## Gerontology

**Original Paper**

# Categorical Organization in Free Recall across Culture and Age

Angela H. Gutchess[a]   Carolyn Yoon[b]   Ting Luo[c]   Fred Feinberg[b]   Trey Hedden[d]
Qicheng Jing[c]   Richard E. Nisbett[a]   Denise C. Park[e]

[a]Department of Psychology, and [b]Ross School of Business, University of Michigan, Ann Arbor, Mich., USA;
[c]The Institute of Psychology, Chinese Academy of Sciences, Beijing, China; [d]Department of Psychology,
Stanford University, Stanford, Calif., and [e]Beckman Institute, University of Illinois at Urbana-Champaign,
Urbana-Champaign, Ill., USA

**Abstract**

*Background:* Cross-cultural differences in cognition suggest that Westerners use categories more than Easterners, but these differences have only been investigated in young adults. *Objective:* The contributions of cognitive resource and the extent of cultural exposure are explored for free recall by investigating cross-cultural differences in categorical organization in younger and older adults. Cultural differences in the use of categories should be larger for elderly than young because categorization is a well-practiced strategy for Westerners, but age-related cognitive resource limitations may make the strategy difficult for elderly Easterners to implement. Therefore, we expect that cultural differences in categorization will be magnified in elderly adults relative to younger adults, with Americans categorizing more than Chinese. *Methods:* Across two studies, 112 young and 112 elderly drawn from two cultures (American and Chinese) encoded words presented in their native language. One word list contained categorically-unrelated words and the other, categorically-related words; both lists were presented in the participants' native language. In experiment 1, the words were strong category associates, and in experiment 2, the words were weak category associates. Participants recalled all the words they could remember, and the number of words recalled and degree of clustering by category were analyzed. *Results:* As predicted, cultural differences emerged for the elderly, with East-Asians using categories less than Americans during recall of highly-associated category exemplars (experiment 1). For recall of low-associate exemplars, East-Asians overall categorized less than Americans (experiment 2). Surprisingly, these differences in the use of categories did not lead to cultural differences in the number of words recalled. The expected effects of age were apparent with elderly recalling less than young, but in contrast to previous studies, elderly also categorized less than young. *Conclusion:* These studies provide support for the notion that cultural differences in categorical organization are larger for elderly adults than young, although culture did not impact the amount recalled. These data suggest that culture and age interact to influence cognition.

Copyright © 2006 S. Karger AG, Basel

Dr. Angela Gutchess
Harvard University, William James Hall 868
33 Kirkland Street
Cambridge, MA 02138 (USA)
Tel. +1 617 495 9031, Fax +1 617 496 3122, E-Mail agutchess@wjh.harvard.edu

## Introduction

Although investigation of cross-cultural differences in cognition is still in its infancy, one of the commonly reported findings is that of cross-cultural differences in the use of categories. Nisbett and colleagues [1–3] suggest that people of Western cultures tend to excel at categorization and that this ability can be traced to the Greek tradition of focusing on rules and linear reasoning. They further posit that Easterners, by contrast, are less facile with categorization because of the greater emphasis in their culture on dialectical reasoning and a focus on functional rather than categorical relationships. In the present studies, we investigate the role of categorization in memory recall across Eastern (i.e. Chinese) and Western (i.e. American) cultures. Further, we explore the changing contribution of culture to cognition across the lifespan.

The empirical evidence to date largely supports the claim that categorization varies with culture. Chiu [4] found, in a sorting task, that American children exhibit a greater tendency to group by shared features or category (i.e. taxonomic relationships) than Chinese children who group items by relationship (i.e. thematic categorization). As discussed by Ji et al. [5], '*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' [p. 57]. For example, when confronted with the triplet *chicken-cow-grass*, Americans pair *chicken* and *cow* together on the basis of their shared category membership (i.e. both are animals) while Chinese pair *cow* and *grass* together on the basis of their functional relationship (i.e. cows eat grass). The same pattern reported in children [4] extends to college students, and is further validated by American and Chinese participants' explanations of the basis for their sorting patterns [5]. The cross-cultural differences in explanations were consistent with the categorical/relational difference first suggested by Chiu [4]. Furthermore, Ji et al. [5] concluded that cross-cultural differences in groupings were an effect of culture rather than language because the cultural differences persisted across multiple Asian populations that spoke a variety of languages.

In a recent study, Unsworth et al. [6] also identified a strong categorical tendency in Westerners using Chiu's sorting task. Furthermore, they investigated priming effects using primes that were associated with targets either relationally (i.e. an item sharing a dependent function with the target item, as in grass-cow) or categorically (i.e. an item with a function in common or from the same category as the target, as in the chicken-cow example). Somewhat surprisingly, East-Asians showed equivalent reduction in reaction times when pictures were primed by either categorical or relational items, whereas Westerners showed priming effects primarily for categorical items. These results suggest that East-Asians may include both relational and categorical associations in their semantic networks, whereas Westerners have a stronger bias towards categorical information.

Cross-cultural differences in categorization influence performance on a broad variety of cognitive tasks, including learning and reasoning. For example, when learning to categorize novel animals into different classes, Americans and Chinese participants performed equivalently when instructed in a similarity-based strategy relying on exemplars [7]. However, Americans outperformed the Chinese when instructed in a rule-based strategy. In the rule-based condition, Chinese attempted to use the exemplar-based strategy, which led to errors based on the misclassification of animals that were similar to the exemplars but did not satisfy the learned rule [7].

