The Effect of Culture on Perspective Taking

Shali Wu and Boaz Keysar

The University of Chicago

ABSTRACT—People consider the mental states of other people to understand their actions. We evaluated whether such perspective taking is culture dependent. People in collectivistic cultures (e.g., China) are said to have interdependent selves, whereas people in individualistic cultures (e.g., the United States) are said to have independent selves. To evaluate the effect of culture, we asked Chinese and American pairs to play a communication game that required perspective taking. Eye-gaze measures demonstrated that the Chinese participants were more tuned into their partner’s perspective than were the American participants. Moreover, Americans often completely failed to take the perspective of their partner, whereas Chinese almost never did. We conclude that cultural patterns of interdependence focus attention on the other, causing Chinese to be better perspective takers than Americans. Although members of both cultures are able to distinguish between their perspective and another person’s perspective, cultural patterns afford Chinese the effective use of this ability to interpret other people’s actions.

Perspective taking is fundamental to social interaction (Decety & Sommerville, 2003; Mead, 1934; Saxe & Kanwisher, 2003). Actions are ambiguous, so people evaluate other people’s beliefs, goals, and intentions in order to interpret their actions. Consideration of mental states is crucial in both competitive and cooperative activities. In competitive settings such as economic “games” (Camerer, 2003), and in cooperative activities such as coordination “games” (Schelling, 1960), one attempts to evaluate another person’s mental state in order to predict his or her future actions. One’s theory of mind provides the ability to infer other individuals’ mental states, to consider their perspective, and thereby to interpret and predict their actions (e.g., Gopnik & Wellman, 1992; Wellman, 1990). We evaluated whether differences between cultures induce systematic differences in the way people consider the other’s perspective during actual interactions.

In principle, considering the other person’s mental states is so important for social action that the human brain might have solved the problem universally, independently of culture. The evidence regarding cognitive development supports this idea. Young children confound their private knowledge with the knowledge of others, failing to understand that others can have a false belief (e.g., Astington, Harris, & Olson, 1983; Dennett, 1987; Perner, Leekam, & Wimmer, 1987). Only after age 4 do children distinguish their knowledge from that of other people (e.g., Perner, 1991; Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). This developmental trajectory is the same across countries and cultures. For instance, Sabbagh, Xu, Carlson, Moses, and Lee (2006) showed that Chinese and American children are the same age when they develop an understanding that other people can have a false belief—despite the fact that Chinese children develop executive functions earlier, which could allow them to inhibit their self-knowledge better, and perhaps distinguish it from other people’s knowledge more effectively. The development of theory of mind does not seem to depend on schooling or literacy. Even children in an isolated, preliterate hunter-gatherer culture show the same trajectory for the appreciation of the other’s mind as American children do (Avis & Harris, 1991). So people’s endowed ability for perspective taking seems universal.

INDEPENDENCE AND INTERDEPENDENCE

Though perspective-taking ability may be universal, the use of this ability to interpret other people’s actions may not be. We investigated the effect of culture on the way people take perspective by comparing people from China and the United States. East Asian culture is often characterized as collectivistic, as opposed to Western culture, which is often characterized as individualistic (e.g., Triandis, 1995; Triandis, Bontempo, Villareal, Asai, & Lucca, 1988). In general, members of collectivistic cultures tend to be interdependent and to have
self-concepts that are defined in terms of relationships and social obligations. In contrast, members of individualistic cultures tend to strive for independence and to have self-concepts that are defined in terms of their own aspirations and achievements (see also Shweder & Bourne, 1984). Markus and Kitayama (1991) described the consequences of this cultural difference to the concept of self. For instance, it suggests that the representation of self is more prominent than the representation of others for Westerners, but that the representation of others is more prominent than the representation of self for East Asians. A study consistent with this idea showed that Americans evaluate the similarity of others to themselves as higher than the similarity of themselves to others (Holyoak & Gordon, 1983). This asymmetry does not hold for Japanese, presumably because the other is more prominent than the self for Japanese (Kitayama, Markus, Tummala, Kurokawa, & Kato, 1990, as cited by Markus & Kitayama, 1991). Members of these two cultures, then, seem to have a fundamentally different focus in social situations.

The strong difference in focus between an independent self and an interdependent self is also reflected in self-descriptions (Brewer & Gardner, 1996). In fact, language can trigger a culture-bound representation of self. Ross, Xun, and Wilson (2002) found that bicultural Chinese-born individuals tended to describe themselves in terms of their own attributes when writing in English, but to describe themselves in relation to other people when writing in Chinese. Self-perception, then, seems to be affected by cultural patterns of independence or interdependence.

