Aging, Culture, and Memory for Categorically Processed Information

Lixia Yang, Wenfeng Chen, Andy H. Ng, and Xiaolan Fu

1Department of Psychology, Ryerson University, Toronto, Ontario, Canada.  
2State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China.  
3Department of Psychology, York University, Toronto, Ontario, Canada.

Objectives. Literature on cross-cultural differences in cognition suggests that categorization, as an information processing and organization strategy, was more often used by Westerners than by East Asians, particularly for older adults. This study examines East-West cultural differences in memory for categorically processed items and sources in young and older Canadians and native Chinese with a conceptual source memory task (Experiment 1) and a reality monitoring task (Experiment 2).

Method. In Experiment 1, participants encoded photographic faces of their own ethnicity that were artificially categorized into GOOD or EVIL characters and then completed a source memory task in which they identified faces as old-GOOD, old-EVIL, or new. In Experiment 2, participants viewed a series of words, each followed either by a corresponding image (i.e., SEEN) or by a blank square within which they imagined an image for the word (i.e., IMAGINED). At test, they decided whether the test words were old-SEEN, old-IMAGINED, or new.

Results. In general, Canadians outperformed Chinese in memory for categorically processed information, an effect more pronounced for older than for young adults.

Discussion. Extensive exercise of culturally preferred categorization strategy differentially benefits Canadians and reduces their age group differences in memory for categorically processed information.

Key Words: Aging—Categorization—Culture—Memory—Reality monitoring—Source memory.

CATEGORIZATION is a process to organize and classify information. Cultural differences in categorization have been widely reported in literature. It has been demonstrated, with both children (Chiu, 1972) and young students (Ji, Zhang, & Nisbett, 2004), that Americans prefer to sort stimuli by shared features or category (i.e., taxonomic categorization), whereas Chinese are more likely to group stimuli by interdependence and relationships (i.e., thematic categorization). For example, when presented with a triplet chicken–cow–grass, Americans likely pair chicken and cow together because they are both animals, whereas Chinese likely group cow and grass together because a cow eats grass. This effect remains true even after controlling for linguistic differences between the two cultures (Ji et al., 2004).

Following this line of research, Unsworth, Sears, and Pexman (2005) reported that Americans were more likely to use categorical classification to sort two objects of a triplet and showed a semantic priming effect for categorical (i.e., chicken–cow) relative to relational (i.e., cow–grass) pairs. In contrast, Chinese were equally likely to engage categorical and relational classification and responded equally fast to categorical and relational pairs (Unsworth et al., 2005). Furthermore, Norenzayan, Smith, Kim, and Nisbett (2002) found that Americans outperformed Asians (i.e., Chinese and Koreans) in classifying novel animals into different categories based on a set of rules. In this situation, European Americans tended to use rule-based formal logic reasoning, whereas East Asians relied more on experience-based intuitive reasoning and thus tended to misclassify animals that were similar to the exemplars but did not satisfy the rules (Norenzayan et al., 2002). This finding supports the ancient philosophical views that Western cultures, in general, follow a Greek philosophy that intended to explain events with universal logic rules, whereas Chinese philosophers, especially Taoists, were more pragmatic and intuitive (Fung, 1952; Norenzayan et al., 2002; Russell, 1945).

Taken together, cross-cultural research generally suggests that people of Western cultures are more skillful at taxonomic categorization than those of East Asian cultures (Nisbett, 2003; Nisbett & Masuda, 2003). However, it is still unclear whether and how these differences may influence memory for information processed in such a way. To our knowledge, only one study has examined cultural differences in categorical clustering of free recall responses (Gutchess et al., 2006). In this study, participants encoded and then recalled a list of words in their native language. The free recall results showed that Chinese categorized less, as indexed by a lower categorical clustering, than did Americans. The cultural effect was only shown for older, but not young adults. This is presumably because categorical clustering is differentially more effortful for Chinese and thus age-related decline in cognitive resources...
limits Chinese older adults’ ability to categorize. In contrast, the prolonged experience utilizing the culturally preferred categorial organization makes it an automatically taken strategy for Western older adults.

In all the aforementioned studies, categorial organization was based on the relationship naturally built in stimuli (e.g., both items in the cow–chicken pair refer to animals by nature). It is unknown whether the cultural differences would appear for stimuli artificially sorted into distinct and mutually exclusive taxonomic categories. Source monitoring paradigm represents a useful manipulation to simulate artificial taxonomic categorization encoding given that most source memory tasks require encoding items into distinct categories/sources (e.g., two distinct speakers). However, Chua, Chen, and Park (2006) failed to detect any cultural differences between Chinese and Americans in source memory, as assessed with a task in which participants identified the correct speaker—out of four—who introduced specific statements during encoding. We note that the item–source pairs were arbitrarily formed, and different speakers were not necessarily mutually exclusive and thus were not as distinctive as taxonomic category cues in common sense. In this situation, categorization may not be an optimal strategy. In this study, we used meaningful and distinct category/source cues. In Experiment 1, we adopted a conceptual source memory task (Rahhal, May, & Hasher, 2002) in which participants encoded photographic faces by sorting them into GOOD or EVIL characters. In Experiment 2, we modified a reality monitoring paradigm (Gonsalves & Paller, 2000; Kensinger, O’Brien, Swanberg, Garoff-Eaton, & Schacter, 2007) in which participants viewed a series of words, each followed either by a corresponding image (i.e., SEEN) or by a blank square within which they needed to form a mental image for the word (i.e., IMAGINED).

