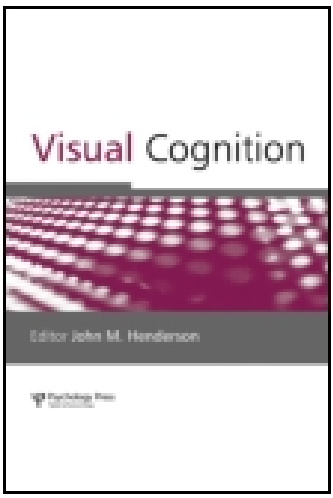

Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pvis20

Gist perception requires attention

Arien Mack [a] & Jason Clarke [a]
[a] New School for Social Research , New York , NY , USA
Published online: 30 Mar 2012.

To cite this article: Arien Mack & Jason Clarke (2012) Gist perception requires attention, Visual Cognition, 20:3, 300-327, DOI: 10.1080/13506285.2012.666578

To link to this article: http://dx.doi.org/10.1080/13506285.2012.666578

VISUAL COGNITION, 2012, 20 (3), 300–327

Psychology Press
Taylor & Francis Group

Gist perception requires attention

Arien Mack and Jason Clarke

New School for Social Research, New York, NY, USA

Four experiments (240 subjects) explored gist perception without attention using the Mack and Rock (1998) cross task. Twelve scenes were flashed under conditions of inattention, divided, and full attention. Subjects described what they saw on critical trials in which a scene was flashed with the cross. In Experiments 3 and 4 subjects also chose the scene from a four scene array. In Experiment 4 the critical scenes were shown twice in the inattention condition. Overall, only 17% reported gist in the inattention condition, 65% did so with divided, and 82% did so with full attention. In Experiment 4 most subjects remained inattentionally blind to the scenes even though they were shown twice, conditions which fostered repetition priming, and we found a suggestion of negative priming. The results of all 4 experiments indicate that gist requires attention.

Keywords: Attention; Inattentional Blindness; Perceived Gist.

Despite the overwhelming evidence of change blindness, the attentional blink, and inattentional blindness, as well as visual neglect in patients with parietal lesions, all of which attest to the centrality of attention for conscious perception, the question of whether anything is consciously perceived without attention remains alive and well. In fact, in the recent past there has been a spate of papers arguing for conscious perception in the absence or *near* absence of attention claiming that attention is not necessary for conscious perception (Koch & Tsuchiya, 2007; Tsuchiya & Koch, 2009; van Boxtel, Tsuchiya, & Koch, 2010; Lamme, 2003). Rather than engage this proposition in its most general form, the experiments presented here focus on one specific claim that is taken as the main support for the more general assertion that attention is not necessary for visual awareness.

This paper directly concerns the question of whether gist is perceived without attention. It is a question that has particular meaning for us because

Please address all correspondence to Arien Mack, New School for Social Research, 80 Fifth Avenue, New York, NY 1011, USA. E-mail: mackarie@newschool.edu

http://www.psypress.com/viscog http://dx.doi.org/10.1080/13506285.2012.666578

it recently has been argued that gist is, in fact, consciously perceived without attention (Koch & Tsuchiya, 2007; Lamme, 2003; Tsuchiya & Koch, 2009), and part of the evidence on which this claim is based is research reported earlier by one of us (Mack & Rock, 1998).

At the point at which the research we are now reporting was almost completed and reported at the VSS, 2011 meeting (Clarke, Mack, Slesar, & Erol, 2011), a paper appeared addressing the question of whether natural scene perception requires attention, (Cohen, Alvarez, & Nakayama, 2011). This paper addresses a question very similar to the one we address here. Using a different procedure, these investigators found no evidence of scene perception when attention is sufficiently engaged. Our results confirm and expand these findings.

In our earlier research (Mack & Rock, 1998, chap. 7), a large photograph either of an outdoor scene (in which there were some people and a dog) or an indoor scene (in which there were people sitting around a table eating) was unexpectedly presented to observers during a series of trials in which they had to report whether the horizontal or vertical arm of a briefly flashed cross, centred at fixation, was longer. On the fourth trial, the cross was superimposed on a picture of a scene. All arrays were displayed briefly and were masked. On trials prior to the critical one, the cross was superimposed on a randomly coloured mosaic block pattern that occupied the same space that subsequently was covered by the photograph. We reported that in most instances subjects were able to report the gist of the scene, but went on to say that, "We believe that it is the size of the scene that attracts attention and that once this occurs, attention may be distributed over the entire array or perhaps those parts of the array that carry most meaning are given priority processing and thus are more likely to be consciously perceived" (p. 169). We did not claim that gist could be perceived without attention although others have taken our results to mean that (e.g., Cohen et al., 2011; Lamme, 2003).

We currently believe that it may not have been the size of the scene that was crucial, but rather the fact that the cross to which subjects were attending was superimposed on the scenes so that the scene was located in the area to which attention was directed.

Our past research, however, is not the only evidence on which the claim that conscious perception can occur in the absence of attention is based. Evidence that scene categorization (e.g., animal vs. nonanimal; vehicle vs. nonvehicle; indoor vs. outdoor) can be done rapidly and in the "near absence of attention" (Li, VanRullen, Koch, & Perona, 2002) is considered another source of support for the claim that conscious perception is possible without attention and, in particular, for the claim that gist perception is possible. The reasoning here is that detecting an instance of a category of a familiar set of objects (e.g., animal) in a scene in which the target instances (animals) are central to the gist can be understood as an important aspect of gist

perception: "The contents of gist could in fact be extended to include information about the presence of a complex target category whose appearance is not known in advance" (Li et al., 2002, p. 9599).

In the Li et al. experiments (2002), which entailed divided attention, the central task was detecting whether there was a single T or L present in flashed arrays of five elements that otherwise consisted of rotated Ls or Ts, and the second task was detecting whether a scene flashed in the periphery during the performance of the central task contained an instance of a prespecified target class, e.g., animal. The central task was difficult and designed to monopolize attentional resources. The fact that highly trained subjects were capable of detecting the presence of the target (e.g., animal or not) under these conditions, with no impairment in their ability to perform the central task, is considered the evidence that high-level perception closely related to gist perception occurs in "the near absence of attention" (Li et al., 2002, p. 9599)

It should be noted that others have questioned whether the ability to detect the presence of an animal, for example, in a briefly flashed scene is in fact a kind of gist perception. Treisman (2006) and Evans and Treisman (2005) suggest that this may only reflect the ability to use a few canonical features associated with the target category to determine whether an instance of it is present. If they are correct, then the Li et al. (2002) results cannot be taken as evidence of gist perception.

The question of what constitutes scene gist is not easily answered with precision. It has been described as "the perceived contents of a scene given a certain amount of viewing time (the appropriate viewing time taken to be about 100 ms)" (Li, Lyer, Koch, & Perona, 2007, p. 22). It has been described as a "spatial representation of the outside world that is rich enough to grasp the meaning of the scene, recognizing a few objects and other salient information in the image, which includes all levels of processing, from low level features (e.g. colour, spatial frequencies) to intermediate image properties (e.g. surface, volume) and high level information (e.g. objects, activation of semantic knowledge)" (Oliva, 2005, p. 251). Frequently, however, the term is used without being defined (see, for example, Potter, Staub, Rado, & O'Conner, 2002; Rousselet, Joubert, & Fabre-Thorpe, 2005; Vo & Henderson, 2010). In the research we will describe, we adopted a lenient criterion for what we accepted as evidence of gist perception. Any description that came close to capturing the main sense of the scene was counted as gist. For example, if the scene was that of two people in front of a hotel desk, we accepted the description, "people" as an instance of gist perception. Our criterion for what counted as gist perception, namely a more or less global sense of the scene (e.g., "people" for scenes in which people were a main aspect of the image, or "mountain" in a scene in which a mountain figured importantly), is consistent with the superordinate and

basic-level categorization distinction on which the Spatial Envelope Model of scene perception rests (Oliva & Torralba, 2001). We did, however, include a control for our criterion for gist, which we will describe later.

