



Published in final edited form as:

*Cogn Psychol.* 2009 June ; 58(4): 487–524. doi:10.1016/j.cogpsych.2008.12.001.

## Of Substance: The Nature of Language Effects on Entity Construal

**Peggy Li,**  
Harvard University

**Yarrow Dunham,** and  
University of California, Merced

**Susan Carey**  
Harvard University

### Abstract

Shown an entity (e.g., a plastic whisk) labeled by a novel noun in neutral syntax, speakers of Japanese, a classifier language, are more likely to assume the noun refers to the substance (plastic) than are speakers of English, a count/mass language, who are instead more likely to assume it refers to the object kind (whisk; Imai and Gentner, 1997). Five experiments replicated this language type effect on entity construal, extended it to quite different stimuli from those studied before, and extended it to a comparison between Mandarin-speakers and English-speakers. A sixth experiment, which did not involve interpreting the meaning of a noun or a pronoun that stands for a noun, failed to find any effect of language type on entity construal. Thus, the overall pattern of findings supports a non-Whorfian, language on language account, according to which sensitivity to lexical statistics in a count/mass language leads adults to assign a novel noun in neutral syntax the status of a count noun, influencing construal of ambiguous entities. The experiments also document and explore cross-linguistically universal factors that influence entity construal, and favor Prasada's (1999) hypothesis that features indicating non-accidentalness of an entity's form lead participants to a construal of object-kind rather than substance-kind. Finally, the experiments document the age at which the language type effect emerges in lexical projection. The details of the developmental pattern are consistent with the lexical statistics hypothesis, along with a universal increase in sensitivity to material kind.

### Introduction

At least since the time of the pioneering work of Benjamin Whorf (1956) and Edward Sapir (1949), linguists, psychologists and philosophers have speculated about the ways in which the differences in the languages we speak might cause differences in the ways we think. Researchers distinguish strong Whorfian effects of language on thought from weaker ones. Strong Whorfian effects reflect learning episodes in which acquiring some linguistic device makes possible the formulation of concepts not previously entertainable (see Carey, in press, for several case studies of strong Whorfian effects of language on thought). Weaker Whorfian effects reflect language's influencing the probability of one construal of the world over another,

---

Correspondence can be sent to P. Li, Laboratory for Developmental Studies, Harvard University, 25 Francis Ave, Cambridge, MA 02138, USA. pegas@wjh.harvard.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

in the face of the capacity to represent both construals (see Boroditsky, Schmidt, & Phillips, 2003 or Casasanto & Boroditsky, 2003, for convincing examples).

One of Whorf's original examples of possible influences of language on thought concerned the origin and cross-cultural universality of the ontological distinction between objects and substances, along with quantificational consequences of categorization of an entity as a member of an object kind vs. as a member of a substance kind (Quine 1960; Lucy 1992; Soja, Carey, & Spelke, 1991; Imai & Gentner 1997, Whorf, 1939/1956). The object-substance distinction is a conceptual one; we could conceive of the *same* perceptual entity (a wooden table) as either a kind of *object* (a table) or as a kind of *substance* (some wood). The distinction is absolutely fundamental to our thought. Two of Aristotle's four explanation categories (formal cause and material cause) implicate it. Completely different types of properties are relevant to categorization in the two cases (e.g., shape dependent function in the case of objects vs. texture and color in the case of substances) and completely different properties of entities are determined by their membership in each kind (e.g., whether one can sit and eat at a given entity or whether it will make a good campfire).

Additionally, construing an entity as a member of an object kind rather than as a member of substance kind has important quantificational consequences. Chopping up a wooden table changes the count from one table to zero tables. In contrast, the chopped up pile persists in being the same amount of wood. Accordingly, in many languages with a count/mass distinction, individuals of most object kinds such as tables are lexicalized as count nouns—they take plurals and can be counted directly (e.g., “one table,” “some tables,” “two tables.” In such languages substances are lexicalized as mass nouns, which cannot be pluralized and cannot be counted unless a measure or kind classifier specifies the individuals to be counted (e.g., \* “one water” vs. “one cup of water” or “one type of water;” \* “woods” vs. “types of wood” or “tons of wood”).

Whorf's provocative claim was that languages with a count/mass distinction (he focused on Indo-European languages) imposed the distinction between object and substance kinds on their speakers as part of a “Standard Average European” or “Aristotelian” ontology, an ontology not shared by speakers of classifier languages (where all nouns are mass nouns) or of speakers of languages like Hopi (where, he claimed, all nouns are count nouns). Thus, Whorf posited the strongest possible effect of language on thought. Without count/mass syntax, he argued, under no circumstances would speakers conceptualize the world in terms of the distinction between individuals and unindividuated entities such as substances. See Quine (1960) for a closely related proposal.

Several considerations militate against Whorf's and Quine's strong linguistic determinism. First of all, all languages, even classifier languages, have the semantic resources to quantify over individuals; they just do so differently from each other. Indeed, studies of linguistic typology have found that languages vary systematically in the types of nouns that can co-occur directly with numerals and that can be pluralized (i.e., that are marked linguistically as individuals), with animate entities at one end of the continuum and non-cohesive substances at the other. This variation determines a strict linguistic hierarchy (Lucy 1992, Croft, 1990, Allan, 1980; Comrie 1981). Cross-linguistic regularities of this sort are usually taken to reflect cross-linguistic conceptual universals that languages select, parametrically, among. Thus, there is no cross-linguistic argument for Whorf's speculation that only speakers of language with a count/mass distinction create the Standard Average European ontological scheme. Second, studies with prelinguistic infants reveal the logical capacity for individuation and tracing numerical identity, as well as other quantificational abilities (Carey, in press; Feigenson, Spelke and Dehaene, 2004; Spelke, Kestenbaum, Simons, & Wein, 1995; Wynn, 1992; Xu & Carey, 1996; see Carey, 1994 for the relevance of these studies to Quine's thesis). Third, prelinguistic infants distinguish objects from non-solid substances, and quantify over these two types of

entities differently (Huntley-Fenner, Carey, & Solimando, 2002). Fourth, prior to the mastery of English language quantifiers, English learning toddlers interpret nouns for newly heard words differently if a word is used to refer to an object (taking it to be a label for object kind or for entities classified by shape) than to a non-solid substance (taking it to be a word for substance kind), providing further evidence that they not only distinguish objects from substances, but represent object kinds and substance kinds (Soja, Carey, & Spelke, 1991). So too do children learning languages without count/mass syntax, i.e., classifier languages (Imai & Gentner, 1997; Imai & Mazuka, 2003, 2007; Lee, 1997).

Although these arguments militate against either Whorf's or Quine's versions of strong linguistic relativity in this case, two lines of research have documented apparent effects of language type on non-linguistic construals of entities as objects or substances (Lucy 1992; Lucy & Gaskins 2001, 2003, Imai & Gentner 1997; Imai & Mazuka, 2003, 2007). The linguistic contrast in both cases is between classifier languages, on the one hand, and languages with a count/mass distinction, on the other. The finding in both cases is that speakers of classifier languages are more likely to construe a given entity as an unindividuated kind of substance than are speakers of count/mass languages.

### Classifier Languages vs. Count/Mass Languages

Three major differences between classifier languages, such as Mandarin, and English, a typical language with the count/mass distinction, illustrate how classifier languages work. First, nouns in classifier languages cannot co-occur directly with numerals. As is the case with English mass nouns, a discretizing unit (i.e., “classifier”) must be specified to numerically quantify any noun, regardless of whether the noun refers to individuated entities (e.g., teacher, table) or non-individuated entities (e.g., sand, water). These “classifiers” provide information such as the shape, animacy, functionality, or the unit of measure of the referent noun. For example, to specify *three pens* in Mandarin requires both the numeral “san” (three) and the classifier “zhi” (stick) as in “san zhi bi” (or “three stick pen”). Second, unlike English count nouns in which number specification via singular-plural marking is obligatory, classifier languages often lack plural morphology. As a result, bare nouns in classifier languages are unspecified in number. If a classifier language has a plural marker, its use is often optional, infrequent, and restricted. Consistent with this and also with the typological individuation hierarchy, Mandarin has an optional and infrequent plural or collective marker (*-men*) that is restricted to nouns for animate beings. Third, nouns in classifier languages, whether referring to objects or substances, typically share the same quantifiers where count/mass syntax languages often have distinct quantifier pairs (contrast *many dogs*, *fewer tables* with *much sand*, *less milk*).

In summary, the sweeping generalization is that classifier languages treat all nouns alike grammatically irrespective of the ontological status of what they denote. More precisely, in classifier languages all nouns pattern like mass nouns in English, and as a result, linguists have argued that classifier languages lack count/mass syntax (e.g., Chierchia, 1994; 1998).

### Evidence for the effects of language type on object/substance construal

In pioneering studies, Lucy (1992) compared speakers of a classifier language, Yucatec Mayans, with speakers of a count/mass language, English, on a similarity judgment task. He showed participants a standard entity (e.g., a cardboard box) and two comparison entities that either matched the standard in shape and object kind or matched in material (a plastic box or a piece of cardboard). When asked which of the comparisons is more similar to the standard, adult English speakers preferred the entity of the same shape and object kind (the box) while adult Yucatec Mayans were split between the two alternatives (box or cardboard). Lucy's subsequent studies replicated the Mayan-English difference with additional sets of stimuli, and

documented that the age of emergence of the cross-linguistic on this task is between seven and nine years of age (Lucy & Gaskins, 2001; 2003).

Data from the second major study addressing this issue showed that Lucy's findings did not depend upon educational differences between his Yucatec-Mayan speaking and his English speaking participants. Japanese is also a classifier language, but Japanese speakers are as educated and urban as American English speakers. Using unfamiliar objects and substances, Imai and Gentner (1997) found differences between English and Japanese speakers in line with Lucy's findings. Following Soja, Carey and Spelke (1991), Imai and Gentner taught English and Japanese speaking adults and children (two to four year olds) a novel label for a novel entity (e.g., "the blicket"). The entities were either non-cohesive (e.g., an omega-shaped portion of Nivea with grape nuts embedded in it), or solid, and if solid, were either complexly shaped artifacts (e.g., an apple corer) or simply shaped objects with no obvious shape dependent function (e.g., a cork half pyramid). Participants were then given two new entities, one matching in shape/object kind but differing in material and one matching in material but differing in shape/object kind. The question was which of these two new entities was the blicket. Imai and Gentner found a robust effect of language type. At each age but the very youngest tested, Japanese speakers made significantly more material-based choices than did English speakers. Speaking a classifier language apparently does make substance construals relatively more likely than does speaking a language with a count/mass distinction. On this task, the age of emergence of the difference due to language type (classifier vs. count-mass) was 2 ½ years.

Many have interpreted these findings as reflecting a Whorfian effect of language on thought. But exactly how and why does speaking a classifier language, rather than a count/mass language, lead to a relatively greater availability of substance construals of material entities (or alternatively, a relatively lesser availability of object kind construals)?

## Interpretations

**Whorfian Interpretation 1: Lucy's lexical noun semantics hypothesis**—Above we argued against the strong Whorfian (and Whorf's own) claim that the quantificational resources of language make the conceptual distinction between individuated and unindividuated entities conceptually available. Lucy (1992) provided the first weak Whorfian hypothesis offered to explain these phenomena. The observation that nouns in classifier languages pattern like mass nouns in English led him to propose that all nouns in classifier languages are semantically unspecified as to its quantificational unit; the nouns refer to "unindividuated essence" or "unformed substance." This contrasts with what count nouns mean in language like English, in that count nouns are already specified as to quantificational unit, thus allowing them to be directly counted without the use of a classifier. These putative differences in noun meanings lead to different habits of construal. In Lucy's own words:

Use of the English lexical items routinely draws attention to the shape of a referent insofar as its form is the basis for incorporating it under some lexical label. Use of the Yucatec lexical items, by contrast, routinely draws attention to the substance composition of a referent insofar as its substance is the basis for incorporating it under some lexical label" (Lucy, 1992, p. 89).

Thus, according to Lucy's analysis, the lexical meanings of nouns differ systematically across these broad language types, and the typical meanings of nouns lead to differences in habitual construals. Lucy's hypothesis leads to the prediction not only that substance construals should be relatively more frequent to speakers of classifier languages, but that the factors that influence construal should be quite different across languages. Nouns in English patently do not pick out "unindividuated material essences" and if Yucatec-Mayan speakers' construal of entities are consisted of such meanings, Mayan speakers should be sensitive to factors that do not figure

much in determining English speakers' construals. The cross-linguistically universal typological hierarchy alluded to above provides evidence against that prediction. Further, even within inanimate entities, in Imai's and Gentner's data, as in Soja et al.'s, the largest effects on entity construal were due to stimulus type (shape based responses were most frequent for complex solids; next for simple solids, and least for non-solid substances), and this ordering was the same for both Japanese and English speakers of every age. This consideration led Imai and Gentner to posit a cross-linguistically universal individuation continuum among entities that is modulated by effects of language type. One goal of the present studies is to put the hypothesis of the universality of the features that influence entity construal to much stronger test, and to adjudicate among competing analyses about what those features might be.

Lucy's linguistic theory for explaining the difference between Yucatec Mayans and English speakers, however, has been challenged. The chief complaint concerns the accuracy of his noun analysis (Imai & Mazuka, 2003, 2007; Mazuka & Friedman, 2000; Barner, Inagaki, & Li, under review). If nouns in classifier languages truly refer to material essence, then nouns for different items made from the same substance would be co-extensional when not modified by a classifier. However, this does not appear to be the case in many classifier languages (e.g., Mandarin and Japanese). For example, what referentially distinguishes a metal fork from a metal spoon is not the choice of classifier but the choice of head noun itself, namely "fork" or "spoon". Confirming this intuition, Colunga and Smith (2005) carried out a rating study on the nouns in both Japanese and English children's earliest vocabularies. Adults rated for each noun whether it picked out entities sharing similar shape or similar material, and whether it referred to solid or non-solid entities. Colunga and Smith found essentially identical distributions of words for object kinds and material kinds, both among solid entities and non-solid entities, in these two corpora. Similar results have also been obtained comparing maternal input in English and Mandarin (Sandhofer, Smith, & Luo, 2000). Thus, the initial nouns that classifier and non-classifier language children *hear* and *produce* are rather similar in content; the nouns typically refer to object kinds rather than substance kinds.

**Whorfian Interpretation 2: The feature weighting hypothesis**—Imai and Gentner offer a different explanation from Lucy's for the effect of language on categorization; see also Smith and colleagues (Samuelson & Smith 1999; Colunga & Smith 2005; Yoshida & Smith, 2003a) for related proposals Imai and Gentner propose that speakers of a count/mass language must choose, for every noun, whether to explicitly mark it as referring to an individuated, directly countable entity, lexicalizing it as a count noun, or not. This obligatory marking makes the features that influence individuatibility (e.g., on their analysis, perceptual complexity and coherence) more salient, and thus weighted more heavily *whenever a representation of the relevant entity is activated*. Speakers of classifier languages would not be subject to this effect. (See Yoshida and Smith, 2003a and Colunga and Smith, 2005 for closely related analyses.) This Whorfian analysis predicts that the effect of language type on entity construal should be observed of *all* tasks, for by hypothesis the relevant features become generally more heavily weighted in every encoding of a given entity. Moreover, the analysis predicts that the effect of language type should be more evident in the construal of some items over others: the effects of language spoken should be greatest for the most ambiguous entities, those for which the features that underlie the hypothesized universal individuation continuum are in balance. In the Soja et al. and Imai and Gentner studies, these are the simple solids. Imai and Gentner found partial support for this prediction -- the effect of language was greater on the construal of simple solids than complex solids, but, contrary to the prediction, they found an equally large effect of language type on the categorization of non-solid substances as on simple solid entities.

**Non-Whorfian Interpretation: The lexical statistics hypothesis**—Lucy's work, and Imai's and Gentner's, is placed in the context of Whorfian effects of language on *thought* (i.e., on cognition in general). In the case of Lucy's work, this is at least partly justified by his task

requiring similarity judgments. But Imai and Gentner's lexical projection task is a *word learning* task. Thus, it is possible that the phenomenon they have discovered reflects an effect of language type on language learning, not an effect of language on thought (Fisher & Gleitman, 2002; Li & Gleitman, 2002; Papafragou, 2005; Slobin, 1996). That is, even though this case has been interpreted as a paradigm example of Whorfian effects on thought, it is possible that it has been misinterpreted as such.

The language one speaks may entail naming conventions, or an attention to different properties of entities *for the purposes of comprehending, producing and learning language*, even if underlying ontological commitments and task neutral feature weightings remain unaffected. In this case, we would expect to see a narrow set of task contexts in which the cross-linguistic differences are observed—namely, only on tasks that centrally involve language. Notice that the lexical statistics hypothesis involves feature weighting. Entity construal is likely to be affected by many factors; the probability that a novel noun is a count noun is only one of these factors and is likely to interact with other features that influence construal. The difference between the Whorfian feature weighting hypothesis and the lexical statistics hypothesis concerns the contextual scope of the effect. The lexical statistics hypothesis predicts that the effect will occur only in the context of lexical projection.

Language on language effects are undisputed when it comes to language learning, where it is generally agreed that children and adults are sensitive to the probabilistic patterns of their language and use such patterns to guide their interpretation of novel words (e.g. Naigles & Terrazas, 1998; Snedeker, Thorpe, & Trueswell, 2001; Trueswell & Kim, 1998, Snedeker & Li, 2000). In the present case, since the majority of the nouns English speakers encounter are count nouns (Samuelson & Smith, 1999), a noun in a neutral context is likely to be interpreted as a count noun. Count nouns, in turn, refer to kinds of individuals. Hence, English speakers might prefer the object kind construal more than Japanese speakers who, because their language has no count syntax, have no comparable probabilistic expectations (see Colunga and Smith 2005 for a simulation that incorporates this assumption in modeling the results of Imai and Gentner).

