Replication of Cohen, M. a, Alvarez, G. a, & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science*, *22*(9), 1165–72. doi:10.1177/0956797611419168

Daniel Birman
dbirman@stanford.edu

Introduction

In everyday life we feel a direct and undeniable connection between attending to something and our clear awareness of it. But is this connection necessary for visual awareness? Despite considerable evidence that visual attention is required for visual consciousness (Mack & Rock, Irvin, 1998) a number of researchers have shown that perception of the gist of a scene is possible without top-down selective attention (Koch & Tsuchiya, 2007; Lamme, 2003; Li, VanRullen, Koch, & Perona, 2002; Reddy, & Koch, 2006; van Boxtel, Tsuchiya, & Koch, 2010). A number of researchers have since taken aim at these findings, showing that under sufficiently constrained conditions awareness of gist perception is indeed degraded by selective attention (Cohen, Alvarez, & Nakayama, 2011; Mack & Clarke, 2012). These more recent experiments have utilized more careful methods, with an emphasis on quantifying the attentional load necessary to impair gist perception. Cohen et al. demonstrated that multiple object tracking and rapid serial visual presentation tasks impaired gist perception only within a range of difficulty—not so difficult that participants cannot complete the task, but sufficiently difficult that their entire focus of attention is on task demands. In addition, the authors reported that many subjects were aware of an event or change occurring (i.e. they were not change blind) but could not report the contents of the change (i.e. they experienced inattentional blindness). Taken together these findings suggest that visual consciousness can be impaired under high attentional load.

In this study we sought to replicate the main finding of Cohen et al. who showed that under sufficient attentional load visual consciousness of scene gist is impaired. Cohen et al. presented participants with a 100 ms natural scene image. While performing a different demanding task (either RSVP or motion tracking) participants were shown an unexpected scene in the background and probed about their awareness of scene gist. Cohen et al. showed that, unlike in previous findings (e.g. Li et al. 2002), participants do not have gist perception in the absence of attention. Mack and Clarke (Mack & Clarke, 2012) explain that this design philosophy has a number of parameters that interact to degrade scene gist perception: (1) scene presentation length, (2) presentation location relative to focal task, (3) scene features (e.g. color, contrast, contents), (4) focal task difficulty, (5) focal task modality. Our expectation was that our

procedure, similar to Cohen et al. experiment 1b, would induce inattentional blindness in ~50% of the study population, supporting previous evidence that inattentional blindness can occur for all forms of visual stimuli.

Methods

Power Analysis

The original study used a sample size of n=30, of which five participants saw each of the six critical trial images. The RSVP task manipulation induced inattentional blindness in 50% of participants while 23% of participants saw and classified the scene immediately. Under the hypothesis that gist perception is always possible despite allocating attention to an alternate task we would expect a rate of inattentional blindness of 0% (n=0), similar to the rate of gist perception when attending to the background scene. Cohen et al. report that when attending to the background participants experienced inattentional blindness 7% (n=30) of the time. We computed a test for equal proportions on 50% and 7% with n=30 and found a 95% CI of [23%, 100%] for the true proportion of inattentional blindness in our sample. This analysis had a power of 99%. For 80% power we would need a sample of n = 12, 90% n = 17, 95% n = 21. Given the strength of this effect, the small sample sizes, and the ease of collecting MTurk data, we saw no reason to use a sample size smaller than n=30.

Planned Sample

We plan to use a sample of 30 mechanical turk participants. We do not plan on doing pre-selection, other than ensuring that participants have a computer screen that is capable of displaying our experiment, e.g. a vertical height of >700 pixels. Participants who report an age under 18 years old will be excluded from the analysis. Participants who leave ‘full screen’ mode and then return to the experiment on or prior to the critical trial will be excluded from the analysis. Participants who leave and return on a regular trial will have that single trial omitted from the analysis.

Materials

Where our materials differ from the original we have highlighted the changes in **bold**.

Experiment 1 Inattentional Blindness script: (Cohen et al., 2011)

At the end of the critical trial, participants provided no response to the primary task and were immediately asked (the original text used “stream”, for clarity we replaced this with “sequence”):

1. Did you notice anything different on that trial?

2. Did you notice something different about the background **sequence** of images?

3. Did you notice that a different type of image was presented in the background that was unique in some particular way?

4. Did you see an actual photograph of a **[omitted: natural]** scene in that **sequence**?

5. If **we** were to tell you that there was a photograph in that sequence, **could** you tell **us** what it was a photograph of?