The claim that gist perception is possible without attention is not trivial. It is the main evidence given for attention-free awareness, and more idiosyncratically for us, because our past work has been taken to support this claim. The experiments we report reexamine this issue.

While our present research shared similarities with our prior gist perception studies and with both the Li et al. (2002) and Cohen et al. (2011) research, it differed from both in important ways. Unlike the Li et al. subjects, who performed a divided attention task and were highly practised on both tasks, having had thousands of training trials (12,000), and were actively searching for instances of a target category, our subjects were (1) not anticipating the presentation of a scene, (2) had no practice in identifying briefly presented scenes, and (3) had as their only task in the inattention condition reporting the longer arm of a cross. The scenes we used were significantly smaller than those used in our earlier research in an effort to prevent stimulus size from capturing attention but large enough to afford gist perception with attention and close in size to those used by Li et al. Additionally, unlike our previous research on this subject where the cross was superimposed on the scene, in the current research the scene and cross were spatially separated.

We believe it is very likely that the 12,000 practice trials given subjects in the Li et al. (2002) experiments made subjects highly proficient at both the central and the peripheral tasks. Proficiency here must mean that both tasks ended up requiring far less attention at the end of training than at the beginning. Consequently, it must have become easier to share attention between the co-occurring tasks, making both tasks much less attentionally demanding (see Hirst, Neisser, & Spelke, 1978, and Spelke, Hirst, & Neisser, 1976, for evidence that extended practice at simultaneous tasks reduces their attentional demands). The authors did run a control task to try to rule this out. The relevant control condition entailed changing the peripheral task from detecting the presence of an animal to detecting whether a single letter in the periphery was a T or L or identifying a red/green disc or a green/red disc, tasks on which subjects performed poorly. The authors took this as evidence that, despite extensive practice, attention remained absorbed by the central task. This may not have been the case, however, since it seems equally possible that if fewer attentional resources were absorbed by the highly practised central task, there still might not have been sufficient attentional resources remaining to perform the unpractised secondary task, which could not be done on the basis of a particular feature.

Our current experiments also shared features with the Cohen et al. (2011) studies but differed from them in significant ways as well. These similarities

and differences are more fully reviewed in the final section of this paper, so we only mention some important differences here. Our studies, unlike theirs: (1) Provide an index of the extent to which gist detection interfered with our distraction cross task when attention was divided; (2) demonstrate the difficulty of detecting gist under conditions of inattention even when the scene is present for longer, 100–200 ms as compared to 67 ms and 100 ms, their presentation times; (3) demonstrate the absence of gist perception even when the target scene is presented twice; and (4) provide some suggestion of negative priming by scenes which are not consciously perceived by subjects.

We chose to use our original inattention procedure because it had been validated many times over, engaged naïve subjects, and allowed us to look at gist perception without attention rather than under conditions of only divided attention. It was also the procedure used in the research reported in *Inattentional Blindness* (Mack & Rock, 1998), which has been construed as demonstrating gist perception without attention. However, unlike our earlier research, in the current studies the cross to which subjects attended was always spatially separated from the target scene so that when the crosses were centred at fixation the scene appeared in the periphery and, conversely, when the scene was centred at fixation, the crosses appeared in the periphery. Thus, the crosses were never superimposed on the scenes and so never fell within the area to which attention was directed by the cross task.

GENERAL METHOD

Participants

Two hundred and forty subjects were recruited from the New School University and Craigslist.org. They were either given course credit or were paid for their participation. Subjects reported normal or corrected-to-normal vision. All subjects were naïve about the purpose of the experiments. Thirty subjects participated in each of the various conditions of these experiments, except for Experiment 1 in which 60 subjects participated in each variant.

Apparatus

Images. The 12 greyscale critical images, all of which were 4.7 by 3.8 degrees, were taken from several sources. Two images were chosen from a dataset of natural scenes (Li et al., 2007) that had been specifically selected for their everyday, naturalistic settings. The other images were selected from Google image archives. We chose indoor and outdoor scenes that were uncluttered and in which gist was easily perceived (to view all images used in this study, please see http://www.newschool.edu/nssr/faculty.aspx?id=10342).

One of the images was chosen because it resembled a picture used by Li et al. (2002), and one was chosen because it resembled the image of people sitting around the table used in the experiments on gist reported by Mack and Rock (1998). There were three different scenes used in each of the first three experiments, and five used in Experiment 4. Each scene served as the critical image in the inattention, divided, and full attention conditions. Ten subjects were tested with each scene order in Experiments 1–3, and five subjects with each scene order in Experiment 4.

Crosses. Different horizontal/vertical crosses were used on every trial. When presented in the periphery, the vertical arm of the cross ranged in length from 4.1 degrees to 3 degrees and the horizontal arm ranged in length from 4.2 degrees to 2.9 degrees. Differences in length between arms ranged from 0.6 degrees to 1.2 degrees. When presented centrally, the vertical arm of the cross ranged in length from 3.8 degrees to 3 degrees and the horizontal arm ranged in length from 3.8 degrees to 2.9 degrees. Differences in length between arms ranged from 0.2 degrees to 0.8 degrees.

Choice arrays. Four images made up the forced-choice arrays in Experiments 3 and 4, one of which was the critical stimulus. The criteria for choosing the other three images were that they should be neither semantically related to nor resemble the spatial layout of the critical stimulus. They were chosen from Google image archives.

Equipment. Subjects were seated in a darkened room with their heads stabilized by a forehead and chinrest. They were positioned approximately 76 cm away from a computer screen. The stimuli were presented on a Mitsubishi Diamond Plus 74SB 14-inch colour monitor driven at 85 Hz, with a resolution of 1024 $\times$ 768 pixels. All experiments were programmed using SuperLab 4.

PROCEDURE

Experimental paradigm

There were two variants of the first and second experiments: One in which the scene was flashed in the periphery and the crosses were flashed centrally and the other in which the reverse was the case, the scene was central and the crosses peripheral. In the third and fourth experiments the cross always appeared peripherally in one of the four quadrants around fixation and the scenes were flashed at fixation. There were three testing conditions in each variant: The *inattention condition* (four trials, of which the first three were noncritical trials and were followed by one critical trial, except in the fourth experiment in which three trials were added and the critical scene was

presented twice, first on the fourth and then on the seventh trial with the cross); the *divided attention condition* (three trials, two noncritical followed by a critical trial), and the *full attention condition* (two trials, one noncritical and one critical trial). These three conditions were preceded by 10 practice trials, and they were always administered in this order since once the subject knew that a scene might be presented, there was no possibility of testing them under conditions of inattention. All in all there were 10 practice trials and nine experimental trials in Experiments 1, 2, and 3, and 12 in Experiment 4. (No scenes were ever presented during practice.) A noncritical trial consisted of a small " + " at fixation for 1500 ms, followed by an array consisting of a rectangular, pixelated pattern (subtending 4.7 degrees by 3.6 degrees) which, when the scene was to be presented centrally, was centred at fixation while a cross was flashed randomly in one of the four quadrants of the screen. The centre of the cross was 2.3 degrees from one of the corners of the centred rectangular pattern and 5.9 degrees from fixation. When the scenes were flashed in the periphery, the cross appeared at fixation and the rectangular pattern appeared centred 5.9 degrees from fixation in one of the quadrants in which the scene was presented. The whole display (cross plus pattern) was presented for 100 ms when the scene was flashed at fixation and the crosses were peripheral and for 200 ms when the scene was peripheral and the crosses were centred at fixation. On all trials, the test array was followed by a pattern mask for 500 ms, and then a blank screen. The subject's task was to tell the experimenter which arm of the cross was longer. On the critical trial, which was the last trial in each condition, an image of a scene replaced the pixelated pattern and was simultaneously flashed with the cross (see Figure 1).