A cultural difference in focus on the self or on the other also suggests a difference in memory and perspective. Cohen and Gunz (2002; Cohen, Hoshino-Browne, & Leung, in press) argued that a focus on the self leads Westerners to adopt an insider’s perspective, but that a focus on the other leads Asians to adopt an outsider’s perspective. They showed that when people were thinking about an event in which they were at the center of attention, Chinese were likely to report the event from a third-person perspective, and Americans were likely to report the event from a first-person perspective. When primed with an emotional memory, Americans tended to project that emotion to an abstract other, whereas Chinese projected the reaction to that emotion to an abstract other. These results clearly show the effect of this cultural difference in focus on how people remember and perceive events.

Like other researchers, we do not assume a categorical distinction between East Asians and Westerners, but only assume that East Asians’ self-representations are more interdependent than Westerners’, and that Westerners’ self-representations are more independent than East Asians’. Of course, individuals can also be more or less interdependent in different situations (Triandis, 1995). In the present study, our goal was to investigate whether the interdependent-self/independent-self cultural difference systematically affects how people interpret other people’s actions. We did this by comparing the performance of Chinese and Americans in a task that required them to distinguish their own knowledge from that of another person.

**CULTURE AND PERSPECTIVE TAKING**

Our focus was on how people use their knowledge about others’ beliefs when they interpret actions. It is possible that this problem is solved in a universal fashion, independently of culture. But if culture does have a systematic impact on perspective taking and its use in interpreting actions, culture could affect perspective taking in two opposing ways. We use the terms **representational hypothesis** and **attentional hypothesis** to refer to these two possibilities.

**The Representational Hypothesis**

Compared with people with independent selves, people with interdependent selves may be more likely to confound their own knowledge and that of another person. It is known that people tend to incorporate the representation of a close other, but not that of a stranger, into their representation of the self (Aron & Aron, 1986; Aron, Aron, Tudor, & Nelson, 1991). Consequently, people make more egocentric errors in reasoning about their friends than in reasoning about strangers. Similar to friends who are interdependent, members of an East Asian, interdependent culture may be more likely to confound their own perspective with that of the other than are members of a Western, independent-selves culture. This hypothesis predicts that Chinese would be worse perspective takers than Americans, behaving more egocentrically.

**The Attentional Hypothesis**

According to this hypothesis, interdependence might focus one’s attention on others and away from the self. Indeed, Markus and Kitayama (1991) explicitly rejected the idea that interdependence involves merging of self and other. Instead, they argued, because the self is defined in relation to others, the role of others becomes more important, inducing a tendency to focus one’s attention on others’ actions, knowledge, and needs. This hypothesis predicts that given their culture of interdependence, Chinese would be better perspective takers than Americans, behaving less egocentrically.

**The Present Study**

To distinguish between the two hypotheses, we used a game involving actual interaction between two individuals. In this game, a person’s successful interpretation of the other person’s actions depends on distinguishing what each person knows (Keysar, Barr, Balin, & Brauner, 2000; Keysar, Lin, & Barr, 2003). A “director” instructs a subject to move certain objects. They sit opposite each other, at a table with objects placed in a grid (see Fig. 1). The director’s role is to say where each object should go, and the subject’s role is to move the objects. The
director’s and subject’s perspectives differ because some objects are occluded from the director’s perspective, preventing the director from seeing them. Crucially, the subject knows that he or she will not be asked to move those objects.

The critical test is exemplified in Figure 1. The target object is the block in the second row, and the director says, “Move the block one slot up.” But this array includes a competitor block visible only to the subject. Given that the subject knows that the director cannot see the second block, this competitor should not affect their understanding. But if the subject does not fully separate the two perspectives, the competitor will confuse the subject—perhaps temporarily, perhaps completely. In this situation, the two hypotheses make opposite predictions. According to the representational hypothesis, Chinese will merge the two perspectives and therefore will show more confusion than Americans. In contrast, according to the attentional hypothesis, Chinese will pay closer attention to the other than Americans do; hence, they will be able to focus on the other’s perspective and will show less confusion. We evaluated these predictions using eye movement measures, as well as behavioral measures.

**METHOD**

**Subjects**

Twenty Chinese, native speakers of Mandarin, and 20 non-Asian Americans, native speakers of American English, participated in the experiment. All subjects were University of Chicago students. The Chinese subjects were born and raised in mainland China and had been in the United States from 2 to 9 months. To minimize the confounding of culture with other variables, we matched the Chinese and the American subjects by age (M = 22 years for both groups), gender (half males, half females), year in school, and major of study. For simplicity, we use the term Americans from here on to refer to non-Asian Americans.