There are data in support of the claim that English-speaking children treat neutral syntax like count syntax. Several studies (Imai & Mazuka, 2003,2007; Gathercole & Whitfield, 2001; Soja 1992) found that children in neutral syntax conditions (e.g. “the blicket, my blicket”) pattern like children in count syntax conditions (e.g. “a blicket, another blicket”), and differently from children in mass syntax conditions (e.g. “some blicket, some more blicket”) who are in turn more likely to construe entities so labeled as portions of substances.

The lexical statistics hypothesis is potentially undermined by results from Imai and Mazuka (2003; 2007), who collected similarity judgments on the Imai and Gentner triads. Participants were shown a standard and told “look at this.” They were then shown the two comparison items (the shape match, the material match) and asked “Which is the same as this?” Among adults, there were no differences in the similarity judgment task and the lexical project task. Thus, the greater tendency for English speaking adults than Japanese speaking adults to construe a novel entity as an individual of a given object kind does not emerge only in the context of applying a noun to it. On its face, this result favors the Whorfian version of the feature weighting explanation of the effect of language type on entity construal. But it is possible that the lexical statistics hypothesis could still apply to this “nonlinguistic” similarity judgment task. The instructions used the deictic pronoun “this” (i.e., “Look at this” and “Which is the same as this?”). The pronoun “this” stands for a noun and is ambiguous in count/mass status. It therefore may be subjected to the same probabilistic expectations as in the case where a novel word label is provided. Just as speakers must decide whether “this blicket” is count or mass, they also

need to decide whether “this”, as the head noun in the expression, is count or mass (Barner Inagaki & Li, under review).

## The Present Studies

Experiment 1 seeks to extend Imai's findings to a comparison between a new classifier language (Mandarin Chinese) and English. Experiments 2 - 6 provide a stronger test than any to date of the hypothesis that the features that determine entity construal that language type modulates are cross-culturally universal, and it contrasts three accounts of what those features are. Experiment 3 seeks to adjudicate between the Whorfian universal feature weighting hypothesis and the lexical statistics hypothesis. Finally, Experiments 4-6 chart the exact developmental course of the effect of language on entity construal. We conclude by showing how these data constrain our account of the representational source of the robust finding that speakers of count-mass languages are relatively more likely to construe novel entities as members of object kinds that share shape and function than are speakers of classifier languages such as Mandarin, Japanese and Yucatec-Mayan.

## Experiment 1: English vs. Mandarin speakers

All of the interpretations offered above require that the effect of language type on construal of material entities emerge in the comparison of all classifier languages with all count/mass languages. However, attempts to replicate Imai and Gentner with other classifier languages have led to mixed success. For example, Gathercole and Min (1997) compared Korean and English speakers using a similar method to Imai and Gentner, and found only a marginal effect of language type, and that was due to English speakers from North Wales and not those from Phoenix Arizona. Subrahmanyam and Chen (2006) report a conceptual replication with quite different stimuli and a different method, in a comparison of Mandarin and English speakers.

Using stimuli and methodology modeled after Imai and Gentner's, Experiment 1 explores whether Mandarin speaking adults behave like Japanese speaking adults, selecting the material match more frequently than do English-speaking adults. Like Imai and Mazuka (2003; 2007), Experiment 1 compared a lexical projection paradigm (Label Condition) with a similarity assessment paradigm (No Label Condition). Both the Whorfian universal feature weighting hypothesis and lexical statistics hypothesis predict that the effect of language, if present, should be the greatest for the more ambiguous simple solids. Furthermore, based on Imai and Mazuka's data, we expect the two sets of instructions (Label vs. No Label Condition) will be treated equivalently by adults. This finding would be consistent with the Whorfian universal feature weighting hypothesis, and inconsistent with the most straightforward version of the lexical statistics hypothesis.

## Method

**Participants**—Thirty-two native English speaking students from Harvard University participated for psychology research credit or \$4.00. Thirty two native Mandarin speaking students from the National Taiwan University in Taipei, Taiwan were recruited in the same manner as were the English speakers in Experiment 1, via postings on school bulletin boards and advertisement through professors. The Mandarin-speaking participants are thus similarly matched in education and socio-economic status to the English-speaking participants. Students were paid 150 NT each (equivalent of about \$4 USD) for their participation. Half of the students from each language group were randomly assigned to the label condition and half to the non-label condition. Each condition had roughly the same number of female and male students.

**Stimuli**—The stimuli were closely fashioned after Imai and Gentner's and Soja et al.'s (see Table 1). Each triad consisted of the standard, a shape-matched alternative, and a material-

matched alternative. Because all items (both solid and non-solid entities) are characterized by shape and material, we use the locutions “shape-match” and “material-match” to characterize responses.

To ensure that the stimuli had the properties we desired, we had ten naïve native English speakers rate them for complexity. We also had them rate each entity for the meaningfulness of its shape to its function. Ratings were made on 7-point Likert Scales, with 7 as the most complex or meaningful and 1 the least. The ratings are shown on Table 1. With respect to rated complexity, the Complex Solids were rated more complex than either the Simple Solids or the Non-solid Substances (all  $p < .02$ , t-tests), and these latter two classes of stimuli did not differ from each other.

The entities clustered differently from the point of view of meaningfulness of shape to function than they did from the point of view of complexity (see Table 1). Complex Solids were rated as having the most function-relevant shape, significantly more so than the Simple Solids, which in turn were judged to have more function relevant shapes than did the Non-solids (all  $p$ 's  $< .01$ , t-tests). On both the complexity and function-relevance of shape dimensions our stimuli were rated similarly to those of Imai & Gentner, ensuring that we indeed made stimuli comparable to theirs.

### Procedure

**Conditions:** Two types of instructions were used to elicit participants' responses. The first (Label Condition) was the lexical projection paradigm of Soja et al. and of Imai and Gentner. Participants were introduced to the standard with a novel word (“Look at this blicket.”), and asked to choose which of the alternatives shared the same name as the standard (“Which is the blicket?”). The words “this” and “the” were chosen to create “neutral” sentence frames, because both determiners could precede either a count noun or a mass noun. The second type (No Label Condition) followed Lucy's and Imai and Mazuka's similarity judgment paradigm. Participants were introduced to the standard through pointing (“Look at this.”), and asked to choose the alternative that is the same (“Which is the same?”). The two types of instructions (Label vs. No Label) were manipulated as a between-subjects factor.

**Presentation of the Stimuli:** The participants were tested in groups of one to three. The instructions (see Appendix A for the Mandarin instructions) were given verbally by the experimenter while presenting the standard (“Look at this blicket” or “Look at this”) and then later while presenting the shape- and material-match side-by-side (“Which is the blicket” or “Which is the same?”). The participants indicated their responses on an answer sheet that listed “left” or “right” sides as the two choices for each question. The presentation order of the twelve triads was randomized creating two list orders, one the reverse of the other. The left and right positions of the shape-match and the material-match were also randomized.

**Practice Trials:** Participants had two unambiguous practice triads before the twelve test triads. These introduced the participants to the procedure with trials that had a different structure from the experimental trials. One of the practice triads had a metal fork as the standard, and the two comparison entities were a plastic fork and a plastic spoon. Thus, there was no material match available and the only rational choice was a shape match. The other practice triad the standard was a rectangular portion of jam and the two comparison entities were two round portions of jam and a triangular portion of play-dough. Here, there was no shape match available and the only rational choice was a material match. All participants made the rational choices on both practice trials, ensuring that they understood the procedure and would choose both shape- and material-matches.



## Results

The data from all experiments are analyzed with respect to two interrelated issues: do we see the effect of language type (classifier vs. count-mass) on entity construal, and if so, under what task conditions and second, what entity features influence entity construal, and are these cross-linguistically universal?

Figure 1 plots the results from the Mandarin speakers for the two instruction types alongside the results from the English speakers. An ANOVA examined the effects of Entity Type (Complex, Simple, and Nonsolid), Instruction (Label, No Label), and Language Type (English, Mandarin) on percentage of shape-match choice.<sup>1</sup> There was no main effect nor interactions involving Instruction (all  $p$ 's > .20); participants' choices were not influenced by whether they were choosing another block like the standard block, or another entity that was the same as the standard. In contrast, Language Type ( $F(1, 60) = 20.6, p < .001, \eta_p^2 = .35$ ), Entity Type ( $F(2, 120) = 157.66, p < .001, \eta_p^2 = .72$ ), and the interaction between Entity Type and Language Type ( $F(2, 120) = 5.62, p < .01, \eta_p^2 = .09$ ) produced significant effects.

The main effect of Language Type is due to English speakers selecting the shape match more frequently than did the Mandarin speakers (58.6% vs. 31.3%). The effect of language type on entity construal observed by Imai and Gentner is robustly replicated, and held for each entity type. English speakers made more shape-based choices when classifying Complex solids (English: 94.5% shape vs. Mandarin: 73.4%,  $t(62) = 3.02, p < .01$ ), Simple solids (English: 59.4% vs. Mandarin: 16.4%,  $t(62) = 11.53, p < .001$ ), and Non-solids (English: 21.9% vs. Mandarin: 3.6%,  $t(62) = 3.51, p < .01$ ).

The main effect of Entity Type reflects the fact that Complex solids elicited more shape-match choices than the Simple solids (84.0% vs. 37.9%  $p < .001$ , pair-wise  $t$ -tests) and in turn Simple solids received more shape-match choices than Non-solids (37.9% vs. 12.9%,  $p < .001$ ). The ordering of stimuli types with respect to percentage shape choice (Complex > Simple > Non-solid) held true for both English speakers alone (94.5% > 59.4% > 21.9%,  $t(31)$ s > 5.9,  $p$ 's < .001 for all comparisons) and for Mandarin speakers alone (73.4% > 16.4% > 3.6%,  $t(31)$ s > 2.6,  $p$ 's < .01 for all comparisons).

Finally, the Entity Type  $\times$  Language Type effect indicates that the difference in the percentage of shape-match choices across the two language groups varies as a function of entity type. Both the Whorfian universal feature weighting hypothesis and the lexical statistics hypothesis predict an Entity Type  $\times$  Language Type interaction, such that the largest effect of Language Type should be observed on the stimuli for which the features that influence entity construal are most evenly balanced: the Simple Solids. An inspection of Figure 1 shows this to be the case. Specifically, the rate of English speakers choosing the shape match was 43% higher than that of Mandarin speakers for Simple Solids, whereas the rates differed only half as much for Complex Solids (21%) and Non-solids (18%).

## Conclusions

Experiment 1 solidly extends Imai and Gentner's results to speakers of a different classifier language: Mandarin. Mandarin-speaking adults were significantly more likely than English-speaking adults to project names on the basis of substance kind. Furthermore, our data revealed the pattern Imai and Gentner predicted but did not find: the effect of language type was twice as great for simple solids than for the other two entity types. The simple solids were the most

<sup>1</sup>For all the experiments reported in this paper, we ran the ANOVAs both as subjects analyses and as items analyses. In almost all cases the pattern of results was the same. In such cases, we report only the subjects analyses. We mention the items analyses only when they showed a different pattern of results from the subjects analyses. We also ran the same ANOVAs using arcsine-root transformation of the dependent variable and the results did not change.

ambiguous stimuli with respect to object kind/substance kind construal (English-speaking participants were at chance); apparently the effect of language type is greatest on ambiguous entities.

Additionally, the labeling manipulation had no effect on speakers' tendency to classify by shape (or object kind). English and Mandarin speakers were just as likely to choose by shape with or without the novel word label. We thus replicate Imai and Mazuka's (2003; 2007) finding with adults. The fact that the effect of native language is found in a context that does not involve word learning potentially militates against the lexical statistics explanation. Experiment 3 will readdress the issue of contextual scope (i.e., in how wide a variety of tasks does language type influence construal) by testing speakers using another paradigm.

In Experiment 1, the most robust effects are those of entity type. Speakers of *both* languages made more shape-match choices for complex solids than for simple solids, and in turn, more for simple solids than for non-solid substances. This finding is consistent with there being universals in the features that influence entity construal in such tasks. However, only a few contrasts in entity types were explored in Experiment 1, and the apparent universality could reflect ceiling and floor effects. Experiment 2 seeks to replicate the effect of language type on construal, to bolster the conclusion that the features that affect construal are universal, and to explore exactly what those features are.

## Experiment 2: Complexity and Function Relevance in Object-Substance Triad Task

Given that there is no effect of whether the entities are labeled or not, Experiment 2 deploys only the lexical projection task. Experiment 2 compares English speaking adults with both Mandarin and Japanese speaking adults to further explore the effects of language type on entity construal. In Experiment 1, like Imai and Gentner (1997) and Soja et al. (1991), the complex solids were judged to have more complex shapes, and to more likely to have shape relevant functions, than did the simple solids or the non-solids. In addition, only the complex solids were real manufactured artifacts. Thus, the contrasts between complex and simple solids confounded several variables. Similarly, the contrast between simple solids and non-solids confounded two variables: solidity and likelihood of shape-dependent function. Thus, Experiment 1 and those previous experiments it is modeled on leave wide open the features that influence entity construal, and whether the finding of cross language type universality would survive tests that separated these confounded variables.

There are several proposals in the literature for what features might influence entity construal of speakers of all languages. Some writers suggest perceptual features: increased complexity of shape is hypothesized to increase the likelihood of shape-based responses and non-solidity is hypothesized to decrease the likelihood (Gentner & Boroditsky, 2001; Soja et al 1991, Yoshida & Smith, 2003). Other writers suggest more conceptual features, such as probability of a shape-based function (Imai & Gentner, 1997). Recently, Prasada and his colleagues (Prasada, 1999; Prasada, Ferenz, & Haskell, 2002; Burger & Prasada, 1997) rationalized these distinct proposals under an elegant and abstract analysis of the cross-linguistically universal features that might underlie the likelihood of construing a given entity as a member of an object kind or as a member of an object kind.

On Prasada's analysis, whether we construe an entity as a kind of object, rather than as a kind of substance, is dependent upon whether its structure (shape, arrangement of parts) is seen as non-accidental, thus unifying under a single analysis a range of features known to affect entity construal. Under this analysis, entities with shape relevant to function will tend to be treated as a member of an object kind while entities with complex irregular shapes might not be. Testing

English speaking adults, Prasada found that shape complexity, if irregular, negatively affects construal as member of an object kind, *unless* there are multiple copies. When presented with a complex-shaped solid entity such as a contorted piece of metal scrap, one might see it as metal. However, when then shown its carefully crafted same-shape duplicate, one might appreciate the low probability that such a match could be coincidental, thus making it appear non-arbitrary, and more likely to be construed as an exemplar of an unknown object kind. Indeed, this is what occurs (Prasada et al., 2002). Similarly, Imai and Mazuka (2003) compared English speaking adults' construal of non-cohesive entities when shown one copy versus multiple copies. Consistent with Prasada's analysis, Imai and Mazuka (2003) found that multiple copies of portions of non-solid substances of a given shape increased construal as a kind of individual (e.g., "pile") rather than as a kind of substance (e.g., "sand").

To better test Prasada's analysis and whether it applies universally, Experiments 2 and 3 use a new set of entities, none of which is a manufactured artifact. The stimuli are thus all more ambiguous than the complex shapes of Experiment 1. Complexity of shape is crossed with probability of shape-based function. This set of stimuli enable us to further explore the effect of language type on entity construal, and if so, to test whether this effect modulates cross-linguistic universal bases of entity construal. Specifically, in Experiment 2, in contrast to Experiment 3 (see below), Prasada's analysis predicts that both shape complexity (as a result of the presence of a shape-matched copy) and a shape's relevance to function should increase shape-consistent categorization.

## Methods

**Participants**—Sixteen native English speakers and 16 native Mandarin speakers who had not participated in the previous experiments were recruited in the same manner as previous experiments. The English speakers were Harvard University college students and the Mandarin speakers were students from Taipei's National Taiwan University and National Taipei Teachers College. Twenty native Japanese speakers were recruited from the local town hall of Kirishima Town of the Kagoshima Prefecture. To be comparable in age and education to the English and Mandarin speaking participants, the Japanese adults were all recent college graduates in their early twenties. All adults were paid approximately \$5 USD for their time. The number of female and male participants was roughly equal for each language group.

**Stimuli**—The current set of stimuli was derived from a larger set consisting of 54 solids and 19 non-solids. We designed these stimuli to be unrecognizable as known artifacts and substances. Unlike the previous experiments (e.g., Experiment 1), none of these stimuli could be directly purchased from stores. Each stimulus had its own unique shape and material, and because each hand-made or cut scraps from purchasable items, none was perfectly regular. Some materials were created by mixing several different materials or disguised by textured paints. Thus, all of these entities are ambiguous, relative to the store-bought complex objects of Imai and Gentner, and Soja et al.

The whole set of new stimuli, in randomized order, were rated on two Likert scales by four naïve English speakers who did not participate in any of the other studies reported in this paper. One scale assessed the complexity of the overall shape and outline of the item, with 1 being the least complex and 7 being the most complex. After rating shape complexity, participants rated the items according to the degree to which the function of the item would depend on its overall shape and outline, with 1 being the least likely to depend on its shape and 7 being the most likely. To ensure that participants understood the task, they first rated two recognizable items (i.e., a metal fork and a pile of jam) before rating the experimental items. These two training trials were presented in the same manner as all the other trials but were always the first two items rated. These test items confirmed the validity of the scale: It was apparent that a fork

has a higher shape complexity (mean = 4.25) than a round pile of jam (mean = 2.25). It was also apparent that the function of a fork depends very much on the shape of the item (mean = 7.0), whereas the function of jam is not dependent on its shape (mean = 1.5). All participants rated the fork substantially higher than the jam on both scales.

