

Cross-cultural similarities in category structure

Christian D. Schunn

University of Pittsburgh, USA

Alonso H. Vera

NASA Ames Research Center, Moffett Field, CA, USA

Categories, as mental structures, are more than simply sums of property frequencies. A number of recent studies have supported the view that the properties of categories may be organised along functional lines and possibly dependency structures more generally. The study presented here investigates whether earlier findings reflect something unique in the English language/North American culture or whether the functional structuring of categories is a more universal phenomenon. A population of English-speaking Americans was compared to a population of Cantonese-speaking Hong Kong Chinese. The findings clearly support the view that functional influences on category centrality are universal (or at least common to Cantonese-speaking Hong Kong Chinese and English-speaking Americans), albeit with specific cross-cultural/cross-linguistic group differences in the particular properties that are considered central to categories.

Previous research on how we categorise suggests that some features of a category are more important or more central than other features (Rosch, 1973). For example, a pen that is not straight is still fairly pen-like, whereas a pen that not used for writing is much less pen-like. Although prototypicality of features (i.e., how often examples of the category happen to have that feature) was initially thought to drive feature centrality, later research showed that there must be more to feature centrality. For example, being curved is very central to the category of boomerangs but not very central to the category of bananas, even though being curved is highly typical of both boomerangs and bananas (Medin & Shoben, 1988; Murphy & Medin, 1985).

Correspondence should be addressed to Christian Schunn, LRDC 821, 3939 O'Hara Street, University of Pittsburgh, Pittsburgh, PA 15260, USA. Email: schunn@pitt.edu

Work on this manuscript was supported by Grant DASW01-00-K-0017 from the US Army Research Institute and Grant N00014-01-1-0321 from the US Office of Naval Research to the first author. We would like to thank the reviewers for their helpful comments.

Considerable research has been devoted recently to determining what underlies category centrality. Some have argued that causal relationships among features are the important factor: features that cause other features are thought to be more central (Ahn, 1998, 1999; Ahn, Kim, Lassaline, & Dennis, 2000; Rehder & Hastie, 2001). For example a disease symptom that causes other disease symptoms is thought to be more central to the category of that disease. Others have argued that the functional role played by the features is the important factor: features that help the object/entity achieve its primary goals are thought to be more central (Schunn & Vera, 1995). For example, having paws is central to the category of cat because paws help cats catch mice and survive; having a seat helps chairs be a good place to sit on. Finally, some researchers have argued that dependency relationships more generally (whether causal, functional, temporal, or other) are the key determining factors of property centrality (Sloman, Love, & Ahn, 1998). While there are clearly some important differences among these accounts, it is also clear that there are very strong inter-relationships among these accounts. Functions are achieved through causal relationships; function and causal relationships are the most common form of dependency relationships; and people may often conceive of causal relationships that mediate apparently acausal relationships among properties. Our current focus is not to attempt to distinguish these three accounts, but instead to provide further evidence for this general type of account across cultures.

A wide variety of previous studies, including experimental studies with novel categories, correlational studies with highly familiar and newly learned categories, and computational modelling work, have all supported this general causal/functional/dependency structure account of feature centrality. However, all of these studies have been done with English speakers in North American culture. Recently, Richard Nisbett and colleagues have presented a wide array of results suggesting that there are many differences between the categorisation, social inference, and general reasoning styles of East Asians versus North Americans. They claim that North Americans view social behaviour as a direct unfolding of dispositions and use analytic, formal-logic, rule-based approaches to categorisation and reasoning, whereas East Asians understand behaviour in terms of complex interactions between person/object characteristics and contextual factors and use holistic, "dialectical" approaches to categorisation and reasoning (Choi, Nisbett, & Norenzayan, 1999; Nisbett, Peng, Choi, & Norenzayan, 2001; Norenzayan & Nisbett, 2000). Moreover, East Asians were reported to be more sensitive to (and more confident in their judgements of) covariations in the environment (Ji, Peng, & Nisbett, 2000a; Ji, Schwarz, & Nisbett, 2000b).

