Journal of Personality and Social Psychology
2004, Vol. 87, No. 1, 57–65

0022-3514/04/$12.00 DOI: 10.1037/0022-3514.87.1.57

Is It Culture or Is It Language? Examination of Language Effects in Cross-Cultural Research on Categorization

Li-Jun Ji
Queen's University

Zhiyong Zhang
Beijing University

Richard E. Nisbett
University of Michigan

Differences in reasoning styles between Chinese and European Americans held even when controlling for the language of testing. Bilingual Chinese organized objects in a more relational and less categorical way than European Americans, whether tested in English or in Chinese. Thus, culture affects categorization independent of the testing language. Nevertheless, language affected some Chinese bilinguals' categorization. The responses of Chinese from the Mainland and Taiwan were more relational when tested in Chinese than when tested in English. Responses of Chinese from Hong Kong and Singapore were equally relational when tested in Chinese and in English. Age and context of learning English are discussed to explain the differential language effects among different Chinese groups. Theoretical and methodological implications are discussed.

Categorization is about organizing the world. Objects can be organized and classified together because they share taxonomic categories or because they share thematic relations. *Taxonomic* categorization (or category-based classification) is made on the basis of similarity of attributes, such as similarities in perceptual properties among objects, whereas *thematic* categorization (or relationship-based classification) is made on the basis of causal, spatial, and temporal relationships among objects (Markman & Hutchinson, 1984). For example, if people are given the triplet *seagull–squirrel–tree* and asked which two go together, the choice *seagull–squirrel* indicates a taxonomic categorization, whereas *squirrel–tree* suggests a thematic categorization. It has been found that American college students sort objects primarily on the basis of taxonomic categories (e.g., Smiley & Brown, 1979). Markman and Hutchinson (1984) argued that thematic, eventlike organizations are a natural way of making sense of the world: "It is children's attention to categorical relations and not their attention to thematic relations that changes most with development" (p. 4).

Culture and Cognition

The above findings in categorization research may not hold for everyone. Research has shown that cognition and reasoning styles differ across cultures. Nisbett and his colleagues have argued that East Asians, and Chinese in particular, reason in a holistic and relational way, whereas Westerners, in particular European Americans, reason in an analytic way (Nisbett, 2003; Nisbett, Peng, Choi, & Norenzayan, 2001). According to their research, East Asians' attention is oriented toward the field and to relationships between objects and events. In contrast, North Americans decontextualize an object from the field and attend to its properties so as to establish category membership in an attempt to understand and predict the object's behavior (Ji, Peng, & Nisbett, 2000).

Research has provided support for this view. For example, Ji, Peng, and Nisbett (2000) presented East Asians (including Chinese, Japanese, and Koreans) and European Americans with a rod-and-frame test (Witkin et al., 1954). They found that Asians' judgments about the verticality of a rod were influenced more by the position of the surrounding frame than were those of European Americans, suggesting that they are more field dependent or sensitive to context. Masuda and Nisbett (2001) presented moving fish on a computer screen. They found that Japanese participants, compared with Americans, recalled more details about the background and more details about relationships between the fish and the background as well as relationships among background elements. East Asians' sensitivity to context is also manifested in conceptual tasks. For example, it has been repeatedly demonstrated that Chinese and other East Asians tend to make more situational attributions for behaviors than do European Americans, who are more inclined to make dispositional attributions (e.g.,

Li-Jun Ji, Department of Psychology, Queen's University, Kingston, Ontario, Canada; Zhiyong Zhang, Department of Psychology, Beijing University, Beijing, China; Richard E. Nisbett, Department of Psychology, University of Michigan.

This article is based largely on a doctoral dissertation submitted by Li-Jun Ji under the guidance of Richard E. Nisbett to the Department of Psychology at the University of Michigan. The research was supported by National Science Foundation Grants SBR 9729103 and BCS 0132074, National Institute of Aging Grant AG15047, the Russell Sage Foundation, a John Simon Guggenheim fellowship to Richard E. Nisbett, and a research initiation grant from Queen's University to Li-Jun Ji. We thank Fong Chan, Emily Chan, Karen Chan, and Amanda Pffifner for their help in data collection.

Correspondence concerning this article should be addressed to Li-Jun Ji, Department of Psychology, Queen's University, Kingston, Ontario, K7L 3N6 Canada. E-mail: lijunji@psyc.queensu.ca

Choi & Nisbett, 1998; Choi, Nisbett, & Norenzayan, 1999; Morris & Peng, 1994).

Given the importance of relationships and contexts to Chinese people and the importance of the object and its properties to Americans, the two groups could be expected to organize the world differently. Chiu (1972) presented American and Chinese children with pictures of objects—for example, a man, a woman, and a child—and asked them to select two of the three objects in the set that were alike or went together. Chinese children tended to group the objects on the basis of shared contextual or functional relationships ("The mother takes care of the child"). In contrast, American children tended to group objects on the basis of shared perceptual features or taxonomic categories ("The man and the woman are both adults").

Culture, Language, and Cognition

Language and Cognition

According to Vygotsky (1962), language plays an essential role in cognitive development, at least from the time the child has attained a certain level of language competence. Language, first developed as a means of social communication, is later internalized and becomes a crucial tool in the shaping of cognitive processes relevant for the elaboration of the abstract symbolic system that will enable the child to organize thought. Recent studies conducted by Boroditsky (2001) have shown that different ways of talking about time in English (as if it were horizontal) and in Mandarin (as if it were vertical) correspond to differences in how English and Mandarin speakers think about time, for both online and long-term processing. For example, Mandarin speakers' judgment that March comes earlier than April was facilitated after they saw a vertical array of objects, whereas English speakers' judgment was facilitated after they saw a horizontal array.