From the ratings of this entire stimulus set, we selected three solids to fill each of the four cells created by crossing two levels of shape complexity (Complex vs. Simple) with two levels of shape functionality (Highly Function-dependent vs. Not highly Function-dependent). We henceforth refer to Complex items as +C and Simple items as -C, items with Highly Function-dependent shape as +F, and Not highly Function-dependent shape -F. The ratings results along with pictures of the stimuli from each of the four cells are shown in Table 2. Complexity ratings for the two Complex sets (+C+F and +C-F) items did not differ from each other ( $t(3)=-.15$ ,  $p=.89$ ). They did differ from two Simple sets -C+F and -C-F ( $p < .05$ ). The Simple sets in turn did not differ from each other on the complexity scale ( $p=.89$ ). For the Functionality ratings, +C+F and -C+F did not differ from each other ( $p = .72$ ). They did, however, differ from +C-F and -C-F ( $p < .01$ ) which in turn did not differ from each other ( $p = .08$ ).

In addition to these solids, four non-solids were randomly chosen from the 19 non-solids to include in the test trials. Non-solidity *per se* militates against construal as a member of a kind of individual (Burger & Prasada, 1997). Furthermore, these non-solids were rated as not having shape-dependent function. These features increase the likelihood that these entities would be construed as kinds of materials. However, there were two copies of the same shape, increasing the likelihood of construal as kinds of individual (Imai & Mazuka, 2003). The four randomly chosen non-solids did not differ from the -F solids on the shape-dependent function scale ( $M=1.69$ ,  $SD=.24$ ), but did differ from the +F solids on this scale ( $p < .01$ ). With respect to shape complexity, the non-solids ratings varied greatly from 1.25 to 4.75 ( $M=3.0$ ,  $SD = .61$ ). Overall though, the shape complexity ratings of the non-solids did not differ from the +C solids ( $p = .15$ ) but differed from the -C solids ( $p = .04$ ). In sum, the non-solid ratings patterned like the +C-F solids in rating. The comparison between non-solids and +C-F solids will speak to the effect of cohesiveness *per se* on entity construal.

After selecting the stimuli, same-shape and same-material alternatives were then created for each.

**Procedure**—The procedure was identical to that of the Label condition in Experiments 1. After presentation of the fork and jam familiarization trials, as in the previous studies, the experimenter presented a standard and labeled it (e.g., the blicket). He or she then presented the two alternatives and asked which one was the blicket. This process was repeated for the entire set of 16 triads.

**Materials**—Altogether, four sets of solids (three entities each) and one set of four non-solids make up the sixteen test triads. The presentation order of the sixteen test triads was randomized creating two list orders, one the reverse of the other. The left and right positions of the shape-match and the material-match were also randomized.

## Results

The analyses address two questions. First, do we again see the cross-linguistic difference in the percentages of shape-match choice between English speakers, on the one hand, and Japanese and Mandarin speakers, on the other? Second, what features affect entity construal?

We calculated the average percentage of shape responses for each entity type in order to examine the effects of the entity's shape complexity and potential shape-dependent functionality on entity construal. Altogether there were five entity types consisting of non-

solids (NS) and four types of solid entities from crossing complexity with functionality (+C +F, +C-F, -C+F, -C-F). The results are plotted in Figure 2.

**Solids**—An ANOVA examined the effects of Language Type (English, Mandarin, Japanese), Shape Complexity (complex, simple) and Functionality (likely, unlikely to have shape-dependent function) on the percentage of shape-match choice. The analysis confirmed the language effects that we have seen in Experiment 1 and in Imai's studies. There was a main effect of Language Type ( $F(1, 49)=258.51, p < .001, \eta_p^2 = .24$ ), with English speakers selecting shape more than Mandarin speakers (English: 65% vs. Mandarin: 35% shaped based judgments;  $t(30) = 3.55, p = .001$ ) and Japanese speakers (English: 65% vs. Japanese: 47%;  $p < .01$ ). Mandarin and Japanese speakers did not differ from each other ( $p = .13$ ).

There was also a main effect of Functionality ( $F(1, 49)=21.99, p < .001, \eta_p^2 = .31$ ). Solids that have shape-dependent functions received more shape-match choices than those that do not (+F: 60% vs. -F: 38%). There was also a main effect of Shape Complexity ( $F(1, 49)=17.60, p < .001, \eta_p^2 = .26$ ). Complex solids receive more shape-match choices than Simple solids (+C: 56%; -C: 42%).

No other significant main effects or interactions were observed. Importantly, although Figure 2 indicates that the effects of Functionality and Shape Complexity were muted in Mandarin and Japanese speakers, the pattern did not differ among the three groups as there was no significant interaction between Language Type and these variables.

**Non-solids**—The percentage shape-match choice for non-solids was submitted to a 3 Language Type (English, Mandarin, and Japanese) univariate ANOVA. There was no effect of Language Type ( $F(1, 49)=2.56, p = .07, \eta_p^2 = .11$ ) though the percentage of shape-match was highest for English speakers (English: 45%; Mandarin: 28%; Japanese: 21%), and this predicted result is significant on a 1-tailed test.

All analyses of the individuation continuum predict that the non-cohesive property of non-solid entities should increase the probability of substance kind construals. To examine the effect of non-cohesion (or non-solidity) on construal, we compared the percentage of shape-match choices for the non-solids to the +C-F solids, which had similar stimuli ratings regarding complexity and shape-dependent functions. A 3 Language Type (English, Mandarin, Japanese) by Solidity (+C-F solids, non-solids) ANOVA found a main effect of Solidity (46% for solids vs. 32% for non-solids,  $F(1, 49)=.36, p < .01, \eta_p^2 = .14$ ) and no other main or interaction effects. The solidity effect is small (a 14% difference between +C-F solids and non-solids). On Prasada's analysis, two factors are at play here regarding entity construal: non-cohesiveness makes it more likely that form is unimportant since form is unstable, but the exact repetition of a given form makes it more likely that the form is important and non-accidental (see also Imai and Mazuka, 2003). Experiment 3 tests the hypothesis that the effect of cohesion on entity construal, as revealed in the comparison with +C-F solids, is more robust when there are no repeated shapes.

## Conclusions

Experiment 2 deployed an entirely differently set of stimuli from Imai and Gentner (1997) and Experiment 1. Nonetheless, it confirmed that adult Mandarin speakers and adult Japanese speakers are less likely to project novel words by shape than English speakers. The effect of language type on entity construal reported by Lucy (1992) and Imai and Gentner (1997) is extremely robust.

Importantly, Complexity, Functionality, and Solidity modulate construal equivalently across the three language groups. These data favor the universality of the features affecting entity

construal. Consistent with Prasada's analysis, both shape complexity and the relevance of shape to function influenced likelihood of construing the ambiguous entity as a member of an object kind. Replicating other studies, solidity also influenced speakers' construal. Experiment 3 will test Prasada's hypothesis that shape complexity per se has no direct role in entity construal, unless also providing evidence that the structure is non-random.

It is now time to turn to the heart of the matter. Is the robust effect of language type on entity construal truly a Whorfian effect? It clearly emerges in a lexical projection context (Experiments 1 & 2, Imai & Gentner, 1997; Gathercole & Min, 1997; Subrahmanyam & Chen, 2006), as predicted by the lexical statistics account of the phenomenon. On this account, English speakers may be influenced by the greater frequency of count nouns than mass nouns to interpret the novel noun heard in neutral syntactic contexts as a count noun, and this interacts with other factors that influence their construal of the entities. Such an effect is what Slobin called “thinking for speaking” (Slobin, 1996). If this were the only context in which the effect of language is observed, it would perhaps be more apt to conceptualize it as an effect of language on language than as an effect of language on thought (Li & Gleitman, 2002; Fisher & Gleitman, 2002; Gleitman & Papafragou, 2005). However, lexical projection is not the only context in which the effect occurs, at least with adults. It also emerges in tasks in which subjects must classify the stimuli with respect to sameness (Experiment 1; Lucy, 1992; Imai & Mazuka, 2003). It is thus possible that the effect of language type has very wide contextual scope, perhaps even universal scope. Such a finding would support the Whorfian feature weighting interpretation of these phenomena. The obligatory marking of the count/mass distinction in English may cause the features relevant to construing entities as individuals of a kind (and hence lexicalized as a count noun, the marked category) be weighted more heavily in all representations of those entities. This would indeed constitute an effect of language on thought.

However, as previously mentioned, the similarity judgment task may not be immune to the influence of lexical statistics. The interpretation of the deictic pronoun “this” in the similarity judgment instruction (“Look at this”) may be subject to the same probabilistic calculations as in the case where a novel word label is provided (“Look at this blicket.”). Experiment 3 therefore introduces a new nonlinguistic task. We explicitly told our participants (English, Mandarin, and Japanese speaking adults) that we are interested in understanding what is it about entities that leads a person to construe it as a kind of object or a kind of substance, illustrating what we meant with examples. Participants then rated how likely they would construe each entity from Experiments 2 as a member of a kind of object or a kind of substance. Notice that although we made it entirely clear what judgment we wanted participants to be making, they were free to use any criteria that were natural to them in deciding about any given entity. The Whorfian feature weighting account holds that speaking English has made adults more sensitive to the features that signal whether an entity should be construed as a member of a kind of individual (solidity, non-accidental shape; shape-dependent function). Therefore, English speakers should be more likely than speakers of Mandarin or Japanese to construe the ambiguous entities as members of object kinds in this task, just as in Experiment 2. However, if the scope of the effect is restricted to contexts in which participants are required to interpret the intent of expressions (“this blicket” or “this”), no effects of language type will emerge on this highly explicit task. This latter finding would favor the lexical statistics interpretation of all previous demonstrations of the effects of language type on construals of entities as members of kinds of individuals vs. kinds of substances.

### **Experiment 3: Complexity and Function Relevance in Object-Substance Rating Task**

English, Mandarin, and Japanese speakers rated on a scale of 1 to 7 how likely they are to construe each of the entities from Experiments 2 as an object or a substance. The fine-grained

scale, opposed to a dichotomous forced choice, provides a sensitive measure of a language effect on construal, if present.

The paradigm depends upon there being words equivalent to “object” and “substance” in all three languages. In that sense it is not “non-linguistic.” But it does not invite establishing a lexical entry or considering which of two entities might be named by the same label. There are, indeed, Japanese and Mandarin equivalents of these English words (see Appendix B for exact verbal protocols in all three languages). However, as a reviewer rightly pointed out, people (philosophers, psychologists, or laymen alike) do not always agree on the usage of the terms “object” and “substance” in any of the three languages. To be sure that our participants understood the terms as we intended, we provided Prasada et al (2002)'s Aristotelian definition alongside a canonical example of an object-kind (*table*) and a canonical substance-kind (*ice cream*). Particularly, we pointed out that when construing something as an object, we think of its form as important. When we construe something as a substance, we disregard its form.

Like Experiment 2, Experiment 3 tests for the effects of shape complexity and functional-relevance of shape on the likelihood that an entity is construed as an object rather than a substance. Given Prasada's results, and those of Experiments 2, we expect to replicate the effect of the function relevance of the shape. Prasada's analysis predicts a different result in this study than in Experiments 2 with respect to complexity. In this study there are no comparison items, and so no evidence that the complex, irregular shapes were not accidental. Furthermore, Prasada et al. (2002) found that irregularity of shape, on its own, influenced participants to construe an entity in terms of *substance* kind. Since the simple shapes were also more regular, we might find that simply shaped solids are judged as *better* members of individuated object kinds than the complexly shaped solids. Finally, because there were no repeated shapes, we expect to replicate Prasada's findings that solidity per se is a powerful determinant of construal. Prasada's findings, as well as those of Imai and Mazuka, predict that the effect of solidity will be stronger in Experiment 3 than in Experiment 2, because it is not contravened by the presence of repeated copies of a single shape.

## Methods

**Participants**—Adult speakers from the three language groups, English, Mandarin, and Japanese were recruited from the same population as Experiment 2. Sixteen English speakers, 16 Mandarin speakers, and 24 Japanese speakers participated. None had been in the previous experiments. All were paid approximately \$5 USD for their time.

**Stimuli**—The stimuli consisted of the 16 standards from Experiment 2.

**Procedure**—The participants were given the instructions to rate entities on a scale of 1 to 7, with 1 being a good object and 7 being a good substance (see Appendix B for instructions in all three languages). However, for ease of comparison with Experiments 2, we report the ratings using an inverted scale so that a higher rating (7 maximum) means a higher likelihood of an object construal and a lower rating (1 minimum) means a higher likelihood of substance construal.

Each standard was placed in its own individual container and presented to the participants. Tested in groups of 1 to 3, the participants recorded their responses on the answer sheets.

To check that the participants understood the questions, we again included the fork and jam training trials prior to the other trials. The two training trials were presented in the same manner as all the other trials and labeled as questions 1 and 2 on the answer sheets. Indeed, participants rated the fork as a good object (mean = 6.97) and the jam as a good substance (mean = 1.45), with no statistical difference across the language groups in how they rated the two items.

To ensure equivalence across languages, the original English instructions were translated into Mandarin by a bilingual native speaker. Translations into Japanese, as in the other experiments reported here, were performed by a native Japanese speaker and back-translated by a second native speaker to ensure accuracy. The translators then discussed the discrepancies and agreed upon a final version.

## Results

Figure 3 plots the ratings of the three groups of language speakers for the different entity types. The important finding from Experiment 3, in contrast to Experiments 1 and 2, is that English speakers were not in the slightest more likely to construe entities as members of object kinds than were Japanese or Mandarin speakers. Also important is the finding that the same features influence entity construal for speakers of all three languages; there is a large effect of whether the shapes appear to have relevance for a potential function of these novel items, and the small effect of shape complexity is in the opposite direction to that of Experiment 2 (Compare Figure 2 & 3).

**Solids**—A  $3 \times 2 \times 2$  ANOVA examined the effects of Language Type (English, Mandarin, Japanese), Shape Complexity (Complex, Simple), and Functionality (Likely vs. Unlikely to have shape-dependent function) on the ratings of how likely each entity is seen as an object. There was no main effect of Language Type ( $F(2, 53)=1.12, p = .33, \eta_p^2 = .80$ ). The mean ratings for English, Japanese, and Mandarin were 4.9, 4.7, and 5.1 respectively.<sup>2</sup> The main effect of shape-dependent functionality was significant ( $F(1, 53)=216.94, p < .001$ ), with higher object ratings for entities seen as likely to have shape dependent functions than as those seen as unlikely to have shape dependent functions (5.88 vs. 3.89). There was, however, a Functionality  $\times$  Language interaction ( $F(2, 53)=5.7, p < .01, \eta_p^2 = .18$ ). To understand the interaction, we conducted pairwise comparisons of the three language groups' ratings for each level of functionality. The only significant difference was between Japanese speakers and Mandarin speakers on the afunctional entities ( $t(38) = 2.64, p = .01$ ); all other comparisons were non-significant ( $p > .18$ ). As can be seen from Figure 3, Mandarin speakers were especially more likely to see -C-F entities as objects than were either the Japanese and English speakers. We do not have an explanation for this effect. Follow-up work from our laboratory suggests that the Mandarin-Japanese difference is not replicable and possibly due to random-sampling (In Experiment 2, the Japanese speakers sampled were slightly more likely than Mandarin speakers to pick the object kind match. In Experiment 3, the reverse is true; See Figures 2 & 3).

There was a main effect of complexity ( $F(1, 53)= 8.58, p < .01, \eta_p^2 = .14$ ) in the subjects analysis that did not emerge from the items analysis ( $F(2(1, 8)=.66, p = .44, \eta_p^2 = .08$ ). Subjects rated the *less* complex items to as *more* likely to be seen as objects (object rating of 5.0 vs. 4.7). Note that the effect is the opposite from that of Experiment 2, where adults in all three languages made more shape-consistent matches for more complex items. As predicted by Prasada's analysis, increased complexity increases object construals when there are more than one version of a complexly shaped object (Experiment 2), for multiple copies greatly decrease the probability that the shape arose accidentally. And also consistent with Prasada's analysis, the simpler shapes, being more regular on average than the complex shapes, were more likely to elicit an object construal when seen alone (Experiment 3), again because regularity of shape is less likely to emerge accidentally.

<sup>2</sup>On the items analysis, this main effect was marginal  $F(2,16)=3.76, p = .05$ . Pair-wise paired t-tests of ratings of the items revealed this main effect to be due to the difference between Mandarin and Japanese speakers; neither differed significantly from English speakers.



**Non-Solids**—The average ratings for the non-solids by English, Japanese, and Mandarin speakers were 2.46, 2.34, and 2.98 respectively. An univariate ANOVA examined the effect of language type on the ratings; there was no main effect of language type ( $F(1, 53) = 2.45, p = .10, \eta_p^2 = .08$ ).

The ratings of non-solids were all significantly below the midline of 4 on the 1-to-7 point scale ( $p < .01$  in each case,  $t$ -tests). In contrast, participants from all language groups rated the solids with shape dependent functions significantly above 4 ( $p < .01$  in each case,  $t$ -tests). The solids without shape-dependent function, whether simple or complex, were clearly ambiguous, with ratings around 4 (see Figure 3). The +C-F solids had ratings for shape complexity and for shape dependent function most like those of the non-solid entities. Thus, to isolate the effect of solidity per se on entity construal, we compared the non-solids with the +C-F solids on subjects judgments that each entity should be construed as a member of an object kind. The 3 Language (English, Mandarin, Japanese) by Solidity (+C-F solids, non-solids) ANOVA revealed only one significant effect: the solids received higher object ratings than non-solids (3.7 vs. 2.6,  $F(1, 53) = 32.32, p < .001, \eta_p^2 = .38$ ). Thus, replicating Prasada, cohesiveness or solidity per se has a large influence on the construal of entities, when the design of the study does not include the repetition of shapes. The effect of solidity was more apparent in Experiment 3 (effect size = .38) than in Experiment 2 (effect size = .14).