Although cross-cultural differences in categorization should be important for a broad range of cognitive processes, the effects on memory are largely unexplored. Categorical clustering during free recall represents a particularly rich task domain for investigating cross-cultural differences in memory in that it requires highly strategic processes utilizing category structure [8]. Categorical clustering [9] occurs when individuals in a free recall task systematically organize their memory output according to categories. In this paradigm, subjects are presented with lists of words that contain exemplars from different categories, with the words from different categories randomly interspersed. When presented with randomly interspersed words belonging to multiple categories such as 'animals' and 'vegetables', individuals show a strong tendency to output the words at recall by category (i.e. recalling all the animals first, and then recalling all the vegetables next, rather than switching between categories). We hypothesize that given the cross-cultural differences observed in categorization, Americans will exhibit more categorical clustering than East-Asians (also see [8] for explication of this hypothesis).

We also investigate age differences across cultures in categorical clustering on a free recall task. Age is a particularly salient variable for the study of cultural differ-

ences in cognition, as the effects of culture may be magnified with age for two reasons [8]. If a strategy (such as categorical clustering) is more effortful for one culture than another, the culture effects should be particularly pronounced with age due to (a) cognitive resource limitations such as decreased speed of processing and working memory capacity that occur with age [10, 11] and make adopting an unfamiliar strategy difficult, and (b) decades of sustained experience utilizing the culturally-preferred strategy for encoding that makes switching to a less-preferred strategy difficult. In other words, both biological constraints (neurobiological aging) and contextual constraints (decades of exposure to a culturally-biased strategy of information processing) operate simultaneously as we age to increase the magnitude of culture effects on information processing.

It is important to note that in investigating category use, we are not suggesting that East Asians *cannot* categorize. A recent study across cultures provided evidence of similar category structure across people from American and Hong Kong cultures; Schunn and Vera [12] uncovered similar category structure across young adults in these cultures based on feature centrality (i.e. importance of functional features in contribution to a category). In another study, young and elderly mainland Chinese participants were readily able to generate exemplars for over 100 categories in a norming study [13], although the category structure differed substantially between East Asians and Americans. However, the norming study [13] did isolate 19 categories that had a similar structure for East Asians and Americans, and we use exemplars from these categories for the present study of categorical clustering. To date, studies that have concluded that East Asians and Westerners organize information differently (based on functional relationships for Asians and categories for Americans) have generally selected stimuli by relying on intuition, or at best, measures of familiarity for the individual items, but they have not accounted for categorical relationships across cultures [5, 6]. We correct for this problem in the present research by utilizing words from categories that have similar properties for mainland Chinese and Americans and for young and elderly adults from these cultures.

Free recall is a particularly interesting domain for the investigation of cross-cultural differences with age because it relies heavily on self-initiated strategies and taxes cognitive resources such as speed and working memory [10]. The literature is unequivocal on the finding of age-related declines in free recall [14], and largely supports age-equivalent use of categorization during free recall

[15]. Park et al. [16] showed that despite reductions in the amount recalled, older adults categorized to the same extent as younger adults (experiment 1, control condition). Wingfield and Kahana [17] also noted the similarity in the organization of recall by young and elderly. When the age groups were equated on degree of learning, equivalent categorical organization during recall was obtained for both high and low prototypical category members [18]. Using inter-response times (IRTs) within and between categories to unconfound clustering with amount recalled (as discussed in [14]), Wingfield et al. [19] found that although older adults have longer IRTs between categories, their recall proceeds in the same manner as for young adults once a category is accessed. Some work suggests more profound differences across age groups in categorical organization, and Hultsch [20] posits that although elderly adults access category-level knowledge, they cannot store as much information under each higher-order category unit. Although the different findings across studies are likely due in part to the use of different clustering measures, differences in procedures and stimuli likely also contribute to discrepant findings. In addition, young adults could be particularly variable in their level of categorization during recall of a single list because they initially tend to adopt a seriation strategy before moving to a categorical one [21].

The lack of age effects in the use of a categorical organization strategy is somewhat surprising given the role of frontal mechanisms in category-based free recall tasks [22]. The frontal lobe undergoes substantial shrinkage with age relative to other regions [23] and decreased frontal function has been associated with normal aging on neuropsychological tests [24, 25]. Category fluency decreases with healthy aging [26, 27], possibly due to its frontal component [27]. One intriguing possibility to explain the findings of intact categorical clustering into old age [16, 18, 19] is that perhaps the emphasis on categorization throughout one's life in Western society protects the strategy against age-related declines. Comparison to a culture with less emphasis on categorization allows for a test of the preservation of category-based strategies with age. The majority of research on aging has been conducted with Western samples, and the degree to which malleable processes and strategies, which can be culturally-determined, contribute to cognitive change is unknown [8, 28].

