Signal Processing: *Image Communication* 39 (2015) 319–327

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Scene masking is affected by trial blank-screen luminance

Tyler E. Freeman [a,1], Lester C. Loschky [a,*], Bruce C. Hansen [b]

[a] *Kansas State University, Manhattan, KS, USA*
[b] *Colgate University, Hamilton, NY, USA*

ARTICLE INFO

Available online 15 April 2015

Keywords:
Visual masking
Stimulus onset asynchrony (SOA)
Scene perception
Scene gist
Rapid scene categorization
Luminance

ABSTRACT

The current study investigated the role of the inter-screen luminance contrast (ISLC) of trial blank screens between target and mask screens in visually masking scenes. Participants performed a scene gist recognition task in which we varied mask strength, blank screen luminance, and stimulus onset asynchrony (SOA). Results showed that the more luminant white, and less luminant black, blank screens produced greater masking than intermediate luminance gray blank screens adjusted to the mean luminance of the target screens, specifically for black screens at SOAs <36 ms and for white screens at all SOAs. Our findings suggest that researchers interested in controlling for 'extraneous factors' should use gray blank screens as they eliminate any contribution of the ISLC component of masking. However, researchers interested in creating and examining differences in processing at early SOAs (<36 ms) should use black blank screens as these were shown to increase variation in the SOA function.

1. Introduction

Recently, there has been a surge in interest in studying real world scene perception [1] often using photographs of scenes as stimuli, and investigating issues such as scene gist perception, attention in scenes, and memory for scenes. These studies often involve the use of visual masks to manipulate the time course of processing. The study of visual masking has a long, rich, and deep history (for reviews, see [2-5]). However, much research on visual masking has used relatively simple stimuli (e.g., letters, disks, or sinusoidal gratings) and so the degree to which basic principles of visual masking derived from these studies scale up to the masking of photographic real-world scenes and meaningful tasks is not well understood (cf. [6-10]). Consequently, the use of visual masks in *scene perception* research often goes unexplained, with little if any rationale given for selecting various temporal or spatial masking parameters. The current exploratory study addresses a rather surprising gap in our knowledge of visual scene masking: Does it matter whether the blank screen shown before and after the target and, after the mask, creates *inter-screen luminance contrast* (ISLC[2]) with the target and mask, and if it does matter, what is the nature of this effect? The answers to these basic questions could potentially influence the interpretation of the results of scene perception studies using visual masks.

In addition, the answers to these questions have implications for real-world technology applications. For example, the recent production of ultra-high-definition displays (e.g., 4K televisions and monitors) rely on increased luminance and contrast levels. The present

* Corresponding author at: Department of Psychology, Kansas State University, 498 Bluemont Hall, Manhattan, KS 66506-5302, USA.
E-mail address: loschky@ksu.edu (L.C. Loschky).
[1] Tyler Freeman, ICF International, Wilmington, NC 28409, USA (present affiliation).

[2] The term inter-screen luminance contrast (ISLC) refers to the discrepancy in mean luminance between target/mask trial screens and the blank screens in a typical masking experiment trial.

http://dx.doi.org/10.1016/j.image.2015.04.004
0923-5965/

research could contribute to a better user experience for consumers of this technology by enriching understanding of how display luminance and contrast impacts users' perception of rapid changes presented on such displays (e.g., scene cuts and briefly flashed imagery). Similarly, emerging 3-D and virtual reality technologies capitalize on stereoscopic vision via rapid alternation of images to each eye. The present study could inform how variance in luminance and contrast of stereoscopic imagery impacts the 3-D and virtual reality user experience. Lastly, some night vision technology (e.g., image-intensifier applications that are often employed by the military) collects and amplifies ambient light to generate the images projected to the user's retina. As such, rapid changes in ambient light (e.g., [17]) may produce similar masking effects observed in scene perception experiments. The present study may therefore inform future development of night-vision technology software in terms of how the equipment processes rapid changes to ambient light to mitigate masking of the image projected to the user's eye.

In a typical scene masking paradigm, the experimental subject is presented with black, white, or neutral gray blank screens during the inter-stimulus intervals (ISI) that occur before, in between and, after stimulus presentations. In fact, whether the ISI blank screens are black, white, or neutral gray is generally not reported in scene perception studies, though it is often illustrated in figures showing the events in a trial. From such evidence, it appears that black blank screens are a common default option [6,7,11,12]. However, many researchers choose an alternative technique of presenting a neutral gray blank screen, set either to the middle gray level (e.g., 127 pixel value of an 8 bit image), or matched to the mean luminance value of the target and mask images [13,14]. Still other perception research has used white backgrounds (e.g., [15,16]). Because no explanation is usually given for choosing to use black, white, or neutral gray blank screens, it appears that the decision to use a neutral gray blank screen is made if the researcher is concerned about the effects of ISLC between the blank screens, target, and mask. However, to our knowledge, no study has been done to determine whether global luminance contrast between the blank screens, target, and mask actually affects scene image masking results, and if so, how.

A black blank screen, together with a higher mean luminance target and mask, may affect scene masking by creating ISLC at the on- and offset of the target, the onset of the mask (i.e., the offset of the black ISI screen), and the offset of the mask (i.e., the onset of a second black screen). The same could be true for white blank screens paired with a lower mean luminance target and mask. It therefore seems plausible that the ISLC produced at these points in time could potentially cause differences in low-level masking processes. Conversely, an intermediate (or mean) luminance gray blank screen, relative to black or white blank screens, would greatly reduce such ISLC. Therefore, any low-level masking differences caused by ISLC would be minimized by using mean luminance gray blank screens. Past research using simpler stimuli [17,18] suggests that the different levels of ISLC produced when using mean luminance gray versus contraluminant blank screens (relative to target and mask screens) could indeed produce important perceptual differences between the two methodologies.