**By Item Cross-Linguistic Comparisons**—Consistent with the universal individuation continuum hypothesis, entity construal by speakers of all three languages appears to be equivalently influenced by factors such as solidity, functionality, and multiple copies of a given shape. That hypothesis can be put to a stronger test. Specifically, there is some variation in ratings among the entities of each type. If similar factors (whatever they may be) are determining whether a given entity is construed as member of an object kind or as a member of a substance kind, irrespective of language type, this similarity should be apparent at even a finer grained level—that of individual test items. To explore whether this is so, for each language group we calculated the average rating for each item.

The result is represented pictorially in Figures 4, which plots the average rating of each item from speakers of Mandarin (Figure 4a) and Japanese (Figure 4b) as a function of the average ratings of that item from English speakers. For completeness, although there is no reason to expect Japanese and Mandarin speakers to differ, Mandarin ratings are plotted as a function of Japanese ratings in Figure 4c. If similar properties of the entities are influencing Mandarin and English speakers' construals, their rating for the likelihood object kind construal should be positively correlated (Figure 4a; see Figure 4b for the comparable Japanese and English comparison).

The three language groups' responses were highly correlated with each other (Mandarin-English: Pearson's  $r = .88, p < .001$ ; Japanese-English: Pearson's  $r = .97, p < .001$ , Mandarin-Japanese  $r = .92, p < .001$ ). The fact that same factors influence participants' construals is therefore evident at the level of individual items.

These data also bear on the effect of language type on entity construal. The degree to which Mandarin or Japanese speakers are less likely than English speakers to treat entities as kinds of object will be reflected in the degree to which the points fall below the diagonal. Paired  $t$ -tests show that English speakers did not differ statistically from either Mandarin speakers ( $t(15) = 1.63, ns$ ; English mean object rating 4.28; Mandarin, 4.59) or Japanese speakers ( $t(15) = 1.56, ns$ ; Japanese 4.09). Indeed, the Mandarin speakers actually rated the entities *more* object-like than did the English speakers (although not significantly so), and they gave reliably higher object ratings than did Japanese speakers (4.59 vs. 4.09) ( $t(15) = -2.84, p = .01$ ). Thus, comparing

the ratings of Mandarin speakers and Japanese speakers to the English speakers by items found no greater tendency for classifier language speakers to construe entities as substance kind.

## Conclusions

The data from Experiment 3 bear on two different issues. First, and most crucially, they support the lexical statistics hypothesis over the Whorfian universal feature weighting hypothesis as an explanation for the effect of language-type on entity construal. This new paradigm eliminated the well-established language type effect. English speaking adults were no more likely than were Mandarin and Japanese speaking adults to construe these ambiguous items as members of object kinds. The effects of language type on object construal do not have universal scope; they do not reflect universally different weightings of the features of perceived entities that in turn inevitably influence construal (Yoshida & Smith 2003a; Imai & Gentner, 1997). Recent findings from our laboratory have extended these results to a Mayan/English comparison (Chen, 2007).

Second, the data support the notion that there are cognitive universals that govern entity construal. The correlation of the ratings by items was extremely high across language groups, indicating that the same stimulus properties affect speakers regardless of their language background. Furthermore, the data provide striking evidence for Prasada's analysis of the factors underlying these universals, confirming previous findings Prasada et al. (2002) obtained with English speakers. In Prasada et al. the dependent measure was whether participants chose to label the entities "a blicket" or "some blicket" (i.e., with count or mass syntax). Of course, in a classifier language one cannot use this dependent measure. The straightforward dependent measure we chose—rating scales of how likely it is to see the entity as an object vs. as a substance, replicates Prasada's findings in great detail. Factors that suggest that a material entity's form is not accidental increase the likelihood that the entity will be construed as a kind of object. These include—being made of a solid rather than non-solid substance (Experiments 2 and 3), complexity of form if there are repeated exemplars of entities with that form (Experiment 2), regularity/simplicity of form if there is only one exemplar (Experiment 3), likelihood that the form has functional consequences (Experiments 2 and 3).

## Experiment 4: Children; Replicating Age Effect in Mandarin

Experiments 1 and 2 confirm previous reports of effects of language type on entity construal, extending them to a Mandarin-English comparison, and to quite different types of entities. The lack of an effect of language type in Experiment 3 supports the lexical statistics account of the results in Experiments 1 and 2. We now turn our attention to the developmental course of these effects. The lexical statistics hypothesis presupposes the object-substance distinction is ontologically available prior to language learning, and that the initial starting point concerning the factors that influence construal of material entities is universal (Soja, Carey, & Spelke, 1991; Imai & Gentner, 1997). According to the lexical status hypothesis, the emergence of the language type effect comes about as a result of learning the statistical probability that a novel noun in neutral syntax is more likely to be a count noun, and after learning how count/mass syntax is correlated with the naming of objects and substances. According to this hypothesis, developmental changes should be driven by an increase of shaped-based categorization by English speakers. Only English speakers will make the default assumption that a newly heard noun is a count noun and in the case of ambiguous stimuli be swayed to treat the noun as referring to an individual of an object kind (see Yoshida & Smith, 2003a).

The existing data present an unclear picture of the developmental pattern. Imai and Gentner found an unusually early emergence of the language difference between Japanese and English speakers (by 2 ½) in contrast to what others have observed (Subrahmanyam & Chen, 2006; Yoshida & Smith, 2003b, Lucy & Gaskins, 2001, 2003). The time of onset found by Imai and

Gentner is unusual given prior developmental work suggesting that learning the correspondence between count/mass syntax and the object-substance distinction occurs at or after three-years of age (Gordon, 1985; 1988; Soja, 1992). The early emergence is also incompatible with Lucy's lexical semantics hypothesis; Japanese 2 ½ year olds have relatively little grasp of the Japanese classifier system (Uchida & Imai, 1999; Matsumoto, 1987; Yamamoto & Keil, 2000; Naka, 1990). In fact, many studies show that the acquisition of classifiers is quite drawn out, and in some cases, even six year-olds may not have fully mastered the classifier system (e.g., Lucy & Gaskins, 2003; Li, Barner, & Huang, 2008. Uchida & Imai, 1999).

Which language group is changing is also unclear. Although Imai and Gentner's data showed an increase in shape-match choice in English speakers over time, Lucy and Gaskins (2001) data did not. Figure 5a plots the average percentage of shape-match choice collapsing across all three entity types for each age in Imai and Gentner's data. The results show a slight increase in English speakers' shape-match choice with age, and a slight decrease in Japanese speakers' shape-match choice. A decrease in shape-match choice indicates an increasing preference for the material choice. In Lucy and Gaskin's data (Figure 5b), in contrast, English speakers did not change much across the ages (and if anything decreased slightly in shape-match choice) whereas older Yucatec Mayans increased dramatically in their preference for the material match choice, developing instead a preference for the material choice.

Although the increase of shape-match choice over time is compatible with the lexical statistics hypothesis, the increase in material-match choice by classifier language speakers is not so straightforwardly explainable. The increase in material-match choice could be explained by Lucy's lexical noun semantics hypothesis that the acquisition of the classifier system reorganizes its speakers' interpretation of nouns as unindividuated essence. However, as indicated above, several findings already militate against Lucy's account. Most convincingly, several studies show that nouns in classifier languages do not differ from nouns in mass-count languages in what they typically denote (e.g., Colunga & Smith, 2005; Barner, Inagaki, & Li, under review), and the present studies, along with those of Imai and Gentner, indicate that the construals of material entities by speakers of classifier languages and by speakers of languages with a count/mass distinction are modulated by the same factors. One question is whether the increase in material-choice by classifier language speakers is replicable.

The next three experiments characterize the developmental pattern. To see if we could replicate Imai and Gentner's finding of an early emergence of the language type effect in a Mandarin-English comparison, Experiment 4 used the stimuli of Experiment 1, contrasting children's responses in the Label condition with those of the No-label condition, as in Experiment 1.

## Method

**Participants**—Participants were 192 children, thirty-two at each of three ages (two-, three-, four-year-olds) for each of the two language groups (English and Mandarin). The number of boys and girls was roughly equal within each age and language group. The average ages for the three groups of English speaking children were 2;5 (1;10-2;10), 3;5 (2;11-3;10) and 4;4 (3;11-4;8). The average ages for the three groups of Mandarin speaking children were 2;6 (1;11-2;10), 3;6 (2;11-3;10), and 4;3 (3;11-4;9).

The English speaking children were recruited from preschools and daycares in the greater Boston area, as well as from a database maintained by Harvard's Laboratory for Developmental Studies. The majority of the Mandarin speaking children were recruited from preschools in Taipei and Chiayi, two cities in Taiwan. A few additional Mandarin-speaking children were recruited from the neighborhood where the experimenter was residing in Taipei. Children from

both language groups were from middle to upper middle class. Children were tested individually in a quiet room. All the children received a small gift for their participation.

**Stimuli and Procedure**—The procedure was identical to that of Experiment 1, except that each child was tested individually in a quiet location at a school, a laboratory room, or the child's home, and the experimenter recorded the child's responses on an answer sheet. Additionally, the training trials were repeated until answered correctly. This was rarely necessary.

Half of the children in each age group in each language (16 of 32) were randomly assigned to the Label condition and half to the No-Label condition.

## Results

Figure 6 plots the percentage of shape-match response for each age group of each language. We collapsed the Label and No Label groups for the Figure, as initial analysis indicates that there was no effect of instruction. Two results are apparent: like adults, children differentiated the entity types, and unlike what was found for adults, there are no systematic effects of language type.

A  $2 \times 2 \times 3 \times 3$  AVOVA examined the effects of Language Type (English, Mandarin), Instruction (Label, No Label), Age (2-year-old, 3-year-old, 4-year-old), and Entity type (Complex, Simple, Non-solids) on percentage of choices of the shape-match. There was a main effect of Entity Type ( $F(2, 360) = 331.3, p < .001, \eta_p^2 = .65$ ), with shape-match choices for Complex Solids (84.8%) greater than Simple Solids (44.7%) and shape-match choices for Simple Solids greater than Non-Solids (20.3%; all pair-wise comparisons significant,  $p$ 's  $< .001$ ). There was no main effect of Instruction, nor any interactions involving this variable

Importantly, there was no main effect of Language Type. ( $p = .36, \eta_p^2 = .01$ ). Mandarin-learning children made no fewer shape-match choices than did English-learning children (49% vs. 51%). There was, however, a significant Language Type  $\times$  Entity Type interaction ( $F(2, 360) = 3.99, p = .02, \eta_p^2 = .02$ ). Pair-wise t-test comparisons of the percentage of shape-match choices of English and Chinese children for each entity type revealed no interpretable results. The two groups did not differ on the Simple solids (English, 48% shape-based judgments; Mandarin, 41% shape-based judgments,  $t(190) = 1.42, p = .16$ ). For the Complex solids English children were *less* likely to choose based on shape (English, 82%; Mandarin, 88%,  $t(190) = 1.82, p = .07, ns.$ ) and for the Non-solids, the trend was in the opposite direction (English, 23%, Mandarin, 17%,  $t(190) = 1.67, p = .10, ns.$ ). None of these trends is significant on its own. We therefore believe this interaction to be spurious, reflecting noise.

There was a main effect of Age ( $F(2, 180) = 4.19, p = .017, \eta_p^2 = .04$ ) and an Age  $\times$  Entity Type interaction ( $F(4, 360) = 5.70, p < .001, \eta_p^2 = .06$ ). The average percentage of shape-match choice for the two-, three-, and four-year-olds were 55%, 48%, and 47% respectively. Pairwise t-test comparisons singled out the two-year-olds as the source of the Age effect. Two-year-olds differed from the three-year-olds ( $t(126) = 2.44, p = .02$ ) and from the four-year-olds ( $p = .01$ ). The three- and four-year-olds did not differ from one another ( $p = .75$ ). Further univariate ANOVAs for each entity type isolated the source of the interaction. The different age groups did not differ for the Complex and Simple Solids ( $p$ 's  $> .53$ ). The two-year-olds differed from the three- and four-year-olds ( $p$ 's  $< .001$ ) for the Non-solids. Two-year-olds were relatively more likely to pick the shape-match for the Non-solids than the three- and four-year-olds (33% shape-match choice for the two-year-olds vs. 15% and 11% for the three- and four-year-olds respectively).

These data are consistent with those of many other studies, including Soja et al.'s original studies and Imai and Gentner's, showing developmental changes between ages 2 and 3 in the categorization of non-solid substances (see also Samuelson, 2002; Smith, Jones, Landau, Gershkoff-Stowe, Samuelson, 2002). Smith and her colleagues argue that the shape bias that is acquired through learning predominantly object kind words is initially over-extended to non-solids because exposure to substance kind words is so rare. Alternatively, many studies show developmental changes in conceptions of materials and material kinds that continue to unfold into late childhood (Au, 1994; Carey 1991; Smith, Carey, Wiser, 1985). Whatever the nature of this effect, the absence of a higher-order interaction with language type indicates that it is a general developmental trend. This complicates the straightforward developmental predictions concerning the locus of the emergence of the effects of language—an increase in object-kind categorization by English speakers relative to speakers of a classifier language (not yet observed by age 4 in these data) will have to be interpreted against a general developmental increase in the availability of substance kind construals, observed equally in English and Mandarin learning children.

To examine whether English speakers increased in shape choices and Mandarin speakers increased in material choice over time, we included the adults from Experiment 1 in an overall analysis. Figure 5c plots the percentage of shape-match choices as a function of age, collapsing across all 12 items. For English speakers, the slope of the regression line is positive ( $t(126) = 1.77$ , one-tailed:  $p = .04$ ), indicating that adults were slightly more likely to choose the shape-match than were children (adults: 58.6% shape-match choice vs. children: 51.0%). For Mandarin speakers the slope is significantly negative ( $t(126) = -3.99$ ,  $p < .001$ ), indicating that adults were less likely to choose the shape-match than were the children (adult: 34.1% vs. children: 48.8%). As in Imai and Gentner's data (Figure 5a), speakers of both language groups changed significantly with age—English speakers towards more shape-match choices, Mandarin toward more substance-match choices, although the divergence was observed earlier in development in Imai and Gentner's data than in the present data.

## Conclusions

Three important findings emerged from Experiment 4. First, the largest effect by far is the effect of entity type. At each age the speakers of the two languages overwhelmingly agree in their construal across the entity types. These data replicate Imai and Gentner, and again militate against a strong linguistic relativity account of the effect of language type on entity construal, for they provide no evidence that young learners of classifier languages ever take nouns to mean something different from what they mean to speakers of a count/mass language. Rather, speakers of both languages we tested have the same two construals available to them: nouns for objects can refer to the kind of object or to the kind of material it is made from, and the same factors influence which construal is made. In the next experiment, using the stimulus set from Experiment 2, we put Prasada's analysis of cognitive universals to further test on children.

Second, unlike the adults, the children did not differ from each other cross-linguistically. They did not differ even at four years of age, the oldest age tested. The result contrasts with Imai and Gentner (1997)'s Japanese-English data, which showed an unusually early language type effect in a lexical projection paradigm (i.e. Label condition). The result is however consistent with the later emergence of the language type effect from the Lucy and Gaskins (2001; 2003), Subrahmanyam and Chen (2006), or Imai and Mazuka (2003)'s No Label condition. In Experiment 5, we seek to pinpoint the exact age of emergence between English and Mandarin speakers with slightly older children and a different set of stimuli (i.e., the set from Experiment 2).

Third, the analysis of preference for shape-match choice over time showed that English speakers slightly increased their shape preference with age, consistent with the lexical statistics

hypotheses. However, Mandarin speakers showed an even greater decrease in shape-match choices over time. Perhaps learning a classifier system may cause the material preference, as Lucy's lexical semantics account would have it. However, our own data from both children and adults and several other noun rating studies (Sandhofer, Smith, & Luo, 2000; Colunga & Smith, 2000; Barner, Inagaki, & Li, under review) already suggest that nouns in Mandarin and Japanese do not more often denote substance kinds than nouns in English, making it unlikely that learning a classifier system would lead to lexical reanalysis of the nouns as denoting unindividuated essence.

Alternatively, the observed developmental changes may have another explanation. It may be that across both languages material becomes more salient with age; children reorganize their understanding of substances between ages 4 and 10 (Au, 1994; Carey 1991; Smith, Carey, Wiser, 1985), and acquire a much larger stock of words referring to substances, both solid and non-solid (Samuelson & Smith, 1999; Prasada, 1993). The effect of age (in speakers of *both* languages) on the increase in percentage of material-matches for the non-solid substances is consistent with this literature. Thus, the developmental functions observed in these studies may reflect two interacting factors: an increase in salience of material with age (in both cultures) together with a language-driven effect of the count syntax distribution in English.

## Experiment 5: Complexity and Functional Relevance for English and Mandarin-speaking Children

Experiment 5 extends the age range tested, seeking to replicate Experiment 4 with a new set of stimuli, and to determine when the effect of language type on entity construal emerges. We use the constructed, ambiguous stimuli from Experiment 2, rather than the store-bought stimuli of Experiment 1. We also seek to extend Prasada and colleagues' experimental work with English speaking children (Burger & Prasada, 1997) to test for universals across the different groups of speakers.

Burger and Prasada (1997) found that children were sensitive to many of the same cues to entity construal as adults, such as whether the entity had a smooth and regular look as opposed to a jagged and torn look, and whether the item's shape appeared function relative. However, the children were insensitive to multiple copies as a cue. Based on these prior studies, we predict that cues to function relevance will have a large effect on likelihood of shape based choices at all ages and in all languages, and complexity per se may not affect children's choices (unlike the adult pattern of response in Experiment 2, reflecting the presence of two copies of the same complex, irregular, shapes).