In the present investigation, we explore the contributions of culture and age to free recall of categorically-related words. We predict that East-Asians will evidence less categorical clustering than Americans, and that cul-

Gutchess/Yoon/Luo/Feinberg/Hedden/
Jing/Nisbett/Park

**Table 1.** Sampling characteristics (means and SDs) for participants in experiment 1

| | Young American | Young Chinese | Elderly American | Elderly Chinese |
|---|---|---|---|---|
| Age | 19.78 (1.52) | 19.97 (1.03) | 67.78 (3.76) | 65.78 (4.12) |
| Years of education | 14.08 (1.24) | 13.44 (0.72) | 15.58 (2.44) | 15.47 (1.43) |
| Health rating[a] | 1.63 (0.61) | 1.84 (0.81) | 2.19 (0.82) | 2.22 (0.87) |
| Pattern matching [31] | 20.48 (4.36) | 18.53 (4.06) | 12.16 (3.46) | 10.84 (3.42) |
| WAIS information [32, 33] | 11.81 (1.35) | 12.13 (1.84) | 12.94 (2.05) | 12.91 (2.41) |

[a] Health ratings reflect how frequently health problems interfere with lifestyle, as assessed on a 5-point scale (1 = 'never'; 2 = 'seldom'; 3 = 'sometimes'; 4 = 'often'; 5 = 'always').

tural differences will be particularly pronounced in East-Asian older adults who may lack the resources to utilize categorical organization to aid in list recall.

## Experiment 1

*Methods*

*Participants.* A total of 128 participants from the US or mainland China participated in the study in exchange for pay or course credit. Young adults (ages 18–22) and elderly adults (ages 60–78) were recruited and tested in Ann Arbor, Mich., USA, or Beijing, China, with 32 participants in each of the groups. All participants were characterized on neuropsychological measures in order to ensure that comparable samples were selected from each culture. Prior work [29, 30] identified appropriate measures of speed of processing and general intelligence for use in cross-cultural research. These studies also established the utility of sampling from equivalent levels of education across samples, even though a particular level of educational attainment may be more selective in one cohort than another. Sample characteristics are presented in table 1. In an ANOVA with age and culture as between-groups variables, the expected effects of age are seen with elderly performing slower than young adults on the pattern matching task [31] (i.e. completing fewer items), $F(1, 124) = 138.69$, $p < 0.001$. Older adults also possess more years of education, $F(1, 124) = 39.78$, $p < 0.001$, and greater world knowledge according to scaled WAIS Information [32, 33] scores, $F(1, 124) = 7.62$, $p < 0.01$. Main effects of culture appear for health ratings, $F(1, 124) = 11.46$, $p < 0.01$, and pattern matching, $F(1, 124) = 5.77$, $p < 0.02$, with Americans reporting slightly better health and completing more items than the Chinese. The only interaction of age and culture occurs for age, with the elderly Chinese slightly younger than elderly Americans, $F(1, 124) = 4.44$, $p < 0.05$. This age difference reflects the shorter lifespan of the Chinese sample, and indicates that the samples are drawn from comparable parts of the age distribution in each population. The overall similar cognitive profile of participants across both cultures suggests that any differences in recall and clustering are unlikely to result from systematic sampling differences of higher ability individuals in one culture over the other.

*Stimuli.* Two lists of 20 words were created: one consisting of unrelated words, and the other consisting of words related by cat-

egory. For the related list, five items were drawn from each of four categories (fruits, internal organs, times of day, chemical elements) deemed to be equivalent across age and culture [13]. The strongest exemplars (e.g. 'apple' for the category 'fruit') were avoided to prevent correct guessing of items when participants could not recall items from the related list. Whenever possible, the same category exemplars were included in the American and Chinese lists; however, the response distributions across cultures made it necessary to use different exemplars across cultures in some cases, but the same categories were always used. Based on the category norming data, the average frequency of generation of the selected list items in the category norming study [13] ranged from 0.34 to 0.38 (SD = 0.18–0.23) for related items – i.e. 34–38% of the individuals in the norming sample generated an item as a category exemplar. The unrelated items were each selected from different category lists from the same norming data. Categorized (i.e. related) and control (i.e. unrelated) lists were equated within each culture for the 85% of the words for which written word frequencies were available for English [34] and for Mandarin [35].

*Procedures.* Participants studied two lists, each containing 20 words arranged in a randomized order for each individual. Words were presented simultaneously aurally, through headphones, and visually, on a computer screen, in the participant's native language (English or Mandarin Chinese). Words were presented visually for 4 seconds each, with a 1-second blank display before the next item. At the conclusion of the list, there was a 1-min retention interval, filled with a written subtraction by 7's task. Participants then had 2 min to recall all the words they could remember, in any order, into a tape recorder. The procedure was then repeated for the second word list. The control list always preceded the list of categorized words. This was done to avoid any possible group differences in the expectation of relationships between the control words, which participants may be more prone to search for if they were to encode the related word list beforehand.