We chose these two tasks for the following reasons. First, studies with Western samples showed that the performance on both tasks was relatively intact with aging (Hashtroudi, Johnson, & Chrosniak, 1989; Rahhal et al., 2002), suggesting that the performance of these tasks may rely heavily on knowledge or cultural experience (Baltes, 1993). Second, the tasks maximally encourage using categorization strategy in Westerners. Considering that the fundamental attribution error (i.e., tendency to attribute a social behavior to internal disposition instead of contextual factors) is more prevalent in Western than in East Asian cultures (Norenzayan & Nisbett, 2000; Ross, 1977), the GOOD–EVIL distinction may differentially encourage Westerners to categorize based on specific facial features. Considering that Westerners rely more on formal logical reasoning, whereas East Asians favor intuition (Norenzayan et al., 2002), the reality–imagination distinction in Experiment 2 will differentially benefit Westerners. Furthermore, guided by their culturally dominant naïve dialecticism, Chinese are inclined to expect changes (e.g., good/true could turn into bad/false) and tolerate contradictions (e.g., coexistence of good/true and bad/false; Spencer-Rodgers, Williams, & Peng, 2010) and thus they are less likely to use categorization to draw a clear-cut GOOD–EVIL or SEEN–IMAGINED distinction. Finally, following previous work (Kensinger et al., 2007; Rahhal et al., 2002), we adopted an intentional encoding in Experiment 1 and an incidental encoding in Experiment 2. This allows us to examine cultural effects across different encoding conditions.

In summary, this study examined whether artificially defined categorial processing of information would lead to cultural differences in memory performance. To maximize cultural effects, we manipulated processing of sources/categories in a conceptual source memory task (Experiment 1) and a reality monitoring task (Experiment 2) in which categorization was more encouraged for Westerners than for Chinese.

**Experiment 1: GOOD–EVIL Source Memory**

**Method**

Participants.—Fifty Canadians (26 young and 24 older), all of Western European descent, and 48 native Chinese (24 young and 24 older) participated in this experiment. The overall \( \chi^2 \) test showed that the two cultures did not differ in gender distribution, \( \chi^2 = 2.10, p = .19 \). Young participants (age range: 19–26 for Canadians; 21–26 for Chinese) were undergraduate students at local universities of their own residence (i.e., Toronto or Beijing). Older participants (age range: 63–77 for Canadians; 65–77 for Chinese) were recruited through internal older participant pools and/or through posters from local communities in Toronto and Beijing. Young Canadians received course credits and all the other groups received monetary compensation for their participation. Four Canadians (three young and one old) were replaced because of health issues or administration errors. Two older Chinese were replaced because they scored lower than 24 on the Mini-Mental State Examination (MMSE, Folstein, Folstein, & McHugh, 1975), a screen for potential dementia-related cognitive impairment. The sample characteristics and related significant effects are displayed in the top panel of Table 1.

Verbal skill was measured to ensure language proficiency. All Canadian participants scored above 26 (out of 40) on the Shipley vocabulary test (Shipley, 1940), with a higher average score for older (\( M = 37.21, SD = 2.26 \)) than for young adults (\( M = 31.68, SD = 3.20 \), \( t = 6.96, p < .001 \)). All Chinese participants scored above 26 (out of 80) on the Vocabulary subtest of the Wechsler’s Adult Intelligence Scale (Gong, 1983), with a higher average score for young (\( M = 62.33, SD = 12.52 \)) than for older adults (\( M = 53.58, SD = 11.85 \), \( t = 2.49, p < .05 \)).

In addition, a demographic questionnaire and the Self-Construal Scale (SCS; Singelis, 1994) were administered to all participants. A 2 (age) \( \times 2 \) (culture) analysis of variance (ANOVA) was conducted on each resulted variable. For
older adults, Canadians had more years of education and rated themselves as healthier than their Chinese counterparts. For young adults, Chinese had slightly more years of education than Canadians, whereas the two cultures did not differ on health ratings. Older Canadians scored higher on MMSE than older Chinese. Replicating previous findings (Singelis & Brown, 1995), Canadians scored higher in SCS-independent self-construal but lower in SCS-interdependent self-construal than did Chinese. For the independent self-construal, the cultural difference was differentially larger for older adults. These SCS results suggest that our samples well represent their own cultural norms, considering that Western individualistic societies highly value independence by attending to the self, whereas Eastern Collectivistic societies prioritize attending to others and harmonious interdependence (Markus & Kitayama, 1991, 2010).