Language serves as an organizer of knowledge (Hamers & Blanc, 2000), and there is reason to believe that aspects of language influence categorization. For example, Markman and Hutchinson (1984) showed that hearing a noun caused (American) children to shift their attention from thematic to taxonomic organization. They suggested that language may "play a direct role in making categorical relations a salient and highly structured mode of organization" (p. 25). Dunham and Dunham (1995) found that the use of certain relational–identity terms (i.e., a single spatial–relational term alone or in combination with a verb or a pronoun or both, such as *down* and *I go out*) at 2 years of age is an antecedent of the tendency to use a thematic strategy in categorization tasks at 3 years of age and that an early tendency to use nouns and adjectives at 2 years of age is an antecedent of the tendency to use a taxonomic strategy in categorization tasks at 3 years of age.

Linguistic Relativity Hypothesis

The relationship between culture, language, and cognition has been debated in anthropology, philosophy, linguistics, and psychology. The famous Sapir–Whorf linguistic relativity hypothesis has been the center of a heated debate. The main theme is that culture, through language, influences people's thinking. Whorf (1956) believed that linguistic patterns (such as grammars) in different languages have impact on people's habitual thinking. According to Whorf, the differences in linguistic structure between languages are reflected in habitual thought and habitual behavior. Certain properties of a given language affect the way people perceive and remember. Whorf also believed that culture and language are not separable.

Logan (1986) offered the provocative proposal that language can be used to account for cultural differences in reasoning styles. He argued that "learning how to read and write with the alphabet has brought us more than literacy and a model for classification. It has provided us with a conceptual framework for analysis and has restructured our perceptions of reality" (p. 18). The phonetic alphabet, Logan argued, is believed to have provided a ground for abstract, logical, and systematic (Western) thought, which helps to explain why science started in the West but not in the East, even though Chinese technology surpassed that of the West from ancient times until at least the 16th century. The absence of Western-style abstractions and classification schemes in Chinese culture is related to the differences in writing systems. The Chinese writing system is based on drawn, concrete characters and reflects itself throughout Chinese thought, discouraging the development of the abstract notions of codified law, abstract science, and deductive logic, which are prerequisite for the development of science. If Logan was correct, then the difference between the alphabetic English language and the nonalphabetic Chinese language would be an important factor producing differences between Americans and Chinese in reasoning styles, including categorization preferences.

Compound and Coordinate Bilinguals

Language is a medium for transmitting and internalizing culture. Culture and language are therefore embedded in each other. It is not easy to separate the two. However, we believe that using bilinguals as participants, we are able to substantially separate the two variables, that is, to study one while controlling for the other.

Ervin and Osgood (1954) suggested that there are two types of bilinguals: compound and coordinate. *Compound bilinguals* have one representation for a verbal label and its translation equivalent, whereas *coordinate bilinguals* have two distinct representations, one for each language. An individual who learned two languages as a child in the same context is more likely to have a single cognitive representation, whereas an individual who learned a second language in a context different from his first language is more likely to have separate representations for two translation equivalents. Lambert, Havelka, and Crosby (1958) showed that learning each language in separate contexts leads to more functional separation between the bilinguals' two codes. Age of acquiring a second language and context of acquisition often go together; for example, learning a second language at an early age often occurs in the same family context, whereas later learning often occurs in a school context distinct from a family context. Compared with their compound counterparts, coordinate bilinguals make more semantic distinctions between a word and its translation equivalent and have two relatively independent association networks for translation equivalents. Experiments with word association techniques demonstrate that compound bilinguals have a higher degree of interdependence in the organization of their two codes than coordinate bilinguals (e.g., Lambert & Rawlings, 1969).

There is an overlap (though not a complete one) between the compound–coordinate dimension and the age of acquisition.

Compound bilinguals more often learn native and second languages simultaneously, whereas coordinate bilinguals tend to learn two languages consecutively. Furthermore, because coordinate bilinguals more often than not learn their languages consecutively, their bilingualism is often not balanced, and they may be more proficient in their native language than in their second language.

There is neurological evidence supporting a distinction between the two types of bilinguals. Kim, Relkin, Lee, and Hirsch (1997) applied functional magnetic resonance imaging (fMRI) to determine the spatial relationship between native and second languages in the cortex and found that for late bilinguals, who acquired their second language in adulthood, the second language is spatially separated from the native language. However, for early bilinguals, who acquired their second language during early childhood, native and second languages are represented in common frontal cortical areas. Chee et al. (1999) used fMRI to examine proficient Singapore Chinese bilinguals who were exposed to both Chinese and English early in life (before age 6) and found that they use common neuroanatomical regions during the conceptual and syntactic processing of a written sentence, regardless of testing language.

In the present research, we hope to add some behavioral evidence for the distinctions between compound and coordinate bilinguals. In addition, we attempt to address a methodological issue in cross-cultural research.

Language Use in Cross-Cultural Psychology

In cross-cultural research, participants are often tested in their native language. The testing materials may be developed in one language (usually in English) and then translated into participants' native language. In this process, called *back-translation*, the effect of the testing language tends to be ignored or treated as random error. Differences found between participants from different countries, speaking different languages, are generally attributed to cultural backgrounds. An inherent problem with this practice, however, is that the language used in testing is confounded with cultural effects, and it is not clear whether any differences found between groups are due to differences in cultural beliefs, norms or values, or the language of testing. This article provides what we believe to be the first attempt to separate language effects from cultural effects in cognition. In this article, *language* refers to the language of testing, and *culture* is defined as shared values, beliefs, and norms among a group of people, who most often speak the same language and live in proximity to each other. Specifically, culture is operationalized as participants' ethnic cultural background.