6. **[Original, rephrased in bold:** If I were to tell you that one of these 6 images was actually presented on that trial, can you pick out which photograph was shown?**] One of these 6 images was actually presented on that trial, can you pick out which photograph was shown? If you aren't sure, but you know whether there was an animal or vehicle present, pick your best guess among those pictures.**

Example Images: (Cohen et al., 2011)


RSVP Task Letters and Digits: (Cohen et al., 2011)

A, *B*, *C*, *D*, *G*, *H*, *K*, *M*, *O*, *P*, *Q*, *R*, *T*, *U*, *V*, *W*, *X*, and *Y*

1, 2, 4, and 5

Procedure

Our methods are largely identical to the original Cohen et al. procedure. Adapting the stimulus to an online setting required making a number of small changes which have been **bolded** in the text below.

“Participants’ task was to count the number of times a digit was presented in a sequence of letters. At points where the characters overlapped with the background, they were reduced to 35% transparency so that 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 sequence 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*, *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 sequence 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.” (Cohen et al., 2011). **Cohen et al. labeled participants according to when during questioning they succeeded (or failed) to identify the unexpected image. We chose to directly report the exact outcome data for all participants, rather than binning them into specific subsets.**

“After the critical trial, participants completed **5 (originally 20)** trials in which they attended to the background while ignoring the RSVP sequence (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.” (Cohen et al., 2011)

In addition, we included 5 trials in which participants performed the RSVP task with background scenes present, to ensure that background scene presentations did not interfere with performance on the RSVP task (Cohen et al. 2011 report a similar control experiment, but do not report the data).

Analysis Plan

The data from this replication were analyzed directly to determine the percentage of participants who experienced total inattentional blindness compared to participants who either saw the scene immediately or classified the scene as either animal or vehicle after being asked questions. By comparing the proportion of participants who experienced inattentional blindness on the critical trial, while attending to the RSVP task, compared to the proportion of participants who experienced inattentional blindness when attending to the background we can determine the influence of attention on gist perception.

Differences from Original Study

The original experiment was conducted in a controlled environment with the experimenter present in the room. This replication will be conducted in an unconstrained environment, using computer monitors where the distance to screen, screen-size, and other factors will be uncertain and the setting is unknown. Despite this, the difficulty of the RSVP task ensures that participants must pay close attention to their screen to complete the task and this should have ensured that there would only be a small difference between our replication and the original study. We expected that these differences would cause an increase in the percentage of participants who experienced inattentional blindness to the critical scene, but that this increase would be marginal or insignificant relative to the overall effect. In addition, we did not expect the change in setting to cause all participants to experience inattentional blindness.

(Post Data Collection) Methods Addendum

Actual Sample

We collected data from 30 subjects on Mechanical Turk. According to the criteria outlined above one subject was excluded for leaving full screen prior to the critical trial. One subject left full screen after the critical trial, and data from that trial was discarded from the analysis. Male/female ratio was 15/14. Age was 34 +- 11.

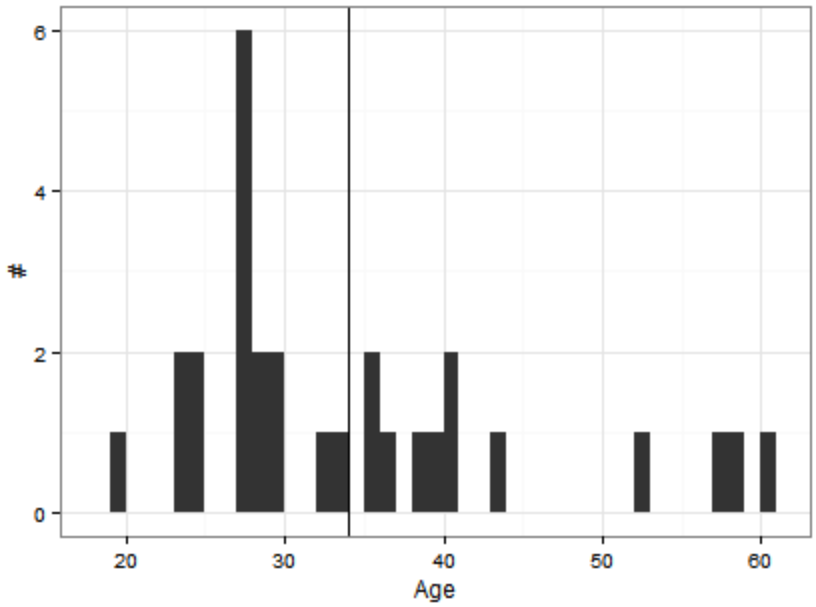

Fig 1: Histogram of participant age in Birman replication.