Following the procedure with adults in Experiment 2, children were only tested in the Label condition.

### Methods

**Participants**—English and Mandarin speaking children who had not participated in the previous experiments were recruited in the same manner and from the same subject pool as Experiment 4. Of the English speaking children, 18 three-year-olds ( $M=3;4$ , range 2;11-3;10), 19 four-year-olds ( $M=4;7$ , range 4;0-4;11), and 21 five-year-olds ( $M=5;4$ , range 5;0-5;11) participated. The children were tested in quiet locations at their preschool or at our laboratory. The Mandarin speaking children were recruited from three preschools at Taipei, Taiwan, and tested individually in a quiet classroom at their preschool. 16 three-year-olds ( $M=3;4$ , range 2,8-3;10), 20 four-year-olds ( $M=4;7$ , range 4;0-4;11), and 16 five-year-olds ( $M=5;6$ , range 5;0-5;11) participated. All children received a small toy for their participation. The number of female and male participants was roughly equal within each language and age group.

**Stimuli and Procedure**—The stimuli and procedure was identical to that of Experiments 2. After presentation of the fork and jam familiarization trials, as in the previous studies, the experimenter presented a standard and labeled it (e.g., the blicket). He or she then presented the two alternatives and asked which one was the blicket. This process was repeated for the entire set of 16 triads.

## Results

As in Experiment 4, two analyses are of interest: the developmental course of the language type effect and the stimulus features that affect children's construal.

**By-Item Cross-linguistic Comparisons and Developmental Shift**—Figures 7a-7d (including adults from Experiment 2) graph each age group's by-item plots as in Experiment 3. The Mandarin speakers' percentage of shape-matches for each entity is plotted as a function of English speakers' percentage of shape-matches for that entity. If speakers of both languages are influenced by the same stimulus features in their construal, at every age, the correlations between the two sets of judgment should be high. Indeed, the correlations by items for each age group are highly significant (all Pearson's  $r > .69$ ,  $p < .01$ ).

If language type effect on construal modulates the universal bases for entity construal such that speakers of a classifier language make relatively more substance construals, this will be reflected in a distribution that falls below the diagonal in Figures 7a-7d. For the three-year-olds, the items fell roughly along the diagonal, but the items fall further away and below the diagonal with increasing age. Paired-t test of the items' shape-match percentages comparing Mandarin and English speakers indicates that at three-years of age, the two language groups do not differ ( $t(15)=-0.76$ ,  $p = .94$ ); however, the older groups did differ ( $p < .05$ , in each case; see Figure 7). English speaking four-year-olds, five-year-olds and adults prefer the shape-match relatively more than do Mandarin speaking participants of these ages.

Experiment 5 finds the difference due to language type already at age four ( $t(15)=2.75$ ,  $p = .02$ ), whereas the four year-old Mandarin and English speaking children did not differ in Experiment 4. The average age for the four-year-olds in this experiment is older than the four-year-olds in Experiment 4 (4;3 vs. 4;7), and so we may have successfully located the onset of the language difference. To examine this more closely, we divided Experiment 5's four-year-olds of each language into a young-four group (4;0-4;5) and an old-four group (4;6-4;11). Divided this way, the young four-year-olds, (consisting of 6 English speakers ( $M=4;3$ ) and 8 Mandarin speakers ( $M=4;3$ )), did not differ ( $t(15)=.44$ ,  $p = .65$ ). In contrast, the older four-year-olds (consisting of 13 English speakers ( $M=4;10$ ) and 12 Mandarin speakers ( $M=4;10$ )) did differ ( $t(15)=3.40$ ,  $p < .01$ ). Thus, the older four-year-olds children carried the effect of language type observed at this age in Experiment 5. The current finding suggests that perhaps the language type effect on entity construal develops between four and five.

Finally, we checked each language group's overall preference for shape-matches across age (see Figure 5d). The change in percentage of shape-match choice from childhood to adulthood is positive (+6.5%) for the English speakers, from 53.6% shape-match choices for the children to 60.1% for the adults. The change is negative (-10%) for the Mandarin speakers (43.2% to 33.2%). The change in percentages again is greater for Mandarin speakers than for English speakers. However, unlike Experiment 4, the trends from linear regressions of age (years old) in predicting percent shape preference (i.e., the slopes of the lines in Figure 5d) were not statistically significant at even the  $p = .10$  level (English:  $p = .17$ , and Mandarin:  $p = .11$ ). Of course, Experiment 5 has older child participants, who were already midway to reaching adult levels.

**Complexity, Functionality, and Solidity**—In order to examine the effects of the entity's shape complexity and potential shape-dependent functionality on children's entity construal, we calculated the average percentage of shape responses for each entity type. Altogether there were five entity types consisting of non-solids (NS) and four types of solid entities from crossing complexity with functionality (+C+F, +C-F, -C+F, -C-F). The results are plotted in Figures 8. For simplicity of presentation, we averaged across the three age groups of children since separate ANOVAs conducted for each age group revealed that the effects of complexity and functionality are the same across the different ages (see below).

The figure shows functionality playing the main role in affecting construal, while complexity plays little or no role.

**Solids:** An ANOVA examined the effects of Age (3-year-old, 4-year-old, 5-year-old), Language Type (English, Mandarin), Complexity (complex, simple) and Functionality (likely, unlikely to have shape-dependent function) on the percentage of shape-match choice given by the children. The analysis confirmed the two effects of language type that we have already seen from the by-items analyses above. First, there was a main effect of Language Type ( $F(1, 104) = 4.86, p = .03, \eta_p^2 = .05$ ), with English speaking children categorizing by shape more than Mandarin speaking children (English: 59.3% vs. Mandarin: 46.5%). There was also a weak Language Type by Age interaction, with larger differences in the older children. The Language Type  $\times$  Age interaction was significant in the items analysis ( $p = .03, \eta_p^2 = .35$ ), but not in the subject analysis ( $p = .56, \eta_p^2 = .01$ ). Paired t-tests corroborated the increasing difference between the two language groups with age. Mandarin and English speaking children did not differ at three years of age ( $t(11) = .56, p = .59$ ), but did differ at four ( $p < .01$ ) and five ( $p < .001$ ).

Next we turn to new information revealed by the analysis. There is a main effect of Functionality ( $F(1, 104) = 148.70, p < .001, \eta_p^2 = .59$ ), and no significant interactions involving that factor. Overall, solids that are likely to have shape-dependent functions received more shape-match choices than those that did not (+F: 67.8% vs. -F: 38.1%). There is a main effect of Complexity ( $F(1, 104) = 1.49, p = .02, \eta_p^2 = .05$ ; although this was not significant by items analysis,  $p = .47, \eta_p^2 = .07$ ) and a Complexity by Age interaction ( $F(2, 104) = 5.87, p < .01, \eta_p^2 = .10$ ). The percentage of shape-matches was 50.6% for complex shapes and 55.3% for simple shapes, in the opposite direction of the adults of Experiment 2. Pair-wise t-tests found only the three-year-olds differed for complex and simple (+C: 49.4% vs. -C: 63.9%,  $t(33) = -3.59, p = .001$ ), but not the four-year-olds (+C: 49.7% vs. -C: 49.4%,  $t(38) = .12, p = .90$ ) and the five-year-olds (+C: 52.6% vs. -C: 52.6%,  $t(36) = .16, p = .88$ ). The data are consistent with three-year-olds, the youngest children, being least influenced by seeing multiple same-shape copies of the stimuli (Burger & Prasada, 1997). That is, the three-year-olds were like the adults who were presented with one copy of the target stimuli in Experiment 3 and subsequently rated the less complex items to more likely be seen as objects. No other significant main effects or interactions were found.

**Non-solids:** A 2 by 3 ANOVA examined the effects of Language Type (English vs. Mandarin) and Age (3-year-old, 4-year-old, 5-year-old) on the percentage of shape-match choices on the non-solid trials. There were no effects of Language Type ( $F(1, 104) = .03, p = .86$ ), Age ( $p = .19$ ), nor any Language Type  $\times$  Age interaction ( $p = .23$ ).

To examine the effect of solidity on construal, we compared the percentage of shape-match choices for the non-solids to the +C-F solids, which had similar stimuli ratings with respect to complexity and likelihood of shape-dependent function. An Age (3-year-old, 4-year-old, 5-year-old) by Language Type (English, Mandarin) by Solidity (+C-F solids, non-solids) ANOVA found no effect of Solidity (37.1% for solids vs. 35.5% for non-solids,  $F(1, 104) =$



04,  $p=.84$ ). The null effect is likely attributable to the small proportion of non-solid trials. The non-solids comprised only four of the 16 test trials, making non-solidity less salient of a variable. However, an Age  $\times$  Solidity interaction ( $F(2, 104)=5.78, p < .01, \eta_p^2 = .10$ ) indicated that children become more sensitive to the entities' material composition with age. The youngest children in both language groups (i.e., three-year-olds) behaved unlike the adults by accepting more shape matches for the solids than the non-solids (+C-F solids: 30% vs. non-solids: 43%,  $t(33) = 2.36, p = .02$ ), while the older children, four- (+C-F solids: 35% vs. solids: 28%, n.s.) and five-year-olds (+C-F solids: 39% vs. 34%, n.s.), began to show the adult pattern. Thus, Experiment 5, like Experiment 4, suggests a developmental increase in sensitivity to a feature that signals non-cohesive substances that holds equally for children in both language groups.

## Conclusions

With a new set of stimuli, Experiment 5 replicated Experiment 4's finding that three-year-old and young four-year-old English learners do not make more shape based or object kind interpretations of a novel noun applied to a novel material entity than do Mandarin learners. It pinpoints four-and-a-half to five years of age as when the effects of language type observed in Experiments 1 and 2 are first seen. Thus, the age of emergence of the language type effect we observe truly is different from that of Imai and Gentner (1997), and more like that of Subrahmanyam and Chen (2006) who did not find robust language type effects on construal between Mandarin and English speaking three-year-olds. To establish whether the effect of language type on entity construal emerges robustly earlier in a Japanese/English comparison than in a Mandarin/English comparison, Experiment 6 tests Japanese learning children with the materials from Experiment 5. The emergence of the effect language type in the Mandarin/English compare at age four and a half is unlikely to be tied to the acquisition of the classifier system, since Li, Barner, and Huang (2008) showed that Mandarin-speaking four-year-olds do not know how classifiers function to mark individuation.

In qualitative terms, Experiment 5 also replicated the pattern of developmental change observed within each language group: English speakers slightly increase the proportion of shape-based categorization with age and Mandarin speakers slightly increase the proportion of material-based categorizations with age. The developmental changes were not as striking as in Experiment 4, due to the fact that the children in Experiment 5 were older and already showing the adult pattern. This developmental pattern, now seen in Lucy's data, Imai and Gentner's, and ours, is not that predicted by the lexical statistics hypothesis, according to which the effect of language type should be driven entirely by speakers of count/mass languages, for only they should be influenced by the probability in their language that a novel noun heard in neutral syntax is a count noun. However, if there is a language independent developmental increase in sensitivity to material kind that affects speakers of all languages, the observed pattern of change is what would be expected, even if the effect of language type is driven entirely by the English speakers. Both Experiments 4 and 5 provided data consistent with a language-independent developmental change in sensitive to material kinds—children speaking both languages become equally more sensitive with age to the cue of non-solidity per se in their entity construals.

Finally, Experiment 5 confirms that the features that influence entity construal, although changing subtly with age, are cross-linguistically universal at each age. We replicate Prasada's findings that young children are not sensitive to the repetition of a complex shape in their entity construal; unlike adults in this triad test, children do not make more shape-based construals for the complex shapes. Our data show that at all ages, in both language groups, the relevance of the shape to potential function is a far more important determinant of a shape-based response

than is the complexity of the shape. Consistent with Prasada's analysis, function-relevance is one demonstrated cue to non-accidental structure.

## Experiment 6: Replication of Experiment 5 with Japanese Speakers

Experiment 6 is a replication of Experiment 5 with Japanese speaking children. At issue is whether the effect of language type is present earlier in a Japanese/English comparison than in a Mandarin/English, as suggested by the comparison of Imai and Gentner's findings and those of Experiment 4.

### Methods

**Participants**—The native Japanese speaking children come from Kirishima Town of the Kagoshima Prefecture. They were recruited from a local preschool, where most children were four years of age or older. Thus, only two groups of children, 14 four-year-olds ( $M=4;7$ , range 3;9-4:9) and 16 five-year-olds ( $M=5;8$ , range 5;0-6;3) were tested. The children were tested individually in a quiet classroom.

**Stimuli and Procedure**—The stimuli and procedure were identical to Experiment 5, except the experiment was conducted in Japanese (see Appendix A for the Japanese wording of the instructions).

### Results

**By Items Cross-linguistic Comparison and Developmental Shifts**—We first plotted the Japanese participants' percentage of shape-match choices for each entity as a function of those of the English participants (See Figure 9). We included the Japanese and English adults in Experiment 2 (Figure 9c). There is a positive relationship between the Japanese and English speakers' proportions of shape-match choices (all Pearson's  $r > .59$ ,  $p$ 's  $< .05$ ). These high correlations, along with those of all previous studies, strongly suggest that speakers of classifier languages, such as Japanese and Mandarin, are guided by the same features as are English speakers in their construals of material entities as members of object kinds vs. substance kinds. The effects of language type, when observed, modulates cross-linguistically universal bases for entity construal.

Paired t-tests (see Figure 9) indicated that the function relating percentages of shape-based responses per item of English and Japanese speakers did not deviate from the diagonal for four-year-olds ( $t(15)=-.86$ ,  $p = .41$ ). Only the older groups (five-year-olds and adults) showed the often observed pattern of more shape-match choices among English speakers (t-tests,  $p < .01$  in each case). Thus, the age of emergence of the language type effects we have seen in our Mandarin-English comparisons was replicated with our Japanese speakers.

Figure 5d displays the change in percent shape-based responses for Japanese speakers. Just as in the Mandarin data, the slope is negative (49.8% to 40.6%), and just as in the Mandarin data, the regression did not differ significantly from 0 slope ( $p = .11$ , 1-tailed). As with the Mandarin data, we attribute the lack of significance to a small sample size and to the fact that the children in this study were even older than those in Experiment 4, and already showing the adult pattern.

**Complexity, Functionality, and Solidity**—To examine the effect of complexity and functionality, we calculated the average percentage of shape responses for each entity type. The results for the children are plotted in Figure 10. We included the English speakers in the Figure for comparison, plotting the data only for the comparably aged four- and five-year-old English speakers. The comparison indicates that the difference between Japanese and English speaking children is small. With respect to Functionality and Complexity, the Japanese data

closely resembled the previous English and Mandarin data. Functionality affects construal at all ages, whereas Complexity only affects construal in adulthood. (See Figure 2 for adult data).

**Solids:** A  $2 \times 2 \times 2 \times 2$  ANOVA examined the effects of Age (4-year-old, 5-year-old), Language Type (English, Japanese), Complexity (complex, simple) and Functionality (likely, unlikely to have shape-dependent function) on the percentage of shape-match choices. The main effect for Language Type was not significant ( $F(1, 67) = .69, p = .41$ ). There was a small Language Type  $\times$  Age interaction, which was not significant in the subjects analysis ( $p = .59, \eta_p^2 < .01$ ), which did reach significance in the items analysis ( $F(2(1, 8) = 5.14, p = .05, \eta_p^2 = .39$ ). The interaction reflected the fact that four-year-olds did not differ as a function of language type (English: 57% by shape, Japanese: 55% by shape,  $t(11) = .61, p = .55$ ), but five-year-olds did (English: 62% by shape, Japanese: 52% by shape,  $t(11) = 3.9, p < .01$ ).

There was no main effect of Complexity ( $F(1, 67) = .46, p = .50, \eta_p^2 = .01$ ). Again, unlike adults, children were relatively unaffected by complexity (Complex: 55.8% by shape, Simple: 57.5% by shape). In contrast, there was a highly significant main effect of Functionality ( $F(1, 67) = 80.75, p < .001, \eta_p^2 = .56$ ), with more shape-match choice for entities rated highly likely that the shape had functional relevance (71.1%) than for those rated as unlikely to have shape-dependent functions (42.3%). No other effects reached significance.

**Non-solids:** A  $2 \times 2$  ANOVA examined the effects of Language Type (English and Japanese) and Age (4-year-old, 5-year-old) on the percentage of shape matches on the non-solid substances. There was no effect of Language Type ( $F(1, 67) = .35, p = .56$ ), Age ( $p = .83$ ); nor was the interaction between Language Type and Age significant ( $F(1, 67) = 3.50, p = .07, \eta_p^2 = .05$ ).

In regards to the effect of solidity, per se, an ANOVA examined the effects of Age (4-year-old, 5-year-old), Language Type (English, Mandarin), and Solidity (+C-F solids, non-solids) on the percentage of shape-matches. As in Experiment 5, the Japanese/English analysis revealed no significant main effects, and unlike in Experiment 5, the interaction between Age and Solidity was not significant (all  $p$ 's  $> .07$ ). Neither English nor Japanese speaking children were more likely to categorize the +C-F solids by shape than the non-solids (41% by shape for solids vs. 37% for non-solids).

## Conclusions

There is a remarkable convergence of the results on Mandarin speakers and the current results on Japanese speakers. The effect of language type, with more shape-matches by English speakers than by speakers of either classifier language, emerged around late four or early five years of age. Our data differ from Imai and Gentner's, but converge with those of Subrahmanyam and Chen (2006), as well as those of Imai and Mazuka (2003; 2007) on their similarity judgment task in not finding a robust effect of language type on entity construal in children under 4. This developmental pattern emerged when tested on the non-manufactured entities of Experiments 2 and 3. It also emerged when tested on the entities of Experiment 1 that were modeled after Imai and Gentner, and Soja et al studies. We have no explanation for why Imai and Gentner found the effect of language type earlier (by age 2 ½); all of the studies taken together suggest that it is not robustly present until children are a couple of years older.