These studies, then, very much raise the question of whether the past findings on the structure of categories and the central role of causal/functional factors will turn out to be a general property of human

categorisation, or whether they are a consequence of how North Americans are taught to think about the world or how English tends to subdivide the world.¹ On the one hand, it may be that the causal/functional theories continue to play a central role in determining category structure for East Asians as well, but only the particular details of what is believed to be causal/functional (and the corresponding category structure) are different across cultures. For example, perhaps the East Asian causal theories about human behaviour include different properties (such as situational vs dispositional factors), and the particular features thought to be central of a category are changed to reflect these differences in theories. This result would be similar to the findings of Atran and colleagues (Atran et al., 2002; Ross, Medin, Coley, & Atran, 2003; Sousa, Atran, & Medin, 2002). They found Yukatek Maya children and urban-USA children have very different biological categories, with Yukatek Maya children not interpreting the biological world anthropocentrically in the way that urban-USA children do. However, at a higher level, both groups of children had essentialist categories that allowed them to project properties onto new biological instances, and differences in the details reflected level of experience with non-human animals.

On the other hand, it may be that East Asian category structure, in contrast to North American category structure, really does not depend so heavily on functional theories. For example, perhaps the East Asian heavier emphasis on context (as found by Choi et al., 1999) heavily dilutes any general, acontextual category structure patterns. Or perhaps the East Asian stronger sensitivity to covariation patterns in the environment (as found by Ji et al., 2000a) produces a more dominant role of property typicality (how frequently a feature is associated with category instances).

The current study seeks to examine whether the prior category structure results generalise to East Asians. We also collect data from the standard North American/English group following the same procedure to make our cross-linguistic/cross-cultural samples directly comparable. We gather property centrality ratings using the standard procedures (which have been shown to correlate well with other methods of gathering centrality ratings; Sloman et al., 1998). We then gather functional ratings and typicality ratings to see whether these factors predict centrality ratings.

Note that the current study does not attempt to tease apart linguistic effects from cultural effects. We compare the case of Hong Kong residents

¹Recent work has shown that name centrality and category centrality are distinguishable (Ahn & Sloman, 1997; Sloman & Ahn, 1999; Sloman et al., 1998), suggesting that the category centrality that has been examined in previous studies cannot be entirely or directly a function of the structure of English. However, it could be that the structure of English has a more generalised effect on thought and thus produces this kind of category centrality structure.

asked about categories in Cantonese with US residents asked about categories in English. Other contrasts may more clearly separate culture and language, but they also run the risk of other problems. For example, fully bilingual participants may have atypical categories through exposure to multiple languages. Similarly, responses from not-fully bilingual participants outside their primary language may suffer from mistranslation problems. The current design provides the broadest generalisation test by varying both language and culture.

SELECTED CATEGORIES

Because possible differences across cultures may reflect differences in general reasoning style (e.g., holistic vs analytic or attention to covariation patterns) or reasoning with respect to human-related situations (e.g., dispositional vs situational/contextual properties), we will examine a variety of kinds of categories to include categories more purely of the natural world (e.g., natural kinds) and categories more purely of human invention (e.g., artifacts, nominal kinds, events). Natural kinds are things that exist independently of human creation or definition such as biological kinds (e.g., lions, sharks). Artifacts are things created by humans for a purpose (e.g., cars, chairs, hammers). Nominal kinds are categories that result from human definitions (e.g., even numbers, uncles). The research literature frequently divides category types along these lines even though modern life contains many boundary cases (e.g., domesticated cats or bioengineered blood).

We selected nine different categories from four different category types to represent a broad range of object and event category types. There were two nominal kinds (fathers, islands²), two natural kinds (cats, coal), two artifacts (ovens, robots), and three events (weddings, thunderstorms, picnics). Because events are a broad class with less well-understood subtypes than objects, we wanted to select instances that roughly correspond to nominal, natural, and artifact categories for objects. But this distinction is less clear for events than objects and thus we do not focus on distinctions within the event type. Within each type, the specific categories were selected such that they have roughly similar meanings and levels of specificity in both American English and in Hong Kong Cantonese.

The experiment reported here investigated what drives property centrality, and collected data roughly simultaneously from American

²Although the category of island has some features of natural kinds, it is most commonly treated as a nominal kind (Keil, 1989) because it has a clear, consensual definition and no internal essence beyond that definition (e.g., no underlying essence like DNA that makes a given exemplar an island, no geological process that makes an island different from continents or peninsulas or submerged islands).

English speakers and Hong Kong Cantonese speakers. All materials were created first in English. Then the English materials were translated into Cantonese by a native Cantonese speaker. To verify the translations, another native Cantonese speaker translated the materials back into English. Discrepancies were addressed and the translation, reverse-translation process was repeated.