Instructions. The main difference among conditions was the instructions given to the subjects. In the inattention condition, the subjects were only asked to report the longer arm of the cross. On the critical inattention trial immediately after the subjects reported the longer arm of the cross, they were asked whether they had seen anything else, and, if so, what. In the divided attention condition, the subjects were asked to report both the longer arm of the cross and anything else they saw on every trial. Since they now knew that something else might be presented, they might well have been on the lookout for it. In the full attention condition, they were asked to ignore the crosses and only report what else they saw. The reports of gist in the inattention condition provided the data about the perception of gist in the absence of attention and were compared with gist reports from the full attention condition. The difference between them was indicative of the contribution of attention to gist perception. The difference between the divided and full attention condition and between the divided and inattention condition provided information about the extent to which shared attention interfered

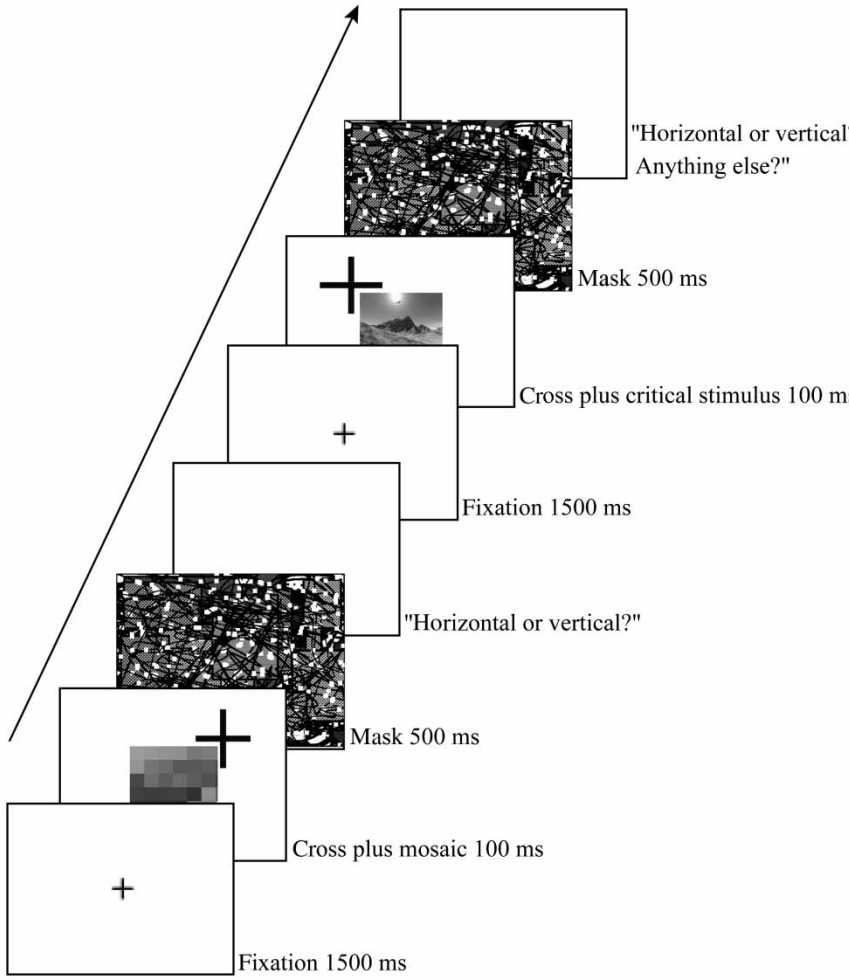

Figure 1. Experimental set-up for a central scene noncritical and critical trial. The position of cross and scene/pixelated pattern are reversed for a peripheral scene presentation.

with gist perception. In Experiments 3 and 4 subjects who failed to report seeing anything new in the inattention and divided attention conditions were asked to choose the image that had been flashed from an array of four images, each with clearly different gist, one of which was the critical image.

Recording gist

In every condition of these experiments, subjects were asked to describe what they saw other than the cross on the three critical trials, namely the trials in

which the image was presented in the three conditions. In the inattention and divided attention conditions, they reported this immediately after indicating which arm of the cross was longer, while in the full attention condition, this description was all that they were asked for. These descriptions were verbally given and unconstrained. If the experimenter did not fully understand the description given by the subjects, he queried them and asked them to say as much as they could about what they thought they had seen. On the critical trials in which the subjects reported nothing new in answer to the question of what else they had seen, this question seemed quite meaningless. The descriptions that were given were carefully recorded. As will be noted, in the third and fourth experiments observers were asked to choose the correct scene from an array of four scenes after reporting what else they had seen, if anything.

Control for gist

Five raters, all from the Mack lab and familiar with the experiments (two of whom are the authors), together reviewed each subject's verbatim descriptions of gist, assigning them a score between 1 and 3. Assigned scores could be either 1, 1.5, 2, 2.5, or 3. As a group, the raters discussed every description and came to a group decision about the most appropriate score to be given in each case. A score of 1 indicated that gist was captured, a score of 2 indicated partial capture of gist, and 3 indicated gist was not described at all. This was a time consuming process, but one that seemed most likely to get closest to a correct judgement of whether subjects' descriptions captured gist.

As an example, the descriptions that the raters accepted as describing the gist of the picture of the skaters: "People skating" or "people playing something" was rated as 1; "people in the subway" or "people in a museum" was rated 2; and "sunset" or "motor vehicle" was rated as 3. Our ratings reflected our acceptance of the Spatial Envelope model of gist perception, in which the superordinate category and the basic-level category are identified as adequate descriptions of the gist of a scene. In this instance, "people", as a description of the skaters, was deemed consistent with the superordinate category, while "people playing something" was deemed consistent with the basic-level category. Hence, both were accepted as describing the gist of the picture. Thus, a report of "people" for either the skater scene or the hotel lobby scene was accepted as gist. Descriptions that were given scores of between 1 and 2 were considered as adequately capturing the gist of the picture.

EXPERIMENT 1A: SCENE CENTRAL

Method

Sixty subjects were tested, all of whom had normal or corrected-to-normal vision.

Five scenes, each of which was flashed at fixation, served as the critical image in the inattention, divided, and full attention conditions for either 10 subjects in the case of the rhinoceroses on a beach, skaters, birthday party, or mountain images or 20 subjects in the case of the farmhouse image.

Results

Of the 60 subjects tested, 15 (25%) reported gist in the inattention condition, and 45 (75%) did not. Six of those 15 reports came from subjects for whom the birthday party scene was the critical inattention stimulus, an image that resembled one of those used in the earlier gist experiments (Mack & Rock, 1998). Of those 15 subjects in the inattention condition who reported gist, 10 correctly reported the longer arm of the cross (67%), and 66% of the subjects who did not report gist did so, which is not a significant difference. If we compare the number of correct cross reports on the noncritical trials in this condition with those on the critical trial, we find that there is no significant drop in the number of correct cross reports, $\chi^2(1) = .03$, $p = .80$.

Forty nine (82%) of the 60 reported gist in the divided attention condition. Of the 49 subjects who reported gist in this condition, 19 (39%) reported the correct cross arm, and six of the 11 (54%) remaining subjects who failed to report gist, did so. This difference is also not significant (Fisher exact probability test $p = .5$). If we compare the number of correct cross reports on the noncritical trials in this condition with those on the critical trial, we do find a significant drop in the number of correct cross reports, $\chi^2(1) = 0.912$, $p < .01$.