Stimuli.—Thirty six photographs of Chinese faces were taken by the researchers and 36 photographs of Western faces were selected from the face database of the Center for Vital Longevity at the University of Texas at Dallas (https://pal.utdallas.edu/facedb/request/index). Within each culture, half the photographs were of young adults (aged approximately 18–30 years, half women and half men) and half were of older adults (aged 60–80 years, half women and half men). Participants in each culture viewed photographs of their own ethnicity. Each photo shows a close-up full face with a neutral expression. The face covers approximately 80% of the photo, posed against a plain light-colored background. Each photo was assigned a unique name, occupation, and residence (i.e., a city/town and a state/province of participants’ own country of residence). The names were randomly selected from telephone books at both sites. All Chinese names contain three characters. The numbers of the same first Pinyin sound (in Chinese) and the same first letter (in English) in family names were matched across all four age × gender categories. Occupations were popular in both cultures and were chosen in English first and then translated into Chinese. Back-translation was done to ensure linguistic comparability. The 36 photos of each culture were divided into three sets that were matched on age, gender, length, and popularity of names, with 3 photos for each age × gender category in each set. They were counterbalanced to be equally likely presented as GOOD, EVIL items at encoding, or as new items at recognition.

At encoding, participants viewed 24 photographs (6 for each age × gender category) sequentially presented at the center of a computer screen. The verbal descriptions were prerecorded, half spoken by a male speaker and the other half by a female speaker in participants’ native language (i.e., Mandarin or English). The male and the female speaker within each culture were equally assigned as the good or evil speaker, counterbalanced across participants. In addition, two different instructions were created for each set of stimuli, one depicted the male speaker as evil and the other described the male speaker as good. At test, participants viewed and responded to 24 old photos, mixed with the 12 new ones by pressing the corresponding keys to indicate whether each face was an old-GOOD, old-EVIL, or a new face.

Table 1. Sample Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Canadian</th>
<th>Chinese</th>
<th>F values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Young</td>
<td>Older</td>
<td>Young</td>
</tr>
<tr>
<td>Age</td>
<td>22.85 (2.22)</td>
<td>69.17 (4.10)</td>
<td>23.42 (2.11)</td>
</tr>
<tr>
<td>Gender ratio (female/male)</td>
<td>18/8</td>
<td>20/4</td>
<td>15/9</td>
</tr>
<tr>
<td>Education*</td>
<td>15.56 (1.97)</td>
<td>17.71 (4.34)</td>
<td>16.77 (1.32)</td>
</tr>
<tr>
<td>Health*</td>
<td>7.52 (1.42)</td>
<td>8.40 (1.37)</td>
<td>7.75 (1.36)</td>
</tr>
<tr>
<td>MMSE</td>
<td>N/A</td>
<td>28.58 (1.28)</td>
<td>N/A</td>
</tr>
<tr>
<td>Independent*</td>
<td>4.83 (0.65)</td>
<td>5.48 (0.51)</td>
<td>4.52 (0.77)</td>
</tr>
<tr>
<td>Interdependent*</td>
<td>4.34 (0.59)</td>
<td>4.37 (0.72)</td>
<td>5.03 (0.58)</td>
</tr>
</tbody>
</table>

Notes. Only significant F values are reported. Each cell (except for those for gender ratios) provides mean score, with standard deviation within the parenthesis. MMSE = Mini-Mental State Examination.
*Education is measured in years of education.
*Health is measured by self-report ratings on a 1–10 Likert-type scale.
*Independent and interdependent self-construal scores on the SCS.
*p < .05. **p < .01.
Procedure.—The stimuli were presented on a 17-in. computer monitor with an E-prime program. During the encoding, participants viewed 24 sequentially presented photographs while listening to a verbal recording by a female or a male voice introducing the name, the occupation, and the residence of the person in each photo. Four practice trials (two GOOD and two EVIL) were provided before the actual encoding task to familiarize participants with the task. The speakers’ volume was individually adjusted during practice trials. The photographs were presented in a pseudorandomized order, with no more than two photos of the same age × gender category appeared consecutively. Each trial started with a centered fixation cross for 1,000 ms and then replaced by a photo presented in a 100% frame (i.e., full screen) for 5,000 ms, along with a simultaneously played audio description. A 1,000-ms blank screen was presented as an interstimulus interval before proceeding to the next trial. Participants were told that the descriptions would be given by either a female speaker called Mary or a male speaker called John. They were also told that one speaker (Mary or John) “is a GOOD person and always describes GOOD people who are socially popular, liked by other people, and easy to get along with” and the other speaker (John or Mary) “is an EVIL person and always introduces EVIL people who are socially unpopular, disliked by most people, and difficult to get along with.” The encoding was intentional in that participants were encouraged to pay attention to all the information for a test later on.