EXPERIMENT

Method

Participants. Across the various conditions and tasks, 86 Carnegie Mellon University undergraduates (all selected to be native English speakers) and 93 University of Hong Kong (HK) undergraduates (all selected to be native Cantonese speakers) participated for course credit ($n = 40, 16, 15,$ and 15 for the generation, centrality, typicality, and functionality conditions in the US, and $n = 40, 18, 17,$ and 18 for the same conditions in HK). Each participant took part in either the generation phase or one rating condition.

Property generation. Participants were each given nine pieces of paper with each category listed at the top and with 20 blank lines down the page. Participants were asked to write as many properties as they could think of for each category, trying to generate at least six properties for each category, and spending about 2–3 minutes on each category.³

Property selection. For each of the nine categories, 18 properties were selected for the next phase of the study using the property generation data (e.g., see Table 1). These 18 properties were selected to represent a range in terms of how typical, definitional, and functional they were, as well as including some properties that seemed high on one dimension (e.g., typicality) and low on another dimension (e.g., functionality). This selection method was used to reduce problems of collinearity for the later regression analyses, although it is impossible to remove collinearity entirely.

Property ratings. There are many ways of measuring property centrality. The most common methods involve property negation, in which participants are asked to consider the relative impact of removing different

³This method of property generation is similar to that used in the working paper (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) that was used as a source of properties for Sloman and Ahn (1999) and Sloman et al. (1998).

TABLE 1

The properties for the Islands category along with frequency-of-mention, centrality, typicality, and functionality means in the US and HK data

<i>Property</i>	<i>US</i>				<i>HK</i>			
	<i>Freq. of mention</i>	<i>Cent.</i>	<i>Typ.</i>	<i>Fn.</i>	<i>Freq. of mention</i>	<i>Cent.</i>	<i>Typ.</i>	<i>Fn.</i>
have beach	39	6.2	6.1	5.3	12	5.5	5.9	5.1
surrounded by sea	38	6.3	6.6	6.1	16	6.6	6.7	6.8
ship as means of transportation	14	5.2	5.7	5.1	18	6.5	5.9	6.0
silent	2	3.4	3.3	2.3	1	6.0	5.3	3.8
low population density	10	3.6	4.9	2.9	1	5.7	4.8	3.8
simple life	3	3.5	4.2	2.6	1	5.5	5.1	3.6
rocky shore	0	3.7	4.2	3.6	1	5.7	6.0	5.1
good camp-site	0	3.0	4.5	2.5	13	4.3	4.8	3.7
few cars	0	4.2	5.0	3.5	1	5.2	5.1	4.1
independent	1	3.4	4.3	3.7	9	4.5	4.9	4.2
not being destroyed	0	3.5	4.2	3.3	1	4.6	3.8	3.8
have hill	0	3.8	4.9	2.7	10	6.2	5.8	4.8
desolate	6	3.4	3.3	2.4	7	3.5	3.9	2.8
few houses	0	3.1	4.0	2.5	1	5.1	4.2	3.2
archaeological value	0	3.0	3.7	1.8	1	4.3	3.9	2.9
green	2	4.3	4.7	3.3	6	5.9	5.3	5.1
swimming place	6	5.6	5.3	4.1	7	4.4	4.7	3.9
village house	1	2.7	4.2	2.6	1	4.8	5.1	3.8

Freq. of mention = frequency of mention; Cent. = centrality; Typ. = typicality; fn. = functionality.

single features (e.g., an apple that is not red, an apple that is not edible, etc.). Centrality is difficult to measure using only a single question because centrality is an abstract concept that does not map perfectly into a single concrete question. One solution is to have different participants answer different questions tightly related to centrality (like surprise, ease of imagination, and similarity to ideal) as do Sloman et al. (1998). Fortunately, these multiple methods tend to correlate highly with one another. Instead we asked participants to jointly consider multiple vectors into centrality at the same time. For example, participants were asked, "Suppose you could change just one property of a thunderstorm so that it is *not cold*. How good a thunderstorm would it be? That is, how good or bad would it then be as an example of its category? How different a kind of thing is it from thunderstorms? How much does removing this property change your concept of it?" Participants made their response on a 7-point Likert scale ranging from "Good example of the category" to "Bad example of the category".