In a classic study, Crawford [17] investigated the time course of dark adaptation, by briefly presenting a target either before or after a luminant disk conditioning-stimulus that served as a mask. Crawford's research revealed several findings relevant to the exploratory predictions of the current study. First, the more rapidly the conditioning field offset followed the test stimulus onset (and vice versa), the stronger the masking. This suggests that closer temporal proximity of luminance changes causes greater masking. Second, increased luminance contrast between the test and conditioning stimuli resulted in stronger masking. Third, Crawford compared the use of a black disk on a gray background versus a white disk on a gray background and found no differences between the two conditions. This suggests that the magnitude of the contrast, not its direction, is responsible for the masking effects. These results were replicated in a follow-up study using detection of objects in natural scenes as the test stimuli following exposure to the same conditioning stimuli as in the original study.

In sum, Crawford [17] attributed the observed masking effects to the time course of the conditioning stimulus offset, as well as the magnitude of the luminance contrast between the conditioning and test stimuli, but not its direction. Crawford [17] argued that, in the case of backward masking, the "relatively strong conditioning stimulus overtakes the weaker test stimulus on its way from the retina to brain and interferes with its transmission" (p. 285). This concept was echoed later by Breitmeyer and Kersey [19] who similarly found that the timing of the offset of the mask relative to the onset of the target could affect masking of the target. Like Crawford [17], they showed that shorter times between target onset and mask offset resulted in more effective masking.

A recent study by Tucker and Fitzpatrick [18] provided physiological evidence (from the primary visual cortex of the tree shrew) of visual masking via stimulus onset and offset. Using single cell recording techniques, the authors found that a sudden increase or decrease in mean luminance of a visual stimulus was accompanied by a decrease in cortical activity. Similar to the performance data from the human psychophysical results of Crawford [17], the authors found that the visual cortex cells were sensitive to the magnitude of a luminance change (i.e., contrast), but not its direction (either more or less luminant). Thus, larger luminance swings produced larger inhibitory responses of the cells, whereas mean luminance stimulus changes did not.

2. Exploratory hypotheses

2.1. Blank screen contrast

Our review of the literature provides a basis for predicting differences in masking between black or white blank screens versus neutral gray blank screens (whose mean luminance is the same as that of the target and mask stimuli). We predict that mean luminance gray blank screens should reduce masking effects by minimizing the ISLC between blank screens, target, and mask. Conversely, maximizing the ISLC between blank screens, target, and mask, by using black or white blank screens, should increase any observable masking effects as the magnitude of the luminance step is larger in these conditions.

2.2. Time course

In line with the findings of Crawford [17] and Breitmeyer and Kersey [19], we further predict that short ISIs, which produce swings in luminance in rapid succession, should cause the most interference and demonstrate stronger masking effects. Importantly, all of the above predictions are based on previous research that used simple stimuli such as luminant disks and gratings, (cf. [16]). Thus, because scenes are a much more complex, and information rich, visual stimulus to process, a key question is whether our predictions will scale up to meaningful real-world scenes used as targets.

For the purposes of the present exploratory study, we ask to what degree the ISLC produced by trial blank screens will interact with processing time. Experiment 1 was designed to test for any effects on masking produced by using black, white, or mean luminance gray blank trial screens with very short target and mask durations (12 ms each). In Experiment 1, we also examined the time course of the effects of trial blank screen luminance on masking efficiency by including a range of SOAs. Experiment 2 is the same as Experiment 1 in all respects except that the presentation durations of both the target and the mask were doubled to increase target accuracy and allow for a greater range of scene gist recognition performance.

3. Experiment 1

3.1. Method

3.1.1. Participants

58 Kansas State University undergraduate students participated for course credit (38 female, $M_{age}=20.9$, $SD=5.09$, age range: 16–46). All participants had normal or corrected near vision of at least 20/25, scored using a Sloan near acuity letter chart.

3.1.2. Stimuli

The study used 300 Gray scale photographs (1024×674) from the Corel Image Database and the internet. The images were from 10 basic level scene categories: Beach, Desert, Forest, Mountain, River, Farm, Home Interior, Market, Pool, and Street, with 30 images per category. Images were displayed on one of four Samsung 957MB CRT monitors (17-in diagonal: 36.51 cm (H) $\times$ 27.30 cm (V), 85 Hz refresh rate) calibrated using a Spyder2Pro colorimeter and calibration software, with a non-linear gamma correction (gamma$=$2.2), which takes account of the perceptual compression of brightness. Black (a gray level of 0) luminance approximated 0.42 cd/m^2, white (a gray level of 255) approximated 80.6 cd/m^2, and mean luminance gray (a gray value of 127) approximated 17.29 cd/m^2. Images were viewed under normal room illumination so that results could be interpreted in the context of typical viewing conditions. Images were viewed from a fixed distance of 53.3 cm using a chin rest, and subtended 37.79° (H) $\times$ 28.71° (V) of visual angle.

All images, including both targets and masks, were equalized in terms of their mean luminance (127 RGB pixel luminance) and their RMS contrast ($=$0.23). Fully phase randomized scenes were used as masks. Examples of a target scene and its associated phase-randomized mask can be seen in Fig. 1 (for details on mask generation procedures, see Appendices of [7]).