Overall, there were 40 (67%) correct cross length reports on the critical trials in the inattention condition and 25 (42%) in the divided attention critical trials. A McNemar pairwise comparison showed this difference to be highly significant ($p < .001$), suggesting that dividing attention is done at the cost of correctly reporting the longer arm of the cross.

Almost all (59 of 60, i.e., 98%) of the subjects reported gist in the full attention condition.

A Cochran's Q comparison of the number of gist reports in each of the three conditions proved significant, $Q(2) = 72.545$, $p < .001$. McNemar pairwise comparisons indicated that the differences in the number of gist reports in the inattention condition were significantly different from those in

the full attention condition ($p < .001$) and from those in the divided attention condition ($p < .001$). There was also a significant difference between the number of gist reports in the full versus divided attention condition ($p < .05$). It is important to note that *all* subjects who did not report gist in the inattention condition were inattentionally blind to the scene, that is, they were completely unaware that the mosaic appearing in the trials prior to the critical one was a scene on the critical trial.

Discussion

While these data indicate that gist is unlikely to be perceived under these conditions when the scene is presented at fixation, they tell us nothing about whether gist would be picked up were the scenes to appear in the periphery as they were in the Li et al. experiments (2002). While this seemed unlikely, it nevertheless warranted exploration, particularly since in our earlier work, we found less inattentional blindness for stimuli presented in the periphery than for those presented at fixation. The purpose of the next experiment was to look at gist perception when the critical stimulus appeared in the periphery.

EXPERIMENT 1B: SCENE PERIPHERAL

Method

A new group of 60 subjects was tested coming from the same population as those who had participated in the prior experiment. Again, groups of 10 subjects were presented with each of the images (with the exception of the farmhouse picture, which was presented to 20 subjects).

This experiment was very much like the previous one except that we exchanged the locations of the crosses and the images and increased exposure time to 200 ms.

Results

Nine (15%) of the 60 subjects tested reported gist in the inattention condition, so 51 (85%) did not. (Note that of these nine subjects, six had been presented with the birthday party scene.) On the critical trial in the inattention condition, 51 of the 60 subjects (85%) correctly reported the longer arm of the cross. Of the nine subjects who reported gist, five correctly reported the longer arm of the cross (55%), and 46 of the remaining 51 subjects (90%) who did not report gist did so. This difference is significant (Fisher exact probability test, $p < .05$). We again found no significant

difference in the number of correct cross arm reports on the noncritical as compared to the critical trial, $\chi^2(1) = 0.21$, $p = .6$.

In the divided attention condition, 30 (50%) of the 60 subjects reported gist. Of the 30 subjects who reported gist, 16 subjects got the cross arm correct (48%), and 14 of the 27 subjects who did not report gist did so (51%), which is not a significant difference, $\chi^2(1) = 0$, $p = 1$. The difference between the number of correct cross arm reports on the noncritical trials with those on the critical trial is significant, $\chi^2(1) = 3.87$, $p < .05$.

If we compare the number of correct arm cross reports on the critical inattention trial with those on the critical divided attention trial, we find a significant difference, ($p < .001$), again revealing the cost of dividing attention.

In the full attention condition 40 (67%) reported gist.

A Cochran's Q comparison of the number of gist reports in each of the three conditions proved significant, $Q(2) = 51.161$, $p < .001$. McNemar pairwise comparisons between the gist results in each of the three conditions confirmed that the number of gist reports in the inattention condition was significantly different from that in the full attention condition ($p < .001$), and also from the divided attention condition ($p < .001$). There was also a significant difference between the number of correct gist reports in the divided and full attention conditions ($p < .05$).

Discussion

The primary finding thus far is that under conditions of inattention, when a scene is flashed in the periphery, subjects are not only unlikely to report gist, they are completely unaware that a scene has been presented regardless of whether the scene was of animals, people, or a building. Since these pictures were visible for almost eight times the duration of the pictures flashed in the Li et al. study (2002), which were also presented in the periphery at about the same remove from fixation as our pictures, our results are in direct opposition to those reported by Li et al.

Our finding that subjects had difficulty reporting gist in both the divided and full attention conditions (only 60% reported gist in the full attention condition) was surprising since in both cases subjects were aware that something other than the cross, namely a scene, would appear and that they were to describe it as best they could. We take these results at face value as indicating that the gist of a peripherally presented scene is not easily picked up even when subjects are trying to perceive it if they are unpractised. Although there was no statistical difference between perceiving gist under conditions of inattention when the scene was shown at fixation or in the periphery, $\chi^2(1) = 1.302$ $p = .25$, we think this could be attributed to the fact that gist reports were very infrequent in both cases. In the divided attention

condition there was a significant decrease in the number of reports of gist when the image was presented peripherally, $\chi^2(1) = 8.665$ $p < .05$, as there was in the full attention condition, $\chi^2(1) = 18.701$ $p < .001$. This difference makes clear that picking up gist is more difficult when the scene is peripheral. That some subjects were even unable to report that animals were present when the rhinoceroses on the beach scene was displayed under conditions of divided or full attention stands in sharp contrast to the Li et al. findings (2002) that the presence of an animal (one of their target categories) was perceptible in the "near absence of attention". A possible reason for this will be proposed later.

Since the birthday party scene, which closely resembled the scene of the young people around a table in our original exploration of gist perception (Mack & Rock, 1998), seemed to allow for easier gist perception than the other scenes (the difference between the frequency of correct gist reports of the birthday scene [60%] and the farmhouse scene, which afforded the next most frequent number of gist reports [22%] in the inattention condition, was significant, $\chi^2(1) = 7.96$ $p < .001$), we wondered whether it was the presence of the faces in the picture that might account for this, as faces are known to be special stimuli. To explore this possibility, we ran a final variant of experiment one in which the three scenes used all containing multiple faces.

EXPERIMENT 2A: FACES IN SCENES (SCENE CENTRAL)

Method

This experiment was identical to Experiment 1A, in which the scenes were central and the crosses were peripheral and flashed for 100 ms, except the scenes were different and all of them contained visible faces. One of the three scenes pictured an audience in a theatre, another depicted three figures standing around a hotel desk, and the third was a group of runners approaching the finish line with arms outstretched. In all three scenes, faces were clearly visible. Because people were the primary component of each of the target scenes, we accepted any description of gist as correct that made reference to having seen people or a person.

A new group of 30 subjects was tested and again groups of 10 subjects were presented with each of the three images in each of the three conditions.

Results

In the inattention condition, four of the 30 subjects tested (13%) reported gist, the others (87%) did not. Twenty-six (87%) reported the correct arm of the cross in the critical inattention trial. Only two of the four subjects who reported gist, got the cross arm correct (50%), and 24 of the 26 subjects who

did not report gist did so (92%), a difference that is not significant (Fisher exact probability test, $p = .07$). There was a significant drop in the number of correct cross reports if we compare those given on the noncritical trials with those given on the critical trial, $\chi^2(1) = 5.14$, $p = .02$.

Eighteen of the 30 subjects (60%) reported gist in the divided attention condition. Nineteen of the 30 subjects (63%) reported the longer arm of the cross in the critical trial in this condition. Of the 18 subjects who reported gist in the divided attention condition, 11 reported the longer arm of the cross (61%), and of the 12 subjects who did not report gist, eight (66%) reported the longer arm of the cross, again a difference, which is not significant (Fisher exact probability test, $p = 1$). In this condition we also find no significant difference between the number of correct cross reports given on the noncritical trial and those on the critical trial, $\chi^2(1) = .01$, $p = .9$.

Again, fewer subjects gave correct cross length reports in the divided than in the inattention condition ($p < .05$).

In the full attention condition, 25 (83%) of the subjects reported gist.

