How we see it: Culturally different eye movement patterns over visual scenes

Julie E. Boland, Hannah Faye Chua, & Richard E. Nisbett
University of Michigan

It is blatantly obvious that culture is a powerful predictor of variance in overt human behavior. The foods we eat and the manner in which they are prepared, the holidays that we celebrate, and the languages that we speak are all strongly influenced by our cultural background. Culturally influenced behaviors like these seem to arise from a combination of conscious choices and environmental exposure. As such, cross-cultural differences in human behavior may be fairly superficial, overlaid onto underlyingly similar cognitive and perceptual processes. Alternatively, it is possible that our cultural backgrounds impact our cognition and perception at a more fundamental level. In this chapter, we explore the possibility that culture impacts basic aspects of visual perception.

Are Cultural Differences Shallow or Deep?

The superficiality of cultural differences is an assumption that is common in cognitive science. Within the Chomskian tradition of linguistics, this assumption is made explicit: we all share the same language organ, endowed with the same Universal Grammar (e.g., Baker, 2003; Chomsky, 1965). The syntactic differences among the world's languages arise from different settings for a small number of grammatical parameters (e.g., whether the verb starts or ends the verb phrase). Within psycholinguistics, the assumption that we comprehend, produce, and learn a language using the same cognitive mechanisms--regardless of the particular language involved--allows us to generalize experimental findings beyond the language in which the experiments were conducted. Yet, one can maintain the Chomskian assumptions about Universal Grammar and still ask, does habitually using English (or Chinese) influence how we think about the world? Whatever cross-linguistic differences do exist (e.g., in how temporal or spatial relations are

expressed) could impact the way that we think within that domain. This is a version of the Sapir-Whorf hypothesis, which, in modified form, continues to capture the interest of cognitive scientists in domains such as color categorization (e.g., Kay & Regier, 2006; Roberson, Davies, & Davidoff, 2000), spatial reasoning (e.g., Majid, Bowerman, Kita, Haun, & Levinson, 2004) temporal comparisons (e.g., Boroditsky, 2001), and grammatical gender (e.g., Lucy, 1992; Sera, Elieff, Forbes, Burch, Rodriguez, & Dubois, 2002).

Of course, language is just one aspect of culture that might influence how we reason about the world, or even how we perceive it. There are numerous dimensions through which cultural differences might impact our thinking, including differences in social structure, educational practices, availability of resources, politeness norms, and so forth. Within the domains of cultural and social psychology, scholars have found that culture impacts a variety of high-level cognitive behaviors such as social judgments and verbal recall of visual scenes (e.g., Chua, Leu, & Nisbett, 2005; Kitayama, Duffy, Kawamura, & Larsen, 2003; Masuda & Nisbett, 2001; Morris & Peng, 1994). Most of this research has contrasted Americans with East Asians, leading to the conclusion that members of the two cultures view the world from quite different perspectives (Nisbett, Peng, Choi, & Norenzayan, 2001). For example, Chua, Leu, and Nisbett (2005) asked participants to recall everyday social events, including personal events and events concerning others. Regardless of the type of event, Americans recalled more events about the main character than East Asians did, and Americans were more likely than East Asians to mention the character's intentions.

Overall, this line of cross-cultural research has shown that Americans are inclined to focus their attention on central objects or entities, analyzing them in terms of their attributes, and categorizing objects in order to reason about them. The causal attributions of Americans tend to

focus on the object in question, ignoring the broader context. In contrast, East Asians are likely to consider the broader context of the situation, noticing relationships and changes. East Asians are more likely to group objects based on family resemblance rather than category membership, and the causal attributions of East Asians tend to emphasize the context rather than unique properties of the central object. However, at just which representational levels culture influences cognitive processing has been unclear: Do East Asians and Americans really perceive the world differently? Or have we just been trained to interpret the task demands of each experiment differently?