Taking the data from Experiments 4, 5, and 6 together, we see that as predicted by the lexical statistics hypothesis, count/mass speakers slightly increased in their likelihood to construe ambiguous novel entities as object kinds with increasing age. On the other hand, classifier language speakers, increased in their likelihood to construe ambiguous entities as substance kinds with increasing age.

Experiment 6, like Experiments 2, 3, and 5 also replicated and greatly extended Imai and Gentner's initial findings that speakers of each language are affected by the same stimulus parameters in their construal of entities. In Experiments 2, 3, 5, and 6, the correlation of the rating or percent shape matches for the 16 items between the speakers of whichever two language groups being compared is always very high, at every age tested.

Our handcrafted, ambiguous stimuli teased apart the relative contribution of shape complexity and function-relevance. The results provided convergent support for Prasada's account of what it means to represent something as a kind of object or substance, and extended what he and his colleagues found of English speaking children and adults to Japanese and Mandarin speaking children and adults (Burger & Prasada, 1997; Prasada et al., 2002). Functional affordance of shape played a strong role in promoting construals in terms of object kinds for speakers of all three languages at every age tested. Complex shapes, if irregular, decreased the likelihood of construal as members of an object kind (Experiment 3), unless there are multiple copies (Experiment 2). Adults, unlike children, are sensitive to the fact that complex shapes being duplicated twice is unlikely to be accidental, and thereby are more likely to treat entities with a duplicated complex shape to be members of a kind of individual. Adults are also more sensitive to non-solidity, per se, than are children. These latter two effects are independent of language spoken.

## General Discussion

The phenomenon Lucy (1992) first documented and that Imai and Gentner (1997) confirmed -- the effect of language type (count/mass syntax vs. classifier languages) on entity classification -- is indeed robust. Experiment 1 extended the Imai and Gentner's findings to a comparison of Mandarin-speaking and English-speaking adults. Shown a conical piece of cork called "the blicket", English speakers are more likely to interpret "blicket" to mean cone than are Mandarin and Japanese speakers, who are more likely to interpret the word to mean cork. Experiment 2 revealed this cross linguistic difference yet again with a new stimulus set that contained no manufactured or familiar items. Thus, the phenomenon has now been demonstrated when speakers of English, on the one hand, and speakers of several different classifier languages, on the other, are asked to classify a wide variety of material entities. Alleviating worries that something about the English speaking participants other than that their language has a count/mass distinction is responsible for the differences between them and Mandarin/Japanese speakers, recent studies show that speakers of other count/mass languages (Spanish and Czech) make an equally high number of shape/object kind matches as do English speakers (Gathercole & Min, 1997; Crane, 2002).

The present studies were designed to elucidate two complementary questions concerning the role of language type on the construal of material entities: First, exactly why does speaking a count/mass language make object kind construals more likely than does speaking a classifier language? Second, is it true that the effect of language type modulates cross-linguistically universal factors that determine entity construal, and if so, what are these factors? We discuss what has been learned with respect to each of these two points in turn.

### Understanding the effect of language type on entity classification

Lucy's landmark books (Lucy, 1992; 1996) reintroduced Whorfian hypotheses into cognitive science, at a time in which the field had come to the consensus that Whorfian effects of language on thought do not exist (see Pinker, 1994, for a relatively recently review of the anti-Whorfian consensus). The effects of language type on construal of material entities as members of kinds of individuated or unindividuated entities was Lucy's empirical example, and many other researchers have taken the effects documented by him, and by Imai and Gentner, as a paradigm example of Whorfian effects of language on thought. However, the present studies, which are

largely consistent with all other extant data, support an alternative position—these are effects of language on linguistic representations, not effects of language on non-linguistic representations. Of course, we must think to speak; our language reflects non-linguistic representations. Slobin (1996) put it well when he pointed out that language differences may well influence thinking for speaking.

The conceptual representation of object and substance kinds is universal across speakers of classifier languages and speakers of count/mass languages. These findings thus militate against a strong Whorfian position that variations in nominal number determine whether a person distinguishes individuated from non-individuated entities at all (Whorf, 1939/1956; Quine, 1960). One goal of the present studies was to test a version of a weak Whorfian interpretation of the phenomena Lucy discovered and Imai and Gentner dramatically confirmed. Perhaps the language we speak can set our ontological boundaries. A proposal well worth entertaining is that the presence of count syntax influences the weighting of whatever features influence construal as a member of a kind of individual (whether these be perceptual or conceptual features), making participants more sensitive to these features whenever they form a representation of any given entity. The essence of this very interesting Whorfian proposal is the scope of the effects across different tasks. If every time an English speaker encoded a material entity, he or she is more likely to construe that entity as a member of an object kind than is a speaker of a classifier language, this would be a clear Whorfian effect of language on thought. Consistent with this Whorfian hypothesis, language type affected construal not only in lexical projection but in tasks requiring similarity comparison (Imai and Mazuka, 2003; Experiment 1).

However, the Whorfian universal feature weighting hypothesis is incompatible with Experiment 3. In Experiment 3, adult speakers, people with prolonged exposure to their native languages, were explicitly asked to classify each entity as a likely instance of an object kind or substance kind. Although the judgment we were asking for was explicit, participants were free to weigh whatever factors underlie their construals however natural to them. Under these straight-forward instructions, Japanese and Mandarin speakers did not differ from English speakers in how they rated the entities. Had language-specific patterns created differential weighting of cues to membership in object kinds, effect of language type should have remained, with English speakers more prone than speakers of classifier languages to construe something as a kind of object. There was no hint of any such effect.

The null effect of language type on construal in the procedure of Experiment 3 has since been replicated in our laboratory using the same rating instructions on a larger set of stimuli (Dunham, Li, & Carey, in prep). We have also replicated the same pattern of findings in a comparison of English speakers with speakers of Tzotzil Mayan, a classifier language similar to the Mayan language studied by Lucy (Chen, 2007). Again, Chen observed the effect of language type on entity construal on the lexical projection task, but not in the task that required a simple rating of each entity as a member or an object kind or a substance kind.

We take the total pattern of findings as supporting the lexical statistics hypothesis (Gleitman & Li, 2002). There is ample evidence that hearing a noun in a count context forces a construal of a novel entity as a kind of individual (Bloom, 2000; Dickinson, 1988; Subrahmanyam, Landau, and Gelman, 1999; Gathercole & Whitfield, 2001; Barner & Snedeker, 2006). In the case of material entities, the relevant individual kinds are most often object kinds. Because most nouns in English are count nouns (Samuelson & Smith, 1999; Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994), a noun heard in neutral syntactic context is probabilistically assigned a count interpretation, influencing classification of entities under that noun on the basis of shared object kind. Especially when trying to interpret the meaning of a noun labeling an ambiguous entity, which itself provides little in terms of cues supporting one construal over

another, English speakers rely heavily on distributional properties of their language when interpreting a novel noun heard in a neutral syntactic context.

A recent study by Barner, Inagaki, & Li (under review) found that English-Mandarin bilinguals, tested in English, were no different than native English speakers in preferring the shape-match for novel nouns in neutral syntax. However, when tested in Mandarin, the bilinguals preferred the material-match, just like monolingual Mandarin speakers. No general ontological boundary has been set by speaking a classifier or count/mass language; the lexical statistics of the language in which a novel noun is encountered affect the meaning assigned to that novel noun.

What, then, of the finding that the effect of language type is observed in similarity judgment tasks as well as lexical projection tasks (Imai and Mazuki, Lucy, Experiment 1)? We see two possible interpretations of this finding. First, the same explanation given to why language type effects are observed on the lexical projection task may apply equally to the similarity judgment task. The pragmatic context in which the experimenter points to an entity (“Look at this; which is the same”) is not all that different than (“Look at this blicket; which is the blicket?”). The deictic pronoun “this”, as the head of the noun phrase, is serving as a linguistic stand in for a noun label. As a result, the similarity judgment task may be subject to the same lexical statistics. This hypothesis might be tested by changing the wording of the triad task such that no pronoun is used (“Look here. Which is more similar?”). If the lexical statistics interpretation of the effects of language type on similarity judgments is correct, the effect of language type should be ameliorated under these instructions. If not, the findings would support a weak Whorfian effect that emerges very widely on tasks where entities are compared to one another for the purpose of classification.

All in all, these studies militate against several versions of the linguistic relativity hypothesis. At least in the case of object and substance kinds, our basic ontological commitments are neither constructed nor powerfully constrained by the language we speak. Nor do our obligatory linguistic choices lead to a universal change in the weighting of the features that enter into these choices. An important lesson from this case study is that linguistic tasks cannot settle a question of whether language affects non-linguistic thought. That we are sensitive to naming conventions in their own language is merely to say we are effective language learners; it is not to say that those conventions inexorably shape our conceptual system.

The distinction between object kind and substance kind is a central piece of our core ontology, integral to our ability to make sense of and navigate a complex and shifting world. As such, it may not be surprising that, far from being highly malleable, it should prove itself quite resistant to linguistic or cultural influence, part of the shared conceptual endowment of our species.

While deflationary with respect to Whorfian effects of language on thought in this case, these data obviously do not rule out other such effects. For example, Carey (in press) argues for strong Whorfian effects in cases of conceptual change, and Boroditsky and her colleagues have provided evidence for many cases of weak Whorfian effects (e.g., of gender marking on representations of inanimate objects, of language on temporal representations, Boroditsky, Schmidt, & Phillips, 2003; Casasanto & Boroditsky, 2003). We take it as shown beyond doubt that language affects thought; our goal in these studies has been to explore exactly how it does so in this particular case.

### **Cognitive Universals**

The linguistic typologists' individuation hierarchy presupposes universal factors that influence entity construal, creating a set of parameters for languages to choose among in deciding which entities, if any, may be lexicalized count nouns (Croft, 1990; Lucy, 1992). All inanimate solid objects are treated alike in the linguists' typological hierarchy. In contrast, Imai and Gentner

suggested that cross-linguistically universal factors may distinguish among inanimate objects with respect to how they are construed, pointing to the fact that speakers of both English and Japanese at all ages gave more shape based responses for complex solids than simple solids, and more for simple solids than for non-solid substances. Experiments 1 and 4 confirmed this empirical pattern. Experiments 2, 3, 5 and 6 put the suggestion of cross-linguistic universality in the features of entities that influence object kind (table) vs substance (wood) construals to a much stronger test: All of the stimuli were ambiguous with respect to their construal, and they varied along many features. Yet the likelihood that a given entity was construed as a member of a kind of object rather than as a kind of substance by speakers of one language strongly predicted the likelihood that the same entity would be so construed by speakers of another language. This pattern held between Japanese speakers, English speakers and Mandarin speakers, and it held at every age.

Besides putting Imai and Gentner's universal individuation continuum to stronger test, these experiments allowed us to explore how that continuum should be conceptualized. There is a tension in the literature as to whether the features that influence shape-based categorization are perceptual features, per se, (e.g., complexity of shape, regularity of shape, cohesiveness), or whether the individuation continuum is better characterized at a more abstract level. Complex solids were judged both more likely to have a shape-based function and more perceptually complex than the other stimuli in Experiment 1 and in Imai and Gentner's studies. Shape complexity is a perceptual feature, whereas whether an unfamiliar artifact is likely to have a shape-based function clearly is not. Previous studies had not allowed exploration of the relative importance of these more abstract, conceptual, considerations and the more perceptual ones.

The entities of Experiments 2, 3, 5 and 6 were all hand-made and somewhat irregular and none had any actual function. We found that the likelihood that the shape deemed relevant to some imagined function, solidity, complexity of shape and repetition of shape all affected construal among adults of all three language groups. The data support Prasada's (1999) Aristotelian account of the universal factors that determine entity construal, in which the distinction between object kinds and substance kinds is highly conceptual and abstract. On Prasada's proposal, what leads people to construe a given entity as a member of an object kind or as a member of a substance kind is a judgment about the arbitrariness or non-arbitrariness of the entity's structure (form, arrangement of parts). Members of object kinds share non-arbitrary structure, whereas the structure of entities that share substance is largely unconstrained, and hence arbitrary. This analysis unifies the specific effects of complexity, repeated exemplars with the same shape, having a function related shape, and solidity found across the present experiments, as well as Prasada's studies. Prasada and his colleagues have also explored other cues to non-arbitrariness of structure, such as knowledge of function (even if it isn't shape dependent) or knowledge of the process that created a given entity, and showed that these also influence English speakers' entity construal. Recent research confirms that these latter two features equally shift entity construal by both count/mass language speakers and classifier language speakers (Gathercole & Min, 1997; Kobayashi, 1999). The ontological distinction between object kinds and substance kinds is a deeply conceptual one, and the factors that influence our construal unite an abstract analysis of it and are universal across speakers.

### The Developmental Story

With respect to the development of the language effect on construal, there are two phenomena we must account for. First, how are we to explain why the effect of language is not observed until ages 4 to 7 (Experiments 4 to 6, Subrahmanyam & Chen, 2006; Yoshida & Smith, 2003b, Lucy & Gaskins, 2001, 2003)? Second, how are we to understand why the greater developmental changes are observed among the speakers of classifier languages?

Clearly, the lexical statistics hypothesis implicates learning. Recent literature documents the slow emergence of other statistically based language-specific biases (e.g., Hohenstein, 2005; Haryu, Imai, Okada, Li, Meyer, Hirsh-Pasek, & Golinkoff, 2005). Similarly, Samuelson and Smith (1999) found that a shape-match bias in generalizing the referents of newly learned nouns is only be detected in children who have acquired a substantial number of nouns in their vocabularies. It stands to reason that it would take time to abstract the regularities from the input that support a bias for interpreting nouns in neutral syntax as count nouns.

With respect to the second question, the lexical statistics hypothesis dictates that the source of the influence of language type on entity construal is to be found in the speakers of count/mass languages, for only they would have probabilistic bases for interpreting nouns heard in neutral syntax as count nouns. And indeed, only speakers of English, a count/mass language, showed a developmental increase in the likelihood to interpret the referent of a newly heard novel noun as a member of a kind of object. However, the larger and more striking developmental change across all of these studies, as well as those of Lucy and of Imai and Gentner, is the increase in material-match choices with increasing age by the classifier-language speakers.

One possible interpretation of this pattern of findings is that these two developmental changes derive from two independent sources. The increase in object-kind construals among speakers of a count/mass language is explained by learning of the lexical statistics. In addition, there are well documented developmental increases in knowledge about substances and words for substance kinds (Au, 1994; Carey, 1991; Smith, Carey, Wiser, 1985) that are most likely cross-culturally universal, leading to greater sensitivity to material matches in the present experiments. On the two source hypothesis, the effect of increasing knowledge and sensitivity to substance kinds, although universally present, is contravened in the present studies in English speakers whose exposure to count syntax contributes to the tendency to interpret novel nouns as referring to object kinds. Obviously, further research would be needed to establish that the two source interpretation of the developmental pattern is correct, but if it is, then, as expected, the effects of language-type on entity classification are indeed due to the speakers of languages with count-mass syntax. This interpretation would be supported if a totally non-linguistic task showed greater sensitivity to material kinds with increasing age equally among speakers of count/mass and classifier languages. One reflection of this general developmental change, which is independent of language type, observed in the present studies is the developmental increase in sensitivity to non-solidity per se in affecting entity construal.

### Concluding Remark

Among other functions, language expresses thought and is the vehicle through which human beings acquire much of their knowledge. In recent years, cognitive scientists have shown that we now have the experimental and analytical tools to address the age-old questions concerning the relations between language and thought (see papers in volumes such as Gumperz & Levinson, 1996; Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003). Many of the experiments reviewed in these volumes have conclusively demonstrated language driven effects on the reasoning styles of speakers in cross-linguistic comparisons. But as all must agree, such findings are merely the first step. We must then seek to understand the exact locus of these observed effects in order to understand the complex interactions between the languages we speak and our mental representations of the world we shape and in which we live.

### Acknowledgments

We thank Sandeep Prasada, David Barner, Lila Gleitman, Anna Papafragou, Jesse Snedeker, Edward Kako, Laura Wagner, Mutsumi Imai, Sandy Waxman, and members of Harvard's Laboratory for Developmental Lab for helpful discussions. We thank Adam Okulicz-Kozaryn at Harvard MIT Data Center for statistical consultation. We thank Tracy Thomas, Leigh Borum, and Eva Xia for helping in data collection, Shu-hwa Lo, Miou-Ching Lo, and Jin-Yin Ung for recruiting children and adults in Taiwan, and the Graduate Linguistic Institute at National Taiwan University



for providing test space. This research was supported by an NIH NRSA Postdoctoral Fellowship F32HD043532 to Peggy Li, NSF ROLE Grant REC-0196471 to Elizabeth Spelke & Susan Carey, and NIH RO1-HD038338 to Susan Carey.

## Appendix

### Appendix A

#### English Instructions

##### Label Condition

On the initial presentation of the standard:

Look at this blicket.

On the initial presentation of the alternatives:

Which is the blicket?

##### No-Label Condition

On the initial presentation of the standard:

Look at this.

On the initial presentation of the alternative:

Which is the same?

#### Mandarin Instructions

##### Label Condition

On initial presentation of standard:

‘kan. zhe shi X.’

Look. This is X.

On the presentation of the alternatives:

‘na yibian shi X?’

Which side is X?

##### No-Label Condition

On initial presentation of standard:

‘kan zheli.’

Look here.

On the presentation of the alternatives:

‘na yibian shi yiyiang-de’

(which side is same-POS?)

Which side is the same?