Following the encoding, participants completed some unrelated filler tasks for 10 min to prevent rehearsal. Then they completed a source memory test in which they viewed 36 photographs (24 old and 12 new) sequentially presented in a pseudorandomized order, with no more than two items of the same age × gender category or of the same response type (i.e., old or new) appeared consecutively. Four practice trials (two old, one GOOD, and one EVIL, from the encoding practice trials and two new) were given before the actual test. Each trial started with a centered fixation cross for 1,000 ms, followed by a photograph presented in a 75% frame at the top of the screen with its corresponding verbal description printed at the bottom. The photo was terminated by a key press response before moving to the next trial. The task was to identify whether each photo was of a GOOD or an EVIL person studied earlier or a new photo they did not see at study. Participants were instructed to press key “z” (labeled as “GOOD”) with left index finger to indicate “GOOD,” key “?” (labeled as “EVIL”) with right index finger to indicate “EVIL,” and the space bar with dominant thumb to indicate “NEW.” They were requested to rest their fingers on the corresponding keys and respond as quickly and as accurately as possible.

Finally, participants completed some paper-and-pencil questionnaires, including the SCS, Shipley test (for Canadians), or the Wechsler’s Adult Intelligence Scale: Vocabulary subtest (for Chinese) and a background questionnaire. Older adults also completed the MMSE. Participants were then debriefed and paid (or granted for course credit). The instructions for the memory task and the background questionnaire were translated and back-translated by two bilingual researchers. All the other tests have a corresponding standard Chinese version.

Results

Item memory.—Item memory was measured with the signal-detection index $d'$ (Green & Swets, 1966), as calculated by subtracting the $z$ scores of false alarm rate (i.e., the proportion of new items recognized as old) from the $z$ scores of hit rate (i.e., the proportion of old items recognized as old; Table 2). Hit and false alarm rates of 0 and 1 were adjusted to 0.01 and 0.99, respectively. The resulted $d'$ were submitted to a mixed-model 2 (age: young vs. older)

Table 2. Item Memory Recognition Scores

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Canadian</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Young</td>
<td>Older</td>
<td>Young</td>
<td>Older</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d'$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1: GOOD</td>
<td>2.41 (1.26)</td>
<td>2.59 (1.31)</td>
<td>2.63 (1.14)</td>
<td>1.69 (0.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1: EVIL</td>
<td>2.78 (1.06)</td>
<td>2.76 (0.81)</td>
<td>2.35 (1.04)</td>
<td>1.73 (1.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2: SEEN</td>
<td>3.61 (1.18)</td>
<td>3.45 (0.83)</td>
<td>3.76 (0.99)</td>
<td>3.02 (0.79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2: IMAGINED</td>
<td>3.07 (1.14)</td>
<td>3.21 (0.92)</td>
<td>2.90 (0.19)</td>
<td>2.18 (1.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hit rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1: GOOD</td>
<td>.79 (.19)</td>
<td>.84 (.13)</td>
<td>.85 (.12)</td>
<td>.67 (.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1: EVIL</td>
<td>.80 (.15)</td>
<td>.88 (.10)</td>
<td>.85 (.12)</td>
<td>.69 (.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2: SEEN</td>
<td>.89 (.20)</td>
<td>.87 (.13)</td>
<td>.93 (.09)</td>
<td>.85 (.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2: IMAGINED</td>
<td>.90 (.10)</td>
<td>.88 (.15)</td>
<td>.84 (.16)</td>
<td>.81 (.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>False alarm rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1: GOOD</td>
<td>.12 (.13)</td>
<td>.16 (.18)</td>
<td>.14 (.16)</td>
<td>.18 (.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1: EVIL</td>
<td>.08 (.15)</td>
<td>.12 (.12)</td>
<td>.20 (.22)</td>
<td>.19 (.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2: SEEN</td>
<td>.02 (.05)</td>
<td>.02 (.05)</td>
<td>.03 (.08)</td>
<td>.07 (.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2: IMAGINED</td>
<td>.10 (.23)</td>
<td>.07 (.14)</td>
<td>.08 (.12)</td>
<td>.22 (.30)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Each cell provides mean score, with standard deviation within parentheses.
two cultural groups (Canadian vs. Chinese) × two item types (GOOD vs. EVIL) ANOVA. The cultural effect was significant, $F(1,94) = 8.22$, $MSE = 1.71$, $p < .01$, $\eta^2 = 0.08$, with a better item memory for Canadians ($M = 2.05$, $SD = 1.21$) than for Chinese ($M = 1.39$, $SD = 1.18$). The age effect was marginally significant, $F(1,94) = 3.57$, $MSE = 1.71$, $p = .06$, $\eta^2 = 0.04$, with a tendency of better item memory for young ($M = 1.95$, $SD = 1.35$) than for older participants ($M = 1.50$, $SD = 1.06$). The age × culture interaction was also significant, $F(1,94) = 5.28$, $MSE = 1.71$, $p < .05$, $\eta^2 = 0.05$. Post hoc analyses suggested that the cultural effect was only significant for older adults ($p < .001$) but not for young adults ($p = .34$). Similarly, the age effect was only significant for Chinese ($p < .05$) but not for Canadians ($p = .66$). Although the samples of the two cultures differed in education, MMSE, and SCS-independent and SCS-interdependent self-construal (Table 1), only MMSE showed significant correlation with item memory ($r = .29$). Given the significant culture × MMSE interaction ($p < .05$), the Homogeneity of Regression assumption of analysis of covariance (ANCOVA) was violated. The mediation analysis was run with MMSE as a mediator for older adults and it revealed a significant cultural effect ($p = .01$), suggesting that MMSE did not account for the cultural effect on item memory in older adults.