Phase-randomized scenes possess the same amplitude spectra of the scenes from which they are based, which makes them more effective at masking scenes than white noise [7,10]. However, they differ from isotropic $1/f$ noise in that they each have the particular spatial frequencies and orientations unique to a given scene image, which arguably makes them better masks for scene perception research (but see [9]). Target and mask images were randomly paired with the constraints that they (1) could not be from the same basic level category so that the phase randomized mask could never match the post cue, (2) each

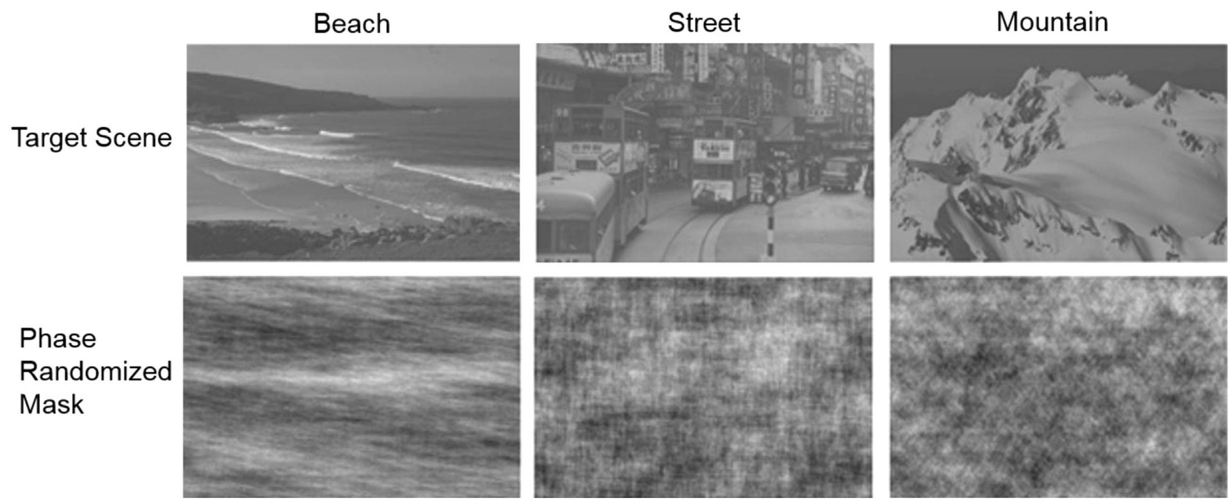

Fig. 1. Examples of target scene stimuli and their associated phase randomized masks.

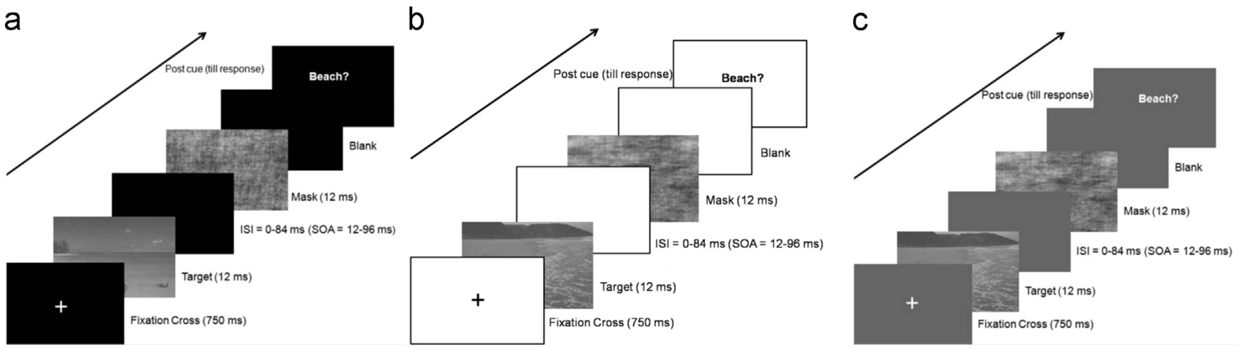

Fig. 2. Trial schematics for Experiment 1 with (A) black (i.e., high contrast) blank screens, versus (B) white (i.e., high contrast) blank screens and (C) mean luminance gray (i.e. low contrast) blank screens. Experiment 2 was the same except that the target and mask durations were 24 ms, thus the minimum SOA was 24 ms, and there was an additional SOA of 84 ms.

category of target was equally often paired with each category of mask, and (3) each target and mask image was used once per participant. The category of the mask never matched the category label shown at the end of the trial (see procedures below) to allow us to ensure that participants were responding to the target category and not to the mask category.

3.1.3. Design and procedure

Fig. 2 illustrates schematics of single trials for the three different ISLC conditions (black, white or gray). On each trial, participants first saw a fixation cross until they pressed a key to start the trial. A period of 750 ms later, the target image flashed for 12 ms, followed by a variable ISI, and a mask for approximately 12 ms (one refresh cycle of the monitor at 85 Hz). Transitions between stimulus screens were immediate with no temporal smoothing operation applied between images (i.e., followed a temporal square-wave function), as is standard practice in studies of the temporal dynamics of visual masking (e.g., [2,3]). Following a 750 ms blank interval, a category label was presented until the participant responded, either by pressing a 'yes' key on a keyboard if the target image matched the label, or a 'no' key if it did not. The target duration of 12 ms is roughly the minimum necessary for above-chance gist recognition performance, and a 1:1 target:mask duration ratio generally produces moderate masking with these stimuli [7,9,10]. A gray, black or white blank ISI was presented between the target and mask images for 0–84 ms, creating SOAs (= target duration + ISI) between target and mask of 12, 24, 36, 47, or 94 ms (varied randomly within subjects, in approximately 12 ms increments dictated by the 85 Hz refresh rate of the monitors[3]). These SOAs were chosen to produce the widest possible range of performance [6,7,9,10]. Of critical importance, the blank screen before the target, during the ISI, and after the mask was either black, white, or a neutral gray set to match the mean luminance of the target and mask images as were the backgrounds of the fixation screen, surrounding the target (as described above in Section 3.1.2), and the post-cue screen. Participants were randomly assigned to one of three conditions: either black, white, or neutral gray blank screens.