The study was approved by the Chinese Academy of Sciences Institute of Psychology and University of Michigan Behavioral Sciences Institutional Review Boards and all participants provided written informed consent at the beginning of the session. Instructions were provided in written English or Mandarin as well as read out loud by the experimenter for each task. Mandarin instructions were translated from the English version by a translator fluent in both English and Mandarin, and the translation was then back-translated and checked by a second translator fluent in
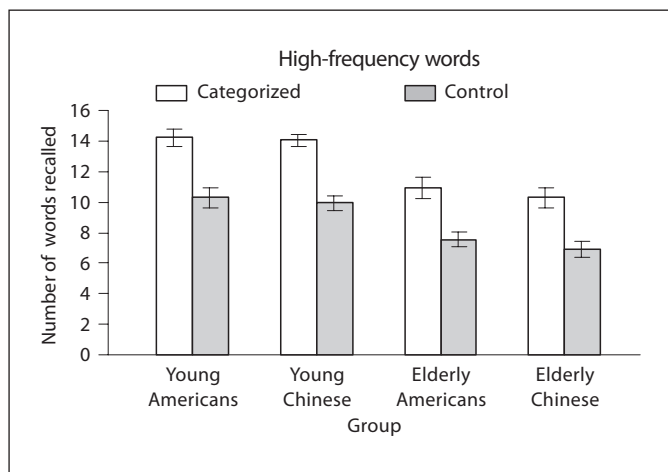
**Fig. 1.** The average number of words free recalled by each group and list type for experiment 1.

both languages. After the recall task was performed, the individual differences measures described in the 'Participants' section were administered.

*Data Coding.* Audiotapes were transcribed by native speakers of each language, including all repetitions of items and intrusions of novel items. For the purposes of analysis, slight modifications of items were considered correct (e.g. mistaken pluralization). The first instance of any repeated items was considered while the latter occurrences were not. The number of correctly recalled items was calculated by counting the number of studied words generated by each participant.

For the clustering analysis, the number of items generated in succession from a single category was coded. If repetitions occurred within same-category clusters, they were omitted from the clustering analysis, but did not mark a break in the cluster. Adjusted ratio of clustering (ARC) scores are a common measure of clustering, considered suitable across a range of numbers of items recalled [36, 37]. An ARC score of 1 denotes perfect clustering (i.e. all items from a single category are recalled together before moving to the next category), whereas a score of 0 denotes clustering at the level of chance (i.e. item order is 'zero-order' or entirely random, given a fixed set of mentioned items). Negative ARC scores (except in rare cases, bounded at –1) are possible when less categorization occurs than by chance (i.e. items from different categories tend to be output adjacent to one another). Comparing across groups using standard statistical techniques requires not only that the means of ARC distributions be the same for any number of mentioned items – and they are, in fact, always zero – but that the standard deviations be the same as well, which they are not. For example, with four categories of five items each, we can compute the 'by chance' ARC distribution for any number of mentioned items; for all 20 mentioned, the ARC standard deviation is 0.148, while for 6 items, it is 0.537, or 3.5 times greater. In this way, ARC scores do, in fact, depend upon the number of items recalled: because conditional distributions can differ across groups when recall performance differs, average ARC should not be used for

between group comparisons when groups differ on number of items recalled.

It is plainly inappropriate to use standard t-type means-based tests, then, on the raw computed ARC score [36, 37]. To address this problem, we transformed the conditional (on number of items recalled) ARC distributions to conform to the same distributional family, one as close as possible to the calculated, empirical ARC distributions.[1] We henceforth refer to these individual-level clustering measures as transformed ARC scores.

*Results*

*Number of Items Recalled.* The number of list items correctly recalled by each participant was subjected to a $2 \times 2 \times 2$ mixed ANOVA with Age (Young/Old) and Culture (American/Chinese) as between-groups variables and List Type (Categorized/Control) as a within-subject variable. As predicted, there was a significant main effect of Age with young adults (mean = 12.13, SD = 3.61) recalling more words than elderly adults (mean = 8.90, SD = 3.74), $F(1, 124) = 46.00$, MSE = 666.29, $p < 0.001$. There was also an expected main effect of List Type, with better recall for categorized list words (mean = 12.37, SD = 3.76) compared to control list words (mean = 8.66, SD = 3.35), $F(1, 124) = 148.83$, MSE = 881.35, $p < 0.001$. However, none of the main effects or interactions involving Culture approached significance, $Fs <1$. See figure 1 for graphs displaying the average recall for each group.

---

[1] The Beta family was an appropriate transformation choice given the unimodal, asymmetric, and bounded (between –1 and 1) values for ARC. Note that inverse Z scores are not an appropriate solution, given the asymmetry and bounded values of the ARC scores, whereas Z is symmetric and unbounded (and therefore not robust, i.e. one extreme Z score could greatly distort average comparisons). We performed simulations to determine the most appropriate transform. For each value of number of items recalled (from 5 through 20), we simulated one million draws of a multinomial distribution, and calculated ARC scores, thus obtaining the conditional pdf (density function) for each distribution; calculated values of the mean and standard deviation accorded closely with theoretical values. For each distribution, we estimated via maximum likelihood the best-fitting member of the Beta family. For larger numbers of items mentioned, the resulting Beta distributions fit exceptionally well, and were very nearly symmetric; for example, 12 items accorded with Beta(8.86, 8.82), and 18 items with Beta(18.96, 18.95). We then aggregated all the runs, and found that the single best value for a symmetric Beta distribution was almost exactly 10 (for our specific application to four groups of five items). We therefore transformed all the conditional distributions to a Beta(10, 10), and used it to calculate individual values for each participant. Finally, we transformed each subject's score so that it is in accordance with the Beta(10, 10) distribution, so that groupwise comparisons of averages would be statistically meaningful and truly reflect differences in category clustering, not number of items mentioned. We note here that substantive results presented based on this analysis were unchanged using Beta(k, k) distributions for other values near k = 10.
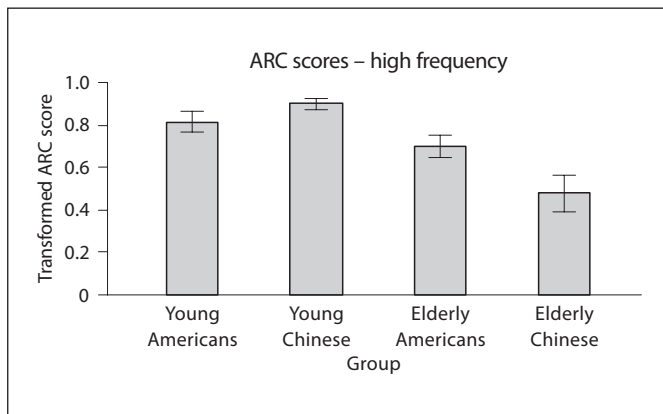
**Fig. 2.** The average transformed ARC clustering score for each group for experiment 1.