Property typicality judgements (also known as category validity or property frequency judgements) were obtained by participants to estimate what proportion of category instances have each property. Participants made their responses on a 7-point Likert scale ranging from "None" to "All".

The ratings instructions for the functionality judgements were tailored slightly for each category. For all the events, participants rated how important each property was for the occurrence of that event. For all the objects except cats, participants rated how important each property was for the object to be functionally useful as an instance of that category (e.g., functionally useful as coal). For cats, which did not have as straightforward a function, participants rated how important each property was for the success of cats as a species. In all cases, participants made their responses on a 7-point Likert scale ranging from "Not important" to "Very important".

Results and discussion

Rating consistency. Four of the US participants and three of the HK participants in the centrality condition appeared to have reversed their ratings such that their ratings consistently correlated negatively with the other participants' ratings. Thus, their data were removed from the analyses. The remaining participants tended to produce quite similar ratings for each item. The mean correlation between each US participant and the average ratings (at the category level) was .77, .75, and .70 for the centrality, typicality, and functionality ratings respectively. The corresponding

numbers for the HK participants were .56, .63, and .65. For the remaining analyses, the mean property ratings for each item were used.⁴

The standard deviations in ratings across participants for a given property varied somewhat by ratings task. For the Hong Kong dataset, the mean standard deviations were 1.29 for typicality, 1.49 for functionality, and 1.68 for centrality. For the US dataset, the mean standard deviations were 1.08 for typicality, 1.61 for functionality, and 1.59 for centrality. That the variability was approximately equally high for the functionality and centrality ratings argues that asking multiple questions for centrality (and not for typicality or functionality) did not cause significant confusion in the participants. As another point of note, the highest consistency for typicality ratings may produce a modest bias overall in favour of typicality over functionality towards being a better predictor of centrality (or any other measure).

Property centrality. Combining all the property judgements for all the categories into one analysis (but separately by dataset), both property typicality and property functionality are strongly correlated with property centrality (US $r = .83$ and $r = .83$, $n = 162$, $ps < .0001$; HK $r = .73$ and $r = .83$ respectively, $n = 162$, $ps < .0001$), and with each other (US $r = .82$, $n = 162$, $p < .0001$; HK $r = .81$, $n = 162$, $p < .0001$). In a multiple regression, both property typicality and property functionality predict independent variance in property centrality (see Table 2). The correlations of typicality with centrality partialling out functionality are significant in both datasets (US partial $r = .46$, $p < .001$; HK partial $r = .18$, $p < .001$), as are the correlations of functionality with centrality partialling out typicality (US partial $r = .46$, $p < .001$; HK partial $r = .60$, $p < .001$).

Placing both datasets into a common multiple regression, we find that there is an interaction between dataset and function, such that the overall relationship between function and centrality is statistically stronger in the HK data than in the US data, $F(1, 316) = 8.48$, $Mse = 0.36$, $p < .005$. Note, however, that in both datasets, function plays a strong role.⁵

The left half of Table 2 presents the beta weights for these regression analyses done by category type for the US data and HK data separately. These regressions are done by category type rather than at the level of individual categories, because even 18 features per category is not a large enough n to

⁴It is not surprising that it was only the centrality condition that included participants who consistently reversed the dimension because it was the only condition with an inverse scale: high numbers meant not very central. By contrast, high numbers meant highly functional or highly typical for the other scales.

⁵Because the size of this interaction is not very large, and the power to detect such small quantitative interactions in correlation relationships is very poor for small ns , we do not formally assess this interaction at the category or category type level.

TABLE 2
Multiple regression standardised weights of typicality and functionality judgements predicting centrality judgements done separately for each category type and each dataset

<i>Category type</i>		<i>Centrality βs</i>	
		<i>Typicality</i>	<i>Functionality</i>
US Data	Overall	0.45 ***	0.45 ***
	Artifact	0.38 *	0.43 *
	Natural kinds	0.60 **	0.39 **
	Nominal kinds	0.26 †	0.61 **
	Event	0.44 **	0.49 **
HK Data	Overall	0.17 *	0.70 ***
	Artifact	0.24 †	0.60 **
	Natural kinds	0.36 *	0.56 **
	Nominal kinds	0.45 *	0.47 *
	Event	0.31 *	0.59 **

Note. *** $p < .0001$, ** $p < .001$, * $p < .05$, † $p < .15$

support reliable multiple regression with the observed levels of collinearity between the two predictors. However, we will discuss the outcomes of the multiple regressions conducted at the level of individual categories to see if there are suggestions of differences. The reader may be tempted to make inferences based on the differential size of the beta weights within each category, but the confidence intervals on beta weights are too large relative to the observed variability to support that category level type of inference.