#### Japanese Instructions

##### Label Condition

On initial presentation of the standard:

‘kore wo mite. kore wa X to iimasu’  
(this-Acc look. This-Top is named X).

Look at this! This is named X.

On presentation of the alternatives:

‘dochira ga buriketto desuka?’  
(which-Subj is X).

Which is X?

## Appendix B

Picture A	Picture B
	

### English

When we see something, we tend to think of it as either an object or a substance. For example, when we see the entity in Picture A, we generally think of it as an object, a table, regardless of what it's made out of. But when we see the entity in Picture B, we generally think of it as a substance, ice cream, regardless of its form.

We're interested in the factors that lead someone to think of something as either an object or a substance, and so we're asking you to rate how you see a series of entities.

Look at each entity and rate your likelihood of seeing that entity as an object or a substance on the 7-point scale. Circle 1 if you would definitely think of it as an object and 7 if you would definitely think of it as a substance.

## Mandarin

當我們看一樣東西時，我們常常會將它看成是一種物體或一種物質。例如，看圖片 A 中的東西時，我們一般將它看作是一件物體，桌子，而不管它是用什麼材料做成的。而在看圖片 B 的東西時，我們一般將它看成是一種物質，冰淇淋，而不管它的體形。  
我們想知道是什麼因素導致一個人將一樣東西看成是一種物體或一種物質，因此我們請您對一系列東西進行評價。  
請看每一樣東西，然後用七點量表對一樣東西是物體還是物質作出評價。如果您非常肯定地將一樣東西看作是一種物體，就在 1 上畫個圈；如果您非常肯定地將一樣東西看作是一種物質，就在 7 上畫個圈。

## Japanese

ものを見る時、そのものを物体か、物質として捕らえることが普通です。例えば、上記の 1 にあるものを見ると、物質が何といつても物体として、つまりテーブルとして捉える。逆に、上記の 2 にあるものを見ると、形が何といつても物質、つまりアイスクリームとして捉える。  
人がものを物体か、物質として捕らえる認知について研究しているので、色々なものを見てその捉え方について教えてもらう。  
ものを一つずつ見て、そのものが物体としてか、物質として捕らえやすいか、と答えてほしい。もし必ず物体として捉えると思つたら、1 に丸をつけてください。もし必ず物質として捉えると思つたら、7 に丸をつけてください。

## References

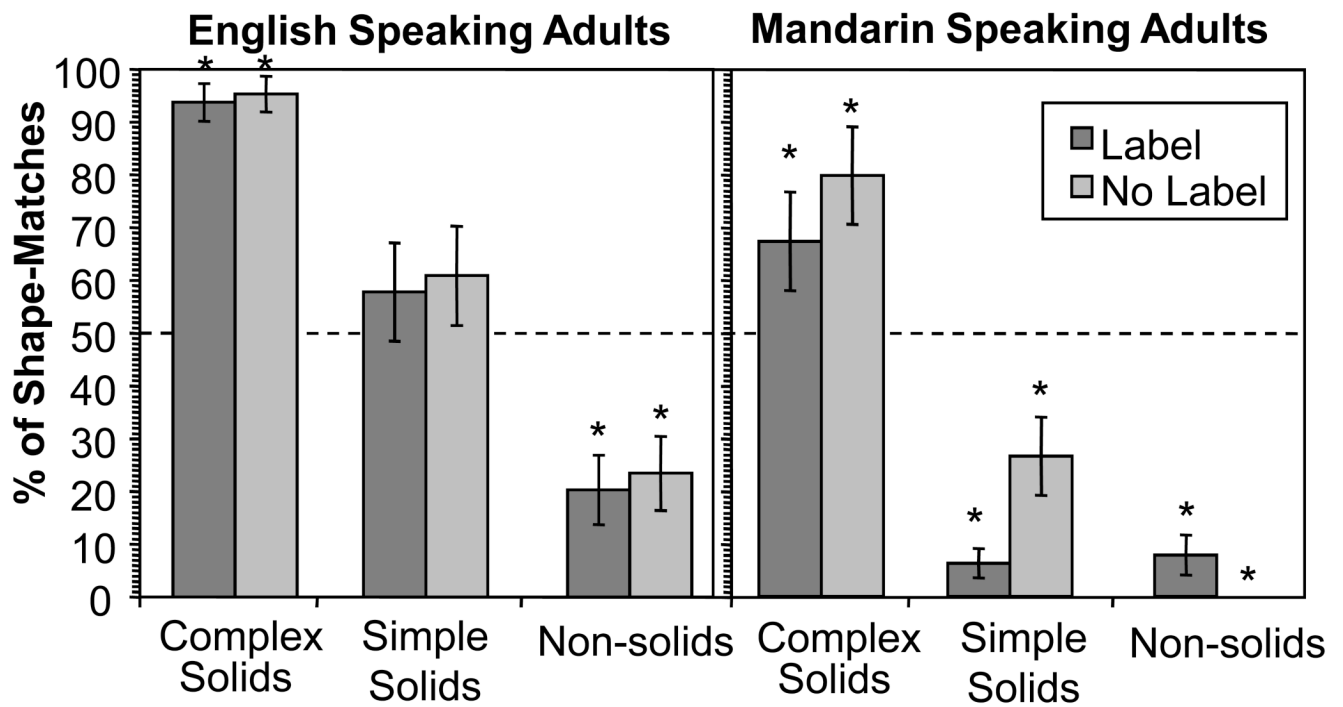
- Au TKF. Developing an intuitive understanding of substance kinds. *Cognitive Psychology* 1994;27(1): 71–111.
- Allan K. Nouns and countability. *Language* 1980;56(3):541–67.
- Barner D, Inagaki S, Li P. Language, thought, and real nouns: Individuation in Japanese, English and Mandarin Chinese. under review
- Barner D, Snedeker S. Children's early understanding of mass-count syntax: individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development* 2006;2:163–194.
- Bloom, P. How children learn the meanings of words. Cambridge, MA: MIT Press; 2000.
- Boroditsky, L.; Schmidt, L.; Phillips, W. Sex, Syntax, and Semantics. In: Gentner, D.; Goldin-Meadow, S., editors. *Language in Mind: Advances in the issues of language and thought*. MIT Press; 2003. p. 61-78.
- Bowerman, M.; Levinson, S., editors. *Language acquisition and conceptual development*. Cambridge: Cambridge University Press; 2001.
- Burger, J.; Prasada, S. Quantification of solid and nonsolid entities: The role of perceived arbitrariness of structure. *Proceedings of the Twenty-first Annual Boston University Conference on Language Development*; Somerville, MA: Cascadilla Press; 1997.
- Carey, S. Knowledge acquisition: enrichment or conceptual change?. In: Carey, S.; Gelman, R., editors. *The epigenesis of mind: essays in biology and cognition*. Hillsdale, NJ: Erlbaum; 1991. p. 257-291.
- Carey S. Does learning a language require the child to reconceptualize the world? *Lingua* 1994;92:143–167.
- Carey S. Bootstrapping and the origin of concepts. *Daedalus* 2004 Winter;:59–68.
- Carey, S. *The Origin of Concepts*. New York: Oxford University Press; in press
- Casasanto, D.; Boroditsky, L. Do we think about time in terms of space?. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*; Hillsdale, NJ: Lawrence Erlbaum Associates; 2003.
- Chen F. A study of language and thought: Object-substance interpretation among the Tsotsil Mayan. 2007 Unpublished undergraduate honors thesis at Harvard University
- Chierchia, G. Syntactic bootstrapping and the acquisition of noun meanings: The mass-count issue. In: Lust, B.; Suner, M.; Whitman, J., editors. *Syntactic theory and first language acquisition: Crosslinguistic perspectives*. Hillsdale, NJ: Erlbaum; 1994. p. 301-318.
- Chierchia G. Reference to kinds across languages. *Natural Language Semantics* 1998;6:339–405.

- Chomsky, N. Noam Chomsky on the generative enterprise: A discussion with Riny Huybregts and Henk van Riemsdijk. Dordrecht: Foris; 1982.
- Colunga E, Smith LB. From the Lexicon to Expectations About Kinds: A Role for Associative Learning. *Psychological Review* 2005;112(2)
- Comrie, B. Language universals and linguistic typology: syntax and morphology. Chicago: University of Chicago Press; 1981.
- Crane S. Effects of grammar on entity construals and word leaning among speakers of English and Czech. 2003 Unpublished undergraduate honors thesis at Swarthmore College
- Croft, WA. Typology and universals. Cambridge: Cambridge University Press; 1990.
- Dickinson DK. Leaning the names for materials: factors constraining and limiting hypotheses about word meaning. *Cognitive development* 1988;3:15–35.
- Feigenson L, Dehaene S, Spelke ES. Core systems of number. *Trends in Cognitive Sciences* 2004;8(7): 307–314. [PubMed: 15242690]
- Fenson, L.; Dale, PS.; Reznick, JS.; Bates, E.; Thal, D.; Pethick, S. Monographs of the Society for Research in Child Development. Vol. 59. 1994. Variability in early communicative development. Serial # 242
- Fisher, C.; Gleitman, L. Breaking the linguistic code: Current issues in early language learning. In: Pashler, HF.; Gallistel, R., editors. *Steven's Handbook of Experimental Psychology, Vol 1: Learning and motivation*. Vol. 3rd. New York: Wiley; 2002. p. 445-496. series ed., vol. ed.
- Gathercole VM, Min H. Word meaning biases or language-specific effects? Evidence from English, Spanish, and Korean. *First Language* 1997;17:31–56.
- Gathercole VCM, Whitfield LC. Function as a criterion for the extension of new words. *Journal of Child Language* 2001;28:87–125. [PubMed: 11258016]
- Gentner, D.; Boroditsky, L. Individuation, relativity, and early word learning. In: Bowerman, M.; Levinson, S., editors. *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press; 2001. p. 215-256.
- Gleitman, L.; Papafragou, A. Language and thought. In: Holyoak, K.; Morrison, R., editors. *Cambridge Handbook of Thinking and Reasoning*. Cambridge: Cambridge University Press; 2005. p. 633-661.
- Gordon P. Evaluating the semantic categories hypothesis: The case of count/mass distinction. *Cognition* 1985;20:209–242. [PubMed: 4064607]
- Haryu, E.; Imai, M.; Okada, H.; Li, L.; Meyer, M.; Hirsh-Pasek, K.; Golinkoff, RM. Noun Bias in Chinese Children: Novel Noun and Verb Learning in Chinese, Japanese, and English Preschoolers. *Proceedings of the Twenty-ninth Annual Boston University Conference on Language Development*; Somerville, MA: Cascadilla Press; 2005. p. 272-283.
- Hespos SJ, Spelke ES. Precursors to spatial language. *Nature* 2004;430:453–456. [PubMed: 15269769]
- Hohenstein J. Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognition and Development* 2005;6(3):403–425.
- Hurford, J. *Language and Number: the emergence of a cognitive system*. Basil Blackwell; Oxford, England: 1987.
- Huntley-Fenner G, Carey S, Solimando S. Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition* 2002;8:2–3. 221.
- Imai M, Gentner D. A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 1997;62:169–200. [PubMed: 9141906]
- Imai, M.; Mazuka, R. Re-evaluation of linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. In: Gentner, D.; Goldin-Meadow, S., editors. *Language in Mind: Advances in the issues of language and thought*. MIT Press; 2003. p. 430-464.
- Imai M, Mazuka R. Revisiting language universals and linguistic relativity: language-relative construal of individuation constrained by universal ontology. *Cognitive Science* 2007;31:385–414.
- Kobayashi H. The influence of adults' actions on children's inferences about word meanings. *Japanese Psychological Research* 1999;41:35–49.

- Muromatsu, K. Classifiers and the count/mass distinction. In: Li, YHA.; Simpson, A., editors. *Functional Structure(s), form and interpretation: perspectives from East Asian languages*. London: Routledge; Curzon: 2003. p. 65-96.
- Lee H. The acquisition of word meaning in Korean: Perceptual properties, ontological categories, and syntactic cues. *Korean Journal of Developmental Psychology* 1997;11(2):107-120.
- Li P, Barner D, Huang B. Classifiers as count syntax: Individuation and measurement in the acquisition of Mandarin Chinese. *Language, Learning, and Development* 2008;4(4):1-42.
- Li P, Gleitman L. Turning the tables: Spatial language and spatial cognition. *Cognition* 2002;83:265-294. [PubMed: 11934404]
- Lucy, JA.; Gaskins, S. Interaction of language type and referent type in the development of nonverbal classification preferences. In: Gentner, D.; Goldin-Meadow, S., editors. *Language in Mind: Advances in the issues of language and thought*. Cambridge, MA: MIT; 2003. p. 465-492.
- Lucy, J.; Gaskins, S. Grammatical categories and the development of classification preferences: A comparative approach. In: Bowerman, M.; Levinson, S., editors. *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press; 2001. p. 257-283.
- Lucy, J. Grammatical categories and cognition: A case study of the linguistic relativity hypothesis. Cambridge: Cambridge University Press; 1992.
- Mazuka R, Friedman R. Linguistic relativity in Japanese and English: Is language the primary determinant in object classification? *Journal of East Asian Linguistics* 2000;9:353-377.
- Matsumoto, Y. Order of acquisition in the lexicon: Implications from Japanese classifiers. In: Nelson, KE.; van Kleeck, A., editors. *Children's language*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1987. p. 229-260.
- Naka M. The acquisition of Japanese numerical classifiers by 2-4-year-old children: The role of caretakers' linguistic inputs. *Japanese Psychological Research* 1990;41(1):70-78.
- Naigles L, Terrazas P. Motion-verb generalizations in English and Spanish: Influences of language and syntax. *Psychological Science* 1998;9:363-369.
- Papafragou, A. Relations between language and thought: Individuation and the count/mass distinction. In: Lefebvre, C.; Cohen, H., editors. *Handbook of Categorization in Cognitive Science*. Elsevier Science; 2005.
- Pinker, S. *Language learnability and language development*. Cambridge, MA: Harvard University Press; 1984.
- Prasada, S. Names for things and stuff: An Aristotelian perspective. In: Jackendoff, R.; Bloom, P.; Wynn, K., editors. *Language, logic, and concepts*. Cambridge, MA: MIT Press; 1999. p. 119-146. p. 119-146.
- Prasada S. Learning names for solid substances: Quantifying solid entities in terms of portions. *Cognitive Development* 1993;8(1):83-104.
- Prasada S, Ferenz K, Haskell T. Conceiving of entities as objects and stuff. *Cognition* 2002;83:141-165. [PubMed: 11869722]
- Quine, WVO. *Word and object*. Cambridge, MA: MIT Press; 1960.
- Quine, WVO. *Ontological relativity and other essays*. New York: Columbia University Press; 1969.
- Samuelson L, Smith L. Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition* 1999;73:1-33. [PubMed: 10536222]
- Sandhofer C, Smith L, Luo J. Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language* 2000;27:561-585. [PubMed: 11089339]
- Sapir, E. *Selected writings in Culture, Language, and Personality*. Berkeley: University of California Press; 1949.
- Slobin, DI. From thought and language "to" thinking to speaking. In: Gumperz, JJ.; Levinson, SC., editors. *Rethinking linguistic relativity*. Cambridge: Cambridge University Press; 1996. p. 70-96.
- Smith C, Carey S, Wisner M. On differentiation: a case study of the development of the concepts of size, weight and density. *Cognition* 1985;21:177-237. [PubMed: 3830547]
- Smith LB, Jones SS, Landau B, Gershkoff-Stowe L, Samuelson LK. Object name learning provides on-the-job training for attention. *Psychological Science* 2002;13(1):13-19. [PubMed: 11892773]

