentional blindness procedure in Experiment I and experimental results. In both experiments, the inattentional blindness procedure consisted of five trials. In Experiment 1a (a), participants performed a multiple-object-tracking (MOT) task in which eight black moving discs were presented against a changing background (new background every 67 ms) consisting of randomly colored checkerboard masks; four of the discs were identified as targets to be tracked. In Experiment 1b (c), participants viewed a rapid serial visual presentation (RSVP) stream of letters and digits presented against a series of changing checkerboard masks (new background every 100 ms); the task was to count the number of times a digit was presented. In both experiments, on the critical, fifth trial (but never on the preceding trials), a natural scene unexpectedly replaced the second-to-last mask, and participants were asked questions about the scene. Following the inattentional blindness procedure, participants completed 20 trials in which they attended to the background while ignoring the moving discs (Experiment 1a) or the RSVP stream (Experiment 1b) and answered questions about the background scenes. The pie charts (b, d) summarize performance on the critical and post-inattentional blindness trials, showing the rates of inattentional blindness, immediate perception of the scene, and correct identification of the scene.

they appeared transparent and the background could be seen through them. The background changed every 100 ms, and with each change in the background, a new letter or digit was presented. As in Experiment 1a, on the fifth, critical trial, a scene unexpectedly replaced the second-to-last mask in the background. In all positions in the stream except for the last two, the character could be either a letter or a digit. On both critical and noncritical trials, the last two characters presented were always letters, never digits. The distractor letters were randomly drawn from the following set: *A, B, C, D, G, H, K, M, N, O, P, Q, R, T, U, V, W, X, and Y*. The target digits were *1, 2, 4, and 5*. No target digit was ever presented more than once in a given trial. Thus, there could be no, one, two, three, or four

target digits presented in a trial. Each stream contained 12 to 17 displays (100 ms per display), so trials lasted from 1,200 to 1,700 ms; the length of each trial was randomly chosen. At the end of the first four trials, participants provided their answer to the digit task by pressing the "0," "1," "2," "3," or "4" key on the computer keyboard. Immediate feedback was given at the end of each trial. On the critical trial, however, participants did not indicate the number of digits that had been presented. Instead, as in Experiment 1a, they were immediately probed about the scene presented in the background on that trial (see the Supplemental Material for the script that was used).

The scenes used in Experiment 1b each contained an animal or a vehicle. We tested scenes showing animals and vehicles

because in previous studies cited to support the claim of scene perception without attention, participants were asked to determine whether each scene contained an animal (Li et al., 2002; Rousselet et al., 2002; Thorpe et al., 1996) or a vehicle (Li et al., 2002; Fei-Fei, VanRullen, Koch, & Perona, 2005). Of the six different images used (see the Supplemental Material), three contained an animal, and three contained a vehicle. Each photo was presented to 5 participants.

In both experiments, participants who accurately identified the scene on the critical trial after being asked if they noticed anything different on that trial were classified as having immediately seen the scene. Participants who correctly labeled the scene after being asked any of the subsequent questions were classified as having classified the scene after questions. Questioning stopped if the participant classified the scene before the last question was asked. Only participants who were asked all questions and were not able to classify the scene correctly when given a five- or six-alternative forced-choice task were classified as having never seen the scene and labeled as having experienced inattentional blindness.

Post-inattentional blindness testing. In both experiments, after the critical trial, participants completed 20 trials in which they attended to the background while ignoring the moving discs (Experiment 1a) or RSVP stream (Experiment 1b). Participants were asked to identify the type of scene that was present, and a scene was present on half the trials. Images were drawn from the same categories as for the inattentional blindness procedure but were entirely different sets of images (i.e. different mountains and beaches for Experiment 1a and different pictures with animals for Experiment 1b). We included these trials to make sure that any observed inattentional blindness was not due to the scenes being imperceptible because of the presence of the MOT or RSVP stimuli.

Results and discussion

In Experiment 1a, 64% of participants experienced complete inattentional blindness on the critical trial (Fig. 1b). Only 18% of participants detected the scene immediately. However, in the post-inattentional blindness testing, in which participants were explicitly asked to detect and classify the scenes, they did so correctly 96% of the time (Fig. 1b).