For the US data, in all category types except nominal kinds, both factors contributed significantly (all $ps < .05$). For nominal kinds, the independent contribution of typicality was much weaker ($p < .15$). Only for natural kinds did typicality predict more independent variance than functionality—this pattern held true for both the biological (cats) and non-biological (coal) categories within natural kinds.

For the HK data, in all category types except artifacts, both factors contributed significantly (all $ps < .05$). For artifacts, the independent contribution of typicality was much weaker and only marginally significant ($p < .15$). In all cases, functionality was the best predictor of centrality.

Although the ns are small enough to produce unstable multiple regression coefficient estimates, the patterns generally held at the individual category level. In the US data, all but three of the categories (cats, robots, and weddings) had statistically significant roles of functionality in the multiple regression ($ps < .05$)—two of those three were marginally significant ($ps < .2$). Because the three categories were all from different category types (natural kind, artifact, and event/nominal kind), it is likely that the

nonsignificance of the independent contributions of functionality can be attributed to noise. In the HK data, once again, all but three of the categories had statistically significant roles of functionality in the multiple regression ($ps < .05$). This time it was a different set of three categories that did not follow the pattern: coal, islands, and picnics. But once again, the three exceptional categories were all from different category types, suggesting that the nonsignificance of the independent contributions of functionality can be attributed to noise rather than to differences in category types.

The focus of our analysis was on the relative strength of typicality versus functionality within a dataset. While it is harder to interpret differences in coefficient size across datasets, we can note some general patterns. Looking across Table 2, it appears that, except for nominal kinds, there is a general pattern for functionality to play a stronger role in the Hong Kong dataset than in the US dataset, and for typicality to play a weaker role in the Hong Kong dataset.

Cross-cultural/cross-linguistic comparisons. A stronger test of the role of function and typicality in category structure involves the predictiveness of function and typicality in predicting differences in category structure across cultures/languages. Overall, US centrality and HK centrality judgements correlated $r = .61$ (36% shared variance). Thus, while the cultures/languages generally agreed in their category structure, there were also many differences.

What predicts these differences in property centrality? Perhaps it is differences in typical instances of each category in each culture. Alternatively, it may be differences in how the categories are perceived to function/be used in each culture/language. A regression analysis was conducted, this time predicting centrality differences using typicality differences and functionality differences as predictors.⁶ Both typicality and functionality differences were independent predictors ($\beta = .33$, $p < .0001$, and $\beta = .32$, $p < .0001$).

We also examined the cross-cultural differences on a case-by-case basis to examine the plausibility of the differences (e.g., possible explanations of differences as they relate to differences in daily life between the US and Hong Kong). Table 3 presents examples that were selected by searching for cases in which a property was near the top in centrality for one culture, but noticeably lower in centrality in the other culture. Note that these examples were not selected for their interpretability. In looking over all the cases of differential feature centrality in our data, in no case could we find a culture

⁶To compute these difference scores, each measure was first standardised by subtracting the mean and dividing by the standard deviation for that measure.

TABLE 3
 Examples of cross-cultural differences (US vs. HK) in ranking of centrality of features, corresponding differences in typicality and functionality, and plausible explanations of the differences

