

NIH Public Access

Author Manuscript

Q J Exp Psychol (Hove). Author manuscript; available in PMC 2009 April 13.

Published in final edited form as:

Q J Exp Psychol (Hove). 2009 February ; 62(2): 276–285. doi:10.1080/17470210802373720.

Scene perception and memory revealed by eye movements and receiver-operating characteristic analyses: Does a cultural difference truly exist?

Kris Evans, University of Massachusetts, Amherst, MA, USA

Caren M. Rotello, University of Massachusetts, Amherst, MA, USA

Xingshan Li, and Vanderbilt University, Nashville, TN, USA

Keith Rayner

University of California, San Diego, San Diego, CA, USA

Abstract

Cultural differences have been observed in scene perception and memory: Chinese participants purportedly attend to the background information more than did American participants. We investigated the influence of culture by recording eye movements during scene perception and while participants made recognition memory judgements. Real-world pictures with a focal object on a background were shown to both American and Chinese participants while their eye movements were recorded. Later, memory for the focal object in each scene was tested, and the relationship between the focal object (studied, new) and the background context (studied, new) was manipulated. Receiver-operating characteristic (ROC) curves show that both sensitivity and response bias were changed when objects were tested in new contexts. However, neither the decrease in accuracy nor the response bias shift differed with culture. The eye movement patterns were also similar across cultural groups. Both groups made longer and more fixations on the focal objects than on the contexts. The similarity of eye movement patterns and recognition memory behaviour suggests that both Americans and Chinese use the same strategies in scene perception and memory.

Keywords

Eye-movements; Memory; Scene; Culture

Several recent studies have focused on how cultural differences influence scene perception and recognition memory. In a typical experiment, American and Asian participants study photographs of a focal object in a scene. In this encoding phase, they are usually asked to rate how much they liked each picture. After a brief delay, recognition of the focal objects is tested. The relationship between the focal objects and the background contexts is manipulated so that participants see an old or a new object in the original or a new context (Chua, Boland, & Nisbett, 2005; Masuda & Nisbett, 2001).

Correspondence should be addressed to Kris Evans, Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA. E-mail: hsiaochu@psych.umass.edu.

The general argument is that, because individualism is strongly emphasized in American culture, Americans attend to the scenes in a more analytical way, and therefore object memory is not affected by a change in the contexts. In contrast, Asian cultures stress the importance of collectivism, and thus people attend to the scenes more holistically. As a result, Chinese are more likely to reject old objects when they appear in different contexts. Chua et al. (2005) monitored eye movements and showed that Americans fixated the focal objects earlier and for a longer time than did Chinese, and Americans looked at the background contexts less than did Chinese. This result led Chua et al. to conclude that cultural effects occurred at the level of perception, especially in the allocation of attention.

Despite this popular view of cultural influences on scene perception, the effect is not consistently observed in eye movement studies. Rayner, Li, Williams, Cave, and Well (2007b) compared the eye movements of native English speakers and native Chinese speakers across a variety of tasks, including reading, face processing, scene perception, and visual search. Except for reading (which differed from the other tasks), the results showed a strong consistency across different tasks for each individual. Moreover, the looking patterns of the Chinese did not differ from those of the Americans while viewing scenes: There were no differences in mean fixation duration or proportion of fixations on the objects. Rayner et al. concluded that there was little evidence of cultural differences on the eye movement patterns during scene perception in their study (see also Rayner, Castelhano, & Yang, 2008).

Boland, Chua, and Nisbett (2008) described three possible explanations for the inconsistency between the Chua et al. (2005) and Rayner et al. (2007b) studies. (a) The number of focal objects in a scene differed; there was only one object in Chua et al.'s study while there were sometimes multiple objects in Rayner et al.'s study. (b) The complexity of the pictures in the two studies might be different from each other. (c) The task differed across the two studies: Chua et al.'s participants rated how well they liked each picture, while Rayner et al.'s participants viewed the scenes in anticipation of a subsequent memory test. We eliminated these factors in the current study by using Chua et al.'s stimuli (plus additional pictures to increase statistical power for the recognition test) and by asking participants to rate how well they liked each picture.