Prior to the experiment, participants completed a category-familiarization task with nine images from each scene category, in order to acquaint them with the scene category labels. Importantly, however, none of the category familiarization scenes were used in the experiment. Because the participants had likely never previously seen the exact images shown in the experiment, their responses would have to be based on their general ability to recognize scenes as members of various categories, rather than their memory for specific features of specific images, thus making the results of the study more generalizable to viewers' ability to rapidly categorize a much larger number of scenes they might encounter in the real world. Participants then carried out 32 practice trials, to become familiar with the experimental task. Likewise, the practice trial images were not used in the main experiment. Trials were self-paced, and participants could take breaks at any time, with the 300 experimental trials taking approximately 15 minutes to finish. Participants wore ear muffs to reduce environmental noise.

3.2. Results

The results for Experiment 1 are shown in Fig. 3. In order to assess the effects of SOA, and blank screen type (black or white/high contrast or mean luminance gray/low contrast), we carried out a 5 (SOA: within-subjects) $\times$ 3 (blank screen: between-subjects) mixed ANOVA. As can be seen in Fig. 3, as expected there was an overall main effect of SOA such that accuracy increased with SOA, $F(4, 220) = 7.39$, $p < .001$, partial $\eta^2 = .118$.

Of central importance to the present study is the fact that the luminance of the blank screens significantly affected masking, $F(2, 55) = 129.59$, $p < .001$, partial $\eta^2 = .825$. Interestingly, as can be seen in Fig. 3, the only blank screen condition in which there was an effect of SOA (i.e., the amount of time that scene information was available on the retina) was the black screen condition. This observation is supported by a significant SOA $\times$ blank screen interaction, $F(8, 220) = 5.53$, $p < .001$, partial $\eta^2 = .168$, and flat SOA slopes for the gray and white screen conditions.

[3] Due to rounding, the fourth SOA in Experiment 1 was 47 ms. The comparable SOA in Experiment 2 was 48 ms. There was a similar rounding issue in with the 94 ms SOA in Experiment 1 and the 95 ms SOA in Experiment 2. Thus, for the purpose of the analyses, the 47/48 ms and 94/95 ms time points were respectively combined into 48 ms and 95 ms levels of the SOA variable.

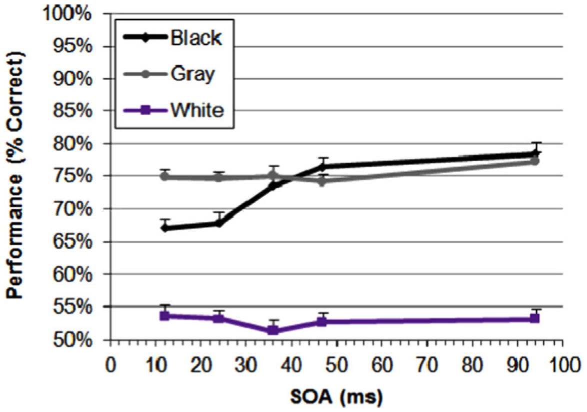

Fig. 3. Experiment 1 results. Scene gist recognition accuracy as a function of stimulus onset asynchrony (SOA) in milliseconds (ms), and inter-screen luminance contrast (ISLC: Black, White or Gray). Target and mask duration = 12 ms. Error bars represent SEM.

t-Tests were performed to examine the effect of blank screens (black, white, or mean luminance gray) at each SOA on scene gist recognition performance. The luminance of the blank screens affected gist at early levels of processing, specifically SOAs < 36. When paired with a 12 ms SOA, black blank screens caused greater masking than gray blank screens, $t(37)=3.96$, $p<.001$, and likewise at 24 ms SOAs, $t(37)=3.54$, $p=.001$. Thereafter (SOAs of 36, 47 and 94 ms) performance was equal in both the black and gray blank screen conditions, all $ps>.05$. As shown in Fig. 3, at early SOAs (SOA < 36 ms), the black screens (i.e., high ISLC) produced stronger masking effects than the mean luminance gray screens (i.e., low ISLC). At 12 ms SOA, the difference in performance between black versus neutral gray blank screen conditions is 8% (.67 vs. .75). Given that our range of potential performance is from 0.5 to 1.0, this effect is relatively large (16% of the performance range). However, at later SOAs (SOA ≥ 36 ms), the ISLC of the black blank screens, relative to the neutral gray blank screens, no longer affected scene gist masking.

The white screen condition produced the worst scene gist recognition performance (i.e., strongest scene masking). We used independent samples *t*-tests to compare the means of the white screen condition with the black and gray blank screen conditions at all SOAs. The comparisons showed highly significant differences at all SOAs, all $ps<.001$. However, these differences may be due a floor effect produced by pairing 12 ms target and mask durations with high-luminance white blank screens. Related to this, the performance function for the white ISLC condition produced near chance performance, which did not significantly differ across SOAs, $F(4, 90)=.363$, $p=.834$, namely, a floor effect. This further supports the interpretation that high-luminance white blank screens produced strong masking effects.

4. Experiment 2

To avoid the floor effect seen when pairing white blank screens with 12 ms target and mask durations, in Experiment 2 we doubled the target and mask durations to 24 ms to improve performance in the white screen condition. This allowed us to more accurately assess the relative effects of using black, white, and gray blank screens, over time, on the masking of natural scenes. Pilot testing showed that this longer target and mask duration improved performance in the white screen condition, allowing for a larger range of scene gist recognition performance.