In Experiment 1b, on the critical trial, 50% of participants experienced total inattentional blindness, and 23% of participants were able to identify the category of the scene immediately (Fig. 1d). On the 20 post-inattentional blindness trials, participants successfully noticed and classified the scenes 93% of the time (Fig. 1d).

The fact that participants could detect and categorize scenes so accurately during the post-inattentional blindness testing in both experiments strongly suggests that participants' inattentional blindness on the critical trial arose specifically because of attentional allocation. Thus, by using a sustained, attentionally demanding task, we were able to induce inattentional blindness

for natural scenes in a large percentage of participants. The combined results of Experiments 1a and 1b stand in stark contrast to Mack and Rock's (1998) results, which did not include a single instance of inattentional blindness for scenes presented with a 30-ms duration (note that we used longer durations of 67 and 100 ms). It is possible that the presence or absence of preceding masks is an important difference between the experiments. That is, the sudden onset of critical stimuli not preceded by masks may have captured attention, allowing the critical stimuli to be perceived in Mack and Rock's study. By this logic, our preceding masks minimized the transient, bottom-up signal caused by the presentation of a scene. In addition, it is likely that Mack and Rock's primary task was not demanding enough to produce inattentional blindness. When task demands are adjusted so that attention is properly and continuously engaged, natural-scene perception is not immune to inattentional blindness.

Experiment 2: Multiple-Object Tracking Impairs Natural-Scene Perception

Previous studies have found that scene perception is unaffected by dual-task interference (Fei-Fei et al., 2005; Li et al., 2002; Rousselet et al., 2002). In Experiment 2, we tested the possibility that these results were obtained because the primary task used was not sufficiently difficult. Thus, we varied the difficulty of the primary task and measured what effects, if any, this had on scene perception. In the dual-task condition, participants simultaneously performed the MOT task from Experiment 1 and a scene-classification task. Difficulty of the MOT task was manipulated by adjusting the speed of the disks (object speed has been shown to lead to variations in MOT performance; Alvarez & Franconeri, 2007).

If scene perception requires attention, but only a minimal amount, tracking and scene perception should show no interference when tracking speed is slow because the easy tracking task should leave enough resources available for the scene perception task (Lavie, 2005; Yi, Woodman, Widders, Marois, & Chun, 2004). However, at faster tracking speeds, few or no resources should be available for the scene perception task, and a dual-task cost should be observed for one or both tasks.

Method

Participants. Eight participants took part in this experiment.

Procedure. The method and parameters of Experiment 2 were identical to those of Experiment 1a except as noted here. Participants knew that scenes would be presented regularly and were instructed to detect and classify those scenes. For the tracking-only condition, participants were instructed to ignore the background stream, and a scene was never presented in the background.¹ When scene detection was the only task, participants were told to ignore the moving discs and to focus only on the background. In the dual-task condition, the two tasks

were performed simultaneously. For the scene-task-only and dual-task conditions, scenes were presented on half of the trials. Two blocks of each task condition (i.e., two single-task MOT blocks, two single-task scene-detection blocks, and two dual-task blocks) were presented for each of two disc speeds, 4.5° per second and 10.5° per second (total of six blocks and 480 trials). As in Experiment 1a, the masks were colored checkerboards that changed every 67 ms. The rate at which the background images changed did not change when the speed of the discs in the MOT task was adjusted. At the end of dual-task trials, participants first clicked on the target dots and then provided an answer for the scene task using the keyboard.

Results and discussion

When the discs moved at the slower speed, scene perception was unimpaired in the dual-task condition, $t(7) = 1.17$, $p = 0.28$ (Fig. 2a). Taken alone, this lack of interference would suggest that scene perception does not require attention. However, when the speed of the discs increased, performance on the scene task was impaired in the dual-task condition, $t(7) = 3.5$, $p < .01$ (Fig. 2a). The same pattern of impairment was found for the MOT task (Fig. 2b). Tracking performance did not differ between the single- and dual-task conditions when the discs moved at the slower speed, $t(7) = 1.09$, $p = 0.31$, but did differ between these conditions when the discs moved at the faster speed, $t(7) = 2.9$, $p < .05$.²

The fact that tracking performance was also affected under dual-task conditions suggests not only that scene perception requires attention, but also that this requirement is substantial, such that attention must be allocated from a competing task. These results suggest that previous results purported to show that focused attention is not needed for scene perception were likely obtained because attention was not sufficiently engaged by the competing task.