<i>Category</i>	<i>Feature</i>	<i>Centrality</i>		<i>Typicality</i>		<i>Functionality</i>		<i>Explanation</i>
		<i>US</i>	<i>HK</i>	<i>US</i>	<i>HK</i>	<i>US</i>	<i>HK</i>	
Island	Have a beach	2 nd	7 th	6.2	5.9	6.1	5.1	In Hong Kong, many beaches are very polluted and thus lose their functional value.
Island	Have swimming place	3 rd	15 th	5.3	4.7	4.1	3.9	In Hong Kong, many beaches are very polluted and thus not good for swimming.
Island	Silent	13 th	4 th	3.3	5.3	3.3	3.8	Relatively-speaking, people in Hong Kong find other islands quite quiet.
Cats	Meowing	6 th	2 nd	6.6	6.5	5.2	5.8	In Hong Kong, there are many feral cats, increasing the functional value of attracting other cats.
Cats	Cat's eye	4 th	1 st	6.5	6.5	5.2	6.3	In Hong Kong, there are many feral cats, increasing the value of seeing predators and prey.
Weddings	Man & woman	10 th	3 rd	6.3	6.2	4.9	6.7	Difference in attitude (at least in college students) towards same-sex marriages.
Picnics	Countryside	10 th	1 st	4.1	6.6	4.2	6.6	In Hong Kong, the city is very crowded and without large parks. By contrast, in the US there are many city parks that are used for picnics.
Picnics	Sunshine	3 rd	7 th	6.0	5.7	6.0	5.7	Hong Kong is very warm and has many cloudy days, and thus it is more common and possible to have a picnic without sunshine.
Thunderstorms	Heavy rain	7 th	1 st	6.1	6.8	5.9	6.5	Very heavy rains are extremely common in Hong Kong, and thus may be perceived as being causal in the production of a thunderstorm.

difference in centrality that was not also accompanied by a corresponding difference in ratings of functionality or functionality and typicality.

The differences shown in Table 3 are all described in terms of cultural rather than linguistic terms (i.e., differences in experience rather than meanings or connotations of words). That we could find cultural explanations for each difference in ratings is suggestive of a cultural basis of the effect. However, we cannot rule out a linguistic basis of the effect because our experimental design confounds language and culture.

GENERAL DISCUSSION

Factors underlying category centrality

Consistent with previous research, both culture datasets revealed a strong role of both function and typicality in predicting property centrality, across natural kinds (things that exist independently of humans and thus with only projected function), artifacts (things that should be defined entirely by function), nominal kinds (things that should be defined by definition and not by typicality or function), and events. In sum, we can find no evidence that East Asians do not also have the strong role of functional factors in their category structures that has been found so often in studies of North Americans. If anything, there may be an even stronger role of functional factors in East Asians, and there is no evidence of a weakening of category structure (e.g., due to strong contextual-driven reasoning) as one might have expected from the work of Ji and colleagues (Ji et al., 2000a, 2000b). Thus, our data suggest that property centrality is strongly determined by functional relationships between the properties and how they typically appear to us, across all categories types. However, future research should examine more categories within each of these category types to show that these relationships are indeed pervasive across these category types.

Sloman and Ahn (1999) argued that typicality (category validity) or variability—which is a transformation of typicality, $\text{typicality} * (1 - \text{typicality})$ —was important for name centrality (what properties are important for whether you would use the given category label) and not for conceptual centrality (whether the properties are important to the ideal conception of the category). By contrast, our own work in this study and in Schunn and Vera (1995) has found that both typicality and function contribute to predicting conceptual centrality. It could be that some of our participants were in fact rating the name centrality of properties rather than the conceptual centrality of properties. However, for basic-level categories, like the ones used in our experiment, both name and conceptual centrality tend to be quite similar (Sloman & Ahn, 1999; Sloman et al., 1998), and thus

one would expect roles of typicality and function in predicting conceptual centrality.

Cross-cultural/cross-linguistic comparisons

In comparing the US and HK data, we found that the categories in both languages/cultures had independent components of typicality and functionality in category centrality. This similarity in structure does not mean that the categories had similar meanings across the cultures/languages, but rather than one could account for differences in feature centrality in a given category by differences in perceptions of feature typicality and function. The cross-cultural comparisons presented here suggest that this bias towards functionality is not specific to categories in the English language or to North American college students.

There were a few differences in the results across the data sets. One general difference was that function appeared to have an even stronger role in the Hong Kong dataset. Interestingly, prior research on East Asian–American differences in reasoning would have predicted the opposite effect (Ji et al., 2000a, 2000b). It may be that the East Asian emphasis on context in reasoning also produces a more utilitarian perspective on category structure. Alternatively, the particular categories selected for this study may have greater functional relevance for the Hong Kong Chinese. But because the differences were quantitative and not qualitative, and function plays such a strong role in both datasets, one would want to see these differences replicated before drawing strong conclusions.