Evidence from Memory and Change-Detection Paradigms

Most of the observed cultural effects rely upon off-line self-report, so it is possible that Americans and East Asians interpret the tasks differently, rather than actually experiencing the world differently. To illustrate, consider Masuda and Nisbett (2001), who asked Japanese and Americans to view twenty-second underwater animations twice and then describe the animations from memory. In the animations, certain fish were considered focal on the basis of size, movement, etc. The animations also contained background fish, as well as other background objects, such as plants. During the recall phase, the Japanese participants mentioned background fish and objects 65% more than the American participants. Furthermore, after all ten videos were observed and tested, participants were given a recognition test in which they saw pictures of fish and were asked whether they had seen that fish in any of the ten videos. Although both Americans and Japanese participants correctly recognized focal fish against the original background, or against a white background, Americans were better at recognizing the fish if it was presented against a novel underwater background. Both the Japanese tendency to report more background fish and their difficulty recognizing a familiar fish on a novel background

could be explained by differences in encoding the pictures: i.e., Japanese participants (more so than the Americans) encoded visual features from the background together with properties of the focal fish, such that properties of the fish could not be recalled without also recalling details about the context. Alternatively, Japanese and Americans might have encoded the same visual features from the scene in the same way, but had different response biases, based on cultural differences in what was deemed to be important and relevant. On this account, the Americans didn't mention the background fish because they didn't consider them important, while the Japanese failed to say that they recognized the familiar fish because the different background was considered relevant in their decision.

In another experiment, Masuda and Nisbett (2006) used a change blindness paradigm, in which two images alternated for up to one minute (or until the change was detected). The images were photographs of complex scenes, such as an airport. The difference between the two images was either a change to a focal object or a change to a background object. Each image appeared for 560 ms, with an 80 ms blank interval in between. East Asian detected changes to the focal object and background objects equally quickly, but Americans were faster to detect focal changes than background changes. In contrast to the memory study, this change blindness paradigm seems fairly on-line, but even here, it is not clear exactly how the East Asian and American participants differed. Did they represent the image differently, with the East Asians encoding more background details? Did they adopt different viewing strategies, with the Americans deciding to focus their attention on the focal object to a greater degree? Were the East Asians more adept at making mental comparisons between the current stimulus and their stored representation of the previous stimulus? Did the Americans have a response bias, such that they felt it was more important to report focal changes than background changes?

At the cognitively “shallow” end of the explanatory continuum, the cultural differences might reflect flexible strategies, capable of adapting to the perceived demands of the situation. In this case, the cultural differences reflect culturally different norms about what is deemed to be relevant and/or important in a given situation. At the cognitively “deep” end of the continuum, early enculturation experiences might lead to automatic and inflexible differences in the mechanisms supporting stimulus perception. For example, in the visual domain, East Asians and Americans might show stereotypically different fixation patterns when presented with the same visual scene. Because material close to the fixation point is known to be encoded in greater detail (e.g., Rayner & Bertera, 1979), this would lead to cultural differences in the amount of visual detail encoded for different regions of the scene. Another possible account, falling in between the “shallow” and “deep” endpoints, assumes that stimuli are perceived in the same manner cross-culturally, but that the representations of the information are structured differently, such that different details are most readily available for retrieval.

While it is not clear whether the observed cultural differences arise from differences in strategies, perception, memory, or something else entirely, the behavioral differences are quite stable. Across a range of tasks, Americans seem to emphasize central objects, representing background information only sparsely, while East Asians seem to bind their representations of focal objects together with rich representations of the background. In the current paper, we summarize recent evidence that the differences in overt behavior reflect cognitively deep differences in how information about the stimulus is perceived, rather than cognitively shallow differences in response patterns.

At this point, it is useful to review some well-accepted facts about scene perception. Within 100 ms of first viewing a scene, people can often encode the gist, e.g., ‘‘picnic’’ or

‘‘building’’ (Potter, 1976). Our mental representation of a scene, even after ample viewing time, is a woefully incomplete rendering of the original scene (e.g., Intraub, 1997; Potter, Staub, Rado, & O’Connor, 2002; Rensink, 2002; Simons & Rensink, 2005). Although the initial eye fixation on a scene is often not related to its contents (but rather where the eye happens to land), the following fixations are on the most informative regions of the scene for the task at hand (Henderson & Hollingworth, 1999). As noted above, fixation positions are important because foveated regions (i.e., the center of the fixated area) are encoded in greater detail than peripheral regions (e.g., Rayner & Bertera, 1979).