Source memory.—Source memory refers to the ability to discriminate between GOOD and EVIL in this study. Considering the discrimination between the two sources could be determined by both correct source attributions (i.e., source hits) and guesses or source misattributions (i.e., source false alarms), we adjusted for guessing in the calculation of source memory score. Specifically, it was calculated by subtracting the source false alarm rate (e.g., the proportion of source misattributions of EVIL items out of EVIL items correctly recognized as old) from the source hit rate (i.e., the proportion of source attributions of GOOD items out of GOOD items correctly recognized as old). Similarly, the guessing-correction approach has been adopted in a recent work in the calculation of item and context memory scores (Craik, Luo, & Sakuta, 2010).

Figure 1 displays the resulted source memory scores. The between-subjects 2 (age) × 2 (culture) ANOVA revealed a significant main effect of age, $F(1,93) = 4.27$, $MSE = 0.07$, $p < .05$, $\eta^2 = 0.04$, with a better source memory for young ($M = 0.23$, $SD = 0.26$) than for older adults ($M = 0.12$, $SD = 0.26$). Neither cultural effect nor age × culture interaction was significant ($p$ values > .28). Despite the lack of interaction, driven by our main theoretical hypothesis, we conducted the follow-up analysis on age differences within each culture. In the expected direction, the age effect was only significant for Chinese ($p < .05$) but not for Canadians ($p = .46$).

Taken together, older Canadians outperformed older Chinese in memory for categorically processed items but not in source memory. Importantly, Chinese, but not Canadians, showed age group differences favoring young adults in both item and source memory.

Discussion

The results of Experiment 1 were largely consistent with our hypothesis by showing a cultural effect favoring Canadians, for older adults specifically, on memory for categorically processed items. Due to life-long exposure to Western cultures and extensive practice of culturally preferred strategies, Canadian older adults remained to be able to effectively use categorization at encoding, and
consequently, they do not show much age group differences in memory for categorically processed information.

**EXPERIMENT 2: SEEN–IMAGINED REALITY MONITORING**

A reality monitoring task (Kensinger et al., 2007) was administered in Experiment 2 to further consolidate the cultural effect found in Experiment 1.

**Method**

*Participants.*—Forty-eight Canadians (24 young and 24 older), all of Western European descent, and 46 Chinese (24 young and 22 older) participated in this experiment. The $\chi^2$ test showed that the two cultures did not differ in gender distribution, $\chi^2 = 0.11, p = .82$. Young (age range: 17–27 for Canadians; 19–29 for Chinese) and older participants (age range: 60–77 for Canadians; 61–74 for Chinese) were recruited and compensated in the same way as in Experiment 1. Four Canadians (three young and one old) were replaced because of health issues or administration errors. Two Chinese were replaced, one older adult for a low MMSE score of 21 and one young adult for withdrawing from the study. The sample characteristics and the related significant effects are displayed in the bottom panel of Table 1.

All Canadian participants scored above 24 (out of 40) on the Shipley vocabulary test, with a higher score for older ($M = 36.75, SD = 2.94$) than for young adults ($M = 29.17, SD = 3.31$), $t = 8.40, p < .001$. All Chinese participants scored above 42 (out of 80) on the Vocabulary subtest of the Wechsler’s Adult Intelligence Scale, with a higher average score for young ($M = 69.00, SD = 5.41$) than for older adults ($M = 62.45, SD = 11.21$), $t = 2.56, p < .05$.

As in Experiment 1, we administered a demographic questionnaire and the SCS. For older adults, Canadians had more years of education than their Chinese counterparts. For young adults, the two cultures did not differ in education. Older adults of the two cultures did not differ on the MMSE score. Replicating Experiment 1, Canadians scored higher on the SCS-independent self-construal but lower on the SCS-interdependent self-construal than did Chinese, suggesting that our samples well represent their own cultural norms.