There are two major reasons for the ubiquitous practice of back-translation. One is that researchers assume that the testing materials, even though in different languages, are equivalent. Another is that researchers believe that culture and language are interconnected, and it is almost impossible to separate the two. In fact, however, back-translation does not guarantee equivalence across two languages, because one word in Language A may correspond to multiple words with slightly different connotations in Language B. For example, *pride* in English can be translated into two different Chinese words, *jiao (1) ao (4)*[1] (which usually has negative connotations, though sometimes can be used positively, depending on context) or *zi (4) hao (2)* (which has positive connotations only). Either word in Chinese would be back-translated into English as *pride*, though they would elicit quite different responses from the native speakers.

Some researchers take an extreme stance on translation issues and believe that there is no way to verify that testing materials are translated properly and equivalently across language (e.g., Boroditsky, 2001). They suggest that the same language, whenever possible, should be used to assess cognitions and behaviors among different cultural groups. However, even the same words can invoke quite different thoughts from two people speaking the same language because of their different life experiences and cultural backgrounds. Thus, cross-cultural researchers are presented with a dilemma: What language should be used in testing?

Present Research

In the present research, we attempted to examine the roles culture and language play in cross-cultural research in general and in research on basic cognition in particular. To do so, we needed to focus on a cognitive task that is verbal, in order to allow language to play a role, and a task that is sensitive to cultural (American vs. Chinese in the present article) influences. In light of the distinctions between holistic and analytic reasoning, a cognitive task with broad implications that might distinguish Americans from Chinese would be one examining categorization or organization of objects or events. We presented participants with a triad categorization task, where sets of three words were presented, and participants were asked to select two out of the three that were most closely related. Such a task allowed us to examine participants' tendency to be categorical (taxonomic) or relational (thematic) in their reasoning.

We examined whether culture and language have relatively independent effects on reasoning by using bilingual participants and testing them in their two different languages. Even though it may be impossible to completely separate language from culture, a design with bilinguals (especially a within-participant design with bilinguals) allowed us to look at the effect of one while controlling for the other.

If cultural beliefs and norms are the primary driving forces underlying any differences we observe between two cultural groups, then the effects should remain regardless of the language used in testing. However, if language plays the key role for the differences, then we should observe a language effect among the bilinguals. That is, bilinguals should think differently using Language A than when using Language B. Given the differences between compound (early) and coordinate (late) bilinguals, we would expect a language effect primarily for coordinate bilinguals, who are supposed to have two different representational systems. For compound bilinguals, the two languages are presumed to share the same representational system, and therefore their thinking should reflect the impact of their cultural beliefs more than the impact of the testing language. Indeed, the testing language should have relatively little effect among compound bilinguals.

[1] The numbers refer to different tones in Chinese.

We recruited participants from Mainland China, Taiwan, Hong Kong, and Singapore who spoke both Chinese and English.[2] Even though they were all ethnic Chinese, they could be categorized into two subgroups on the basis of their experience with English and with Western cultures. In Hong Kong and Singapore, English is learned early, beginning in kindergarten, and is used frequently in daily communication. In Singapore, English is one of the official languages. In Hong Kong, 80% of the secondary schools use English instead of Chinese as the primary medium of instruction (Hong, Chiu, & Kung, 1997). The modal age of learning English is around 3 years in Hong Kong (C. Y. Chiu, personal communication, December 19, 2001). In addition, both Hong Kong and Singapore were British colonies for over 100 years and are more Westernized than Mainland China and Taiwan. In contrast, in Mainland China and Taiwan, English is typically not learned until after elementary school and is rarely used in communication outside the English classroom. These differences in cultural and language experience led us to believe that Hong Kong and Singapore Chinese would be more likely to be compound bilinguals and Mainland and Taiwan Chinese would be more likely to be coordinate bilinguals. Thus, we expected a stronger language effect among Mainland and Taiwan Chinese than among Hong Kong and Singapore Chinese.

Study 1

In Study 1, we examined whether there were any cultural differences in categorization independent of language and whether there were any language effects independent of culture.

Method

Participants. One hundred nineteen Chinese students at Beijing University and 43 European American and 131 Chinese students from Mainland China, Taiwan, Hong Kong, and Singapore at the University of Michigan were recruited. Table 1 shows the sample information for each group and for each language condition. European American participants received course credit. Chinese participants in the United States had been in the United States for less than 5 years. They were paid or received course credit, and Chinese participants in China were paid. Before the end of the study, Chinese participants in the United States were given a questionnaire measuring their language skills, such as Test of English as a Foreign Language (TOEFL) scores, how often they spoke English to their parents and friends, how often they spoke English in childhood and adolescence, and how often they read and watched movies and TV in English.

Table 1
Sample Information for Study 1

Nationality and language	Sample size	
	Male	Female
European American		
English	22	21
Mainland and Taiwan Chinese in the United States		
English	10	18
Chinese	9	8
Mainland Chinese in China (Beijing)		
English	33	27
Chinese	32	27
Hong Kong and Singapore Chinese in the United States		
English	13	21
Chinese	26	26

Note. Language data refer to the language of testing.

It is worth noting that all the Chinese bilinguals recruited in the United States were undergraduate or graduate students at the University of Michigan. In order to get into the university, they had to pass certain English tests, and they do all their schoolwork in English. Thus, there should be no doubt that they understood the simple task and the simple words used in Study 1. The Chinese participants in China were recruited from one of the top universities in China, to enter which all students had to do well on several subjects, including English. In addition, all these students should have been taking English classes every semester (and every week) since they got into junior high school. Given that both the instructions and the words in this study were very simple, these top Chinese university students should have had no difficulty understanding the task or the words.

Materials. We presented participants with sets of three words (in one of three random orders) and asked them to indicate which two of the three were most closely related and why.[3] We used very simple words in the task so that it was easy for bilingual Chinese to understand the task in English. There were 10 sets of test items and 10 sets of fillers. The three words in each testing set could be grouped on the basis of thematic relations, categorical relations, or neither. Participants' groupings were coded as relational if they suggested an object–context or subject–object relationship, such as *monkey* and *bananas*, *shampoo* and *hair*, or *conditioner* and *hair*. Groupings were coded as categorical if they suggested shared features or category memberships, for example, *monkey* and *panda* or *shampoo* and *conditioner*. Similarly, participants' explanations were coded as either relational (e.g., "Monkeys eat bananas") or categorical (e.g., "Monkeys and pandas are both animals"). Examples for filler items included *child–teenager–adult* and *Monday–Wednesday–Friday*.