Can Eye-tracking resolve the “Shallow vs. Deep” Question?

In a recent experiment (Chua, Boland, & Nisbett, 2005), we asked graduate students at the University of Michigan to look at a series of pictures with a single focal animal or object and a richly detailed background. Two examples are in Figure 1. About half the students were raised in China, and about half were Americans of European descent. The Chinese and American students were drawn from a similar range of academic fields and all received the experimental instructions in English. We wanted a task that would be culturally neutral, so we asked the students to rate the pictures in terms of how much they liked each one. We tracked their eye movements for three seconds while they looked at the picture and verbally reported a rating. After the participants had seen and rated each of the pictures, they were given an object recognition test, to confirm that we replicated the primary finding from Masuda and Nisbett (2001).

Insert Figure 1 about here

As reported in Masuda and Nisbett (2001), the Chinese participants were less likely than American participants to recognize one of the original focal objects if it was presented against a different background. We were happy to replicate the previous study, but were most interested in the pattern of eye movements during the initial exposure to the stimuli. We categorized each fixation as an object fixation or a background fixation, based on predefined segmentation of the images. Then, for each participant, we calculated several dependent variables: the target of the first saccade, the latency of the first look to the object, the number of saccades ending in a background fixation over a three second interval, the mean fixation duration for both object fixations and background fixations over the three second interval, and the probability of an object or background fixation across the three second interval. Because people have quite limited awareness of their own fixation patterns, we assumed that these dependent measures capture largely unconscious, fairly automatic behavior. We also kept track of two overt judgments: how much participants reported liking the images (the Chinese participants reported liking them more), and how familiar participants rated the images to be in a post-test (Americans rated the pictures as more familiar, although we had taken pains to use primarily culturally neutral images, with a few specifically Chinese and specifically American scenes, in equal numbers).

Perhaps surprisingly, the eye movement patterns revealed cross-cultural differences in nearly every dependent measure we considered. Before the image came on the screen, we had participants look at a central fixation cross. So the first measurable fixation after the image appeared was usually in the center of the screen, where the fixation cross had been. Because our focal objects tended to be off center, this fixation was on the background portion of the scene more often than not. But on the very first saccade, Americans were more likely than Chinese participants to shift their gaze to the focal object. Because of this tendency to immediately orient

visual attention toward the focal object, Americans had shorter latencies for the initial fixation on the object than the Chinese participants. The Chinese made more background fixations than the Americans, but because the Chinese fixations were generally shorter than the Americans, the Chinese did not spend significantly more time fixating the background than the Americans. Compared to the Chinese participants, the Americans had substantially longer fixations on objects than on backgrounds.

When we examined the pattern of object and background fixations across the three second interval, we found that differences in the proportion of object and background fixations for American and Chinese participants began to emerge about 420 ms after stimulus onset. The period during which Americans were more likely to be looking at the focal object than the background spanned 420 to 1100 ms after stimulus onset. There was no temporal interval in which the Chinese participants were more likely to fixate the focal object than the background.

In sum, we did indeed observe cultural differences in the fixation patterns of Chinese and American participants, suggesting that we literally see the world differently. Our findings are consistent with a cognitively deep account of the much-reported cross-cultural differences in American and East Asian behavior (Chua, Nisbett, & Leu, 2005; Miyamoto, Nisbett, & Masuda, 2006; Nisbett, 2003; Nisbett & Masuda, 2003; Nisbett & Miyamoto, 2006). To be clear, we are not suggesting that there are cross-cultural differences in object recognition (i.e., recognizing the central object as a cow) or color perception, or the way in which visual information travels through the visual pathways in the brain. We assume that all of those basic vision processes are universal. The point is that the same stimulus is attended differently by people from different cultures. On one hand, this is not so surprising, because it is well-known that visual attention is influenced by high-level cognition, even if we are not consciously aware of how we are fixating

a stimulus; in brief, we fixate regions of the stimulus which are deemed most relevant given the current task demands (Henderson & Hollingworth, 1999). On the other hand, differences in fixation patterns have deep implications for cognition, because only the foveated areas of a visual scene are represented in rich detail. We can't retrieve or reason about details that we don't encode in the first place.