Experiment 3: Performing an RSVP Task Impairs Natural-Scene Perception

In our next experiment, we sought to extend the findings from Experiment 2 using the RSVP paradigm of Experiment 1b. We did this to see if the dual-task costs previously observed were due to the use of an attentionally demanding task or the use of a primary task that involved motion. In the dual-task conditions, participants counted the number of digits presented in the RSVP stream while simultaneously performing a detection-and-categorization task on the background scene.

Method

Eight participants performed the RSVP task and the two natural-scene perception tasks (i.e., five-categories task, animal-vehicle task; see the Supplemental Material for example trials) both in single-task conditions and in dual-task conditions that combined each scene task with the RSVP task. New sets of scenes were selected for both the five-categories task and the animal-vehicle task. In the five-categories condition, the scene task was to classify each scene as being a mountain, beach, highway, indoor, or building scene. In the animal-vehicle condition, the scene task was a discrimination task in which participants reported whether they saw a scene that contained an animal, a scene that contained a vehicle, a scene without either target, or no scene. For the stimuli without an animal or vehicle, we selected scenes in which an animal or vehicle might easily have been present but simply was not (e.g., a forest or highway scene; see the Supplemental Material for example stimuli).

Each RSVP stream contained 12 to 17 displays (100 ms per display), so trials lasted from 1,200 to 1,700 ms. The length of trials was uniformly distributed so that all lengths occurred the same number of times for each subject. The five-categories condition and the animal-vehicle condition each comprised

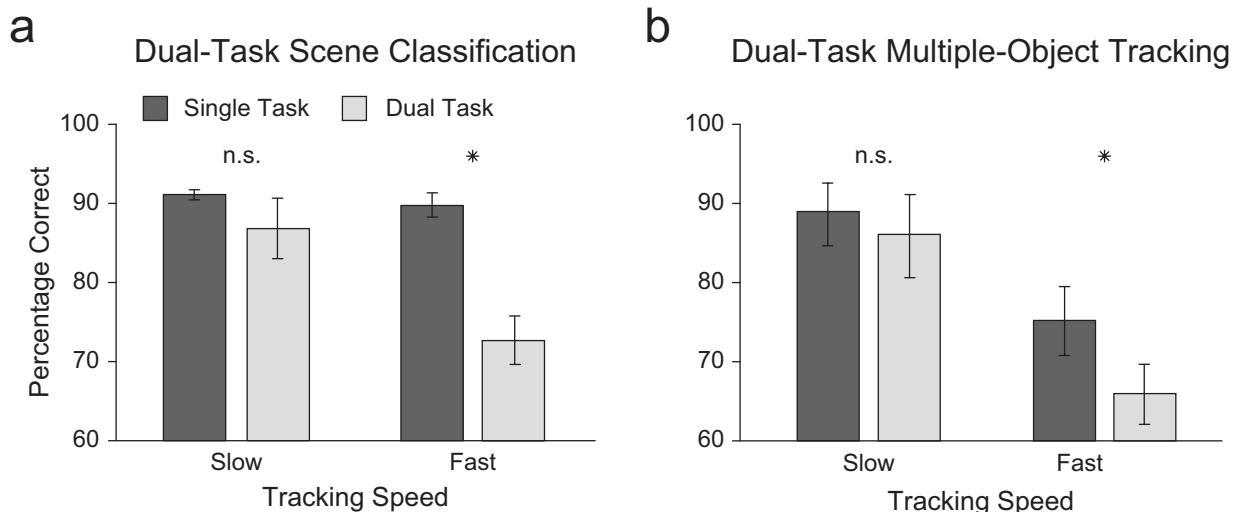


Fig. 2. Experiment 2 results: (a) natural-scene classification and (b) tracking performance in the single-task and dual-task conditions as a function of the speed of the moving discs in the tracking display. Error bars represent standard errors of the mean. Asterisks indicate significant differences between conditions ($p < .05$, within-subjects t tests).

three blocks: a single-task block of the scene task, a single-task block of the RSVP task, and a dual-task block. Participants completed 120 trials in each block of the five-categories condition and 100 trials in each block of the animal-vehicle condition. The order in which the six blocks were completed was counterbalanced across subjects. In all blocks, a scene was presented 50% of the time. Within each block, the various scene categories were presented equally often. On dual-task trials, subjects first provided a response to the RSVP task and then to the scene task.