Chua et al. (2005) also reported that Chinese participants showed weaker memory than American participants for focal objects tested in a new context. Their conclusion was based only on participants' hit rates (the probability of calling a studied object "old") in each condition. However, hit rates cannot be interpreted in the absence of the corresponding false-alarm rates (the probability of calling a new object "old"; Macmillan & Creelman, 2005), both of which vary with participants' willingness to say "old". We used *receiveroperating characteristic* (ROC) curves to independently assess the effects of culture on memory accuracy and response bias (Macmillan & Creelman, 2005), which can be seen at a glance. ROCs can be constructed from simple confidence ratings, as we describe later.

We also monitored participants' eye movements during the test phase. Eye movements were recorded only in the encoding phase by Chua et al. (2005). Even though Chua et al. proposed that cultural influences occurred early in perceptual processing, without recording eye movements in the test phase, one cannot fully exclude the possibility that culture also affects later processes, such as memory retrieval. Monitoring eye movements in both the encoding and the test phase allows the current study to thoroughly investigate the entire time course of cultural differences from perception to memory recognition.

EXPERIMENT

Method

Participants—A total of 22 native English speakers (referred to as Americans) and 22 native Chinese speakers (referred to as Chinese) participated in the experiment. Most were graduate students at the University of Massachusetts, Amherst. The mean age of the Americans was 25.6 years, and all of them were Caucasians born in the United States. The mean age of the Chinese was 26.9 years, and all of them were born either in mainland China or in Taiwan.1 Participants were paid \$10.00.

Apparatus—Eye movements were recorded via an EyeLink 1000 tracker. Only the right eye was recorded. Stimuli were presented on a 19-inch CRT monitor controlled by a Dell PC. The display resolution was set to 1024×768 pixels. A chin-rest was used to minimize head movements and to keep the viewing distance constant (57 cm from the monitor). Participants responded by pressing keys on a keyboard.

Materials—Materials and design closely resembled those of Chua et al. (2005). A total of 50 real-world pictures of scenes were used in the encoding phase, 36 of which were adopted from Chua et al.2 A total of 50 additional real-world pictures were also used in the test phase. All of the scenes consisted of a single focal object in a complex but distinctive context (see Figure 1); for all pictures, the focal object was created and pasted onto a background using PHOTOSHOP software (Adobe Systems, San Jose, CA).

The test phase had a 2 (focal objects: old vs. new) \times 2 (context: old vs. new) design, which yielded four conditions: 25 old objects in the original contexts, 25 old objects in new contexts, 25 new objects in old contexts, and 25 new objects in new contexts. The size (range $= 3-45\%$ of picture, mean $= 22\%$) and position of the old objects were held constant for encoding and test. For the eye movement analysis, we defined the objects by their exterior contours. The objects could appear in any plausible location in a scene, although almost half of them were centred in the image. All participants saw the same pictures in a different random sequence.

Procedure—In the encoding phase, participants were told to view a series of pictures one at a time and to judge the degree to which they liked each picture. Before each picture was presented, a small black circle (the fixation dot) appeared at the centre of the display, and participants were asked to fixate it. Then, a picture appeared on the screen, and participants could move their eyes to examine it. After 3 seconds, the picture disappeared, a grey display showing a rating scale appeared, and participants entered a number from 1 to 6 to indicate the degree to which they liked the picture (1 for don't like at all; 6 for like very much; these data were not analysed). Three practice trials were conducted to make sure that the participants understood the task. No memory test was mentioned. After the encoding phase, participants were brought to another room and performed an irrelevant task (judging how typical a series of line drawings were for a corresponding word set) for about 5 minutes.

Participants were then brought back to the encoding room for a memory task in the test phase. They were told to view the pictures again and were tested for their memory of the focal objects in the pictures. If they believed they had seen the focal object during the

 $¹$ The experimenter was a native Chinese speaker, and all of the Chinese participants were very much tied to their native culture. They</sup> all had completed their undergraduate studies in mainland China or Taiwan before coming to the US for their graduate studies. They spoke Chinese at home, and when they entered the lab for the present study they conversed with the experimenter in Chinese.
²Although the original versions of the test stimuli were lost in a computer crash, Hannah Chua p materials.

encoding phase, they were asked to make an "old" response. If not, they were asked to make a "new" response. The procedure in the test phase was the same as that in the encoding phase (with the scene presented for 3 s) except that participants made an old–new judgement with a confidence rating from 1 to 6 on the keyboard (1 = sure new; 2 = probably new; 3 = maybe new; $4 =$ maybe old; $5 =$ probably old; $6 =$ sure old). Participants were told to make the old–new response after the picture disappeared (at which point a grey display appeared with the confidence rating scale). Three practice test trials were presented, in which participants were asked to point out where the focal objects were before making their judgements, to make sure that they understood the task. Eye movements were recorded in both the encoding and the test phase.