4.1. Method

4.1.1. Participants

51 Kansas State University undergraduate students participated for course credit (30 female, $M_{age}=19$, $SD=1.04$, age range: 18–22). All participants had normal or corrected near vision of at least 20/30, scored using a Sloan near acuity letter chart.

4.1.2. Stimuli

The stimuli were the same as used in Experiment 1, as were the target-mask pairing constraints and monitors. The only difference was that the target and mask durations were doubled to 24 ms resulting in SOAs of 24, 36, 48, 83, and 95 ms (the same as Experiment 1, except that the minimum SOA changed from 12 ms to 24 ms, and there was an additional SOA of 83 ms).

4.1.3. Design and procedures

The design and procedures were the same as Experiment 1 except for the above-mentioned changes.

4.2. Results

The results for Experiment 2 are presented Fig. 4. In order to show the effects of the 24 ms target and mask durations relative to the 12 ms durations used in Experiment 1, this figure combines the data from both experiments. To assess whether doubling our target/mask durations (from 12 ms to 24 ms) had an effect on the data, we conducted a 4 (SOA: within-subjects) × 3 (blank screen ISLC: between-subjects) × 2 (Experiment 1 vs. 2: 12 vs. 24 ms target/mask durations: between subjects) mixed

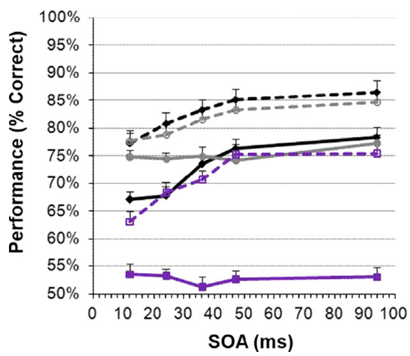

Fig. 4. Experiment 2 results combined with the results from Experiment 1 showing the scene gist recognition accuracy as a function of stimulus onset asynchrony (SOA) in milliseconds (ms), and inter-screen luminance contrast (ISLC: Black, White, or Gray), and target and mask duration (12 or 24 ms). Error bars represent SEM.

Table 1
Three-way mixed ANOVA results for SOA (within-subjects) × blank screen ISLC (between-subjects) × Experiment (target/mask durations; between subjects).

	df	*F*	*p*	*η*2
Main effects				
ISLC	2, 102	108.16	< .001	0.680
SOA	3, 102	33.21	< .001	0.246
Exp (target/mask durations)	1, 102	84.96	< .001	0.454
2-Way interaction effects				
ISLC × SOA	1, 102	5.59	.005	0.099
ISLC × Exp (target/mask durations)	2, 102	9.89	< .001	0.162
SOA × Exp (target/mask durations)	1, 102	14.43	< .001	0.124
3-Way interaction effect				
ISLC × SOA × Exp (target/mask durations)	2, 102	7.60	.001	0.130

ANOVA, the results of which are reported in Table 1. The SOA factor was limited to those SOAs shared between the two experiments (24, 36, 48, 95)[2]. As shown in Fig. 4 and Table 1, there was a main effect of Experiment (1 vs. 2), with performance being significantly better with the 24 ms target/mask durations than the 12 ms durations. We also found significant main effects of SOA, and, most importantly, blank screen ISLC. However, the main effect of blank screen ISLC was qualified by a significant 2-way interaction between ISLC × Experiment (1 vs. 2) such that the main effect of ISLC was greater with the shorter, 12 ms target and mask durations in Experiment 1, $F(2, 55)=129.09$, $p<.001$, $\eta^2=.824$, than with the longer, 24 ms target and mask durations in Experiment 2, $F(2, 47)=19.812$, $p<.001$, $\eta^2=.457$. This can be seen in Fig. 4, in which the large effect of ISLC in Experiment 1 is due to the floor effect for the white screen ISLC with the 12 ms target/mask durations. The effect of ISLC was smaller in Experiment 2 with the 24 ms target/mask durations because there was no longer a floor effect in the white screen ISLC, though it still produced considerably worse performance than the gray and black screen ISLCs at all SOAs (all *ps* < .001). We also found a significant 2-way interaction between SOA × Experiment (1 vs. 2), such that the effect of SOA was greater in Experiment 2 with the longer 24 ms target/mask durations than in Experiment 1 with the shorter 12 ms target/mask durations. As shown in Fig. 4, the 12 ms target/mask durations produced flat SOA slopes for both the gray and white screen ISLCs, but the longer 24 ms target/mask durations produced positive slopes for all three blank screen ISLCs. Finally, we also found a significant 2-way interaction between ISLC × SOA. Importantly, however, that 2-way interaction was qualified by a significant three-way (ISLC × SOA × Experiment (target/mask duration)) interaction.

The nature of this 3-way interaction was further explored by a 2-way ANOVA for the data from Experiment 2, specifically a 5 (SOA: within-subjects) × 3 (blank screen ISLC: between-subjects) mixed ANOVA. This showed significant main effects of both SOA, $F(4, 188)=27.78$, $p<.001$, $\eta^2=.371$, and blank screen ISLC, $F(2, 47)=18.53$, $p<.001$, $\eta^2=.994$. More to the point, it explained the above-mentioned 3-way interaction. Specifically, the 2-way interaction between blank screen ISLC × SOA was *not* significant, $F(8, 188)=1.05$, $p=.396$, $\eta^2=.043$, which is in contrast to the significant 2-way interaction between these factors found in Experiment 1. Thus, Fig. 4 suggests that increasing the target/mask durations in Experiment 2 eliminated the differences in masking observed in Experiment 1 between the black and gray ISLC conditions at 12 and 24 ms SOA. This observation was supported by *t*-tests of scene recognition performance differences across blank screen ISLC conditions (black, white, or mean luminance gray) at each SOA in Experiment 2. The results showed no significant differences in masking by the black and gray blank screens at any SOA (all *ps* > .05), which is in contrast to the clear differences between black and gray blank screens at SOAs < 36 ms in Experiment 1, and between both the black and gray blank screens and the white screens at all SOAs.