Within each of the 10 testing sets, there were 3 possible ways for participants to select two items. In total, there were 30 possible ways of grouping, 14 of which were coded as relational (such as *policeman* and *uniform*, and *postman* and *uniform*) and 11 of which were coded as categorical (such as *policeman* and *postman*). Thus, the stimuli were biased toward relational grouping.

The Americans were tested in English, and the Chinese were tested in either English or Chinese, as randomly assigned. Ideally, it would be very informative if we could have recruited American bilinguals who could read and write in both English and Chinese, but that turned out to be an almost impossible task because of the great difficulty of finding such people, even on the campus of a large U.S. university.

Results

Cultural effect. The main dependent variable was the difference between frequency of relationship-based grouping and frequency of category-based grouping, with positive numbers indicating preference for relationships over categories and negative numbers indicating preference for categories over relationships. Figure 1 shows the comparisons among all the groups. The omnibus analysis of variance (ANOVA) shows a main effect of

[2] Mandarin is spoken in Mainland China, Taiwan, and Singapore, and Cantonese is spoken in Hong Kong. However, the written language is more or less the same, though the simplified version is used in Mainland China and Singapore whereas the traditional version is maintained in Taiwan and Hong Kong.

[3] The English word *related* is translated into Chinese as *guan (1) lian (2)*.

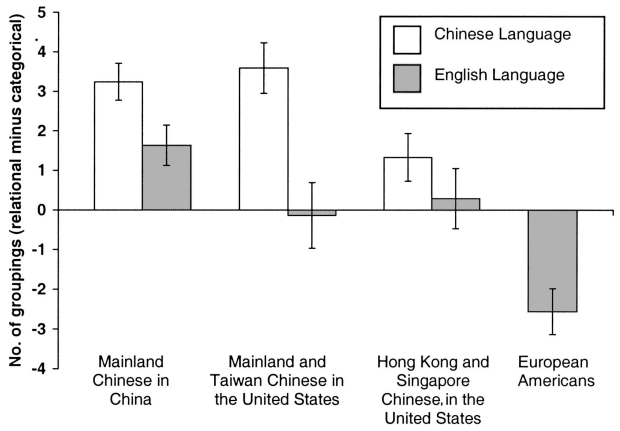

Figure 1. Relational versus categorical groupings by Mainland/Taiwan Chinese, Hong Kong/Singapore Chinese, and European Americans. Dependent variable is the frequency of relational groupings minus that of categorical groupings. Bars represent standard errors (Study 1).

group, $F(6, 286) = 10.92$, $p < .001$. European Americans showed a clear preference for categories, $t(42) = 4.39$, $p < .001$. None of the Chinese groups showed a preference for categories, and indeed, most of them showed a clear preference for relationships. For example, the Chinese in China showed a clear preference for relationships, whether they were tested in Chinese, $t(58) = 6.93$, $p < .001$, or in English, $t(59) = 3.20$, $p < .005$.

The European Americans differed from Mainland/Taiwan Chinese tested in English in the United States, $F(1, 69) = 6.05$, $p < .02$, and from Hong Kong/Singapore Chinese tested in English in the United States, $F(1, 75) = 9.24$, $p < .005$. Because these groups were all tested in English in the United States, the difference between European Americans and the Chinese in grouping preferences indicates a robust cultural effect independent of any effect having to do strictly with language or country–context effect.

When only Chinese bilinguals were compared on their grouping preferences, an ANOVA test on sample (Chinese in China [Beijing] vs. Chinese from Mainland and Taiwan tested in the United States vs. Chinese from Hong Kong/Singapore tested in the United States) and language of testing (Chinese vs. English) revealed a significant main effect of sample, $F(2, 244) = 4.01$, $p < .05$, and a significant effect of language, $F(1, 244) = 14.33$, $p < .001$. There was a trend for the interaction effect between sample and language, but it did not reach significance, $F(2, 244) = 1.65$, $p < .20$. A further analysis showed that Chinese in China were significantly more biased toward relationship than Hong Kong/Singapore Chinese in the United States, regardless of language of testing, $t(203) = 2.62$, $p = .01$. Mainland and Taiwan Chinese in the United States also showed a stronger preference for relationship over categories than Hong Kong/Singapore Chinese, but only when tested in Chinese, $t(46) = 2.58$, $p = .01$.

Language effect: Mainland and Taiwan Chinese. A significant language effect was found for Mainland/Taiwan Chinese in the United States, $F(1, 43) = 10.09$, $p < .005$, and for Chinese in China, $F(1, 117) = 5.36$, $p < .05$. Both groups showed stronger relationship preference when tested in Chinese than when tested in English. When tested in the United States in Chinese, Mainland and Taiwan Chinese showed a strong preference for relationships over categories, $t(16) = 5.64$, $p < .001$. However, when tested in English, they did not show any preference for relationship-based grouping, $t(27) < 1$. Thus, the Chinese tested in English shifted away from the relationship-focused Chinese pattern and moved closer to the American response pattern. These results indicate that for the bilinguals from Mainland China and Taiwan there is a language effect independent of culture. Similar language effects were found in the explanation data (see Figure 2) for Mainland/Taiwan Chinese in the United States, $F(1, 43) = 12.64$, $p < .001$, and for Chinese in China, $F(1, 117) = 5.36$, $p = .02$.[4]

Why does language of testing affect categorization? There are at least two explanations. One could be that structural differences in English and Chinese lead to different reasoning styles, such that certain features of the Chinese language make people think in a relational way whereas certain features of the English language make people think in a categorical way. The other explanation does not pertain to language per se. Instead, it is likely that the language used for a task makes certain ways of reasoning more accessible by activating representations that are common in a particular culture. An examination of the effects of the same two languages with different participants, presented in the following section, helps to address this issue.