Challenges to Chua, Boland & Nisbett (2005)

The finding that Chinese participants make more and shorter fixations than the American's has now been replicated by Rayner, Li, Williams, Cave, and Well (to appear, see also Li, Williams, Cave, Well, and Rayner, this volume), who found that relative fixation durations are consistent across different visual tasks. Rayner et al. compared American and Chinese participants (similar to our participant groups) as well as English-dominant participants with some knowledge of Chinese. However, Rayner et al. reported mixed results with respect to cultural differences in looks to focal objects versus looks to the background. For the 18 scenes in which they were able to define (multiple) focal objects, they report no such cultural difference. For the six scenes[1] that they deemed most like those of Chua et al. (2005), they report a cross-cultural pattern that is similar to Chua et al. We will first outline our hypotheses about why they failed to replicate all of our effects in the 18 scene analysis.

Although the overt experimental task differed in the two studies (Rayner et al. (to appear) told participants that they would be given a memory test on the scenes), we doubt that the task difference was responsible for the diverging findings. While it is plausible that the memory instructions would have caused participants to view the pictures differently than they would have under our instructions, we assume that cultural differences in what is deemed relevant would

remain, regardless of a main effect of task. Rather, we believe that differences in our findings are due to differences in the visual stimuli and in how the spatial regions of interest were defined.

In our study, the scenes always had exactly one focal object and its corresponding region of interest followed the contours of the object itself, to minimize the amount of background information that was included in the focal region of interest. In contrast, the 18 scenes that Rayner et al. analyzed had from one to more than 30 objects that they deemed to be foreground and treated as equivalent to our focal objects. The regions of interest (which never numbered more than six) could contain any number of objects and were always rectangular; thus their foreground regions sometimes contained considerable information that we would have defined as part of the background. For example, there is a kitchen scene in which a little boy is standing at a counter. We would have defined just the boy as the focal object, but Rayner et al. included (as separate foreground regions of interest) the overhead light, and various objects that were on the counter or stove. Another scene, set in an art museum, has just one region of interest, but it includes about 20 people and 15 statues or friezes. The only part of the image not included as foreground is the ceiling and the blank wall above the artwork.

In short, Rayner et al.'s (to appear) images are simply not comparable with ours. The total area of their foreground regions tended to be larger than our focal object region, contain many more individual objects, and included portions of the region that we would have defined as background. Correspondingly, they report a dramatically higher proportion of focal region fixations than Chua et al. (2005), for both Americans and Chinese. On the first fixation (Figure 4 of Rayner et al.), both the Chinese and American participants were fixated within a focal region about 80% of the time, while they fixated the background region only about 20% of the time.

[1] For this analysis, Rayner et al. redefined some of their foreground regions as background, in order to make their definition of foreground more like our definition of focal object. Thus, the six scene analysis is not simply a subset

Looks to the focal objects increased on the next fixation for both groups, then fell and leveled off at around 70% looks to the foreground (and 30% looks to the background) for each group. In stark contrast, the first fixation in our study (Chua et al., Figure 4)[2] was on the focal object only about 30% of the time, for both Chinese and American participants, and our American participants spent roughly equal amounts of time on the focal object and background across the 3-second interval, while our Chinese participants spent less time on the foreground than the background.

Clearly, differences in our images and how we defined the regions of interest caused this dramatic difference in the relative proportion of time spent looking at the background. We suspect that these differences are at the heart of Rayner et al.'s (to appear) inability to replicate all of the cultural differences that we reported in Chua et al. (2005). When considering the 18 scenes used in their primary analysis, there seems to have been a difference in the information content of the backgrounds between the two experiments. In designing the Chua et al. stimuli, we took pains to ensure that the backgrounds were sufficiently interesting to attract a fair amount of attention. This was motivated by our theoretical assumption that Chinese, more so than Americans, gather details from the context that inform their interpretation of focal objects and actions. In other words, we would not predict a cultural difference in looks to the focal object versus looks to the background if the background was a blank wall, or a uniform texture. While the Rayner et al. backgrounds were not featureless, most of the visually interesting features of the images, including rocks, mountains, and distant buildings, were counted as foreground. Regardless of our intuitions about the images, the relatively few background fixations reported

---

of the data from the 18 scene analysis.