After the test phase, participants filled out a demographic questionnaire that focused on their age and language history. They were asked what they noticed about the experimental pictures (all of them noticed that the same focal objects appeared in different contexts). None of them were aware the purpose of the experiment before the test phase started. They also had no difficulty in finding where the focal objects were in each picture except that a few of the participants were not certain where the focal object was in one picture from Chua et al. (2005), though all of them guessed correctly.

Results

Eye movements

Encoding phase—The duration of the initial fixation in every encoding and test trial was excluded from analysis because it was a response to the fixation dot rather than the scene. The number of fixations on the focal objects and the background contexts and average fixation durations are presented in Table 1, separately for the American and Chinese participants. The data were submitted to an analysis of variance (ANOVA) with cultural group as a between-participant factor and interest area (object vs. background) as a withinparticipant factor. For all participants, the average number of fixations on the focal object was greater than that on the context ($M = 5.28$ and 4.28, respectively), $F(1, 42) = 19.58$, $MSE = 1.13$, $p < .001$. Likewise, the average fixation duration on the focal object was longer than that on the context ($M = 248$ and 236 ms, respectively), $F(1, 42) = 11.85$, $MSE =$ 264.23, *p* < .001. More importantly, neither the main effect of cultural group nor the interaction between interest area and cultural group was significant for either measure $(F < 1)$ for both tests).

We analysed separately the first real fixation (the one following that on the fixation dot) in each trial to see whether there was an early effect of culture. The probability that that fixation was on the focal object was .79 for Americans and .73 for Chinese, *t*(42) = 2.37, *p* < .05. Although this difference suggests a cultural effect, a supplementary analysis indicated that it may be attributable to the fact that the Chinese participants tended to undershoot the target more often than the Americans. For the Americans, 12% of the first saccades were made in the direction of the object but fell short; for the Chinese, it was 15%. The two groups did not differ in the time it took to fixate the object for the first time in a trial $(M =$ 393 ms for Americans, 408 for Chinese), $t(42) = 0.544$, $p > 0.5$, suggesting that the Chinese participants quickly adjusted their fixations to land on the object. Therefore, we conclude that this small effect of culture is one of shorter fixations that are likely caused by an oculomotor tendency to undershoot the target, rather than a cognitive difference.

A final way that cultural differences might emerge is in the probability that participants refixate on the object. People who process scenes holistically may tend to look at the background context after they look at the object or vice versa, thus binding the object with the context. On the other hand, people who process scenes analytically may tend to refixate

the object rather than alternate between the object and background. We calculated these transitions as the probability that fixation $N + 1$ was on the focal object, given that fixation N was the first fixation on the object. No difference was found between Americans and Chinese participants ($M = .79$ and .75 respectively), $t(42) = 1.01$, $p > .1$, suggesting that most of the time, participants refixated the objects immediately after they first looked at them. Overall, the eye movement data were similar across American and Chinese participants. The results suggest that cultural differences did not modulate how eye movements were distributed on the focal objects and background contexts in the encoding phase (see also Rayner et al., 2007b).