No language effect: Hong Kong and Singapore Chinese. Data from Hong Kong and Singapore Chinese indicate that grouping preferences may have little to do with differences in language structures per se. As may be seen in Figure 1, the grouping preferences of Chinese from Hong Kong and Singapore were not affected by language, $F(1, 84) = 1.15$, *ns.* They showed a slight preference for relationship over categories in their groupings, $t(85) = 1.95$, $p < .06$. There was also no significant language effect for explanations (see Figure 2), $F(1, 84) = 2.39$, *ns*.[5]

[4] We expected no differences between Mainland Chinese and Taiwan Chinese in their groupings and explanations. Thus, they were recruited as one group, and their data were combined. When analyses were done including Mainland Chinese participants only, the findings remained the same. Their groupings were more relational when tested in Chinese ($M = 4.64$, $SD = 2.20$) than when tested in English ($M = -0.50$, $SD = 4.54$), $F(1, 23) = 11.83$, $p = .002$, and their explanations were also more relational when tested in Chinese ($M = 4.18$, $SD = 2.14$) than when tested in English ($M = -1.14$, $SD = 4.33$), $F(1, 23) = 13.88$, $p = .001$. Most of the 20 Taiwan Chinese were randomly assigned to the English condition. This does not allow for a meaningful comparison between the two language conditions, but the trends were the same as for Mainland participants.

[5] We expected no differences between Hong Kong Chinese and Singapore Chinese in their groupings and explanations. Thus, their data were combined. Similar patterns were obtained when analyses were conducted for Hong Kong Chinese data only and when conducted for Singapore Chinese data only. For Hong Kong Chinese, the language of testing had no impact on groupings (Ms = 1.31 and 0.50 for Chinese and English conditions, respectively), $F(1, 63) < 1$, or on their explanations (Ms = 1.10 and −0.15 for Chinese and English conditions, respectively), $F(1, 63) = 1.60$, $p > .20$. For Singapore Chinese, there were no language effects on groupings (Ms = 1.38 and −0.38 for Chinese and English conditions, respectively), $F(1, 19) < 1$, *ns*, or on their explanations (Ms = 1.62 and −0.50 for Chinese and English conditions, respectively), $F(1, 19) < 1$, *ns*.

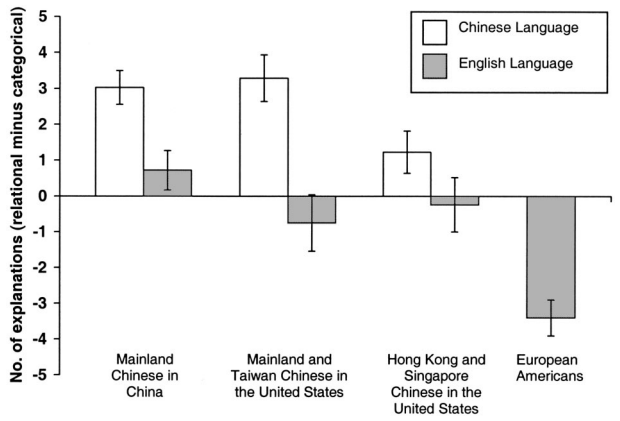

Figure 2. Reported explanations for groupings by Mainland/Taiwan Chinese, Hong Kong/Singapore Chinese, and European Americans (Study 1).

Effect of testing location. When Mainland Chinese (excluding Taiwan Chinese) tested in the United States were compared with Mainland Chinese tested in China, we found a significant effect of testing location. The language effect observed was much stronger for those in the United States (*M*s = 4.64 and −0.73 for Chinese and English conditions, respectively) than for those in China (*M*s = 3.24 and 1.63 for Chinese and English conditions, respectively), $F(1, 141) = 5.23$, $p < .03$, for the interaction effect of language and location. This suggests that being in the United States provided a stronger cue for Western reasoning styles, resulting in greater sensitivity to the language of testing by bilingual Chinese. In the meantime, consistent with our predictions, both groups demonstrated language effects, as we would expect from coordinate bilingual groups.

Language skills of bilingual Chinese in the United States. Among those Chinese participants who reported their TOEFL scores ($n = 101$), there was no correlation between TOEFL scores and participants' grouping preferences ($r = .04$, *ns*). In addition, grouping preference was found to be uncorrelated with any of the language experience questions we asked. Thus, it is likely that the language effect we observed was due to the language of the task and not to language proficiency. This is consistent with findings by Norenzayan, Smith, Kim, and Nisbett (2002), who observed no effect of English proficiency on East Asians' performance in a category learning task.

As to language differences between Mainland/Taiwan Chinese and Hong Kong/Singapore Chinese, we found that the two groups reported equivalent TOEFL scores (*M*s = 607 and 598; 600 on TOEFL is considered to be very high), though the latter group rated their own overall English skill better than did the former group on a 1 (*know little*) to 5 (*fluent*) point scale (*M*s = 3.53 vs. 4.10), $t(105) = 3.55$, $p = .001$. In addition, Hong Kong/Singapore Chinese reported higher frequency than did Mainland/Taiwan Chinese for speaking English with their parents; speaking English in childhood, adolescence, and adulthood; reading English magazines; and watching Chinese movies. The two groups did not differ in the frequencies for speaking English with friends, watching English movies, or reading Chinese magazines (as seen in Table 2).

In sum, we found marked language effects with Mainland and Taiwan Chinese bilinguals but literally no language effect for Hong Kong and Singapore Chinese. Apparently, this was not due to differences in language abilities, because these groups reported similar English test scores. This differential language effect may have to do with different cultural and language experiences that different Chinese groups have had, a possibility we discuss in the General Discussion section.