*Categorical Clustering.* For the categorized list, each individual's transformed ARC score was subjected to a 2 (Age: Young/Old) × 2 (Culture: American/Chinese) univariate ANOVA. There was a main effect of Age with young adults (mean = 0.86, SD = 0.23) exhibiting higher clustering scores than older adults (mean = 0.59, SD = 0.42), $F(1, 124) = 20.89$, MSE = 2.31, p < 0.001. Although the main effect of Culture did not reach significance, $F(1, 124) = 1.29$, MSE = 0.14, p = 0.26, the interaction of Age and Culture was significant, $F(1, 124) = 6.65$, MSE = 0.74, p < 0.02. As displayed in figure 2, the young Chinese (mean = 0.90, SD = 0.15) had transformed ARC scores that did not differ significantly from the young Americans (mean = 0.82, SD = 0.28), $t(62) = 1.51$, p = 0.14, while the elderly Chinese (mean = 0.48, SD = 0.50) had disproportionately lower transformed ARC scores relative to the American elderly (mean = 0.70, SD = 0.31), $t(62) = 2.11$, p < 0.04.

*Discussion*

The interaction between Age and Culture on transformed ARC scores supports our prediction that elderly Chinese rely less on categories as an organizational structure to support recall. Interestingly, this pattern diverges from the recall data within each age group in which both Americans and Chinese performed nearly identically. This suggests that although elderly East Asians may use a categorical organization strategy less than Americans, categorization strategy does not necessarily correspond to recall performance. Although cross-cultural differences in the use of categories could exist across age groups, they did not exert an effect on recall in this task.

## Experiment 2

The surprising disconnect between free recall and categorical clustering found in experiment 1 could reflect the strong categorical associations of the list items. The strong associations could make the categorical organization apparent and evident as a possible strategy, consistent with the finding of high transformed ARC scores. Even though the organizational scheme may not be commonly employed as a spontaneous strategy in some groups, the salience of the manipulation may have contributed to the finding in experiment 1. In order to address this possibility, in experiment 2 we selected words less strongly associated with the categories and assessed memory and categorical clustering in new samples across age and culture. This approach afforded the benefit of a more subtle manipulation that might allow cultural biases to emerge more clearly.

*Methods*

24 participants from each of four Age by Culture groups participated in the study for course credit or pay. Participants were recruited from the same population samples and through the same procedures as experiment 1. The sample characteristics are presented in table 2.

There were main effects of Age on pattern matching, education, and age. Elderly adults had more years of education than young adults, $F(1, 92) = 9.73$, p < 0.01; reported worse health,[2] $F(1, 91) = 5.50$, p < 0.05; and were slower on the pattern matching task, $F(1, 92) = 84.01$, p < 0.001. There was a marginally significant effect of culture on scaled WAIS Information scores, with a trend for higher performance by the Chinese than for the Americans, $F(1, 91) = 3.12$, p < 0.10. All other effects, including those involving interactions of culture and age, did not approach significance. As in experiment 1, samples are well matched on these measures. None of the small differences across cultures in pattern matching (experiment 1) or WAIS scores (experiment 2) extend across both studies, suggesting that there are not systematic differences in the samples on cognitive ability.

Stimuli were selected in the same manner as for experiment 1, although the lists consisted of lower frequency items with a maximum frequency of occurrence of 0.22 according to the category norms [13]. Based on these norms, the average frequency of occurrence of the selected list items ranged from 0.07 to 0.08 (SD = 0.04–0.06) for related items and from 0.07 to 0.08 (SD = 0.04–0.05) for unrelated items across the four Age by Culture groups. In addition, the category 'units of time', isolated as an age and culturally-equivalent category [13], was substituted for 'times of day' (experiment 1) in order to have a better distribution of lower frequency responses. The testing and data coding procedures were otherwise identical to those in experiment 1.

---

[2]   The health rating for one elderly American and the WAIS Information score for one young American are unavailable.