**Procedure**

The American subjects played the game in English with a female director who was a native English speaker, and the Chinese played the game in Mandarin with a female director who was a native Mandarin speaker. We made sure the instructions in English and in Mandarin were comparable by translating from English to Mandarin and back to English, as is standard procedure with cross-cultural research (Brislin, 1970).

In order to keep the critical instructions consistent across subjects, we used confederate directors. The directors were trained to behave just as a regular subject would, and provided their instructions in a natural, conversational manner. It is important to note that the subjects believed that the director was a naive subject.

The experiment started with two practice grids. So the subject would clearly understand the role of the director, the two players switched roles for the second practice grid. In addition, several “cues” were included to convince the subject that the director was a real subject. For example, the confederate director made some errors during practice and feigned unfamiliarity with some objects (e.g., by saying, “What is this called?”) during the experiment. In each round, the experimenter placed a grid between the two players and gave the director a picture showing the desired final state of the objects. The picture was taken from the perspective of the director, so it showed the occluded slots as blocked. Then, the director instructed the subject to move objects around in the grid so that the final arrangement corresponded to the picture. To maintain uniformity across subjects, we scripted the critical instructions to move the target objects, but the instructions for all the other objects were unscripted, so as to maintain naturalness.

The setting allowed for a fairly natural interaction, as subjects could talk whenever they wanted and move as they pleased. The only restriction was that prior to the instructions, the director said, “Ready?” and subjects had to fixate their gaze on the center point of the grid. As soon as the director started providing the instructions, subjects were allowed to move freely.

**Materials**

The experiment included 10 different target objects, which appeared in five different grids. Each grid included two target objects. One target object appeared with an occluded competitor, and the other had no occluded competitor. Thus, for each subject, 5 of the target objects had a competitor, and 5 did not. Which items had a competitor was counterbalanced across subjects. In addition, to make sure that the competitors were not systematically better referents than the targets, we switched each target object and its competitor for half the subjects. Each grid included four occluded slots, but their location varied across grids. The grids were presented in a random order.

**Equipment**

We used an SMI (Berlin, Germany) iView X head-mounted eye-tracking system to follow the subjects’ eye movements. The gear was mounted on a lightweight helmet and was relatively unobtrusive. An eye camera recorded the movement of the eye with respect to the head, and a magnetic sensor provided information related to the head position.
about head movement with respect to the world. Together, this information determined eye fixation on objects. A scene camera recorded the array, and a gaze cursor indicating the computed gaze position was overlaid on the image of the scene. Overlays were recorded as MPEG videos at a temporal resolution of 30 Hz, and a computer running SMI software digitally stored the real-value coordinates of gaze at a rate of 60 Hz. A microphone placed near the director recorded her instructions into the MPEG videos. Videos were filmed from the subject’s point of view, and both the director and the grid were visible in the videos.

Coding and Measures
To evaluate confusion, we considered both eye-tracking measures and behavior. Eye gaze is a sensitive measure of comprehension, indicating what object the subject is considering even before the subject acts. We used the following two measures: (a) the number of fixations on the competitor object (to evaluate the extent to which the subject considered the competitor as a potential target) and (b) the latency of the last fixation on the target before reaching toward it (to evaluate the extent to which the presence of the competitor interfered with the subject’s ability to identify the target). We defined a window of observation starting at the first sound of the identifying term (e.g., “b” in “block”) and ending with the selection of the target, defined as the subject reaching for it. Within this window, we coded for eye fixations. To count as a fixation, the gaze had to remain in the same slot of the grid for at least 100 consecutive milliseconds. To evaluate the effect of the competitor on eye gaze, we compared the eye data when the competitor was present with the eye data in the baseline condition, when the competitor was replaced by an unrelated object.

A Chinese-English bilingual and a native American-English speaker, both undergraduates at the University of Chicago, coded the digital video data files. They were both blind to the hypotheses of the experiment. For latency, we used the median latency for each subject in each condition in order to avoid skewing the data with unusually long reaction times.

The eye-tracking measures were able to reveal any temporary confusion, indexed by a delay in finding the target. We also examined cases in which the confusion was not resolved at all by considering the tendency of subjects to ask for clarification. For example, if they asked, “Which block?” it was probably because they thought the director could have had either block in mind. This indicated that they did not distinguish between the target block, which was visible to the director, and the competitor block, which was visible only to them. In addition, whenever the subject moved the competitor, he or she was not taking into account the fact that the director could not see it. We counted both the clarification questions and movement of the competitor as failure to consider the mental state of the director.

