

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, 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. We replicated Experiment 1b from Cohen et al, in which a main RSVP task, with letters/digits appearing at fixation, was presented in repeated short trials. The main task required considerable focal selective attention to complete. Throughout the trial an aggressive mask image was continually updated in the background, replaced by a natural scene for 100 ms at the end of the critical fifth trial. 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 replication substantiates these previous findings. Our expectation was that our procedure, identical 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% participants, while 23% participants saw and classified the scene immediately. Under the null hypothesis that gist perception is always possible despite

allocating attention to an alternate task we expected a rate of inattentive blindness of 0% ($n=0$), while Cohen et al. reported a rate of 7%. We computed a test for equal proportions on 50% and 7% with $n=30$ and found a confidence interval of 23% and 100% for the true proportion of inattentive blindness in our sample with $p < .001$. This analysis had a power of 99%. For 80% power we needed a sample of $n = 12$, 90% $n = 17$, 95% $n = 21$. Given the strength of this effect and the small sample sizes necessary we saw no reason to use a sample size with a power less than 99%, or $n=30$.

Planned Sample

We plan to use a sample size of 30 people to achieve a power of 99%. We do not plan on doing pre-selection, other than ensuring that participants have a computer screen that is capable of displaying our experiment (see **Procedure**).

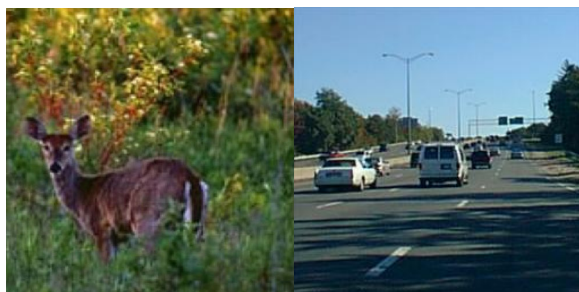
Materials

Experiment 1 Inattentive 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:

1. Did you notice anything different on that trial?
2. Did you notice something different about the background stream 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 natural scene in that stream?
5. If I were to tell you that there was a photograph in that stream, can you tell me what it was a photograph of?
6. 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?

Example Images: (Cohen et al., 2011)



Sample 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

“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 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, 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.” (Cohen et al., 2011).

“... after the critical trial, participants completed 20 trials in which they attended to the background while ignoring the [...] 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.” (Cohen et al., 2011)

In addition, participants were tested to determine whether their computer monitors were capable of displaying a short (<33ms) frame, to ensure that the critical scene was visible. Prior to being accepted into the experiment participants watched a

brief stimulus in which a single frame (33ms, 30hz) was changed to a highly visible letter, if they detected the single frame they were allowed to proceed with the experiment.

Analysis Plan

The data from this replication were analyzed directly to determine the percentage of participants who experienced total inattention blindness compared to participants who either saw the scene immediately or classified the scene after being asked questions. Unlike the original study in which the experimenter was present at all times we needed to ensure that participants were completing the RSVP task successfully. To this end we will collect and analyze data only for participants who succeeded at detecting one or more of the target digits on every RSVP task trial preceding the critical trial (with the exception of trials in which no target digit was present).

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 un-tested computer monitors where the distance to screen 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 inattention 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 inattention blindness.

(Post Data Collection) Methods Addendum

Actual Sample

sample size, demographics, data exclusions based on rules spelled out in analysis plan

Differences from pre-data collection methods plan

Any differences from what was described as the original plan, or “none”.

Results

Data preparation

Data preparation following the analysis plan.

Confirmatory analysis

The analyses as specified in the analysis plan

Exploratory analyses

Any follow-up analyses desired (not required).

Discussion

Summary of Replication Attempt

Open the discussion section with a paragraph summarizing the primary result from the confirmatory analysis and the assessment of whether it replicated, partially replicated, or failed to replicate the original result.

Commentary

Add open-ended commentary (if any) reflecting (a) insights from follow-up exploratory analysis, (b) assessment of the meaning of the replication (or not) - e.g., for a failure to replicate, are the differences between original and present study ones that definitely, plausibly, or are unlikely to have been moderators of the result, and (c) discussion of any objections or challenges raised by the current and original authors about the replication attempt. None of these need to be long.