A Cochran Q comparison of the number of gist reports in the three conditions was significant, $Q(2) = 29.158$, $p < .001$. McNemar pairwise comparisons revealed significant differences between the inattention and full attention condition ($p < .001$), between the inattention condition and divided attention condition ($p < .001$), and between the divided and full attention conditions ($p < .05$).

EXPERIMENT 2B: FACES IN SCENES (SCENE PERIPHERAL)

The scenes with faces again served as the critical stimuli but they now were presented in the periphery while the crosses were presented centrally. A new group of 30 subjects was tested.

Results

Only one of the 30 subjects (3%) reported gist in the critical inattention trial, meaning that 29 (97%) did not. Twenty-nine of the 30 subjects (97%) reported the correct arm of the cross on the critical inattention trial. The one subject who reported gist on the inattention critical trial did not get the cross arm correct, but all the remaining 29 subjects did. Again if we compare the number of correct cross reports on the noncritical trials in this condition with those on the critical ones, we find there is no significant difference, $\chi^2(1) = .03$, $p = .8$.

In the divided attention condition 14 of the 30 subjects tested (47%) reported gist. Of those 14 subjects, nine got the cross arm correct (64%). Of the 16 subjects who did not report gist, 11 (68%) reported the longer arm of the cross, which is not a significant difference (Fisher exact probability test,

$p = 1$). There also was no significant drop in the number of correct cross reports if we compare those given on the noncritical with those given on the critical trials, $\chi^2(1) = 2.78$, $p = .1$.

There were significantly fewer correct cross reports given by the subjects on the critical trial in the divided attention condition than in the comparable trial in the inattention condition ($p < .05$).

On the full attention critical trial, 21 of the 30 subjects (70%) reported gist.

A Cochran Q comparison of the number of gist reports in each of the three attention conditions proved significant, $Q(2) = 30.900$, $p < .001$. McNemar pairwise comparisons between the gist results in each of the three attention conditions confirmed that the number of gist reports in the inattention condition was significantly smaller from those in the full attention condition ($p < .001$) and the divided attention condition ($p < .001$). The difference between the results in the divided and full attention conditions was also significant ($p < .05$).

Discussion

The results from this experiment do not support the conjecture that the presence of faces in the birthday party scene alone could account for the greater number of gist reports when subjects were shown this scene in the inattention condition. This leaves us with an unexplained outcome, but with still no evidence that gist is perceived without attention.

The results of Experiments 1 and 2 indicate that picking up gist is more difficult in the divided and full attention conditions when the scene is presented in the periphery than when it is presented centrally, $\chi^2(1) = 9.57$, $p < .001$. Similarly, a chi-square analysis of the difference between the number of gist reports in the full attention condition, central versus peripheral, was also significant, $\chi^2(1) = 18.76$, $p < .0001$. The difference does not quite reach significance in the inattention condition where gist perception was infrequent in both locations, $\chi^2(1) = 3.33$, $p = .06$ (see Figure 2).

In our next attempt to find evidence of gist perception without attention, we used an altered testing procedure that involved scene recognition as well as recall of the gist of the critical scenes. After subjects reported whether they had seen anything other than the cross on the critical trials, whether or not they reported seeing anything, they were told that a scene had been present and to try to select it from among four scenes. Of course, if the observer had already reported gist, asking them to select the scene was redundant. The same forced choice procedure was also used in the divided and full attention conditions.

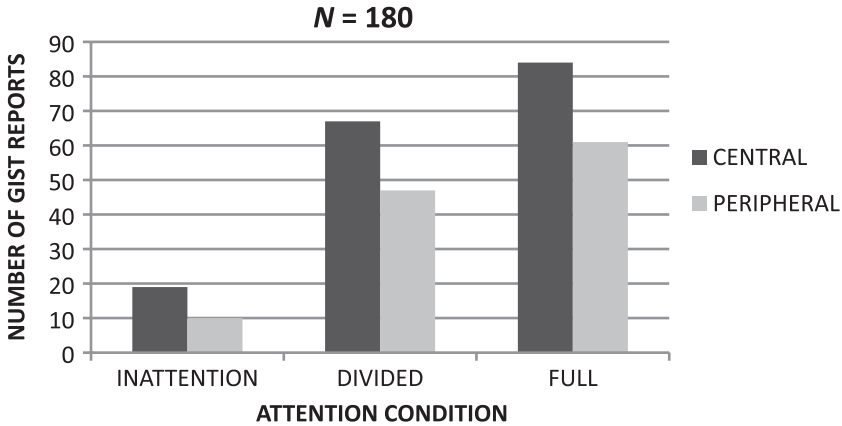

Figure 2. Combined results of Experiments 1–2.

Although we recognize that picking up the gist of a scene and being able to identify it from among others are *not* identical tasks (identifying a scene might require more information), this was not the case in our experiments where only the correct target image and none of the other three choice images represented anything close to the gist of the target scene. Thus, had subjects picked up gist, they would easily have been able to choose the correct image. Moreover, since recognition is less demanding than recall, we thought by using a forced-choice recognition procedure after questioning subjects about whether they had seen anything other than the cross, we might increase the possibility of finding evidence of scene/gist perception. In addition, asking observers to select the flashed scene from among four scenes is not only less demanding, it also has the possibility of revealing priming by the flashed scene even if it is not consciously perceived.

EXPERIMENT 3: GIST RECOGNITION

Method

Thirty new participants were tested.

All scenes were presented at fixation with the crosses flashed for 100 ms in the periphery. Three scenes used earlier were the target scenes, namely the skaters, the hotel lobby, and the mountain landscape, each of which was presented in each of the three viewing conditions to groups of 10 subjects. The choice arrays consisted of the target scene and three scenes selected for their relative simplicity and because they shared no obvious similarities with the target scene either in terms of the gist they conveyed or their layout.

These arrays were shown only on critical trials immediately after subjects reported whether they had seen anything else and what it was. When they had been blind to the flashed scenes, they of course thought that they were simply guessing and, as it turned out, they were.

Results

Of the 30 subjects tested only three reported gist in the inattention condition, (10%) and the remaining 27 were inattentionally blind. Of the inattentionally blind subjects, only four chose the correct scene from the choice array (see Figure 3). Of the three subjects who reported gist in this condition, two got the cross arm correct (66%), and 23 of the 27 subjects who did not get gist did so (85%), which is not a significant difference (Fisher exact probability test, $p = .5$) There was no significant drop in the number of correct cross reports given on the noncritical as compared to the critical trials in this condition, $\chi^2(1) = .06$, $p = .8$.

Seventeen of the 30 subjects (57%) reported gist in the divided attention condition and of the remaining 13 subjects only one chose the correct picture from the array. Of the 17 subjects who reported gist in the divided attention condition, 10 got the cross arm correct (58%), and of the remaining 13 subjects, eight did so (61%). Again this difference is not significant (Fisher exact probability test, $p = 1$). There was no significant drop in the number of correct cross reports between the critical and noncritical trials, $\chi^2(1) = 2.78$, $p = .9$.

There were more correct reports of the longer arm of the cross on the critical trial in the inattention condition (25 out of 30 subjects) (83%) than in

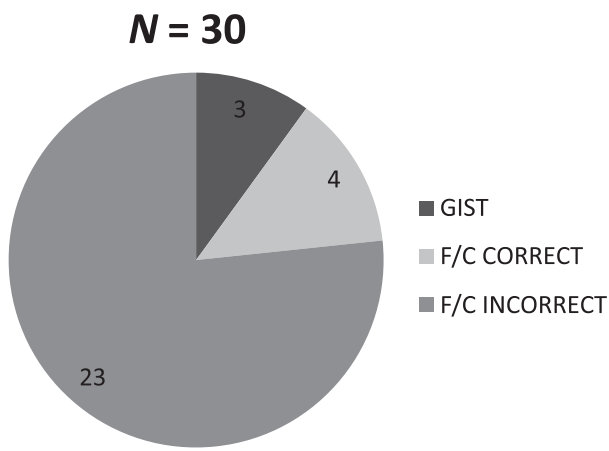

Figure 3. Experiment 3: Results of forced-choice (F/C) task. Number refers to subjects in the inattention condition.

the divided attention condition (18 out of 30 subjects) (60%) ($p < .05$), which is additional evidence that picking up gist draws attention away from the cross task. If gist perception did not require attention, this should not have occurred.