*Stimuli.*—Sixty concrete nouns were selected from a picture naming database normed with American and Chinese young and older adults (Yoon et al., 2004). All the selected names refer to easily recognized common objects (e.g., airplane, pineapple) that are relatively familiar to people of both cultures. The familiarity ratings ranged 2.04–5 based on a 5-point scale, with larger numbers indicating higher familiarity. The average familiarity within each age × culture group ranged 4.02–4.85. All the chosen names showed concept agreement across the four age × culture groups. For each selected object name, we obtained a colored photographic image of the corresponding object from Google images. All images were modified to be of the same size and each depicted a typical image of the object, presented against a plain white background.

In the learning phase, participants viewed 40 words, half were followed by a corresponding image (i.e., SEEN condition) and the other half followed by a blank square in which an image should be imagined for the word (i.e., IMAGINED condition). The 60 words were divided into three sets that were counterbalanced to be presented as SEEN, IMAGINED items at encoding, or as new items at recognition. The three sets of stimuli were roughly matched on relevant linguistic characteristics (i.e., imagability, concreteness, familiarity, word frequency, and word length), if available at http://www.psych.rl.ac.uk/MRC_Psych_Db.html. They were also matched on the normed familiarity rating (Yoon et al., 2004). In the test phase, participants viewed 60 words (40 old and 20 new words) and responded through pressing the corresponding keys to indicate whether each word was an old-SEEN, old-IMAGINED, or a new word.

*Procedure.*—The stimuli were presented on a 17-in. computer monitor with an E-prime program. During the encoding, participants learned 40 words sequentially presented in a pseudorandomized order, with no more than two words of the same condition (SEEN or IMAGINED) appeared in a row. Different from Experiment 1, this was an incidental encoding, because participants were instructed that the task was to measure their mental imagery ability. Six practice trials (three SEEN and three IMAGINED) were provided before the actual encoding. Each trial started with a centered fixation cross for 1,000 ms, followed by a word presented for 3,000 ms. Participants were instructed to read each word out loud. For the SEEN trials, each word was followed by a corresponding picture presented for 3,000 ms. For the IMAGINED trials, each word was followed by a blank square for 3,000 ms, and participants were instructed to form a vivid mental image of the object depicting the word. In both conditions, participants were asked to determine whether the object image (SEEN or IMAGINED) was longer horizontally (i.e., width) or vertically (i.e., height), or equal in the two dimensions by pressing the correspondingly labeled keys: key “z” (labeled with a horizontal arrow), key “.” (labeled with a vertical arrow), and the space bar were, respectively, pressed with left index finger, right index finger, and the dominant thumb to indicate “horizontally longer,” “vertically longer,” and “equal.” The labels were presented as location cues at the bottom of the screen. The screen went blank for 1,000 ms (interstimulus interval) before proceeding to the next trial.

Following the encoding phase, participants completed some unrelated filler tasks for 10 min to prevent rehearsal. Then they completed a source memory test in which they viewed 60 words (40 old and 20 new) sequentially
Presented in a pseudorandomized order, with no more than three words of the same response type (old-SEEN, old-IMAGINED, and new) appeared consecutively. Six practice trials (two SEEN, two IMAGINED, and two new) were provided before the actual test. Each trial started with a centered fixation cross for 1,000 ms, then replaced by a word presented for a maximum of 6,000 ms or terminated by a key-press response. Participants were instructed to press key “z” (labeled as “SEEN”) with left index finger for old-SEEN, key “x” (labeled as “IMAGINED”) with right index finger for old-IMAGINED, and the space bar with dominant thumb to indicate a NEW word. These labels were also showed as location cues at the bottom of the screen. They were requested to rest their fingers on the corresponding keys and respond as quickly and as accurately as possible. The instructions for the memory task were translated and back-translated by two bilingual researchers.

Finally, participants completed the same paper-and-pencil questionnaires as in Experiment 1: the SCs, Shipley (for Canadians), or the Wechsler’s Adult Intelligence Scale: Vocabulary subtest (for Chinese) and a background questionnaire. Older adults completed the MMSE. Participants were then debriefed and paid or granted for course credit.