Study 2

There are alternative explanations for some of the findings in Study 1. First, it is possible that the Chinese participants in the United States were highly self-selected, and the group from Mainland and Taiwan who chose to go to the United States may be markedly different in some respect from the group from Hong Kong and Singapore who chose to go to the United States. Second, most of the Mainland and Taiwan participants recruited in the United States in Study 1 were graduate students, but most of the Hong Kong and Singapore participants were undergraduate students. Therefore, age and education differences may be confounded. To rule out these alternative explanations, college students in Hong Kong and in Mainland China were tested in both English and Chinese. Study 2 used a within-participant design so as to provide a particularly sensitive test of the effect of language.

Method

Participants. Fifty-nine (52 women) Hong Kong University students and 57 (29 women) Beijing University students were recruited. Each participant performed two grouping tasks, one in English and one in

Table 2
Language Skills of Bilingual Chinese in the United States (Study 1)

	Mainland and Taiwan Chinese		Hong Kong and Singapore Chinese	
Measure	*M*	*SD*	*M*	*SD*
TOEFL score	607	70	598	85
Overall English ability	3.5	0.7	4.1	0.8
Speak English to your parents	1.7	1.0	2.7	1.5
Use English in childhood and adolescence	2.3	1.1	3.6	1.3
Use English as an adult	3.8	1.1	4.3	0.9
Speak English with friends	3.6	1.4	3.8	1.2
Speak Chinese with friends	4.5	1.2	4.7	1.1
Read English magazines	3.6	1.2	4.1	1.2
Read Chinese newspapers and magazines	3.8	1.3	3.8	1.3
Watch TV and movies in English	4.4	1.0	4.4	1.1
Watch TV and movies in Chinese	3.0	1.1	4.1	1.3

Note. All the above questions (except the first two) were answered on a scale ranging from 1 (*never*) to 6 (*always*). Participants were asked to rate their overall English ability on a scale ranging from 1 (*know little*) to 5 (*fluent*). TOEFL = Test of English as a Foreign Language.

Chinese. The two tasks were given on 2 different days, with about 2 weeks in between the two tests. All participants were volunteers.[6]

Materials and procedure. In Study 1, participants had more opportunities to group objects on the basis of relationships than on the basis of categories within each set. For example, selecting *postman* and *uniform* or *policeman* and *uniform* would both be considered relational, whereas there was only one way to be categorical, that is, selecting *postman* and *policeman*. This resulted in a possible bias toward relational grouping overall. In Study 2, we presented more balanced sets so that within each set only one relational and one categorical grouping could be made. We designed two grouping tests in the same format as the test used in Study 1. In each test, there were 8 test items, such as *carrot–rabbit–eggplant* and *teacher–doctor–homework*, and 10 filler items. We counterbalanced the language used in testing, the test versions, and the testing order.

Results

For relationship preference over categories in grouping, repeated ANOVAs revealed a main effect of sample (Hong Kong Chinese vs. Mainland Chinese), $F(1, 101) = 6.67$, $p < .05$, indicating that Mainland Chinese showed greater relationship preferences than Hong Kong Chinese. There was also a main effect of language, $F(1, 101) = 14.16$, $p < .001$, suggesting that testing in Chinese led to greater relationship preferences than testing in English. In addition, the interaction effect between sample and language was significant, $F(1, 101) = 10.00$, $p = .002$, indicating a stronger language effect among coordinate bilinguals in Mainland China than among compound bilinguals in Hong Kong (as seen in Figure 3). Similar results were obtained for the explanation data. There were significant main effects for sample, $F(1, 101) = 8.68$, $p = .004$, and for language, $F(1, 101) = 12.86$, $p = .001$, and their interaction was significant as well, $F(1, 101) = 8.59$, $p = .004$.

Within-sample analyses were consistent with the results in Study 1. The Hong Kong Chinese showed a strong relationship preference in their groupings in both the Chinese condition, $t(53) = 3.58$, $p < .001$, and the English condition, $t(51) = 3.04$, $p < .005$. There was, however, no language effect for the Hong Kong Chinese, $t(45) < 1$. No language effect was found in their explanations, either (see Figure 4), $t(45) < 1$.

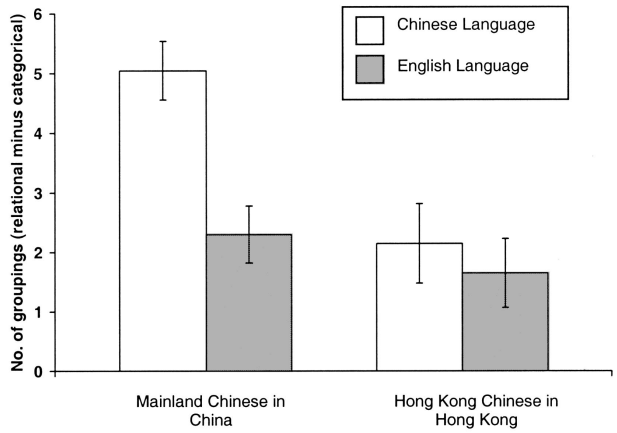

Figure 3. Relational versus categorical groupings by Chinese in Mainland China and in Hong Kong (Study 2).