5. Discussion

The current study addressed a surprising gap in our knowledge about the visual masking of scenes by answering the following question: Does it matter whether the blank screens in a typical masking paradigm create luminance contrast with the target and mask, and if so, what effect(s) does it have? Based on previous research [17–19] that used simpler stimuli (e.g., disks and gratings), we hypothesized that black and white blank screens that create ISLC with the target and mask would create stronger masking than mean luminance gray blank screens that minimize ISLC. We also hypothesized that such ISLC effects would be greatest at relatively short ISIs, due to the close temporal proximity of the targets to the luminance contrasts [17,19]. Both hypotheses were supported by the results of Experiments 1 and 2. In Experiment 1, which used 12 ms target/mask durations, we found stronger masking for black than mean luminance gray blank screens at SOAs < 36 ms, and very strong masking by white blank screens at all SOAs. In Experiment 2, which strengthened the target perception by increasing target/mask durations to 24 ms, we found an effect of ISLC only when white blank screens were used. In sum, the effect of black screens is only present with 12 ms target durations at SOAs < 36 ms. When the target is made stronger by doubling its duration, the effect of black blank screens disappears. These findings suggest that the effect of the black blank screens (relative to the effect of white blank screens) is weak whereas the effect of white blank screens on scene masking is strong and temporally robust.

Furthermore, the flat performance slopes of both the gray and white blank screens in Experiment 1 warrant additional discussion. There are at least two possible explanations for the lack of an effect for SOA in the gray screen condition with 12 ms target/mask durations: (1) phase-randomized scenes simply make weak masks, or (2) a 1:1 target to mask duration ratio using 12 ms durations produces weak masking. Given that Loschky et al. [7,10] have shown that phase randomized masks make *stronger* masks than white noise masks (a common choice of visual mask), we are led to the conclusion that the weak masking shown in the gray blank screen condition in Experiment 1 is due to our specific target and mask duration parameters. In fact, when the target and mask durations were increased in Experiment 2, the SOA × performance slope steepens significantly for the gray, $t(36)=2.306$, $p=.027$, and the white $t(34)=5.215$, $p<.001$, ISLC conditions relative to their slopes in Experiment 1 that used 12 ms target/mask durations. Although there was an increase in scene gist recognition performance in the black ISLC condition in Experiment 2, the slope of the accuracy × SOA function for the black ISLC was similar in both experiments, $t(32)=1.235$, $p>.05$.

6. Implications for scene perception research

It is important to note that while we only found effects of the black blank screen ISLC at short SOAs (< 36 ms), several recent studies of scene gist recognition [7,9,10,13,14,20] have shown that important processing goes on at these early processing times. For example, differences in processing between the superordinate and basic levels of scene categorization are clearly shown at these early SOAs [13,20]. In addition, the effects of the black blank screen ISLC with 12 ms target/mask durations in Experiment 1 were fairly substantial, accounting for 16% of our possible performance range. These early differences in masking due to the ISLC of the black blank screens produced globally different accuracy-by-SOA masking functions, with the black blank screen ISLC producing a steeper slope and the gray ISLC producing a flat slope when target and masks had 12 ms durations, as shown in Figs. 3 and 4. Even more impressively, the effect of the white screen ISLC was enough to create a floor effect at these early SOAs as well as later ones extending out to roughly 100 ms. Furthermore, this stronger masking by white ISLC was also found even after the floor effect was eliminated by increasing the target (and mask) duration. Clearly, such differences in masking by the luminance of the ISLC of the blank screen could affect one's interpretation of whether a variable of interest has an effect on scene gist recognition or not. Therefore, if one is interested in probing scene perception, particularly at early levels of processing (< 36 ms SOA), using 12 ms target and mask durations, the nature of the blank screen ISLC with the stimuli is very important to consider.

Two approaches to making use of the current study's findings seem apparent. Many scene perception researchers may want to eliminate any "extraneous" factors (such as ISLC) that could influence their results. From this perspective, gray blank screens may be optimal since they eliminate any contribution of the ISLC component to masking.

However, other researchers may utilize these findings based on a more pragmatic perspective. Specifically, if what one is interested in is creating differences in processing time, and seeing how those differ between different conditions [e.g., [20], which compared the Natural/Manmade task vs. Basic level task], then having an SOA function that shows a greater range of variation is better. Specifically, our black blank screen condition showed greater variation in the accuracy-by-SOA function at 12 and 24 ms SOAs that was not present in the gray or white screen 12 ms target/mask conditions. The white screens paired with 24 ms target/mask durations showed the most variation in the performance × SOA function though performance was globally lower in this condition. Thus, the use of black or white blank screens could uncover variation at early SOAs that may elude researchers using gray blank screens, though the use of white causes decrements in performance relative to the black.

However, our findings also suggest that if researchers are using target/mask durations of 24 ms or more, whether they choose to use black or gray blank screens should not affect their results. Our findings further indicate that whether the use of white blank screens is preferable or not depends on whether researchers are looking for good variation in the performance × SOA function or overall better performance. White screens may be preferable in the former case but not in the latter. Such information makes an important methodological contribution to the scene perception and image processing and communication literatures. We suggest future scene perception research using visual masks always report the luminance (in candelas) of the trial blank screens.