**Results**

**Item memory.**—As in Experiment 1, $d’$ was calculated in the same way to index item memory performance. The 2 (age) × 2 (culture) × 2 (item category) mixed model ANOVA on $d’$ showed significant main effects of age, $F(1,90) = 4.73, MSE = 1.39, p < .05, \eta^2 = 0.05$; culture, $F(1,90) = 4.59, MSE = 1.39, p < .05, \eta^2 = 0.05$; and item category, $F(1,90) = 29.77, MSE = 0.61, p < .001, \eta^2 = 0.25$. Item memory was better for young adults ($M = 2.92, SD = 1.23$), Canadians ($M = 2.98, SD = 1.21$), and SEEN items ($M = 3.47, SD = 0.99$) than for older adults ($M = 2.41, SD = 1.24$), Chinese ($M = 2.35, SD = 1.23$), and IMAGINED items ($M = 2.85, SD = 1.09$), respectively. The item category × culture interaction was also significant, $F(1,90) = 4.08, MSE = 0.61, p < .05, \eta^2 = 0.04$. Post hoc comparisons showed that cultural effect was only significant for the IMAGINED ($p < .01$) but not for the SEEN items ($p = .55$). The age × culture interaction was significant, $F(1,90) = 4.30, MSE = 1.39, p < .05, \eta^2 = 0.05$. Post hoc tests showed that the cultural effect was only significant for older ($p < .001$) but not for young adults ($p = .85$). Similarly, the age effect was significant for Chinese ($p < .01$) but not for Canadians ($p = .91$).

**Source memory.**—Similar to Experiment 1, source memory was calculated by subtracting the source false alarm rate from the source hit rate in order to correct for guessing. Figure 1 displays the resulted source memory scores. The 2 (age) × 2 (culture) ANOVA revealed significant effects of age, $F(1,90) = 5.17, MSE = 0.07, p < .05, \eta^2 = 0.05$; culture, $F(1,90) = 17.59, MSE = 0.07, p < .001, \eta^2 = 0.16$. Source memory was better for young adults ($M = 0.58, SD = 0.28$) and Canadians ($M = 0.63, SD = 0.24$) than for older adults ($M = 0.46, SD = 0.30$) and Chinese ($M = 0.41, SD = 0.30$), respectively. The age × culture interaction was not significant ($p = .45$). Guided by our main theoretical interest, the follow-up analysis showed that the age effect was marginally significant for Chinese ($p = .06$) but not significant for Canadians ($p = .23$).

Although the samples of the two cultures differed in education, SCS-independent and SCS-interdependent self-construal (Table 1), only education showed significant correlations with item ($r = .36$) and source memory ($r = .30$). To determine if education difference accounts for the cultural effect on memory performance, we ran ANCOVA with education as covariate. The results showed that the cultural effect was approaching significance in item memory ($p = .09$) and remained significant in source memory ($p < .01$).

Taken together, an age effect favoring young adults and a cultural effect favoring Canadians were revealed in both item and source memory. For item memory, only older adults, not young adults, showed cultural effect.

**Discussion**

The results of Experiment 2 largely replicated Experiment 1 and are consistent with a previous finding (Gutchess et al., 2006). Going beyond Experiment 1 in which the cultural effect was only seen in item memory, Experiment 2 revealed cultural effects in both item and source memory. A possible explanation for this is that the task in Experiment 2 greatly encourages logical realistic reasoning that favors Canadians and guides them to attend to both items and sources. In addition, the easily recognizable stimuli and the deep semantic encoding of Experiment 2 may also increase the possibility of engaging the culturally favored encoding strategies.

**General Discussion**

The primary goal of this study was to examine cultural differences in memory for categorically processed information in both young and older adults. Overall, the results from both intentional source memory (Experiment 1) and incidental reality monitoring (Experiment 2) tasks showed consistent memory advantage favoring Canadians for information processed in their culturally preferred categorization manner. The cultural effect was more pronounced for older than for young adults. The findings make a novel and meaningful addition to the literature on cultural differences in cognition (Goh et al., 2007; Gutchess, Hedden, Ketay, Aron, & Gabrieli, 2010; Gutchess, Schwartz, & Boduroğlu, 2011; Gutchess et al., 2006). The reported cultural effect largely supports the notion that Westerners use taxonomic categorization more often (Ji et al., 2004), and consequently,
they rely more on categorization as a memory organization strategy than East Asians (Gutchess et al., 2006). Going beyond previous work, this study extended the cultural effect to a condition in which categories were artificially assigned based on a taxonomic categorization encoding.

The findings of this study (particularly the cultural effect on source memory performance in Experiment 2) suggest that source memory could vary with cultures and thus challenge the dominant view of source memory as a basic culture-invariant “hardware of the mind” (Chua et al., 2006; Park & Huang, 2010). This, however, was inconsistent with some previous work that did not find any cultural differences in source memory (Chua et al., 2006). We propose that when the encoding process maximally encourages utilization of culturally preferred strategies, culturally relevant experience or strategies would be recruited to promote encoding (Baltes, 1993) and thus lead to a memory advantage.