[2]The percentages from the two figures are only roughly equivalent, because Chua et al. (2005) graphed the data by time from picture onset, whereas Rayner et al. (to appear) graphed the data in terms of fixation number. As a result,

by Rayner et al. raises the concern that there was not sufficient content in the background to elicit a culturally different pattern of fixations. Support for this assertion comes from their analysis of six scenes similar to ours, which revealed similar gaze patterns to those reported in Chua et al.

We cannot dismiss the possibility that the Chua et al. (2005) eye movement results are limited to scenes with a single focal object. While this might seem to be an undesirable constraint, we designed our stimuli in this way because scenes with a single focal object allow for the clearest predictions for cultural differences, as motivated by the theoretical views outlined in Nisbett (2003). When there are multiple focal objects, each object plays a dual role; a given object may be considered focal to the scene, but it is also part of the context for another focal object. To illustrate, consider Li et al.'s (this volume) example of their scene with three bears near a mountaintop. In the 18 scene analysis, the foreground regions included all three bears and the mountaintop. In the six scene analysis, only the bears were considered focal. Nonetheless, one of the most important aspects of the context for each bear is that it is in the company of two other bears. In a culture, such as Chinese, that stresses relationships over individuals, the perceived relationships (e.g., familial, aggressive, etc.) among the bears may be the most important aspect of the picture. On the other hand, Americans would tend to look at the three bears for the same reason that they looked at a single bear in our study—because the bears are the focal subject matter of the scene. In such cases, Nisbett's theory would not predict any cultural differences in the proportion of looks to the bears versus looks to the mountaintop, sky, and grass.

---

the Chua et al. values sum to less than 1 (because at a given timepoint, some viewers were saccading rather than fixating), while Rayner et al.'s values sum to 1.

Some insight into the generalizability of our results can be gleaned from Rayner et al.'s (to appear) six scene analysis. The scenes had from one to three focal objects: a distant silo on the horizon, an elderly man and woman sitting across from each other at a table, a boy on a dock, the rear view of a car parked in front of a house, the three bears described above, and two rams on a snowy mountain. For five of the scenes, regions considered as foreground in the primary analysis were redefined as background for this analysis. Thus, these scenes all had richly detailed, interesting backgrounds and participants divided their gaze fairly evenly between foreground and background elements, as in Chua et al. (2005). Nonetheless, three of the six scenes had more than one focal animal/person, potentially diluting the cultural differences. We are not surprised therefore, that the predicted interaction was nonsignificant (though they report that their Chinese participants spent significantly less time fixating on the focal objects and made significantly more fixations on the background, compared to the Americans.

In sum, the cross-cultural difference that Rayner et al. (to appear) most clearly replicated from Chua et al. (2005), was that Chinese participants made shorter, more numerous fixations than Americans. Rayner et al. seem to view this as an effect of experience (perhaps language experience) rather than an effect of culture, but it is not clear to us how (or why) one could separate experience from culture. Unfortunately, the research described in Li et al. (this volume) and Rayner et al. was not designed to test cross-cultural theories of cognition. Their failure to replicate some of our findings is neutral with respect to Nisbett's theory. Rather, the primary contribution of their very ambitious study is in providing novel insights concerning the stability of eye movements over various tasks. We see this as an important finding, especially given the cross-cultural sampling of their participants.

Open Questions and Related Research

Of course, the Chua, Boland, and Nisbett (2005) study leaves many intriguing questions unanswered. What is it about Western and Eastern cultures that gives rise to the differences we observed? As noted at the outset, culture encompasses an amalgam of inter-related phenomena, and merely noting that there is a cross-cultural difference in some behavior is far from offering an explanation about why the differences arise, or how the differences might be generalized to other cultural groups.