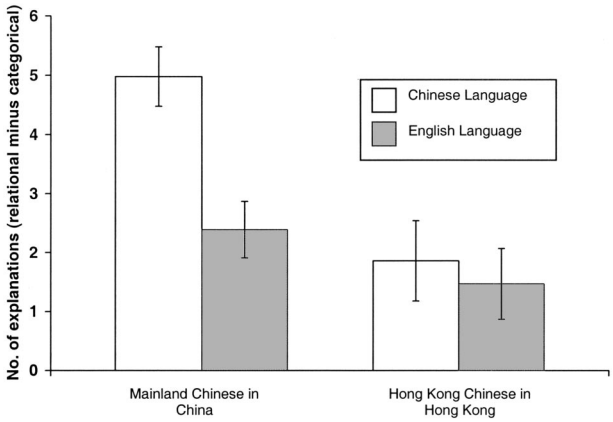

Figure 4. Reported explanations for groupings by Chinese in Mainland China and in Hong Kong (Study 2).

Consistent with the language effect found in Study 1, the Chinese in Mainland China showed a stronger relationship preference in the Chinese condition than in the English condition, $t(56) = 5.12$, $p < .001$. A similar language effect was found in their explanations, $t(56) = 4.88$, $p < .001$, as seen in Figure 4.

In summary, consistent with Study 1, there was no language effect for the compound bilinguals in Hong Kong on the grouping task, whereas a strong language effect was found for the coordinate bilinguals in Mainland China.

General Discussion

We found that culture had a substantial effect on the way participants grouped objects. Regardless of the language or location of testing, Chinese participants grouped objects more on the basis of relationships and less on the basis of category membership than did European Americans. The data suggest that European American organization of the world is based to a substantial degree on taxonomic categories, whereas Chinese organization is based more on relationships. These findings are consistent with the view that Westerners' reasoning is relatively analytic (including a tendency to focus on categories) and that Chinese reasoning is relatively holistic (including a tendency to focus on relationships). More important, our data suggest that the cultural differences between European Americans and Chinese are not an artifact of the testing language.

We found a substantial language effect among Mainland and Taiwan Chinese. When tested in Chinese, they grouped objects primarily on the basis of relationships, but when tested in English, their groupings were much less relationship based. This was true regardless of whether they were tested in the United States or in

[6] Marginally significant gender effects were found for Hong Kong/Singapore Chinese in the United States (Study 1) and Mainland Chinese in China (Studies 1 and 2), such that women showed a greater preference for relationships over categories than men. No significant interaction effect involving gender was found. Results for Study 2 were entirely similar whether all Mainland and all Hong Kong participants were compared or whether only the women (for whom there was an adequate sample in both groups) were compared.

China. However, there was no language effect observed with Hong Kong and Singapore Chinese, whether tested in the United States or in Hong Kong. The fact that bilingual Chinese from Hong Kong and Singapore did not show any language effect suggests that language effects on object grouping are not due to any inherent structural differences between Chinese and English. Otherwise, we should have seen a language effect for Hong Kong and Singapore Chinese.

Why Are the Different Chinese Groups Different?

We believe that the difference between Chinese groups has to do with the age at which people learn English as a second language and the context in which it is learned. On the basis of the available research on the distinctions between compound and coordinate bilinguals, it is reasonable to assume that Hong Kong and Singapore Chinese, who learn English at a much earlier age, are more likely to be compound bilinguals, whereas Mainland and Taiwan Chinese are more likely to be coordinate bilinguals. If so, for Hong Kong and Singapore Chinese, the two languages—Chinese and English—are presumed to share the same representational system, resulting in no effect of language of testing. In contrast, Mainland and Taiwan Chinese, being coordinate bilinguals, may have distinctive representations associated with English and Chinese. Therefore, using Chinese may prompt them to think in a relatively more Chinese fashion than using English.

The age of learning English may not be solely responsible for our findings. Learning English at an early age is also an indicator of the environment where the children grow up. Hong Kong and Singapore are more Westernized than Mainland China and Taiwan. A mixture of English and Chinese languages and a mixture of Chinese and Western ways become the reality that Hong Kong and Singapore Chinese live. Indeed, Hong Kong and Singapore participants showed less relationship preference than the Mainland and Taiwan participants when tested in Chinese (in both studies), suggesting that their early and common experiences with Western culture and language may have made them more Westernized. Our results also indicate that when different languages are integrated into a single system of cultural practices, as in the case of Chinese bilinguals in Hong Kong and Singapore, the language effect is minimal.

Theoretical Implications

Our findings have at least two important theoretical implications. First, people from different cultures tend to focus on different things when thinking about objects. In addition, the findings indicate that cultural differences in object grouping cannot be accounted for by differences in the language of testing. In other words, it is culture (independent of the testing language) that led to different grouping styles. But then, how exactly does culture shape the way people reason and organize the world? Researchers have found that Chinese people value interpersonal relationships and pay more attention to the social environment than do Americans (e.g., Fung, 1983; Hedden et al., 2000; Hsu, 1981; Ji, Schwarz, & Nisbett, 2000). This tendency to attend to the social environment and interpersonal relationships is reinforced in family and school and then may be carried over to all environments and relationships in general (Markus & Kitayama, 1991). In contrast, Americans value individual autonomy and freedom. Such values are reflected in schooling and home education even for infants. Both Bornstein and colleagues (1990) and Fernald and Morikawa (1993) found that, when playing with toys together with their children, the American mothers focused on the objects and their attributes ("It's a truck. See, it's got nice wheels"), whereas the Japanese mothers focused on social routines and relationships ("See, it's a vroom-vroom. I give it to you. You give it to me. Yes! Thank you.").

Second, language apparently plays a critical role in "tuning" categorizations among Mainland and Taiwan Chinese bilinguals. Our studies present behavioral evidence to support the distinctions between compound and coordinate bilinguals. For Mainland and Taiwan Chinese, the language used affected the way they reasoned, even when culture and testing location were controlled for, suggesting that different representations are associated with different languages and that language can serve as a cuing effect for reasoning style. If true, this indicates that reasoning styles are flexible, and can be modified through learning another language and another culture, as late as early adolescence.

Methodological Implications: Are We Misled by Different Languages Used in Cross-Cultural Testing?