It should be noted that the cultural effect on item memory was significant in both experiments, but the cultural effect on source memory was only significant in the reality monitoring task. This may be due to the different item-source instruction orientation of the two tasks. In Experiment 1, the task requires a distinction between GOOD and EVIL personality character categories. The fundamental attribution errors, more prevalent in Western than in East Asian cultures (Norenzayan & Nisbett, 2000; Ross, 1977), may differentially guide Canadians to pay close attention to and thus engage deep elaborative processing of items (e.g., the specific facial features), but face features did not directly prime or cue the associated categories; thus, the cultural effect was seen in item but not in source memory. The reality monitoring task primarily relies on the distinction between reality and imagination. In this task, the words (items) share semantic meaning with the SEEN or IMAGINED images (categories), so memory for items directly primes memory for sources and thus cultural effects were expressed in both item and source memory. Furthermore, the recognition task in Experiment 2 may be primarily based on semantic processing of concrete images or words and thus easier than the recognition task in Experiment 1 where the recognition may largely involve perceptual feature-based recognition of faces. Cross-experiment comparison analyses confirmed that item and source memory were better in Experiment 2 than in Experiment 1 ($p$ values < .001). Most importantly, the between-experiment difference in source memory was qualified by an experiment × culture interaction ($p < .05$). The deep semantic processing at both encoding and retrieval in Experiment 2 may make it more likely to engage culturally preferred experience or strategies, leading to a more pronounced cultural effect on source memory.

In addition, we also found that the cultural differences in memory benefits from categorical processing were primarily shown for older adults but were very limited for young adults. This somewhat contradicts the cultural convergence theory proposed by Park, Nisbett, and Hedden (1999), according to which cultural effects on resource-demanding source memory, if any, ought to decline with age due to that culturally universal biological decline with aging that may constrain the usage of culturally preferred strategies in older adults. However, the finding was largely supportive to Gutchess and coworkers (2006) by showing cultural differences in categorization for older adults but not for young adults. Following the interpretations offered by Gutchess and coworkers (2006), this interaction may presumably be driven by the following factors: First, considering the neurocognitive evidence that culturally nonpreferred processing is more cognitively demanding than culturally preferred processing (Hedden, Ketay, Aron, Markus, & Gabrieli, 2008); taxonomic categorization is generally not favored by Chinese; and using this strategy is likely more resource demanding and effortful for Chinese than for Canadians. Therefore, the decreased cognitive resources with aging may have limited Chinese older adults’ ability to use categorization encoding strategy. Second, the life-long practice of using culturally preferred categorization may make it less effortful for Canadian adults. Thus, older Canadians are able to effectively use this strategy to maintain their memory performance. In this perspective, the current finding supports the idea of Park et al. (1999) that prolonged immersion in a specific culture (e.g., with aging) can magnify the expression of cultural biases in information processing.

The substantial age effects on item and context memory replicated the well-reported finding of age-related decline in episodic memory (for reviews, see Light, 2000; Zacks, Hasher, & Li, 2000). However, this age effect tends to be significant only for Chinese but not for Canadians. The lack of age-related decline in conceptual source memory and internal–external source monitoring has been demonstrated with Western population (Rahhal et al., 2002; Hashtroudi et al., 1989). Our findings suggest that this aging-spared memory for categorically processed information is specific to Western cultures in which categorization is likely to be automatically engaged and remained to be effective across life span.

This study also has some limitations. First, the samples between the two cultures were not well matched. Older Chinese had lower level of education than older Canadians. In Experiment 1, older Chinese also rated themselves as less healthy and scored lower on MMSE relative to their Canadian counterparts. The lower education of older Chinese may be a result of Chinese history, considering that almost all older Chinese of this age went through 10 years of Culture Revolution that terminated or significantly disrupted their formal education. Nevertheless, we should note that the cultural effects remained significant even after controlling for these differences in mediation or covariance analyses. In addition, substantial cultural effects were shown in Experiment 2, despite the older samples of the two cultures were matched on MMSE and the young samples
of the two cultures were matched on all critical variables. Taken together, we believe that the cultural effect found in this study is not likely driven by the sample differences. Second, the observed cultural effect may be confounded with the cohort effect. The lack of cultural effect in young groups could be alternatively due to the recent trend of westernization of Chinese young generations in major cities such as Beijing.

Despite the limitations, this study provides the first empirical evidence for a consistent cultural effect favoring Westerners in memory for categorically processed information. The cultural effect was shown in both an intentionally encoded source memory and an incidentally encoded reality-monitoring task. The cultural effect is more pronounced in older adults, as a result of accrued years of cultural experience and life-long practice of culturally favored cognitive strategies.

**Acknowledgments**

We thank all the research assistants (e.g., Shuyuan Fan; Shurui Li, Tianwei Liu, Dana Greenbaum, Connie Kuan, Karen Lau, Sasha Mallya, Gina Polsinelli, Linda Truong, and Mandi Torris) for their help in data collection and scoring.

**Correspondence**

Correspondence should be addressed to Lixia Yang, PhD, JOR918, Department of Psychology, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada. E-mail: lixiay@ryerson.ca.

**References**