One of us has written extensively about differences in eastern and western ways of thinking, and has proposed a general theory about why East Asians tend to behave more holistically in cross-cultural experiments, whereas Americans tend to be more analytic (e.g., Nisbett, Peng, Choi, & Norenzayan, 2001; Nisbett, 2003). The essence of the account is that East Asians depend, to a large degree, on social networks, and that attention to context is crucial for functioning effectively within these complex social structures (Markus & Kitayama, 1991). Westerners, on the other hand, function more independently, in less constraining social environments.

Nisbett's (2003) theory provides an account at a societal level, but it does not explain how a generic brain gets taught to exhibit the Chinese fixation pattern or the American fixation pattern. Our thoughts on this matter are quite speculative, but we do believe that the differences must arise from very early childhood experiences, prior to schooling. Childrearing practices are clearly influenced by societal differences, and reported differences in child-directed language could contribute to the fixation differences that Chua, Boland, and Nisbett (2005) observed. For example, Fernald and Morikawa (1993) found that American mothers tend to label objects for very young children, while Japanese mothers use more verbs, and use objects to engage very

young children in social routines. Tardif, Gelman, and Xu (1999) found similar East/West differences when comparing English speaking mothers and Mandarin Chinese speaking mothers. Tardif and her colleagues also reported that the predominance of nouns in the early vocabularies of American children (and many other western cultures) is not mirrored in the early vocabularies of Chinese children, which contain a larger proportion of predicates. Crucially, predicates often specify relations among objects. If western children are encouraged from an early age to focus their attention narrowly on objects, while East Asian children are encouraged to understand the relations among different entities from an early age, this might be an important factor leading to the differences in orienting and sustained visual attention that we observed in the eye movement study.

Finally, we confess to having done little since our paper was published to explore how the Chua, Boland, and Nisbett (2005) data generalize to eye movements for other cultural groups. (Though we have found that relatively interdependent southern Italians group objects more thematically, and less categorically, than more independent northern Italians (Knight & Nisbett, in press)). If our speculations above have any merit, then an analysis of child-directed speech might lead to clear predictions about which cultural groups should exhibit the American fixation pattern and which cultural groups should exhibit the Chinese pattern. More generally, one could predict that cultures that rely heavily on complex social networks should exhibit the Chinese pattern, and cultures with a more individualistic emphasis should exhibit the American pattern.

There is some support for the more general prediction: Bailey, Chuah, Siebold, Bailey, and LaBianca (2007) recently replicated our American pattern of eye movements with both American and Caribbean participants, and replicated our Chinese pattern with both Korean and

Latin American participants. As in our study, there was a single focal object against a richly featured background that comprised the majority of the image, e.g., a koala in a forest. And as in our study, Americans divided their time about equally between the focal object and the background regions during the three-second viewing interval. Thus, the Bailey et al. experiment provides an excellent testing grounds for Nisbett's (2003) account of cross-cultural differences. The really interesting novel finding is that a western culture (i.e., the Latin Americans) patterned like the Asian groups that have featured most prominently in Nisbett's account. Because Hispanic culture is often viewed as more wholistic than other American cultures (Hofstede, 1980; Hampden-Turner & Trompenaars, 1993; Triandis, 1995; Triandis, Marin, Lisansky, & Betancourt, 1984), these new results suggest that social and familial organization may be the most important factors driving the early experiences that lead to the cultural differences in how we allocate our visual attention.

In sum, there is growing evidence that cultural differences impact cognition in surprisingly deep ways. A clear implication of the Chua et al. (2005) finding, and related research by Bailey et al. (2007), is that differences in socialization, experience, or expertise can impact even highly automated, unconscious behaviors, such as eye movements, expanding the domain in which cultural variation is understood to impact cognition and perception.

References

Bailey, K. G. D., Chuah, M., Siebold, L. C., Labianca, Ø. & Bailey, R. N. (2007, May). Cultural background affects eye movements during scene perception. Poster presented at the 79th Annual Meeting of the Midwestern Psychological Association, Chicago, IL.

Baker, M. C. *The Atoms of Language: The Mind's Hidden Rules of Grammar*. Oxford University Press, 2003.