Differences from pre-data collection methods plan

None

Results

Data preparation

Data was converted into an R-readable format, identifiable information (workerID, IP address) were stripped, and stored at https://github.com/dbirman/repo254/production-results. The following analysis was split according to trial type:

RSVP trials were marked as “correct” or “incorrect” based on whether participants chose the correct number of digits. This is identical to the procedure used by Cohen et al.

Background trials were analyzed in two streams. We checked whether participants got the correct category information as well as whether they identified the actual image seen. marked as either: “Saw the correct image immediately” if participants responded “yes” to the first question and then identified the image, “saw something, but chose the incorrect image” if they responded “yes” to the first question but then chose wrong, “chose correct image after questions” if participants responded “no” to any of the first questions but ultimately chose the right image, or “blind” if they responded no to the questions and then chose the incorrect image in a six-option forced choice array. Note: this splitting mirrors the groupings used by Cohen et al.

Confirmatory Analysis:

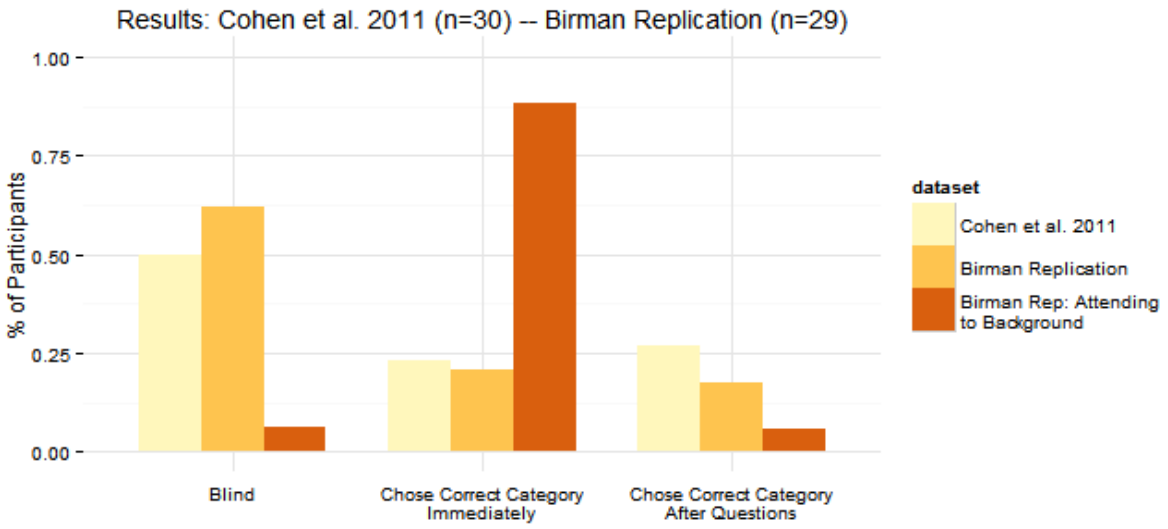

Fig 2: Comparison of Critical Trial Results in Cohen et al. with Birman replication. We found similar magnitude effects for all participant categories on the critical trial. Approx. half of participants were unable to correctly categorize the background image. Of the approx. one in four participants who claimed to have seen the background image appear all correctly identified its category. A small portion of participants chose the correct image category despite claiming not to have seen the image appear. Note that if participants were perfectly blind and picking randomly then the “Blind” and “Chose Corr Cat After Questions” categories would be of equal magnitude. We found here instead a small bias towards picking the *incorrect* image category. Note that these are actual percentages of participants in each grouping on the critical trial not estimated statistics.