Results and discussion

In the five-categories condition, performance in the dual-task condition was inferior for both the scene task, $t(7) = 7.05$, $p < .001$, and the RSVP task, $t(7) = 4.43$, $p < .01$ (Fig. 3a). In the animal-vehicle condition, performance in the dual-task condition was again inferior for the scene task, $t(7) = 4.76$, $p < .01$, and the RSVP task, $t(7) = 4.96$, $p < .01$ (Fig. 3b). Thus, dual-task costs are not observed only when motion processing is involved, but appear to be observed when the secondary task is attentionally demanding in any of a variety of ways. Given that we followed previous studies in using a task that required the detection and classification of animals and vehicles within a scene, the difference in results between the current study and previous studies (Li et al., 2002; Rousselet et al., 2002; Thorpe et al., 1996) cannot be attributed to a difference in the task used to assess scene perception. In fact, we found that the five-categories task was more resistant to dual-task interference (an 11% decrease in performance) than the animal-vehicle task was (a 26% decrease). This result suggests that the amount of attention required to identify objects in scenes is greater than the amount required for identifying a scene's gist. However, whether this is generally true cannot be firmly concluded without further research.

General Discussion

The relationship between attention and awareness is one of the most hotly debated issues in neuroscience and psychology (Block, 2005; Koch & Tsuchiya, 2007; Lamme, 2003; Mack & Rock, 1998; O'Regan & Noë, 2001). The predominant view is that certain classes of stimuli—in particular, natural scenes—can be perceived without attention—that is, “preattentively.” What reason is there to believe in awareness without attention? Initially, this position was supported by the finding that basic visual features pop out in visual search (Treisman & Gelade, 1980). However, it has since been shown that attention plays an important role in feature perception, as features can be missed during the attentional blink (Joseph et al., 1997) and pop-out does not occur when spatial cues direct attention away from the target location (Theeuwes, Kramer, & Atchley, 1999).

More recently, natural-scene perception has been repeatedly cited as providing the primary evidence for awareness without attention (Koch & Tsuchiya, 2007; Lamme, 2004; Tononi & Koch, 2008). In the experiments reported here, we tested this view using multiple types of attentionally demanding tasks. We found that observers can be rendered inattentionally blind to natural scenes and that scene perception can be impaired under dual-task conditions. Thus, perception of natural scenes does require attention.

It should be noted that although performance on the scene tasks dropped in dual-task conditions, it remained well above chance. It is possible that our primary tasks were not sufficiently difficult and that there were still attentional resources that could be directed toward the scene images. However, it is also possible that some, but not all, aspects of scene perception require attention and that it was those aspects that were affected in our study. Additional research will be needed to identify if there are in fact any particular aspects of scene perception that are systematically immune to attentional interference.

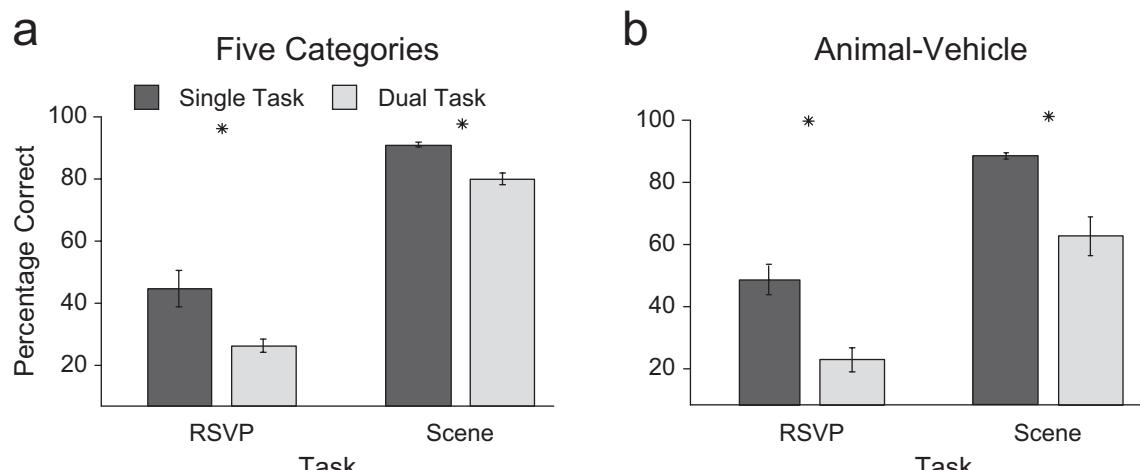


Fig. 3. Experiment 3 results: single- and dual-task accuracy on the rapid serial visual presentation (RSVP) task and the scene task in the (a) five-categories condition and (b) animal-vehicle condition. Error bars represent standard errors of the mean. Asterisks indicate significant differences between conditions ($p < .05$, within-subjects t tests).