Of the 50 pictures in the study phase, 36 were the original stimuli from Chua et al. (2005). To evaluate whether the specific scenes used by Chua et al. yielded the same pattern as the overall set of pictures, the same kind of analyses were performed only on those 36 pictures. The eye movement patterns turned out to be almost identical to those from all 50 pictures. Participants made more fixations on the focal objects than on the background contexts (*M* = 5.27 and 4.23, respectively), *F*(1, 42) = 21.06, *MSE* = 1.11, *p* < .001, and they looked at the focal objects longer than the background contexts ($M = 251$ and 236 ms, respectively), $F(1)$, 42) = 14.22, $MSE = 331.08$, $p < .001$. Neither the main effect of cultural group nor the interaction between interest area and cultural group was significant $(F < 1$ for both tests). An early difference of culture was found in the probability that the first real fixation (after the one on the fixation dot) was on the focal object (.81 for Americans and .73 for Chinese), $t(42) = 2.59$, $p < .05$. However, there was no difference in how quickly two groups fixated the object ($M = 385$ ms for Americans, 408 for Chinese); $t(42) = 0.84$, $p > .4$. As for the overall stimulus set, we interpret the first-fixation probability as reflecting Chinese participants' slightly shorter saccades (undershoots) than those of Americans, rather than as a difference in the processing strategy for the image. There was no effect of culture on the probability of refixating the object immediately after first fixating it (*M* = .79 for Americans and .75 for Chinese); $t(42) = 1.01$, $p > .3$. Because the results based on the 36 pictures used by Chua et al. were consistent with those of the entire set of 50 pictures, we conclude that no general cultural difference was found in the eye movements regardless of the stimulus set. Thus, we failed to replicate Chua et al. even with their own stimuli.

Test phase—The eye movement patterns in the test phase are also presented in Table 1. For all participants, most of the fixations landed on the focal object rather than on the background context ($M = 7.16$ vs. 1.53), $F(1, 42) = 1,111.28$, $MSE = 0.628$, $p < .001$, and the average fixation duration on the focal object was also longer than that on the background context (*M* = 283 vs. 232 ms), *F*(1, 42) = 141.87, *MSE* = 397.33, *p* < .001. However, neither the main effect of cultural group nor the interaction between interest area and cultural group was significant $(F_s < 1.18$ for both tests). The two cultural groups differ neither in the probability of the first fixation on the object (*M* = .88 for Americans, .86 for Chinese); *t*(42) $= 0.79, p < .437$, nor in how quickly they fixated the focal object ($M = 303$ for Americans, 279 ms for Chinese), $t(42) = 1.63$, $p > 0.1$. Thus, no early effect of culture was observed. The analysis of transition probabilities was uninformative because most of the fixations landed on the focal object. Overall, and consistent with the data from the encoding phase, the eye movement patterns of Americans and Chinese were similar to each other. Therefore no effect of culture was found with regard to the eye movements in the test phase.

Recognition memory

The recognition memory judgements were examined in two ways. First, we report the hit and false-alarm rates. Next, we use the confidence ratings to construct ROCs and evaluate memory accuracy as well as response bias.

Response proportions—Both American and Chinese participants distinguished studied and new objects; their hit rates were higher than their false-alarm rates in all conditions (see Table 2). Although our data show a larger effect of background status on the hit rate than that reported by Chua et al. (2005), we replicated their observation that the Chinese participants' hit rates to focal objects tested in novel contexts were particularly low (Chinese: *M* = .491; Americans: .558). However, their false-alarm rate was also slightly lower numerically, suggesting that their response bias may differ from that of the Americans (Chinese: $M = .158$; Americans: .165). To evaluate this possibility, we turn to the ROC data.

ROCs—ROC curves plot the hit rate against the false-alarm rate, as a function of response bias or decision confidence. The first (leftmost) point on the curve reflects the highest confidence responses ("sure old"), the next point adds slightly lower confidence responses to those ("sure old" + "probably old"), and so on. The final point on the ROC necessarily falls at (1, 1) when all responses have been cumulated. ROCs have a few convenient properties. First, curves that fall higher in the space reflect greater memory accuracy, because the hit rate (correct responses) is greater for any particular false-alarm rate (error responses). Second, points that fall on the same theoretical curve all reflect the same level of memory accuracy, but different response biases; points to the upper right along an ROC reflect more liberal biases because both the hit and false-alarm rates are increased relative to points that fall to the lower left along the same curve.