Some previous work on cross-linguistic category structure found that similarity judgements of objects—common containers—have similar structure across languages even when naming patterns differ across languages (Malt, Sloman, Gennari, Shi, & Wang, 1999). Subsequent studies found that asking people to provide verbal encodings of events changed their representations of similarity structure in linguistic-specific ways (Gennari, Sloman, Malt, & Fitch, 2002). However, as those authors argued, the verbal encodings likely directed attention to certain aspects of events. In other words, the linguistic effects were not directly on how categories were structured but rather through how events are encoded, followed by a common event-structuring process. Thus, these findings are consistent with our finding that the conceptual core of categories follows a common process that is linguistic-independent, even when cultural and linguistic experiences lead people to name things differently, and even across cultures that have significant differences in their cognition about causality (Choi et al., 1999). In particular, unlike the Ji et al. (2000b) findings of greater emphasis on covariation in reasoning in East Asians, the current study found the same or possibly less emphasis on simple covariation in category structure.

On functional theories of category centrality

As stressed in the introduction, the goal of this paper is not to advance the functionality theory of category centrality, nor to tease apart the different related theories of category centrality. Schunn and Vera (1995) should be consulted for a more detailed discussion of the functionality theory. However, a few words are in order here, as functionality is a complex construct (just as causality is a complex construct). In particular, functionality clearly has a different meaning across category types. For example, functionality for biological kinds must be different from that of artifacts (as reflected in the differences in our functionality ratings instructions across category types) and prior research on category structure of biological kinds necessitates a different treatment (Keil, 1989; Springer & Keil, 1991). The functionality theory acknowledges this difference and states that functionality, however best defined for a given general category type, is the core component of category centrality. This complication makes functionality theory less parsimonious at one level than causality or mutability approaches to category centrality; but the mind is not always a parsimonious entity. Future work should be directed at further teasing apart these different approaches to category centrality.

Manuscript received 25 September 2003
Revised manuscript received 12 April 2004

REFERENCES

- Ahn, W. K. (1998). Why are different features central for natural kinds and artifacts?: The role of causal status in determining feature centrality. *Cognition*, *69*, 135–178.
- Ahn, W. K. (1999). Effect of causal structure on category construction. *Memory & Cognition*, *27*(6), 1008–1023.
- Ahn, W. K., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, *41*(4), 361–416.
- Ahn, W. K., & Sloman, S. A. (1997). Distinguishing name centrality from conceptual centrality. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Atran, S., Medin, D., Ross, N., Lynch, E., Vapnarsky, V., Ek, E. U. et al. (2002). Folk ecology, cultural epidemiology, and the spirit of the commons: A garden experiment in the Maya lowlands, 1991–2001. *Current Anthropology*, *43*(3), 421–450.
- Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*, 47–63.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*(1), 49–79.
- Ji, L. J., Peng, K., & Nisbett, R. E. (2000a). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, *78*(5), 943–955.

- Ji, L. J., Schwarz, N., & Nisbett, R. E. (2000b). Culture, autobiographical memory, and behavioral frequency reports: Measurement issues in cross-cultural studies. *Personality and Social Psychology Bulletin*, *26*(5), 585–593.
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory & Language*, *40*, 230–262.
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*, 158–190.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.
- 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., & Nisbett, R. E. (2000). Culture and causal cognition. *Current Directions in Psychological Science*, *9*, 132–135.
- Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, *130*, 323–360.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and acquisition of language* (pp. 111–144). New York: Academic Press.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). *Basic objects in natural categories* (Working Paper No. #43). UC Berkeley, CA: The Language Behavior Research Laboratory.
- Ross, N., Medin, D., Coley, J. D., & Atran, S. (2003). Cultural and experimental differences in the development of folkbiological induction. *Cognitive Development*, *18*(1), 25–47.
- Schunn, C. D., & Vera, A. H. (1995). Causality and the categorization of objects and events. *Thinking & Reasoning*, *1*(3), 237–284.
- Sloman, S. A., & Ahn, W. K. (1999). Feature centrality: Naming versus imagining. *Memory & Cognition*, *27*(3), 526–537.
- Sloman, S. A., Love, B., & Ahn, W. K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*(2), 189–228.
- Sousa, P., Atran, S., & Medin, D. (2002). Essentialism and folkbiology: Evidence from Brazil. *Journal of Cognition and Culture*, *2*(3), 195–223.
- Springer, K., & Keil, F. C. (1991). Early differentiation of causal mechanisms appropriate to biological and nonbiological kinds. *Child Development*, *62*, 767–781.