In Experiments 1 and 2, we showed using white blank screens produced much stronger masking than using black or gray blank screens. While we predicted white blank screens would produce greater masking than the gray screens, the large difference in scene recognition performance between the black and white screen conditions was unexpected. Although the goal of the current study was to address a basic methodological question of whether trial blank screen luminance affected scene masking (it clearly does), we will offer a few possible explanations of our data in order to guide future research aimed at more fully addressing the underlying masking mechanisms that may be active here. To do so, the remainder of this discussion is organized into five sections. In the first three sections we offer possible explanations for the current results. These sections are followed by suggestions for future research and our concluding remarks.

6.1. Monitor attributes

First it is necessary to rule out a possible explanation of the increased masking by the white ISLC over the black ISLC based on the phosphor persistence of the monitors. If the monitor's phosphor luminance rise time is faster than its decay time, then the transition from gray to white would be faster than from gray to black, producing greater masking for the white screen condition. However, the P22 phosphor used in our monitors has a decay time ranging from 10 μs to 1 ms. Thus, such an explanation can be ruled out.

6.2. Stimulus intensity

Another possible explanation for the increased masking by white ISLC in our data is provided by Piéron's [21,22] psychophysical research on variation of processing times as a function of stimulus intensity. Specifically, Piéron [21,22] showed that reaction times decreased as a function of light intensity. In other words, people were faster to respond to bright flashes of light relative to dimmer ones. This has been taken as evidence that higher luminance stimuli are processed faster relative to lower luminance stimuli (see [23], p. 417). Thus, in the context of the current study, speeded perception of white blank screens, relative to black and gray blank screens, could explain the increased masking observed in that condition. This interpretation is supported by the work of Chichilnisky and Kalmar [24] on retinal processing asymmetries between ON and OFF channels.

6.3. Retinal processing asymmetries between ON and OFF channels

Another possible account of the stronger masking with white ISLC may be in terms of the retinal processing asymmetries between the ON and OFF channels, which respond to light and dark stimuli respectively, before they are believed to interact at the cortical level. Chichilnisky and Kalmar [24] demonstrated temporal asynchronies between the two pathways that might help explain the current results. Specifically, they found that the ON pathway reached peak response magnitude about 10–20% sooner than the OFF channel. Several other psychophysical studies have pointed to this difference in time-to-peak response to explain various motion illusions. For example, Del Viva, Gori, and Burr [25] cited this difference to account for a motion illusion observed with glass patterns made up of white and black dots; however, the estimated lag they reported was quite small, only about three milliseconds. Thus, the fact that the white blank screens produce additional masking for SOA's up to 96 ms (and significantly stronger than black blank screens up to that duration) makes this temporal processing notion seem unlikely.

Chichilnisky and Kalmar [24] point to another, perhaps more attractive asymmetry that may possibly account for the current results, namely that the ON cell receptive fields were 20% larger than OFF cell receptive fields. Thus, the ON cells' receptive fields *may* have higher "full-field" sensitivity. In addition, Chichilnisky and Kalmar [24] also report that, at threshold, the ON cells were capable of signaling increments as well as decrements, whereas the OFF cells were only capable of signaling decrements. Zaghloul, Boahen, and Demb [26] have demonstrated this latter point in the Guinea pig retina (i.e., ON cells signaling both increments and decrements) however, they were only testing near threshold stimulation in vitro. Lastly, this group also demonstrated a further asymmetry, in that the inhibition between the two pathways (prior to cortex) is unidirectional. Specifically, the ON pathway seems to have more inhibitory control (at threshold) over the OFF pathway than vice versa. Since the data reported here are from suprathreshold stimuli (including the ISI blanks), one is left with an explanation for the stronger masking by white blank screens as possibly arising from an ON pathway that has higher full-field sensitivity (i.e., larger receptive fields) and reaches a critical level faster than the OFF pathway which could, on theoretical grounds, lead to stronger masking effects with white blank screens.

6.4. Future directions

In the current study, the blank screens appeared in the trial sequence before and after the target and mask (as well as at fixation and post-cue screens). Because of this, we cannot be sure if the current results are due to forward masking, backward masking, or a combination of the two and therefore, further research may wish to examine the independent contributions of forward or backward masking to the effects of ISLC we have shown here. Additionally, another direction for future research may be to assess whether or not the current findings hold with other stimuli and tasks. The current study used stimuli that were limited to real-world scenes in a category matching task and thus, the current results may not generalize to studies where simpler stimuli (e.g. disks or sinusoidal gratings) and tasks are used. Further experimentation will be necessary to answer this question.

7. Concluding remarks

While the current study is not sufficient to address the full implications of blank screen ISLC for broader masking theories (e.g., 3, [27]), it is nevertheless the first study to address an important methodological question: In scene perception studies using backward masking, does it matter whether the blank screens are black, white or mean luminance gray? The current results suggest that it does matter, and that the effects differ for black, white and gray screens as a function of both target/mask duration and SOA. When probing very early scene processing by using 12 ms target and mask durations and SOAs of 12–24 ms, there are important differences between using black, white or gray blank screens. However, with 24 ms target and mask durations only white, not black blank screens, differ from gray screens in their ability to affect masking. Because many early scene perception processes occur during the early processing times where blank screen ISLC makes a difference, the current study makes an important methodological contribution to the area of scene perception research.