Figure 2A shows the ROCs for American participants when the objects were tested in either an old or a new background context, and Figure 2B shows the analogous ROCs for the Chinese participants. For each participant, we quantified memory accuracy by calculating A_g , a measure of the area under the ROC (Pollack & Hsieh, 1969). A_g equals 1 when performance is perfect and .5 when memory is at chance. The old-context ROCs fall higher in the space than the new-context ROCs for both groups, reflecting higher accuracy for items tested in old contexts (American: mean $A_g = .89$ vs. .78; Chinese: .85 vs. .74). This effect of background context status was statistically reliable, $F(1, 42) = 141.2$, $p < .001$, $MSE = 0.253$. The ROCs produced by American participants fall slightly higher in the space than do those for the Chinese participants, reflecting slightly more accurate memory overall, $F(1, 42) = 4.66$, $p < .05$, $MSE = 0.034$. Most importantly, the effect of background context status was similar for both groups: There was no interaction $(F < 1)$.

Participants' response biases are also visible in the location of the points along each ROC curve: Points to the upper right reflect a greater willingness to say "old" than do points to the lower left. Figure 2 clearly show that participants were more willing to call objects "old" when they were tested in an old context than when they were tested in a new background context. To quantify this effect, we used a standard measure of response bias,

$$
cz = -\frac{1}{2}(z(HitRate) - z(FalseAlarmRate))
$$

where *z* is the standard normal transformation. Positive values of *c* reflect conservative responding, 0 reflects unbiased responding, and negative values indicate more liberal response biases. Both groups of participants had a more liberal bias for objects tested in old rather than in new contexts (American: mean $c = -0.24$ vs. 0.45; Chinese: -0.07 vs. 0.55). The main effect of background status was significant, $F(1, 42) = 202.31$, $p < .001$, $MSE =$ 9.40, but neither the effect of cultural group, $F(1, 42) = 1.95$, $p > .15$, $MSE = 0.399$, nor the interaction $(F < 1)$ was reliable.

Discussion

The similarity between the eye movement patterns of the Americans and of the Chinese, as well as the similarity of their recognition memory data, suggests that cultural differences do not play a significant role in scene perception and memory. When viewing scenes, both Americans and Chinese looked at the focal objects more and longer than the background contexts. This pattern held true across the encoding phase and test phase. The eye movement patterns did not vary with cultural groups either in the entire stimulus set or in the subset of stimuli used in the Chua et al. study (2005). The results from the recognition memory task showed that memory accuracy for objects was reduced in both the American and the Chinese groups when objects were tested in the new background contexts compared to the old contexts. Both groups also showed a shift in response bias in that they were less willing to say "old" to objects tested in new background contexts. Furthermore, the magnitudes of sensitivity decrease, and response bias shift did not differ with culture. Our findings thus suggest that Americans and Chinese are sensitive to the contexts to a similar degree and that both groups use the same strategies in scene perception and memory. Using different types of materials, Chua, Chen, and Park (2006) also observed that American and Chinese participants did not differ in their ability to recognize the source of a test item.

In our study, the hit rate for object memory was lower when the old objects were tested in new background contexts than in old contexts; this was true for both Americans and Chinese. In contrast, only Chinese showed such a hit rate decrease in Chua et al. (2005). Our ROC analysis provides a more appropriate measure of memory accuracy; it revealed that there were no differences between the two groups in terms of sensitivity decrease and response bias shift. This leads us to suspect that American participants in Chua et al. may have adopted a more liberal criterion than Chinese participants in the test phase. Therefore the higher hit rate for Americans when old objects were tested in new contexts could mainly have resulted from response bias, rather than true sensitivity. This fact illustrates the importance of using appropriate measures in recognition memory, such as estimates from the ROC curves.

Our findings concerning fixations on focal objects and background contexts are consistent with Rayner et al. (2007b) even though the stimuli used in the two studies differed in complexity and in the number of focal objects. They also found that a much greater proportion of fixations landed on the focal objects than on the backgrounds. Despite this good agreement, our findings and those of Rayner et al. (2008; Rayner et al., 2007b) are inconsistent with Chua et al. (2005), who reported that both Americans and Chinese made more fixations on the contexts than on objects. Actually, the only finding in Chua et al. that was replicated by the current study is that both cultural groups looked at focal objects longer than background contexts. The only cultural effect in our data is the probability that the first fixation is on the object, which was higher for the American participants than the Chinese. Because the time it took participants to fixate the object did not differ with culture, and because Chinese participants showed a slightly greater tendency to undershoot the object on their first saccade, we do not believe that the first-fixation probability reflects cognitive factors. In any case, the first-fixation probability effect is inconsistent with Chua et al.'s results in which a cultural difference was observed later in time (after 420 ms of viewing). Thus more research needs to be done to investigate the early effect of culture. It is clear that the complexity of stimuli, the number of focal objects, and differences in tasks cannot account for the incongruence between our findings and Chua et al.'s. At best, we suspect that the Chua et al. results should be viewed with some caution.