---

**Table 2.** Sampling characteristics (means and SDs) for participants in experiment 2

|  | Young American | Young Chinese | Elderly American | Elderly Chinese |
|---|---|---|---|---|
| Age | 20.08 (1.79) | 20.08 (1.02) | 68.08 (4.89) | 66.58 (5.71) |
| Years of education | 14.04 (1.50) | 13.54 (.66) | 15.19 (1.85) | 14.63 (2.48) |
| Health rating[a] | 1.92 (0.78) | 2.04 (0.81) | 2.35 (0.83) | 2.42 (0.93) |
| Pattern matching [31] | 18.73 (4.57) | 17.67 (3.94) | 10.98 (3.91) | 10.56 (3.37) |
| WAIS information [32, 33] | 11.91 (2.04) | 12.46 (1.56) | 11.50 (2.25) | 12.46 (2.36) |

[a] Health ratings reflect how frequently health problems interfere with lifestyle, as assessed on a 5-point scale (1 = 'never'; 2 = 'seldom'; 3 = 'sometimes'; 4 = 'often'; 5 = 'always').
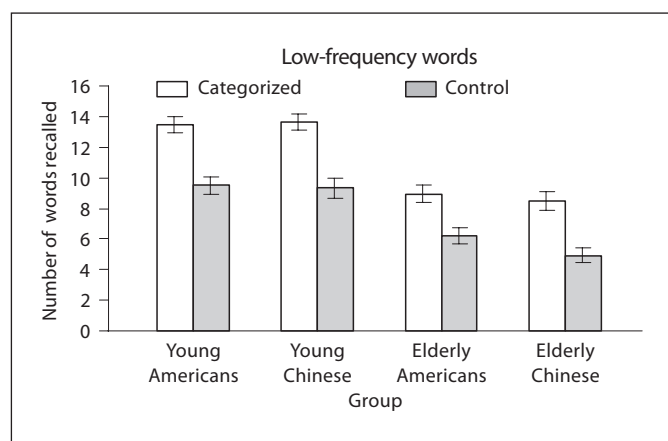


**Fig. 3.** The average number of words free recalled by each group and list type for experiment 2.
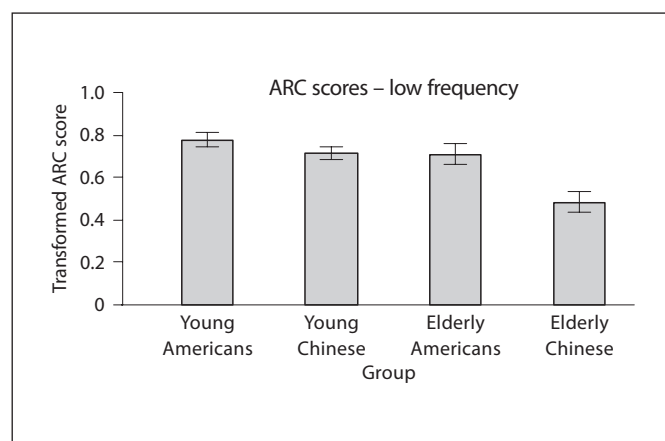


**Fig. 4.** The average transformed ARC clustering score for each group for experiment 2.

*Results*

*Number of Items Recalled.* The number of list items correctly recalled by each participant was subjected to a 2 × 2 × 2 mixed ANOVA with Age (Young/Old) and Culture (American/Chinese) as between-groups variables and List Type (Categorized/Control) as a within-subject variable. As in experiment 1, there was a significant main effect of Age, $F(1, 92) = 78.65$, MSE = 905.67, p < 0.001, and a significant main effect of List Type, $F(1, 92) = 162.94$, MSE = 634.38, p < 0.001, such that young (mean = 11.48, SD = 3.45) recalled more than elderly (mean = 7.14, SD = 3.17), and recall was higher for the categorized list (mean = 11.13, SD = 3.64) than for the control list (mean = 7.49, SD = 3.40). The only other effect to approach significance was the interaction of List Type with Age, $F(1, 92) = 2.96$, MSE = 11.51, p < 0.09. Neither the main effect nor the interactions involving culture approached significance, $Fs < 1$. Results are shown in figure 3.

*Categorical Clustering.* For the categorized list, each individual's transformed ARC score was subjected to a 2 (Age: Young/Old) × 2 (Culture: American/Chinese) univariate ANOVA. Data are presented in figure 4. The main effect of Age was significant, with the young (mean = 0.75, SD = 0.34) performing better than the elderly (mean = 0.60, SD = 0.38), $F(1, 92) = 4.35$, MSE = 0.54, p < 0.05. The main effect of Culture also reached significance, with higher clustering scores for the Americans (mean = 0.74, SD = 0.33) than the Chinese (mean = 0.60, SD = 0.38), $F(1, 92) = 4.02$, MSE = 0.49, p < 0.05. The interaction of Age and Culture did not reach significance, $F(1, 92) = 1.31$, MSE = 0.16, p = 0.26.

*Discussion*

The overall pattern of results of experiment 2 mirrors those of experiment 1. Cultural differences do not emerge in the recall of list items, even though East-Asians relied on a clustering strategy less than Americans. Although the main effect of culture suggests that young and elderly East-Asians rely less on categories to organize low frequency words, the difference is slight for young adults but more pronounced for elderly, as in experiment 1.