This game allowed us to evaluate perspective taking by looking at the extent to which the competitor object confused the subject. The representational hypothesis predicted that the presence of the competitor would confuse the Chinese much more than the Americans because Chinese people’s representation of the other is confounded with their representation of the self. The attentional hypothesis predicted that because a collectivistic culture directs one’s attention to the other’s knowledge and perspective, the Chinese would be less confused by the competitor than the Americans would be. However, if people solve the perspective problem in a culture-independent way, then the performance of these two groups would not differ.

RESULTS
The results showed a substantial effect of culture, and overwhelmingly support the attentional hypothesis. Americans considered the occluded competitor much more than Chinese did. On average, they fixated on the competitor more than twice as often as they fixated on the neutral, baseline object ($M_s = 1.85$ vs. $0.80$), $t(19) = 5.54, p_{rep} = .99, d = 2.53$. In contrast, the Chinese subjects fixated on the competitor only slightly more than they fixated on the baseline object ($M_s = 0.86$ vs. $0.54$), but not significantly more, $t(19) = 1.30, p_{rep} = .71, d = 0.59$ (see Fig. 2, top panel). The fixation data showed a significant interaction between culture and presence of the competitor, $F(1,38) = 5.353, p_{rep} = .92, \eta^2 = .123$.

This tendency to consider the occluded competitor dramatically delayed the Americans’ selection of the target, as indicated

![Fig. 2. Mean number of fixations on the competitor and the baseline object (top panel) and mean latency of the final fixation on the target when the competitor was present and when it was absent (baseline; bottom panel). Results are shown separately for American and Chinese subjects.](Image)
by the latency of the final fixation on the target before reaching for it. It took Americans 3,799 ms, on average, to finally identify the correct target when the competitor was present, compared with 2,785 ms to identify the target when the competitor was replaced by the baseline object, $t(19) = 3.34, p_{rep} = .98, d = 1.53$. Thus, the competitor caused a 1,014-ms delay (see Fig. 2, bottom panel). Indeed, the great majority of American subjects (80%) showed this pattern of delay. In contrast, the competitor caused virtually no delay for the Chinese. It took them a mere 68 ms longer to identify the target in the presence of the competitor than in the presence of the baseline object ($M_s = 1,621$ and 1,553 ms, respectively), $t(19) = 0.65, p_{rep} = .49, d = 0.30$. The interaction between culture and presence of the competitor was significant, $F(1,38) = 18.04, p_{rep} > .99, \eta^2 = .322$. Clearly, the Chinese were much more attuned to the perspective of the other than were the Americans.

The Chinese subjects did not accomplish their superior perspective taking by reflecting more about the director's perspective. Such a reflective strategy would have slowed them down overall; instead, they were consistently faster than the Americans. Yet this result raises a potential confound: Given that IQ is positively correlated with performance speed, it is possible that our Chinese subjects performed better because they were smarter than our American subjects. But our data show that speed did not predict the extent of interference. An analysis of covariance that adjusted for overall differences in speed showed that the Chinese were indeed 1,336 ms faster than their American counterparts when the competitor was present, at every level of latency of the final fixation on the target in the baseline condition, $F(1,38) = 11.27, p_{rep} > .99, \eta^2 = .229$ (Fig. 3). Thus, controlling for speed in the baseline condition, the Chinese were faster than the Americans to detect the target object and were much less distracted when the competitor was present.

An alternative explanation might be that the Chinese subjects showed better perspective taking because they were in a novel environment, which required them to pay attention to their surroundings. If this explanation were correct, one would expect their superior performance to wane with time in the United States. To evaluate this possibility, we compared Chinese subjects who recently arrived (2–3 months) with those who were at the end of their first year of study (9 months). The two groups did not differ in number of fixations on the competitor or in latency to detect the target (all statistical tests nonsignificant).