With more and more studies on the relationship between culture and cognitive processing, it seems clear that cultural differences can be systematically observed in eye movements in

certain tasks. In reading, it has been well established that culture (or more precisely, the writing system) affects eye movements (see Rayner, Li, & Pollatsek, 2007a). However, the relationship between culture and other tasks, such as scene perception and memory, seems to be a bit more controversial. Certainly, in contrast to Chua et al. (2005), the current study suggests that cultural differences have little impact on how people view and remember scenes. However, this does not mean that there are not cultural influences on other modes of thought and cognitive processing. Indeed, we suspect that there are such differences (Nisbett, 2003; Nisbett & Miyamoto, 2006).

Acknowledgments

This research was supported by Grant HD26765 from the National Institute of Health to K.R. and by Grant MH60274 from the National Institute of Mental Health to C.M.R. Portions of the data were presented at the 2008 Visual Sciences Society meeting. We thank Martin Fischer and three reviewers for their helpful comments on an earlier draft.

References

- Boland, JE.; Chua, HF.; Nisbett, RE. How we see it: Culturally different eye movement patterns over visual scenes. In: Rayner, K.; Shen, D.; Bai, X.; Yan, G., editors. Cognitive and cultural influences on eye movements. Tianjin, China: Tianjin Peoples Press; 2008.
- Chua HF, Boland JE, Nisbett RE. Cultural variation in eye movements during scene perception. Proceedings of the National Academy of Sciences. 2005; 102:12629–12633.
- Chua HF, Chen W, Park DC. Source memory, aging and culture. Gerontology. 2006; 52:306–313. [PubMed: 16974102]
- Macmillan, NA.; Creelman, CD. Detection theory: A user's guide. 2. Mahwah, NJ: Lawrence Erlbaum Associates; 2005.
- Masuda T, Nisbett RE. Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. Journal of Personality and Social Psychology. 2001; 81:922–934. [PubMed: 11708567]
- Nisbett, RE. The geography of thought: How Asians and Westerners think differently … and why?. New York: The Free Press; 2003.
- Nisbett RE, Miyamoto Y. The influence of culture: Holistic vs. analytic perception. Trends in Cognitive Science. 2006; 9:467–473.
- Pollack I, Hsieh R. Sampling variability of the area under the ROC-curve and of *d' e*. Psychological Bulletin. 1969; 71:161–173.
- Rayner K, Castelhano MS, Yang J. Eye movements when looking at unusual/weird scenes: Are there cultural differences? 2008 Manuscript submitted for publication.
- Rayner K, Li X, Pollatsek A. Extending the E-Z Reader model of eye movement control to Chinese readers. Cognitive Science. 2007a; 31:1021–1034. [PubMed: 21635327]
- Rayner K, Li X, Williams CC, Cave KR, Well AD. Eye movements during information processing tasks: Individual differences and cultural effects. Vision Research. 2007b; 47:2714–2726. [PubMed: 17614113]

Figure 1.

Example scenes used in this study. The focal objects are the same in (A) and (B) but the contexts are different. The focal objects are different in (C) and (D) but the contexts are the same. To view this figure in colour, please see the online issue of the Journal.

Figure 2.

Recognition memory for the object, plotted with receiver-operating characteristic (ROC) curves for American (A) and Chinese participants (B) when objects were tested in an old background context and in a new context.

Table 1

Means for number of fixations and fixation durations on the objects and contexts in the encoding and test phase, separately for Americans and Chinese Means for number of fixations and fixation durations on the objects and contexts in the encoding and test phase, separately for Americans and Chinese
group

Note: Standard errors in parentheses.

Table 2

Means for hit rates and false-alarm rates for the object recognition memory in the old and new contexts

Note: Standard errors in parentheses.