In the full attention condition, 27 of 30 the subjects (90%) reported gist correctly and spontaneously, but none of the three other subjects chose the correct picture.

A Cochran Q comparison of the number of gist reports in the three conditions was significant, $Q(2) = 36.333$, $p < .001$. McNemar pairwise comparisons revealed significant differences between the inattention and full attention condition (McNemar, inattention vs. full attention, $p < .001$), the inattention condition and divided attention condition (McNemar, inattention vs. divided, $p < .001$), and between the divided and full attention conditions (McNemar, divided vs. full attention, $p < .001$). We did not consider the four correct identifications of the target image from the choice array by the 27 inattentionally blind subjects as evidence of gist perception, since 6.75 of these subjects should have chosen this image by chance and the difference between four out of 27 and 6.75 out of 27 is not significant ($p = .09$).

Once again even using a less demanding index of gist perception, we found no evidence of gist perception without attention. There was no evidence of positive priming since this would have been revealed by a significant increase in the number of correct scene selections even when subjects failed to report seeing anything new other than the cross and this did not happen. Because fewer "blind" subjects chose the correct image from the choice array than was expected by chance, although not significantly fewer, we thought this might suggest the possibility of negative priming and that it merited further investigation. To that end we ran a final experiment in which the critical stimulus was presented twice in the inattention condition, which, by virtue of repetition priming, should enhance priming, whether negative or positive, and, of course, should also facilitate gist detection. Of course, presenting the critical scene twice risked the possibility that it would be noticed and, had this occurred, the results would no longer be relevant to the question of gist perception without attention.

EXPERIMENT 4: REPETITION PRIMING

The difference between this experiment and the previous one is that the critical scenes were presented at fixation twice rather than once in the inattention condition. As usual the critical image was flashed with the cross on the fourth inattention condition trial, but now was followed by two more trials in which the cross was presented again with the mosaic pattern, and

finally on the seventh and last trial in this condition, the critical image was again flashed with the cross. The subjects were asked *only* after the last trial whether they had seen anything new and if so, what it was and if not, to choose a scene from the choice array. After they had done so, they also were asked whether they had seen a picture on any earlier trial.

We tested a group of 30 new subjects using seven scenes as target stimuli, three of which were used earlier (the skaters, the mountain, and hotel lobby) and four of which were new and depicted a plane on a runway, a train, a dog, and an open road.

Results

In the inattention condition five of the 30 subjects tested (17%) gave reasonable accounts of gist, meaning that 25 (83%) did not. Only one of the 30 subjects reported having perceived the target scene on the first of the two trials on which it was presented and this was one of the subjects who correctly reported gist in the critical inattention trial. Of the 25 subjects in the inattention condition who did not see anything new on the critical trial, only three chose the correct picture from the choice array (see Figure 4), which again was fewer than expected by chance, but again not significantly so ($p = .064$). Twenty-two subjects (73%) reported the correct arm of the cross in the fourth inattention trial (a trial in which the scene appeared with the cross but was not enquired about), and 14 (47%) reported the correct arm of the cross on the seventh and critical inattention trial. Of the five subjects who reported gist in this condition, two got the cross arm correct (40%), and 12 of those 25 who did not get gist did so (48%) This difference is

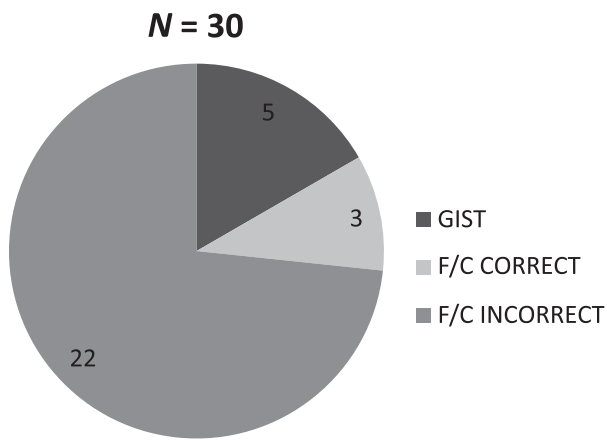

Figure 4. Experiment 4: Results of forced-choice (F/C) task. Number refers to subjects in the inattention condition.

not significant (Fisher exact probability test, $p = 1$). We do find a significant decrease in the number of correct cross arm reports if we compare the number given on the critical trial with those on the noncritical trials, $\chi^2(1) = 4.24$, $p = .03$.

Twenty-four of the 30 subjects (80%) reported gist in the divided attention condition. Of the six subjects who did not report gist in the divided attention condition, two chose the correct image from the choice array, which did not differ significantly from what would be expected by chance. Eighteen subjects (60%) correctly reported the longer arm of the cross on the critical trial in this condition. Of those 24 who reported gist in the divided attention critical trial, 15 got the cross arm correct (62%), and of the six who did not report gist, three did so (50%), a difference which again is not significant (Fisher exact probability test, $p = .6$). There again was no significant drop in the number of correct cross reports if we compare the critical with the noncritical trials, $\chi^2(1) = 0.94$, $p = .33$. However, in this experiment, for reasons we cannot explain there were fewer reports of the correct cross arm on the critical inattention trial than in the divided attention condition.

In the full attention condition, 24 (80%) of the subjects reported gist. Of the six subjects who did not report gist, four chose the correct image from the choice array, and this is greater than what one would have expected by chance.

A Cochran Q comparison of the number of gist reports in the three conditions was significant, $Q(2) = 38.000$, $p < .001$. McNemar pairwise comparisons revealed significant differences between the inattention and full attention condition ($p < .001$), between the inattention condition and divided attention condition ($p < .001$), but not between the divided and full attention conditions ($p = 1.000$).

To summarize, showing the scene twice under conditions of inattention did not facilitate gist perception (three out of 30 subjects in Experiment 3 reported gist in the inattention condition and five out of 30 did so in Experiment 4, a difference which is not significant, $\chi^2(1) = 0.144$, $p = 0.7$. If we combine the results from Experiments 3 and 4, which would mean that together seven out of 52 subjects who were inattentionally blind in these two experiments chose the correct picture from the choice array, then we find that this number is significantly different and less than what could be expected by chance ($p < .05$). This suggests that there may be a tendency for unseen target scenes to produce negative priming. If this were, in fact, the case, this would mean that scenes that are not reportable might nevertheless be encoded, but encoded with some kind of inhibitory tag making it less likely that the subject will choose the correct image (Kuvaldina, 2011). However, if this is a case of negative priming, the way it manifests itself is distinctly different from the normal way in which negative priming is

observed, namely as an increase in reaction time when responding to the previously ignored stimulus.

GENERAL DISCUSSION

Evidence obtained in these experiments is abundantly clear: Gist perception, even given our lenient criteria for what constitutes gist, infrequently occurs under conditions of inattention. This is so when the scene is flashed for 100 ms at fixation and for 200 ms in the retinal periphery and even when the critical image is presented twice at fixation under conditions of inattention. This is evident too from our summary finding from all 4 experiments that there were significantly fewer correct reports of the longer cross arm in the critical divided attention trial than in the inattention critical trial signifying that there is an attentional cost for dividing attention, $\chi^2(1) = 30.52$, $p < .0001$.