Exploratory Analysis: Results for Individual Image Identification

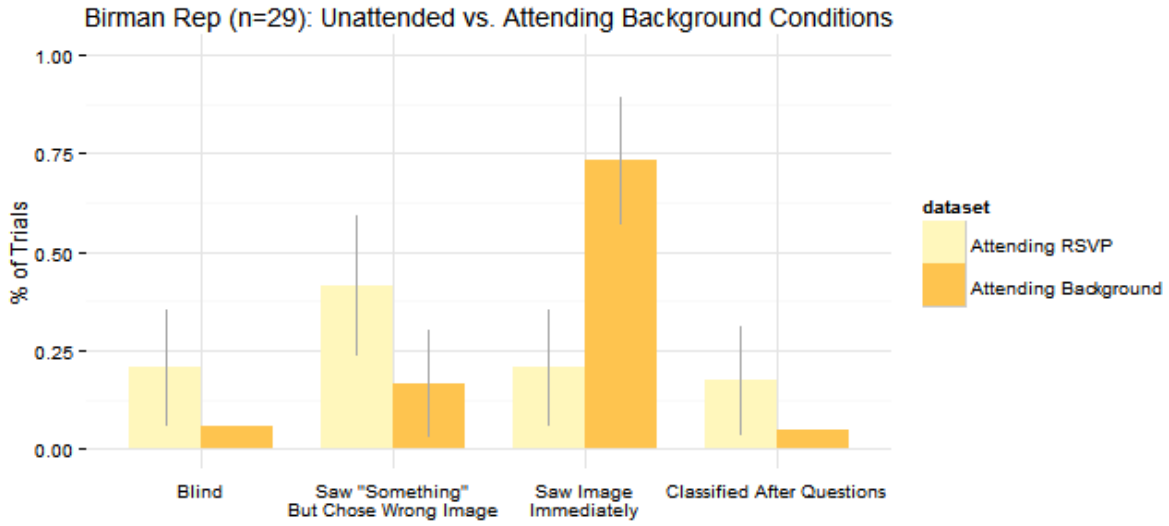

Fig 3: Birman Replication Unattended vs. Attending Background Conditions. We found that when attending to the background images participants correctly identified the specific image that appeared from a six-image array on 75% of trials. In the remaining trials they were only very rarely blind, but often made mistakes. In contrast when attending to the RSVP task on the critical trial participants were blind 75% of the time and chose the correct image immediately only 25% of the time. Often participants “saw” that something appeared but were unable to choose the correct image from the six-image array. The similar magnitudes of the “classified after questions” and “blind” categories for the Attending to RSVP trial suggests that participants were not choosing images from the six-image array randomly, but were biased towards making the correct choice. Note that the confidence intervals for two of the groups were near zero.

Control Tasks

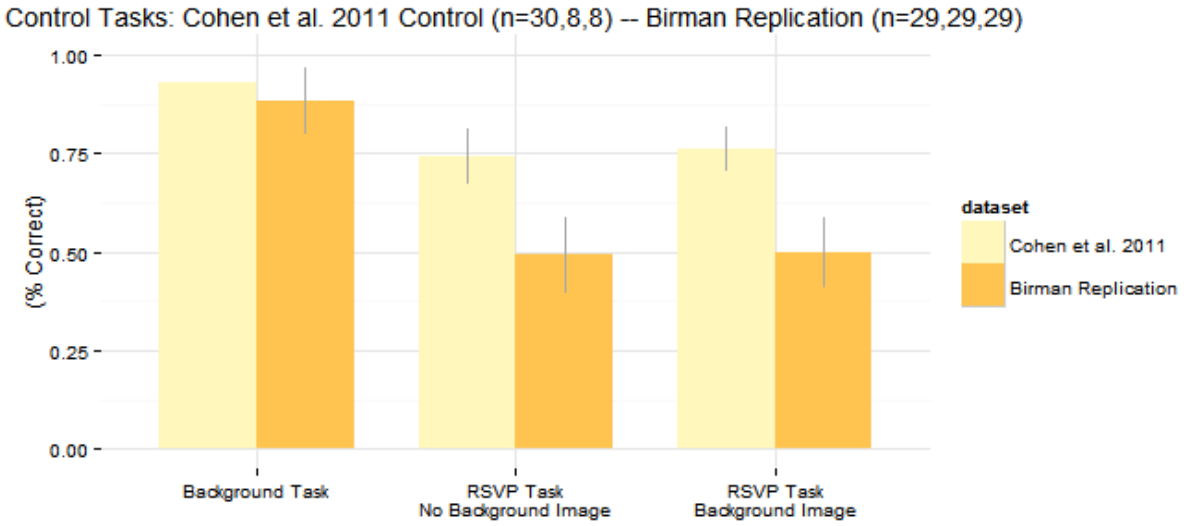

Fig 4: Control Tasks. Cohen et al. reported a very high task accuracy for identifying the category of the background images when attending to the background. Our results show a similar magnitude effect, participants identified the background images correctly 88.30% of the time (note the original data did not include statistics for this result, so we could not estimate their confidence interval). In contrast, participants in our replication showed diminished performance on the RSVP task relative to the Cohen et al. subjects. Task performance in our subject pool was ~25% worse relative to the Cohen et al. results.