Acknowledgments

This collaborative effort was conducted at Kansas State University and supported by funds from the Kansas State

University Office of Research and Sponsored Programs (USRG 1983), and by the NASA Kansas Space Grant Consortium (Activities at Kansas State University, 2007-8: Scene Classification by Humans and Artificial Vision Systems). The authors wish to acknowledge Zach Maier and Kwang Park for their contributions to initial discussions and development of this study, and the work of Katie Gibb, Jeff Burns, John Caton, Stephen Dukich, Ryan Eshelman, Nicholas Forristal, Kaci Haskett, Hannah Hess, Rebecca Millar, Ryan Ringer and Merideth Smythe who helped carry out the experiments. We also wish to thank Bruno Breitmeyer for helpful discussions of the study.

References

[1] Real-World Scene Perception, in: J.M. Henderson (Ed.), Psychology Press Ltd, New York, NY, 2005.
[2] U. Ansorg, G. Francis, M. Herzog, H. Ogmen, Visual masking and the dynamics of human perception, cognition, and consciousness: a century of progress, a contemporary synthesis, and future directions, Adv. Cogn. Psychol. 3 (2007) 1–8.
[3] B.G. Breitmeyer, H. Ogmen, Visual Masking: Time Slices through Conscious and Unconscious Vision, Clarendon Press, Oxford, 2006.
[4] J.T. Enns, V. Di Lollo, What's new in visual masking? Trends Cogn. Sci. 4 (2000) 9.
[5] N.V.S. Graham, Visual Pattern Analyzers, Oxford University Press, New York, 1989.
[6] N. Bacon-Mace, M.J. Mace, M. Fabre-Thorpe, S.J. Thorpe, The time course of visual processing: backward masking and natural scene categorisation, Vis. Res. 45 (2005) 1459–1469.
[7] L.C. Loschky, A. Sethi, D.J. Simons, T. Pydimari, D. Ochs, J. Corbeille, The importance of information localization in scene gist recognition, J. Exp. Psychol. Hum. Percep. Perform. 33 (6) (2007) 1431–1450.
[8] J.W. Rieger, C. Braun, H.H. Bulthoff, K.R. Gegenfurtner, The dynamics of visual pattern masking in natural scene processing: a magnetoencephalography study, J. Vis. 5 (3) (2005) 275–286.
[9] B.C. Hansen, L.C. Loschky, The contribution of amplitude and phase spectra defined scene statistics to the masking of rapid scene categorization, J. Vis. 13 (13) (2013) 1–21, http://dx.doi.org/10.1167/13.13.21.
[10] L.C. Loschky, B.C. Hansen, A. Sethi, T. Pydimari, The role of higher-order image statistics in masking scene gist recognition, Attent. Percep. Psychophys. 72 (2) (2010) 427–444.
[11] F. Li, R. VanRullen, C. Koch, P. Perona, Rapid natural scene categorization in the near absence of attention, Proc. Natl. Acad. Sci. U.S.A. 99 (14) (2002) 9596–9601, http://dx.doi.org/10.1073/pnas.092277599.
[12] R. VanRullen, L. Reddy, F. Li, Binding is a local problem for natural objects and scenes, Vis. Res. 45 (25-26) (2005) 3133.
[13] F. Li, A. Iyer, C. Koch, P. Perona, What do we perceive in a glance of a real-world scene? J. Vis. 7 (1:10) (2007) 1–29, http://dx.doi.org/10.1167/7.1.10.
[14] Oliva, A. (2006). From zero to gist in 200 msec: the time course of scene recognition. In: Paper presented at the Scene Understanding Symposium (SUNS) 2006, MIT Department of Brain and Cognitive Sciences.
[15] I. Biederman, Recognition-by-components: a theory of human image understanding, Psychol. Rev. 94 (2) (1987) 115–147.
[16] N.P. Costen, D.M. Parker, I. Craw, Effects of high-pass and low-pass spatial filtering on face identification, Percep. Psychophys. 58 (4) (1996) 602–612.
[17] B.H. Crawford, Visual adaptation in relation to brief conditioning stimuli, Proc. R. Soc. London Ser. B, Biol. Sci. 134 (875) (1947) 283–302.
[18] T.R. Tucker, D. Fitzpatrick, Luminance-evoked inhibition in primary visual cortex: a transient veto of simultaneous and ongoing response, J. Neurosci. 26 (52) (2006) 13537–13547.
[19] B. Breitmeyer, M. Kersey, Backward-masking by pattern stimulus offset, J. Exp. Psychol. Hum. Percep. Perform 7 (5) (1981) 972–977.
[20] L.C. Loschky, A.M. Larson, The natural/man-made distinction is made prior to basic-level distinctions in scene gist processing, Vis. Cogn. 18 (4) (2010) 513–536.
[21] H. Piéron, Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices, L'Annee Psychologique 20 (1914) 17–96.
[22] H. Piéron, Nouvelles recherches sur l'analyse du temps de latence sensorielle en function des intensités excitatrices et sur la loi qui relie ce temps a l'intensite de l'excitation, L'Annee Psychologique 22 (1920) 58–142.
[23] J. Hedgé, Time course of visual processing: coarse-to-fine processing and beyond, Prog. Neurobiol. 84 (2008) 405–439.
[24] E.J. Chichilnisky, R.S. Kalmar, Functional asymmetries in ON and OFF ganglion cells of primate retina, J. Neurosci. 22 (7) (2002) 2737–2747.
[25] M.M. Del Viva, M. Gori, D.C. Burr, Powerful motion illusion caused by temporal asymmetries in ON and OFF visual pathways, J. Neurophysiol. 95 (6) (2006) 3928–3932.
[26] K.A. Zaghloul, K. Boahen, J.B. Demb, Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities, J. Neurosci. 23 (2003) 2645–2654.
[27] G. Francis, M.H. Herzog, Testing quantitative models of backward masking, Psychonom. Bull. Rev. 11 (1) (2004) 104–112.