Interpretation of cross task results

Although we found that overall there was not a significant difference between the number of correct cross reports on noncritical trials as compared to the critical trial in the inattention condition across all four experiments, $\chi^2(1) = 3.37$, $p = .06$, we did find that when we combined the results from all four experiments, subjects who reported gist in this condition were significantly less likely to get the cross arm correct than those who did not, $\chi^2(1) = 8.92$, $p < .01$, which suggests that there is an attentional price to pay for gist detection, such that when gist is reported, it is at the cost of cross-task performance.

If we analyse the cross performance results from the divided attention conditions of all four experiments combined, we find that there was a significant decrease in the number of correct cross arm reports in the critical as compared to the noncritical trials, $\chi^2(1) = 4.4$, $p < .01$, suggesting that when the scene appeared on the critical trial, it taxed attention even if scene gist was not perceived. But, if we compare the number of correct cross reports of all those who reported gist with all those who did not in this condition, the difference is not significant, $\chi^2(1) = 0.39$, $p = .5$, which we take to mean that because attention was already divided, there was no additional cost for perceiving gist (see Figure 5).

Our results are consistent with those reported by Cohen et al. (2011), who also found no gist perception without attention, but are at odds with the Mack and Rock (1998) results, which seemed, at least to others, to suggest that gist could be picked up without attention.

The difference between the results we report here and those we reported earlier (Mack & Rock, 1998) can, we think, be attributed to procedural

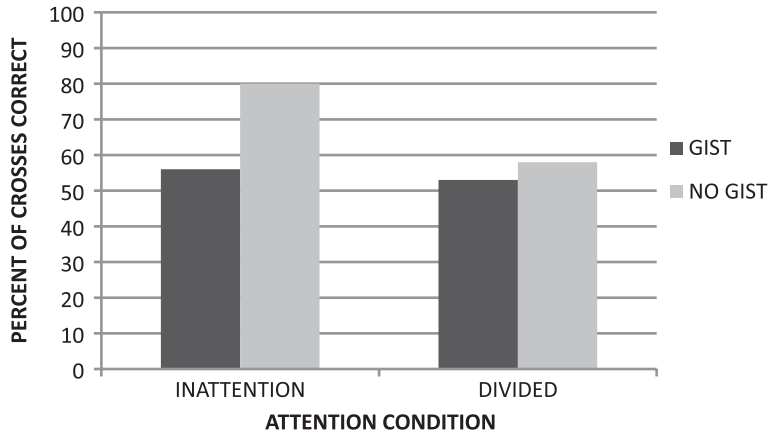

Figure 5. Comparing percent of crosses correct for those who got gist with those who did not across all 4 experiments. Inattention condition significant ($p < .01$). Divided attention condition not significant ($p = .5$).

differences. In our earlier work the cross to which the subjects attended was superimposed on the scene, whereas in the current experiments the two were spatially separate, making it less likely that the focus of attention would include the target scene and, therefore, making the current experiments a cleaner test of gist perception without attention. It should be noted, however, that in the Cohen et al. (2011) studies the elements composing the distraction task, and therefore the elements to which the subjects attended were superimposed on the target scene and, nevertheless, gist was generally not perceived. This may be because their distraction task, multiple object tracking or an RSVP task, was more demanding than ours, which, if so, would have more fully engaged attention, so that even though the target scene was located within the attended region, it might not benefit from it. This reasoning is in line with a perceptual load account of inattentional blindness which proposes that the greater the load, the fewer the attentional resources available for task irrelevant stimuli (Cartwright-Finch & Lavie, 2007).

These results do not support the claim that gist is picked up without attention (e.g., Tsuchiya & Koch, 2009; van Boxtel et al., 2010), nor are they consistent with the conclusion that gist, even if one equates it with a superordinate category like inside/outside or a basic-level category like "mountain" or "animal" (Loschky & Larson, 2010), is perceived without attention (Li, VanRullen, Koch, & Perona, 2005).

There were, of course, major differences in procedure between our studies and the Li et al. (2002) studies, although there were also some shared aspects. Our scenes were of approximately the same size and composition as Li

et al.'s, and on half of our trials in Experiments 1 and 2, the scenes were also flashed in the periphery at about the same distance from fixation when the cross was at fixation. Our scenes were also masked, but our exposure times were much longer than theirs (200 ms when the image was in the periphery and 100 ms when the scenes were centred at fixation, whereas theirs were flashed for only 27 ms in the periphery). This should have made detection considerably easier for our subjects. Although this may have been offset by the fact that our images were greyscale whereas theirs were coloured and colour may help to facilitate gist.

Most importantly, although our procedure did include a divided attention condition, which was the only condition used by Li et al. (2002), we also included an inattention condition, in which subjects had neither the expectation of seeing a scene nor the intention to search for a particular kind of scene, one with or without an animal, for example. In addition, there was a full attention condition, in which the subject's only task was to report what else they saw ignoring the cross. Our observers (unlike theirs, who were trained over 15 days with daily training sessions lasting an hour) were naïve and untrained. Our subjects received only 10 practice trials, which consisted of reporting the longer arm of the cross. There was only a single critical trial in each condition. It might also be noted (as mentioned earlier in this paper) that our distraction task, reporting the longer arm of the cross, was easier than their central attention task, reporting the presence of a T among Ls or an L among Ts in a briefly flashed array but, even so, gist was not perceived without attention.

There is an additional important difference between their procedure and ours, which we mentioned in passing earlier. In ours, observers were never given a search target but were simply asked what else they had seen other than the cross. In Li et al.'s (2002) and virtually all other research examining the capacity to report the presence of some target, which has frequently been taken to be synonymous with gist perception, subjects are given a target category to search for (e.g., Greene & Oliva, 2009; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Rousselet et al., 2005).

These differences in procedure are important and we believe that they may account for the dramatic differences in results. Since we believe that our procedure actually addressed the question of whether gist is perceived without attention, whereas theirs (Li et al., 2002) asks whether it can be detected under conditions of divided attention, our results are more relevant to the question at issue, namely whether gist is perceived in the near absence of attention. It should be remembered that under conditions of inattention, not only did most of our observers most of the time fail to report gist, they failed even to be aware that anything new had appeared on the critical trial. In other words, they were truly inattentionally blind.

It should be noted that asking observers whether a scene contains an animal or is an indoor view clearly clues them about what to look for and primes the set of features that distinguish these scenes from others. In the absence of these kinds of search categories, the task is more difficult, since there are no preexisting constraints on what might be present, which we think is why observers sometimes failed to report gist even with full attention and why they were able to perform the search tasks assigned by other researchers, who argue that gist is perceived under conditions of divided attention. As Evans and Treisman (2005) have argued, detecting the presence of an animal or a vehicle, the task used by Li et al. (2002), may require only detecting the presence of a single canonical feature.

A word about the failure of Li et al. (2002) to find that observers who could report the presence of an animal in a briefly flashed peripheral scene while carrying out a difficult central task were unable to distinguish between an L or T or red/green versus green/red circle. Their failure to be able to do so allowed the researchers to conclude that gist, if equated with, for example, identifying whether an animal is present in a scene, occurred in the near absence of attention. However, it seems clear that if one is searching for the presence of an animal, as Evans and Treisman (2005) note, there are a set of canonical features which distinguish them (e.g., legs, fur, ears), whereas when searching for a T or an L, it is only the way in which the lines are conjoined that distinguishes them from each other and the same is true when distinguishing a green/red from a red/green circle. As Evans and Treisman point out, this conjunction requires focused attention. This could account for why their observers were able to perform one search task but not the other while carrying out a concurrent attention-demanding task.