Although attention may be necessary for the conscious perception of natural scenes, several pieces of evidence suggest that processing to high levels can still occur in the absence of attention, awareness, or both. Recently, Serre, Oliva, and Poggio (2007) have shown that a purely feed-forward architecture can accurately predict human observers' level and pattern of performance on a task requiring rapid categorization of stimuli as animals or nonanimals. In addition, it has been repeatedly demonstrated that participants who are presented with two scenes and instructed to look quickly toward the scene that contains an animal, a vehicle, or a face can make ultrarapid saccades toward the target scene within only 100 to 130 ms (Crouzet, Kirchner, & Thorpe, 2010; Kirchner & Thorpe, 2006). Thus, certain behavioral decisions concerning the nature of a scene can be biased in one way or another without conscious awareness (i.e., the subject is predisposed to looking, say, toward the right even without awareness of the scene to the right). This idea is consistent with research showing that subconscious information can influence subsequent behaviors and decisions (Naccache & Dehaene, 2001). It should be emphasized that this notion is in no way inconsistent with the present findings: Although unconscious, high-level scene processing can occur without attention, attention is necessary for the natural scene to reach conscious awareness.

In previous studies (Li et al., 2002; Rousselet et al., 2002), task difficulty was set so that primary-task performance was well below 100% correct (i.e., ceiling). The logic behind this decision was that if primary-task performance was below ceiling, attention was fully engaged by this task, and consequently, a task that could be completed in addition to the primary task was completed without attention. However, the present study presents a challenge to this logic by showing that adjusting primary-task performance to below ceiling is not enough to ensure that attention is entirely occupied by this task. In the single-task condition of Experiment 2, participants always performed below ceiling on both the tracking and the scene-classification tasks (Figs. 2a and 2b). Nevertheless, it was only with the higher tracking speed that there was a mutual dual-task cost for both tracking and scene performance (Figs. 2a and 2b). The insight that task performance cannot be used as a one-to-one index of available attentional resources has important implications for future research, particularly if the goal is to understand the attentional requirements of perceptual processes. It is possible that there may be an excess of attentional resources even when the primary task is quite challenging (Norman & Bobrow, 1975). Indeed, there might even be an actual inability to allocate more resources to a task despite their availability (Kahneman, 1973). This idea is merely speculative at this point, and future research will have to address this important topic. For our present purposes, the important point is that below-ceiling performance does not necessarily indicate that attentional resources are fully consumed.

In sum, the present results indicate that natural-scene perception is so efficient and requires so little attention that the perceptual system must be properly taxed if this attentional

cost is to be identified. Researchers who propose models of awareness that dissociate attention and awareness (Block, 2005; Koch, & Tsuchiya, 2007; Lamme, 2003) cannot use gist perception or the perception of objects, such as animals, in natural scenes as examples of awareness that does not require attention. Indeed, if there is no evidence of natural-scene perception without attention, no evidence of awareness without attention currently exists. Although there is good reason to believe in attention without awareness (Naccache, Blandin, & Dehaene, 2002), there is no evidence for awareness without attention. In conjunction with previous research, the present results support the claim that visual attention is required for visual awareness.

Acknowledgments

We thank Justin Jungé, Arien Mack, Rufin VanRullen, Maryam Vaziri-Pashkam, and Jeremy Wolfe for helpful conversations on the project, and Ron Rensink and an anonymous reviewer for comments on the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by a National Science Foundation Graduate Research Fellowship to M. A. C., a National Science Foundation Career Award (BCS-0953730) to G. A. A., and a National Institutes of Health grant (RO1 EY01362) to K. N.