**General Discussion**

Overall, our results provide some support for the notion that East-Asians use categories as an organizational strategy based on taxonomic categories less than do Americans. Whereas past studies demonstrated this effect in young adults [5–7, 38] the difference was demonstrated primarily in elderly adults in the present experiments. The careful selection of words, equated across culture for both familiarity of items as well as overall structure of the category, may have prevented the expression of cultural differences in the young. Although cohort effects could be a candidate to explain why cross-cultural differences were selectively identified for elderly but not young adults, this is an unlikely explanation. In previous literature, young Americans exhibit greater use of categories compared to a number of different young East-Asian populations. These cross-cultural differences are robust, evident when young are tested in different cultures [5] and even extend to Asians spending substantial time in Western cultures [6]. These data suggest that our reported cultural differences do not result from factors affecting only older adults, such as greater emphasis on rote memory in the Chinese educational system during the period when elderly adults attended school.

The emergence of cultural differences in our present research for elderly adults, but not young, provides evidence to support Park et al.'s [8] hypothesis that prolonged absorption of a culture can lead to greater expression of cultural biases in information processing. When tasks are well practiced, such as the use of a categorization strategy by Americans, elderly will continue to use that automatic strategy. When a strategy is less familiar and thus requires more cognitive resource, elderly will not adopt it, as is the case for categorical organization for the East-Asian elderly. On the other hand, sufficient cognitive resource allows young adults from either culture to adapt the categorical organization strategy, even if it is not a preferred one.

One caveat of the results from the present study is that they are based on a fixed list order, with the unrelated list always preceding the related list. A fixed ordering allows us to have a comparable assessment of memory for the control words across groups, avoiding potential contamination from higher-order interactions if cultures differ in their expectations of categorical relationships once a related list is presented. On the other hand, the presentation of the unrelated list could make it more difficult to detect categories in the subsequent lists, which could impact the Chinese groups more than the American groups. Although this finding would be a result of cultural differences in attention to categories, it would suggest that the present findings exaggerate the magnitude of the cultural differences for elderly adults. With larger samples, memory for related and unrelated lists could be separately assessed in a between-subjects design.

Somewhat surprisingly, the use of categories as an organizational strategy across cultures did not directly correspond to the success of free recall. This suggests that clustering is likely one of many strategies and processes that contribute to successful recollection. Although use of categories contributes to cross-cultural differences in reasoning and learning [7, 38], the effect on memory may be obscured by the fact that multiple strategies can contribute to successful encoding. Elaborative processes emphasizing the distinctive features of list items can enhance encoding as much as relational processes emphasizing the associations between list items [39]. Although the clustering measure reveals cultural differences in the use of categories, there is no single organizational strategy with which to assess East-Asians' encoding strategy in the present study. If the list was structured so that both categorical and functional relationships were viable organizational strategies that could be measured, the two strategies could be pitted against each other to elucidate the strategy preference for each culture. Manipulation of lures in either a recognition or cued recall paradigm (e.g. false recognition of categorical or relational associates as in the Deese-Roediger-McDermott paradigm [40], or memory for related word pairs) might allow for discrimination of differences in semantic organization and strategy selection.

Even though clustering differences were not accompanied by memory differences across cultures, reduced clustering by older adults likely contributed to their reduced recall of both high and low frequency words. The finding of reduced categorical clustering in older adults contradicts findings of equivalent categorical organization with age [16, 18, 19]. This may be due in part to the

abstract nature of the words and categories used in the present study (e.g. chemical elements, times of day) which could make it more difficult to access or store exemplar information under the higher-order category labels [20] compared to past studies that tested clustering of concrete objects (e.g. animals, fruit). In addition, past studies used raw (untransformed) ARC scores to compare clustering across younger and older adults. Because older adults generally recalled less information, their raw ARC scores were likely inflated. The procedure used in this article corrects for this tendency, and allows us to compare degree of clustering across different levels of recall.

In conclusion, the present studies suggest that cultural differences in the use of categories as an organizational strategy are more prominent for older adults, with East Asians less likely to organize their recall by category. It is unlikely these effects are due to cohort differences due to careful sampling and similar findings of cross-cultural differences in young adults [5, 6]. Instead, our findings reflect the bias to process information less categorically in East-Asians, with the effect manifest in the older adults who have accrued years of experience in the culture. Despite the overwhelming effects of neurobiological aging, elderly adults express the signature of their culture in their use of cognitive strategies.

## References

1 Nisbett RE: The Geography of Thought: How Asians and Westerners Think Differently…and Why. New York, The Free Press, 2003.

2 Nisbett RE, Masuda T: Culture and point of view. Proc Natl Acad Sci USA 2003;100: 11163–11170.

3 Nisbett RE, Peng K, Choi I, Norenzayan A: Culture and systems of thought: holistic versus analytic cognition. Psychol Rev 2001; 108:291–310.

4 Chiu LH: A cross-cultural comparison of cognitive styles in Chinese and American children. Int J Psychol 1972;7:235–242.

5 Ji LJ, Zhang Z, Nisbett RE: Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. J Pers Soc Psychol 2004;87:57–65.

6 Unsworth SJ, Sears CR, Pexman PM: Cultural influences on categorization processes. J Cross Cult Psychol 2005;36:662–688.

7 Norenzayan A, Smith EE, Kim BJ, Nisbett RE: Cultural preferences for formal versus intuitive reasoning. Cogn Sci 2002;26:653–684.

8 Park DC, Nisbett R, Hedden T: Aging, culture, and cognition. J Gerontol [B] 1999;54B: P75–P84.

9 Bousfield WA: The occurrence of clustering in the recall of randomly arranged associates. J Gen Psychol 1953;49:229–240.