Some psychologists are concerned that researchers might be misled by comparisons based on different testing languages. But are we? Our data suggest that for compound bilinguals or bilinguals who grew up in a mixed cultural environment, it does not matter which language is used in testing. For coordinate bilinguals or bilinguals who grew up in an environment in which one mode of thinking is dominant, however, it does matter. Different results could be obtained depending on the testing language. So are we misled by the grouping results when Americans were tested in English and Chinese were tested in Chinese? We might be, because, as seen in Figure 1, an enhanced cultural effect was found when Chinese participants were tested in Chinese rather than in English. So testing in native languages may make any cultural differences look larger than they are.

However, do we know which language condition represents the reality? It is quite possible that the English condition may have reduced the real cultural effect, which exists only, or primarily, when the native language is used in testing. If we presume that there is only one dominant way of thinking for each individual, then it may be reasonable to assume that the native language is more likely to elicit it than the second language. However, if we presume that learning a foreign language may not only bring a new way of thinking to the individual, but also modify their original way of thinking, then the issue becomes more complicated. In future research, investigations and comparisons should be made between bilinguals and monolinguals in order to find out whether learning a foreign language (such as English) might have changed the bilinguals' reasoning even in their native language (such as Chinese). If such an effect exists, to what extent does it vary depending on the age of learning a foreign language?

It is important to keep in mind that we have obtained cultural differences between European Americans and all of the Chinese groups, regardless of the language of testing. This suggests that even though language of testing may lead to changes in bilinguals' responses, such changes occur in a limited range and do not

necessarily threaten the conclusion pertaining to cultural comparisons.

A question to address in future research concerns the scope of the language impact. Is there a language effect also for memory, perception, logical inference, or decision making, at least among coordinate bilinguals? Other questions concern whether similar language effects could be observed for bilinguals who speak two languages that are not as distinct as English and Chinese, or bilinguals who are influenced by two cultures that are not as distinct as the U.S. and China. We believe that the language effects we have observed in coordinate bilinguals may only be obtained when the two languages are associated with two distinctively different thinking styles.

In conclusion, cultural backgrounds affect reasoning, independent of the testing language. In addition, the language of testing may also affect thinking, depending on when and how the language is learned.

References

Bornstein, M. H., Toda, S., Azuma, H., Tamis-LeMonda, C., & Ogino, M. (1990). Mother and infant activity and interaction in Japan and in the United States II: A comparative microanalysis of naturalistic exchanges focused on the organization of infant attention. *International Journal of Behavioral Developmental Psychology, 13,* 289–308.

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

Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W. L., Thiel, T., & Weekes, B. (1999). Processing of visually presented sentences in Mandarin and English studies with fMRI. *Neuron, 23,* 127–137.

Chiu, L. -H. (1972). A cross-cultural comparison of cognitive styles in Chinese and American children. *International Journal of Psychology, 7,* 235–242.

Choi, I., & Nisbett, R. E. (1998). Situational salience and cultural differences in the correspondence bias and actor-observer bias. *Personality and Social Psychology Bulletin, 24,* 949–960.

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin, 125,* 47–63.

Dunham, P., & Dunham, F. (1995). Developmental antecedents of taxonomic and thematic strategies at 3 years of age. *Developmental Psychology, 31,* 483–493.

Ervin, S. M., & Osgood, C. E. (1954). Second language learning and bilingualism. *Journal of Abnormal and Social Psychology, 49,* 139–146.

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

Fung, Y. (1983). *A history of Chinese philosophy*. Princeton, NJ: Princeton University Press.

Hamers, J. F., & Blanc, M. H. A. (2000). *Bilinguality and bilingualism* (2nd ed.). Cambridge, England: Cambridge University Press.

Hedden, T., Ji, L. J., Jing, Q., Jiao, S., Cui, Y., Nisbett, R. E., & Park, D. C. (2000, July). *Culture and age differences in recognition memory for social dimensions.* Poster session presented at the 16th Biennial Meeting of the International Society for the Study of Behavioral Development, Beijing, China.

Hong, Y., Chiu, C., & Kung, T. (1997). Bringing culture out in front: Effects of cultural meaning system activation on social cognition. In K. Leung, Y. Kashima, U. Kim, & S. Yamaguchi (Eds.), *Progress in Asian social psychology* (pp. 135–146). Singapore: Wiley.

Hsu, F. L. K. (1981). *Americans and Chinese: Passage to differences*. Honolulu: University of Hawaii Press.

Ji, L. J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology, 78,* 943–955.

Ji, L. J., Schwarz, N., & Nisbett, R. E. (2000). Culture, autobiographical memory, and social comparison: Measurement issues in cross-cultural studies. *Personality and Social Psychology Bulletin, 26,* 586–594.

Kim, K., Relkin, N., Lee, K., & Hirsch, J. (1997, July 10). Distinct cortical areas associated with native and second languages. *Nature, 388,* 171–174.

Lambert, W. E., Havelka, J., & Crosby, C. (1958). The influence of language-acquisition contexts on bilingualism. *Journal of Abnormal and Social Psychology, 56,* 239–244.

Lambert, W. E., & Rawlings, C. (1969). Bilingual processing of mixed-language associative networks. *Journal of Verbal Learning and Verbal Behavior, 8,* 604–609.

Logan, R. F. (1986). *The alphabet effect.* New York: Morrow.

Markman, E., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16,* 1–27.

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.

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.

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently . . . and why.* New York: Free Press.

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

Norenzayan, A., Smith, E. E., Kim, B., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science, 26,* 653–684.

Smiley, S. S., & Brown, A. L. (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology, 28,* 249–257.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Whorf, B. L. (1956). *Language, thought, and reality.* Cambridge, MA: Technology Press of MIT.

Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B., & Wapner, S. (1954). *Personality through perception.* New York: Harper.

Received September 2, 2003
Revision received February 23, 2004
Accepted February 26, 2004 ■