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by-supplemental-data>

Notes

1. A control experiment was run to ensure that single-task tracking is unaffected by the presence or absence of a scene in the background. A group of 8 participants completed 60 trials of the MOT task in which half of the trials contained no natural scene in the background stream and the other half had a scene from one of the five categories presented as the second-to-last image in the stream. There was no difference in performance between trials in which a scene was present ($M = 76\%$ correct, $SEM = 2.8$) and trials in which no scene was present ($M = 74\%$ correct, $SEM = 3.6$), $t(7) = 0.35$, $p = .74$.
2. In a control experiment, 6 additional participants performed the MOT and scene tasks in both single- and dual-task conditions. The only differences from the main experiment were that only the faster tracking speed was used, and for the scene task, participants had to say whether a scene that contained an animal, a scene that contained a vehicle, a scene with no target, or no scene had been presented. For this control experiment, a new set of stimuli was selected. We found a significant drop in performance from the single- to the dual-task condition for both the MOT task, $t(5) = 3.65$, $p < .05$, and the scene task, $t(5) = 2.72$, $p < .05$.

References

- Alvarez, G., & Franconeri, S. (2007). How many objects can you attentively track? Evidence for a resource-limited tracking mechanism. *Journal of Vision*, 7(13), Article 14. Retrieved from <http://www.journalofvision.org/content/7/13/14/>
- Biederman, I. (1972). Perceiving real world scenes. *Science*, 177, 77–80.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral & Brain Sciences*, 18, 227–247.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9, 46–52.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Crouzet, S., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*, 10(4), Article 16. Retrieved from <http://www.journalofvision.org/content/10/4/16/>
- Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1476–1492.
- Fei-Fei, L., VanRullen, R., Koch, C., & Perona, P. (2005). Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. *Visual Cognition*, 12, 893–924.
- Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a “preattentive” feature search task. *Nature*, 387, 805–808.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, 46, 1762–1776.
- Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences*, 11, 16–22.
- Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, 10, 204–211.
- Lamme, V. A. (2004). Separate neural definitions of visual consciousness and visual attention: A case for phenomenal awareness. *Neural Networks*, 17, 861–872.
- Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, 9, 75–82.
- Li, F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, USA*, 99, 9596–9601.
- Mack, A., & Rock, I. (1998). *Inattentional blindness: Perception without attention*. Cambridge, MA: MIT Press.
- Marois, R., Yi, D.-J., & Chun, M. M. (2004). The neural fate of perceived and missed events in the attentional blink. *Neuron*, 41, 464–472.
- Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, 13, 416–424.
- Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, 80, 215–229.
- Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, 7, 480–494.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44–64.
- O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral & Brain Sciences*, 24, 939–973.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences, USA*, 91, 7398–7403.
- Potter, M. C. (1975). Meaning in visual scenes. *Science*, 187, 964–967.
- Plylyshyn, Z., & Storm, R. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 1–19.
- Rensink, R. A., 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, 368–373.
- Rousselet, G., Fabre-Thorpe, M., & Thorpe, S. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience*, 5, 629–630.
- Rousselet, G., Thorpe, S., & Fabre-Thorpe, M. (2004). Processing of one, two or four natural scenes in humans: The limits of parallelism. *Vision Research*, 44, 877–894.
- Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences, USA*, 104, 6424–6429.
- Simons, D., & Levin, D. (1997). Change blindness. *Trends in Cognitive Sciences*, 1, 261–267.
- Slagter, H. A., Johnstone, T., Beets, I. A. M., & Davidson, R. J. (2010). Neural competition for conscious representation across time: An fMRI study. *PLoS ONE*, 5(5), e10556. Retrieved from <http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0010556>
- Theeuwes, J., Kramer, A. F., & Atchley, P. (1999). Attentional effects on preattentive vision: Spatial precues affect the detection of simple features. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 341–347.
- Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: An update. *Annals of the New York Academy of Sciences*, 1124, 239–261.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Walker, S., Stafford, P., & Davis, G. (2008). Ultra-rapid categorization requires visual attention: Scenes with multiple foreground objects. *Journal of Vision*, 8(4), Article 21. Retrieved from <http://www.journalofvision.org/content/8/4/21>
- Wolfe, J. M. (1999). Inattentional amnesia. In M. Coltheart (Ed.), *Fleeting memories* (pp. 71–94). Cambridge, MA: MIT Press.
- Yi, D.-J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience*, 7, 992–996.