10 Park DC, Smith AD, Lautenschlager G, Earles J, Frieske D, Zwahr M, Gaines C: Mediators of long-term memory performance across the life span. Psychol Aging 1996;11: 621–637.

11 Salthouse TA: The processing-speed theory of adult age differences in cognition. Psychol Rev 1996;103:403–428.

12 Schunn CD, Vera AH: Cross-cultural similarities in category structure. Think Reason 2004;10:273–287.

13 Yoon C, Feinberg F, Hu P, Gutchess AH, Hedden T, Chen H, Jing Q, Cui Y, Park DC: Category norms as a function of culture and age: comparisons of item responses to 105 categories by American and Chinese adults. Psychol Aging 2004;19:379–393.

14 Kausler DH: Learning and Memory in Normal Aging. San Diego, Academic Press, 1994.

15 Light LL: Memory and aging: four hypotheses in search of data. Annu Rev Psychol 1991; 42:333–376.

16 Park DC, Smith AD, Dudley WN, Lafronza VN: Effects of age and a divided attention task presented during encoding and retrieval on memory. J Exp Psychol Learn Mem Cogn 1989;15:1185–1191.

17 Wingfield A, Kahana MJ: The dynamics of memory retrieval in older adulthood. Can J Exp Psychol 2002;56:187–199.

18 Kahana MJ, Wingfield A: A functional relation between learning and organization in free recall. Psychon Bull Rev 2000;7:516–521.

19 Wingfield A, Lindfield KC, Kahana MJ: Adult age differences in the temporal characteristics of category free recall. Psychol Aging 1998;13:256–266.

20 Hultsch DF: Adult age differences in retrieval: trace-dependent and cue-dependent forgetting. Dev Psychol 1975;11:197–201.

21 Witte KJ, Freud JS, Brown-Whistler S: Adult age differences in free recall and category clustering. Exp Aging Res 1993;19:15–28.

22 Becker S, Lim J: A computational model of prefrontal control in free recall: strategic memory use in the California Verbal Learning Task. J Cogn Neurosci 2003;15:821–832.

23 Raz N, Gunning FM, Head D, Dupuis JH, McQuain J, Briggs SD, Loken WJ, Thornton AE, Acker JD: Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. Cereb Cortex 1997;7:268–282.

24 Butler KM, McDaniel MA, Dornburg CC, Roediger HL, Price AL: Age differences in veridical and false recall are not inevitable: the role of frontal lobe function. Psychon Bull Rev 2004;11:921–925.

25 Glisky EL, Polster MR, Routhieaux BC: Double dissociation between item and source memory. Neuropsychology 1995;9:229–235.

26 Chan AS, Poon MW: Performance of 7- to 95-year-old individuals in a Chinese version of the category fluency test. J Int Neuropsychol Soc 1999;5:525–533.

27 Troyer AK, Moscovitch M, Winocur G: Clustering and switching as two components of verbal fluency: evidence from young and older healthy adults. Neuropsychology 1997; 11:138–146.

28 Gutchess AH: Aging, culture, and cognition; in Park DC, Schwarz N (eds): Cognitive Aging: A Primer, ed 2. Philadelphia, Psychology Press, in press.

29 Hedden T, Park DC, Nisbett R, Ji LJ, Jing Q, Jiao S: Cultural variation in verbal versus spatial neuropsychological function across the life span. Neuropsychology 2002;16:65–73.

30 Park DC, Hedden T, Jing Q, Jiao S, Lautenschlager G, Nisbett RE: Culture and the Aging Mind. Unpublished manuscript, University of Illinois, Urbana-Champaign, IL.

31 Salthouse TA: General and specific speed mediation of adult age differences in memory. J Gerontol [B] 1996;51:P30–P42.

32 Wechsler D: Wechsler Adult Intelligence Scale – Revised. San Antonio, Psychological Corporation, 1981.

33 Gong Y: Revision of Wechsler's Adult Intelligence Scale in China. Acta Psychol Sin 1983;15:362–370.

34 Kucera H, Francis WN: Computational Analysis of Present-Day American English. Providence, Brown University Press, 1967. Accessed via the web: http://www.psych.rl.ac.uk/MRC_Psych_Db_files/psych.htm

35 Wang W, Chang BR, Li YS, Lin LH, Liu J, Sun YL, Wang ZW, Yu YX, Zhang JW, Li DP (eds): Xiandai Hanyu Pinlu Cidian [Modern Chinese frequency dictionary]. Beijing, Beijing Language Institute Press, 1986.

36 Puff CR (ed): Handbook of Research Methods in Human Memory and Cognition. New York, Academic Press, 1982.

37 Roenker, DL, Thompson, CP, Brown, SC: Comparison of measures for the estimation of clustering in free recall. Psychol Bull 1971;76:45–48.

38 Choi I, Nisbett RE, Smith EE: Culture, categorization and inductive reasoning. Cognition 1997;65:15–32.

39 Hunt RR, Einstein GO: Relational and item-specific information in memory. J Verb Learn Verb Be 1981;20:497–514.

40 Roediger HL III, McDermott KB: Creating false memories: Remembering words not presented in lists. J Exp Psychol Learn Mem Cogn 1995;21:803–814.