All of the data and code necessary to reproduce these results is held in a git repository: http://github.com/dbirman/repo254

Discussion

Summary of Replication Attempt

Cohen et al. 2011 reported that visual awareness of scene gist, sometimes called “gist perception”, does not occur independently from top-down directed attention as claimed by a number of researchers (see e.g. Koch & Tsuchiya, 2007). Their results show that with sufficient attentional load gist perception suffers and a significant proportion of participants experience “inattentional blindness” for the gist of scenes. This replication effort shows that not only can this effect be produced in the constrained environment of the laboratory but it can be extended to online experiments in unconstrained viewing environments. Effect size magnitudes were similar in both environments despite differences in participant performance on the attentional load RSVP task. This replication supports the conclusion that the gist perception cannot be used as an example of awareness in the absence of attention.

Commentary

We extended the analyses of Cohen et al. 2011 by examining whether the choice of using category information vs. actual image identification influenced the analysis. Our results (Fig. 3) suggest that the results are similar in both cases, but also reveal additional insights into the timeline of gist perception. As expected, a percentage of participants experience inattentional blindness in both cases. But we found that in addition ~40% of participants attending to the RSVP task reported seeing "something" but choosing the incorrect image out of the array. These participants fall somewhere in the middle of the range of 'gist perception', between total blindness and complete awareness. It appears that the majority of participants know that something happened in the background but do not have full access, or for some reason failed to fully process, the visual input in the background.

What do these collective findings suggest for our understanding of attention and awareness? We think psychologists have been misled by research findings suggesting that certain types of visual input bypass top-down attention on their path into awareness. Results such as those reported by Cohen et al. 2011 and the successful replication presented here implicate attention in even the fastest processing pathways in the brain. But what role does attention play? Evidence from monkey physiology (see Knudsen, 2007 for a review) suggests that attention modifies different neural activity in vastly different ways. One possibility is that only certain kinds of attention, or certain modifications of neural activity, represent modulations of awareness. Understanding the mechanistic role of attention will require models that can link behavioral measures of awareness with neural measurements of attention. We hope that future analyses of this kind can be used in combination with the cheap and fast methodology outlined in this report to clarify the roles of "attention" and "awareness" in visual perception.

References

Cohen, Michael a, George a Alvarez, and Ken Nakayama. "Natural-Scene Perception Requires Attention." *Psychological Science* 22, no. 9 (September 2011): 1165–72. doi:10.1177/0956797611419168.

Knudsen, Eric I. "Fundamental Components of Attention." *Annual Review of Neuroscience* 30 (January 2007): 57–78. doi:10.1146/annurev.neuro.30.051606.094256.

Koch, Christof, and Naotsugu Tsuchiya. "Attention and Consciousness: Two Distinct Brain Processes." *Trends in Cognitive Sciences* 11, no. 1 (2007). doi:10.1016/j.tics.2006.10.012.

Lamme, Victor a.F. "Why Visual Attention and Awareness Are Different." *Trends in Cognitive Sciences* 7, no. 1 (January 2003): 12–18. doi:10.1016/S1364-6613(02)00013-X.

Li, Fei Fei, Rufin VanRullen, Christof Koch, and Pietro Perona. “Rapid Natural Scene Categorization in the near Absence of Attention.” *Proceedings of the National Academy of Sciences of the United States of America* 99, no. 14 (July 2002): 9596–9601. doi:10.1073/pnas.092277599.

Mack, Arien, and Jason Clarke. “Gist Perception Requires Attention.” *Visual Cognition* 20, no. 3 (March 2012): 300–327. doi:10.1080/13506285.2012.666578.

Mack, Arien, and Irvin Rock. *Inattentional Blindness.* The MIT Press, 1998.

Reddy, Leila, Lavanya Reddy, and Christof Koch. “Face Identification in the near-Absence of Focal Attention.” *Vision Research* 46, no. 15 (July 2006): 2336–43. doi:10.1016/j.visres.2006.01.020.

van Boxtel, Jeroen J a, Naotsugu Tsuchiya, and Christof Koch. “Consciousness and Attention: On Sufficiency and Necessity.” *Frontiers in Psychology* 1, no. December (January 2010): 217–217. doi:10.3389/fpsyg.2010.00217.