Boroditsky, L. (2001). Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology*, **43**, 1-22.

Chomsky, N. *Aspects of the Theory of Syntax*. MIT Press, 1965.

Chua, H. F., Boland J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Science*, **102**, 12629-12633.

Chua, H. F., Nisbett, R. E., & Leu, J. (2005). Culture and diverging views of social events. *Personality and Social Psychology Bulletin*, **31**, 925-934.

Fernald, A. & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development.* **64**, 637-656.

Hampden-Turner, C., & Trompenaars, A. (1993). *The seven cultures of capitalism: Value systems for creating wealth in the United States, Japan, Germany, France, Britain, Sweden, and the Netherlands.* New York: Doubleday.

Henderson, J. H. Hollingworth, A. (1999). High level scene perception. *Annual Review of Psychology*, **50**, 243-271.

Hofstede, G. (1980). *Culture's Consequences.* Beverly Hills: Sage.

Intraub, H. (1997). The representation of visual scenes. *Trends in Cognitive Sciences*, **1**, 217-222.

Kay, P. & Regier, T. (2006). Language, thought and color: Recent developments. *Trends in Cognitive Sciences*, **10**, 51-54.

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J.T. (2003). Perceiving an object and its context in different cultures: A cultural new look. *Psychological Science*, **14**, 201-206.

Knight, N., & Nisbett, R. E. (in press). Culture, class and cognition: Evidence from Italy. *Cognition and Culture*.

Li, X., Williams, C. C., Cave, K. R., Well, A. D., & Rayner, K. (this volume). Eye movements, individual differences, and cultural effects.

Lucy J (1992) *Grammatical categories and Cognition: a Case Study of the Linguistic Relativity Hypothesis.* Cambridge,*UK:* Cambridge University Press.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, **8**, 108-114.

Markus, H. R. Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, **98**, 224-253.

Masuda, T. & Nisbett, R. E., (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, **81**, 922-934.

Masuda, T. & Nisbett, R. E. (2006). Culture and change blindness. *Cognitive Science*, **30**, 381-389.

Morris, M. W. & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology* **67**, 949-971.

Miyamoto, Y., Nisbett, R. E., & Masuda, T. (2006). Culture and physical environment: Holistic versus analytic perceptual affordances. *Psychological Science*, **17**, 113-119.

Nisbett, R. E. (2003). *The Geography of Thought: How Asians and Westerners Think Differently… And Why*. The Free Press, New York.

Nisbett, R.E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review*, **108**, 291–310.

Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Science*, **100**, 11163-11175.

Nisbett, R. E., & Miyamoto, Y. (2006). The influence of culture: Holistic vs. analytic perception. *Trends in Cognitive Sciences*, **9**, 467-473.

Potter, M. C. (1976). Short term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning*, **2**, 509-522.

Potter, M. C., Staub, A., Rado, J., & O'Connor, 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.

Rayner, K. & Bertera, J. H. (1979). Reading without a fovea. *Science*, **206**, 468-469.

Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Will, A. D. (to appear) Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research.*

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, **53**, 245-277.

Roberson, D. Davies, I. & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, **129**, 369-398.

Field Code Changed

Sera, M. D., Elieff, C. Forbes, J. Burch, M. C., Rodriguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: an analysis of grammatical gender and classification. *Journal of Experimental Psychology: General*, **131**, 377-397.

Simons, D.J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, **9**, 16-20.

Tardif, T. Gelman, S. A. & Xu, F.(1999). Putting the `noun bias ' in context: A comparison of English and Mandarin. *Child Development*, **70**, 620-635.

Triandis, H. C. (1995). *Individualism and collectivism.* Boulder: Westview Press.

Triandis, H. C., , H. (1984). Simpatia as a cultural script of Hispanics. *Journal of Personality and Social Psychology*, **47**, 1363-1375.

Figure Captions

Figure 1. Two example scenes from Chua, Boland, and Nisbett (2005) illustrate the range in size and placement of our focal objects and the salience of the background objects. In the train-yard picture, the train in the foreground was considered the focal object. In the bird/bridge picture, only the bird was considered the focal object.

Figure 1.