As already noted, we also found that in general, under all our testing conditions, observers had more difficulty reporting gist when the scene was presented in the periphery for 200 ms than when it was at fixation for only 100 ms, except for the inattention condition where there was a not quite significant difference in gist perception, probably because gist perception in this condition was so poor. These results are in stark contrast with the Li et al. (2002) results, which were said to demonstrate gist perception when the scenes were presented peripherally.

The fact that we tended to obtain a consistently greater number of correct reports of the longer arm of the cross on the critical trials in the inattention conditions than in the divided attention conditions across all the experiments (McNemar pairwise comparison, $p < .001$) is additional evidence that gist perception uses attentional resources. When attention is entirely focused on the cross task, subjects are more likely to be correct than when they have to divide their attention between the cross and the now expected additional stimulus, suggesting that the added task which includes gist detection is not attention free.

Since the published results of Cohen et al. (2011) have already demonstrated the difficulty of perceiving gist without attention, what do our experiments add to what we know beyond serving as a confirmation of their results using a different inattention procedure, which in itself may be useful? First of all, our results show that even with longer exposure times (we used presentation times of 100–200 ms as compared to the 67 and 100 ms exposure used by Cohen et al. *and* what they describe as a less attentionally demanding distraction task; Cohen et al., 2011, p. 1168), the large majority of our subjects are inattentionally blind to the scenes. Eighty-three per cent of all 240 subjects tested in our experiments were inattentionally blind, that is saw nothing new on the critical inattention trial. If we eliminate the birthday party scene from our analysis because it seemed to allow for easier gist perception for reasons we are yet to determine, then we find that 90% of our subjects were inattentionally blind to the scenes across all the inattention critical trials. In contrast, Cohen et al. report that 57% (this is an average of the 64% inattentionally blind subjects they report in Experiment 1a and the 50% inattentionally blind subjects they report in Experiment 1b) of their subjects demonstrated inattentional blindness. And if, as they conjectured, our cross task carries a lighter perceptual load than the multiple object tracking and RSVP tasks they used, this makes that point even stronger. Despite longer presentation times and perhaps a lighter perceptual load, we found greater inattentional blindness. However, their images were much larger than ours (subtending 21 degrees whereas ours were only 4.7 by 3.6 degrees) and were coloured, whereas ours were greyscale, and if size and colour facilitate gist perception, this might have offset the difficulty caused by the speculated increase in perceptual load of their task.

Another significant difference between the Cohen et al. (2011) results and our own stems from the nature of the different forced-choice procedures which they used and which we used in Experiments 3 and 4. In one condition of their first experiment, when their subjects were inattentionally blind to the scene, they were asked to choose a word from a list of words that best described what had been there which they had not seen. In another condition, instead of choosing a word, subjects chose the scene from an array of six images, three of which always were from the same object category as the target scenes (either of vehicles or animals) and three from the other category. When our subjects were inattentionally blind to a scene, they were asked to choose the scene from an array of four pictures, which included the target scene and in which the gist and layout of the other three pictures were distinctly different from that of the target scene, but even with

our forced-choice procedure, where if the slightest gist had been picked up identification would have been easier than in the Cohen study, we still found the correct image was chosen less often than would have been expected by chance. We note that there was no tendency for subjects to choose any one wrong picture more than any other, which would seem to confirm this.

In addition to these differences, our fourth experiment, which was meant to explore the possibility of priming, either positive or negative, adds additional force to the conclusion that gist is not perceived without attention, since even when the critical scenes are presented twice and thus have been visible for a total of 200 ms at fixation in the inattention condition, they are not seen, i.e., inattentional blindness persists. As noted earlier, the combined number of inattentionally blind subjects who failed to choose the correct image from the choice array in Experiments 3 and 4 was significantly less than what would have been expected by chance, hinting at the possibility of some sort of negative priming. The possibility of encoding of the unreportable scenes, which would be necessary were priming to occur, is consistent with the findings of Marois and Chun (2004), that scenes not reported due to the attentional blink activate regions of the medial temporal cortex involved in high-level scene representations. This, however, requires further investigation.

We note that like Cohen et al. (2011), we too found apparent differences in the ease with which gist was perceived among the 12 different pictures that served as our critical stimuli. In Experiment 1, one of these differences was significant, namely the difference between the birthday party scene and the other images used in that experiment. We thought this particular difference might have been due to the presence of faces in the party scene, but Experiment 2 failed to confirm this. The failure in Experiment 2 to find a decrease in inattentional blindness with images containing faces is nevertheless interesting given that faces are special stimuli and have been shown to be able to capture attention under some conditions (e.g., Borrmann, Boutet, & Chaudhuri, 2003; Mack & Rock, 1998). It might be noted that none of the earlier work on scene perception without attention or in the near absence of attention involved images containing faces, making our failure to find that their presence facilitates gist perception of additional interest.

In conclusion, we believe there are strong grounds on which to assert that gist perception does not occur without attention, even when target scenes are shown twice. This removes the main basis for the claim that conscious perception occurs in the absence of attention.

REFERENCES

Borrmann, K., Boulet, I., & Chaudhuri, A. (2003). Spatial attention favors faces over non-face objects in an attentional cueing task. *Journal of Vision*, *3*(9), Article 818.

Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition*, *102*(3), 321–340.

Clarke, J., Mack, A., Slesar, C., & Erol, M. (2011). No gist perception without attention. *Journal of Vision*, *11*(11), Article 156.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural scene perception requires attention. *Psychological Science*, *22*(9), 1165–1172.

Evans, A. K., & Treisman, A. (2005). Perception of object in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1476–1492.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*(4), 464–472.

Hirst, W. C., Neisser, U., & Spelke, E. S. (1978). Divided attention. *Human Nature*, *1*, 54–61.

Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*, 3256–3297.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Science*, *11*(1), 16–22.

Kuvaldina, M. (2011). Negative choice in inattentional blindness. *Journal of Vision*, *11*(11), Article 158.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Science*, *7*, 12–18.

Li, F.-F., Lyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real world scene? *Journal of Vision*, *7*, 1–29.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Science*, *99*, 9596–9601.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2005). Why does natural scene categorization require little attention? Exploring attentional requirements of natural and synthetic stimuli. *Visual Cognition*, *6*, 983–924.

Loschky, L. C., & Larson, A. M. (2010). The natural/man-made distinction is made before basic-level distinctions in scene gist processing. *Visual Cognition*, *18*, 512–536.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: Bradford Books, MIT Press.

Marois, R., Yi, D., & Chun, M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron*, *41*, 466–472.

Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Encyclopedia of neurobiology of attention* (pp. 251–256). San Diego, CA: Elsevier.

Oliva, A., & Torralba, A. (2001). Modeling the shape of a scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*, 145–175.

Potter, M. C., Staub, A., Rado, J., & O'Conner, D. H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1163–1175.

Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get the "gist" of real world natural scenes? *Visual Cognition*, *12*, 852–877.

Spelke, E. S., Hirst, W. C., & Neisser, U. (1976). Skills of divided attention. *Cognition*, *4*, 215–230.

Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, *14*, 411–443.

Tsuchiya, N., & Koch, C. (2009). The relationship between consciousness and attention. In S. Lawreys & G. Tononi (Eds.), *The neurology of consciousness: Cognitive neuroscience and neuropathology* (pp. 63–79). London, UK: Elsevier.

Van Boxtel, J. J. A., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: On sufficiency and necessity. *Frontiers in Psychology*, *1*(217), 1–13.

Vo, M. L., & Henderson, J. L. (2010). The time course of initial scene processing for eye movement guidance in natural scenes. *Journal of Vision*, *10*, 1–13.

Manuscript received July 2011
Manuscript accepted February 2012