Compelling evidence for the attentional hypothesis comes not only from cases in which perspective taking was temporarily delayed, but also from cases of complete failure to take the director's perspective. If a subject did not consider the perspective of the director at all, then the instructions to move, for example, “the block” would have been ambiguous because from the subject’s own perspective, there were two blocks. If a subject processed the instructions with no regard to the director’s knowledge or mental state, then he or she might have resolved the ambiguity by asking the director for clarification (e.g., “which block?”). Such clarification requests are clear evidence for failure to consider the director’s perspective. In addition, if subjects moved the competitor, they were clearly not considering the director’s perspective. Despite the obvious simplicity of the task, the majority of American subjects (65%) failed to consider the director’s perspective (i.e., asked for clarification or moved the competitor) at least once during the experiment. In contrast, only 1 Chinese subject asked for clarification—and only once ($p_{rep} > .99$, Fisher’s exact test). On average, Americans failed to consider the director’s perspective 24% of the time, whereas Chinese subjects were able to quickly identify the object the director had in mind without asking for clarification ($p_{rep} > .99$, Fisher’s exact test). This difference is particularly striking because the subjects had all the relevant information readily accessible to them. They did not need to ask “which block,” as it was clear that the director could see only one block. Although the Chinese quickly and effectively made use of the perspective information to solve the problem, the Americans had substantial difficulty with this task.

**DISCUSSION**

We found strong support for the attentional hypothesis. In interpreting the actions of the director, Chinese subjects were almost unaffected by potential competitors from their own perspective. In contrast to the Americans, who were delayed in...
finding the target, the Chinese showed no delay. Most important, the Chinese were almost never “egocentric” in the sense that they failed to distinguish the director’s perspective from their own. In stark contrast, the majority of Americans showed such failure at least once. We therefore demonstrated that cultural differences induce different patterns of perspective taking: Chinese culture, which emphasizes interdependence, focuses attention on other people, whereas American culture, which emphasizes independence, focuses attention on the self. Consequently, compared with Americans, Chinese are better at solving perspective-taking problems, make fewer errors in assessing the intentions of another person, and are less distracted by their own private perspective.

Our subjects had a very simple task: following instructions to move everyday objects, such as blocks. One would expect the human brain to process such simple instructions in a universal manner, but members of different cultures processed this information very differently. A culture that promotes self-focus leads people to look for what “block” means to them, and a culture that promotes other-focus leads people to look for what “block” means to the other.

There is no reason to suspect that Chinese and Americans have a different understanding of the role of mental states in people’s actions. In fact, the appreciation of the mind of the other, or theory of mind, has an identical developmental trajectory for Chinese and Americans. By 5 years of age, both can begin to use another person’s knowledge, distinguishing it from their own knowledge and showing appreciation for the role of another person’s knowledge in predicting what he or she will do (Sabbagh et al., 2006). On the surface, then, our results are strange because they might suggest that our American subjects had lost this ability by the time they reached adulthood. This is not what we mean to imply, however.

We make a distinction between having perspective-taking ability and using this ability (Keysar et al., 2003). Both Chinese and American children show clear ability to reflect upon the mental states of other people. But using this ability to spontaneously and unreflectively interpret the actions of another person is a different matter. It seems that culture has its effect here at the level of use, not ability. It takes prolonged exposure to cultural patterns that reinforce attention to the other to induce a mode of interpretation that is not egocentric. Apparently, the interdependence that pervades Chinese culture has its effect on members of the culture over time, taking advantage of the human ability to distinguish between the mind of the self and that of the other, and developing this ability to allow Chinese to unreflectively interpret the actions of another person from his or her perspective. Americans do not lose the ability to reflect on and reason about another person’s mental state. They can accurately judge that another person cannot see occluded objects. But years of exposure to a culture that values independence and does not promote other-orientation does not provide the tools to unreflectively interpret actions from the perspective of the other.

This caused our American subjects either to show disregard for the director’s perspective (“which block?”) or to take more time and effort overcoming their own perspective in order to understand what the director actually meant (see also Epley, Morewedge, & Keysar, 2004).

As Mead (1934) suggested, perspective taking is indeed crucial for any social interaction. People’s behavior is ambiguous because it can be motivated by a variety of underlying intentions. Therefore, the interpretation of another person’s actions depends on the ability to consider that person’s mental states. We have shown that unreflective perspective taking is very much a function of cultural patterns. Unreflective perspective taking is more natural for members of a culture that emphasizes interdependence than for members of a culture that emphasizes independence.

Acknowledgments—Funding for this study was provided by National Institutes of Health Grant R01 MH49685-06A1 to Boaz Keysar. We thank Clifton Emery, Linda Ginzel, Susan Goldin-Meadow, Shiri Lev-Ari, and Rick Shweder for helpful comments on the manuscript; Michael Stein for comments and advice on statistical analysis; and Travis Carter, Jennifer Flores, Erica Kees, Chen Yang, and Kenny Yu for technical help.

REFERENCES