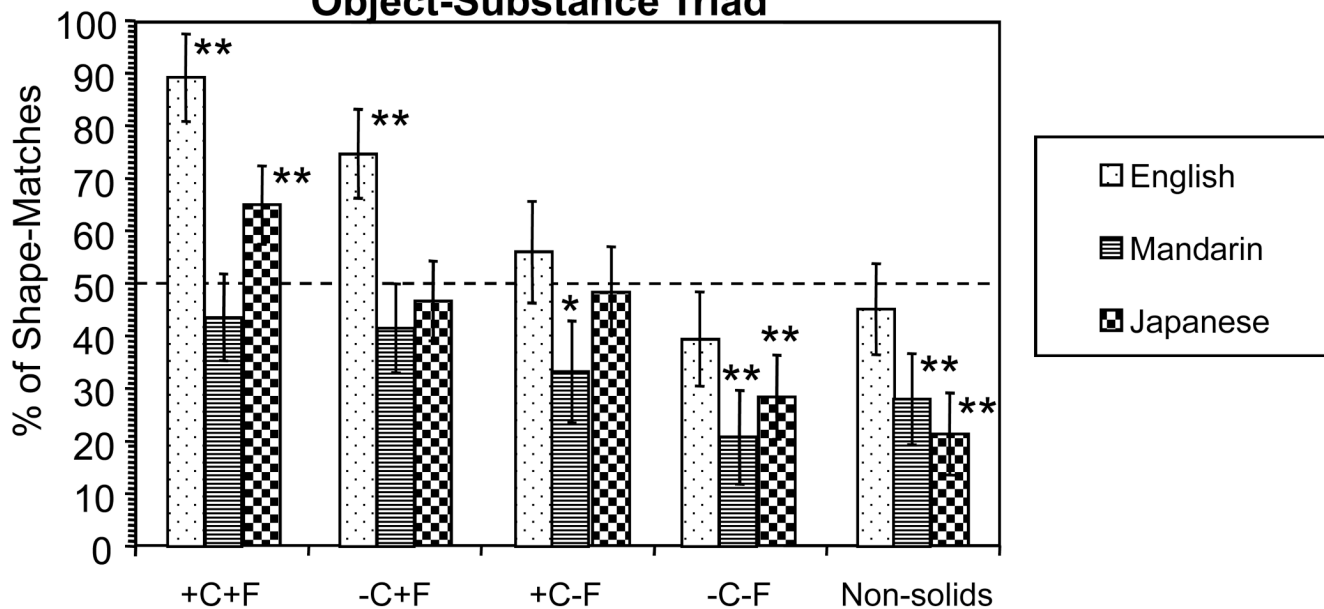
- Snedeker, J.; Li, P. Can the situations in which words occur account for cross-linguistic variation in vocabulary composition?. In: Tai, J.; Chang, Y., editors. Proceedings 7th International Symposium on Chinese Languages and Linguistics. 2000.
- Snedeker, J.; Thorpe, K.; Trueswell, J. On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. Proceedings of the 23rd Annual Conference of the Cognitive Science Society; 2001. p. 964-969.2001
- Soja NN. Inferences about the meanings of nouns: the relationship between perception and syntax. *Cognitive Development* 1992;7:29–45.
- Soja N, Carey S, Spelke E. Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition* 1991;38:179–211. [PubMed: 2049905]
- Spelke, ES. What makes humans smart?. In: Gentner, D.; Goldin-Meadow, S., editors. *Language in Mind*. Cambridge, MA: MIT Press; 2003.
- Spelke ES, Kestenbaum R, Simons D, Wein D. Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology* 1995;13:113–142.
- Subrahmanyam K, Chen HN. A cross-linguistic study of children's noun learning: the case of object and substance words. *First Language* 2006;26(2):141–160.
- Subrahmanyam K, Landau B, Gelman R. Shape, material, and syntax: Interacting forces in children's learning in novel words for objects and substances. *Language and Cognitive Processes* 1999;14:249–281.
- Trueswell JC, Kim AE. How to prune a garden-path by nipping it in the bud: Fast-priming of verb argument structures. *Journal of Memory and Language* 1998;39:102–123.
- Uchida N, Imai M. Heuristics in learning classifiers: The acquisition of the classifier system and its implications for the nature of lexical acquisition. *Japanese Psychological Research* 1999;41(1):50–69.
- Whorf. The relation of habitual thought and behavior to language. In: Carroll, JB., editor. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, MA: The MIT Press; 1939. p. 134-159.1956
- Wynn K. Addition and subtraction by human infants. *Nature* 1992;358:749–750. [PubMed: 1508269]
- Xu F, Carey S. Infants' metaphysics: the case of numerical identity. *Cognitive Psychology* 1996;30:111–153. [PubMed: 8635312]
- Yamamoto K, Keil F. The acquisition of Japanese numerical classifiers – Linkage between grammatical forms and conceptual categories. *Journal of East Asian Linguistics* 2000;9:379–409.
- Yoshida H, Smith L. Shifting ontological boundaries: how Japanese- and English-speaking children generalize names for animals and artifacts. *Developmental Sciences* 2003a;6(1):1–34.
- Yoshida H, Smith L. Known and novel noun extensions: Attention at two levels of abstraction. *Child Development* 2003b;74(2):564–577. [PubMed: 12705573]

## Experiment 1



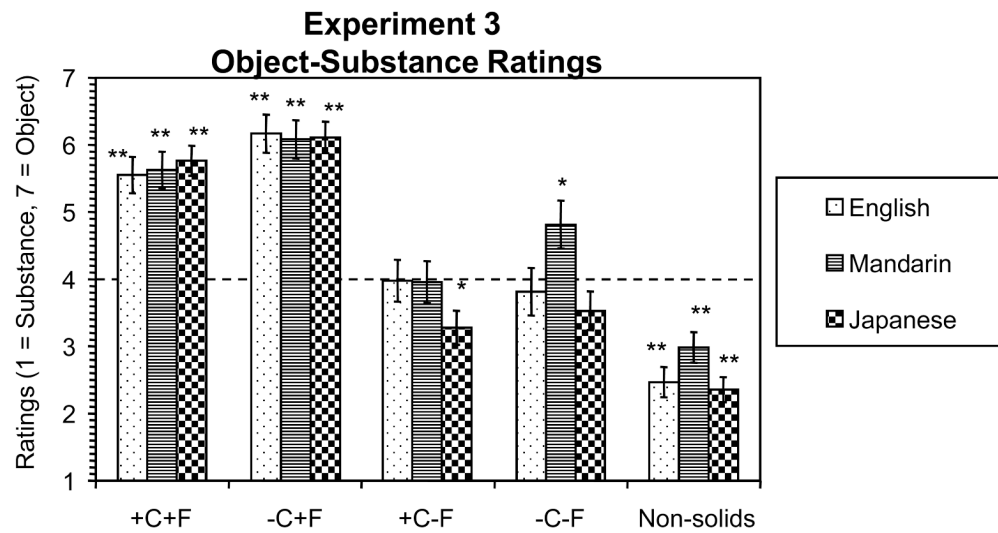
**Figure 1.** Effect of Entity Type and Instruction on English and Mandarin speaking adults (Experiment 1). Percentages significantly above or below chance (i.e., 50% shape-match) are marked: \* for  $p < .05$

## Experiment 2 Object-Substance Triad



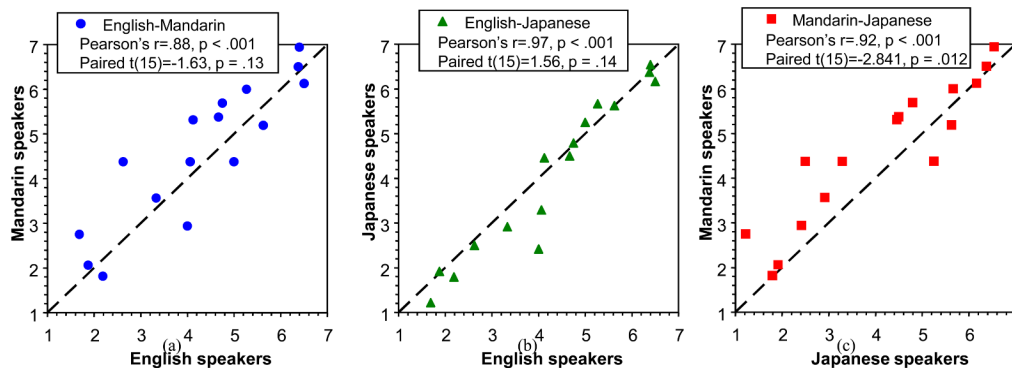
**Figure 2.** Effect of Entity Type on English, Mandarin, and Japanese speaking adults (Experiment 2). Percentages significantly above or below chance (i.e., 50% shape-match) are marked: \* for  $p < .05$ , and \*\* for  $p < .01$ .



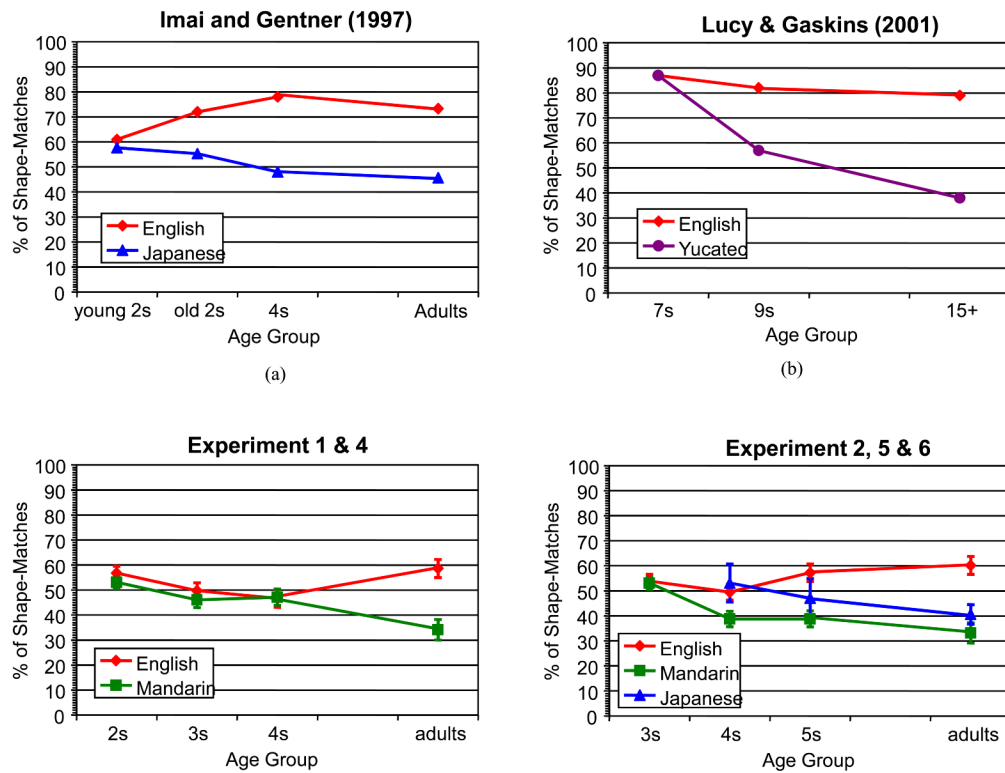


**Figure 3.** Effect of Entity Type on English, Mandarin, and Japanese speaking adults (Experiment 3). Ratings significantly above or below 4 (i.e., the mid-point of the scale) are marked: \* for  $p < .05$ , and \*\* for  $p < .01$ .

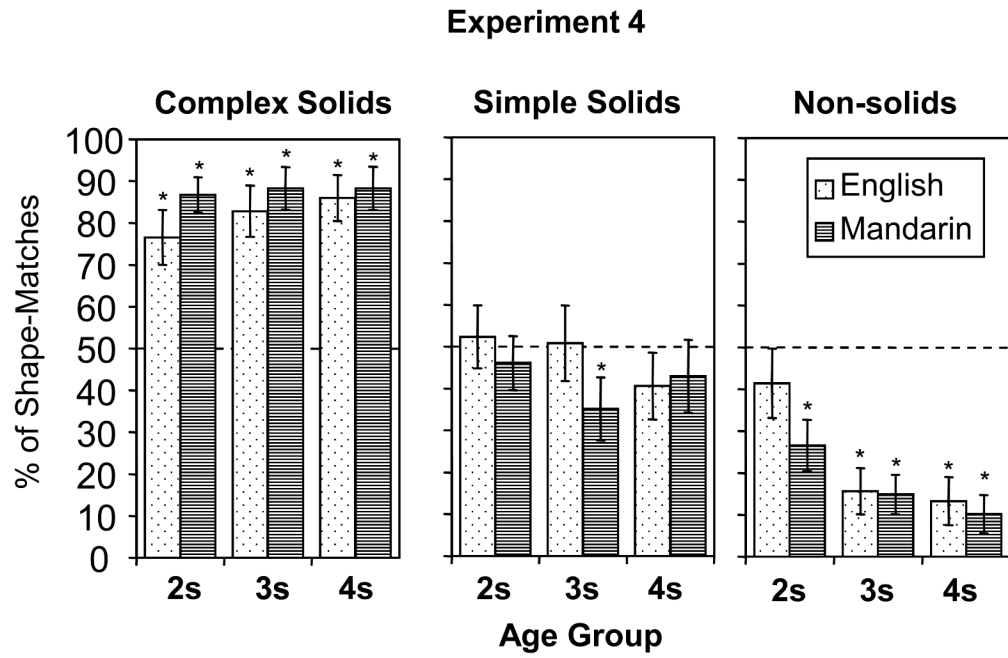
Experiment 3



**Figure 4.** By-item scatter plots for English, Mandarin, and Japanese speaking adults (Experiment 3). Pearson's  $r$ 's indicate the degree to which the two language groups' ratings correlate across items. Paired- $t$  tests indicate the degree to which English speakers gave higher object ratings than Mandarin speakers across items.

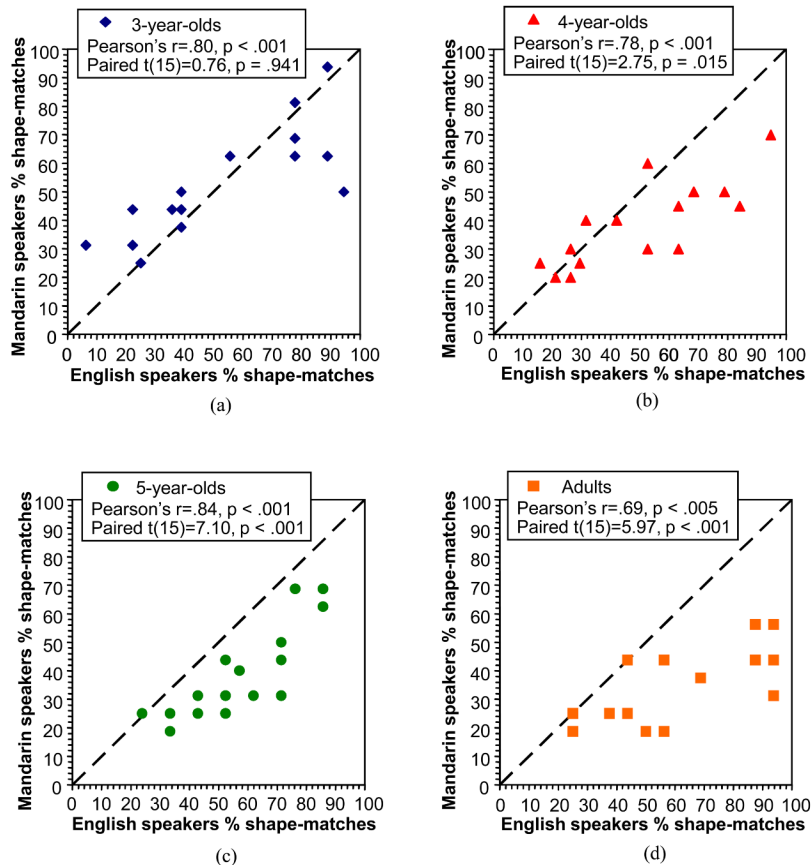


**Figure 5.** Developmental change across studies: (a) Japanese and English speakers from Imai and Gentner (1997); (b) Yucatec Mayan and English speakers from Lucy and Gaskins (2001); (c) Experiments 1 and 4 of current study; (d) Experiments 2, 5, & 6 of current study.

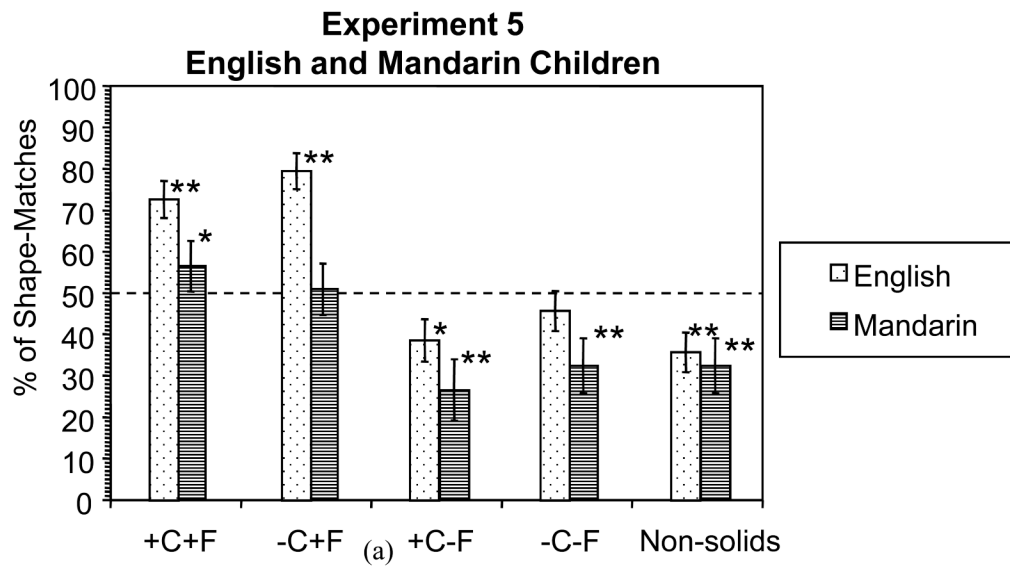


**Figure 6.** Effect of Entity type on English and Mandarin speaking children using Experiment 1 stimulus set (Experiment 4). Percentages significantly above or below chance (i.e., 50% shape-match) are marked \* for  $p < .05$ .

Experiment 5

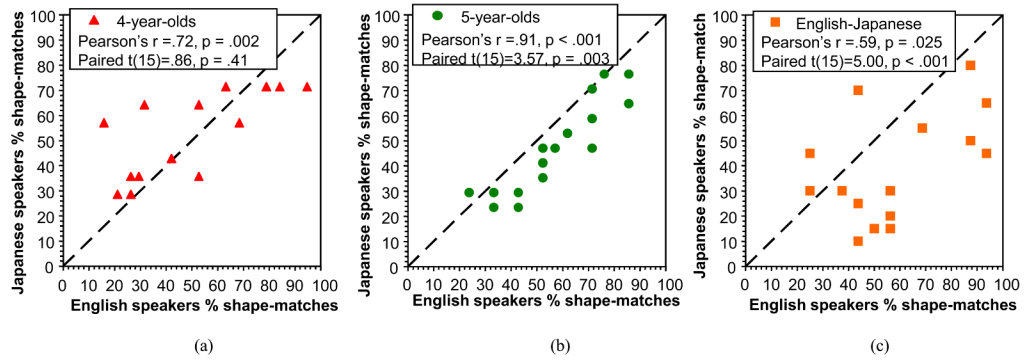


**Figure 7.**  
**a-d.** English and Mandarin By-item scatter plots for Experiment 2 & 5. Pearson's  $r$ 's indicate the degree to which the two language groups' shape-match percentages correlate across items. Paired- $t$  tests indicate the degree to which English speakers prefer more shape-matches than Mandarin speakers across items.

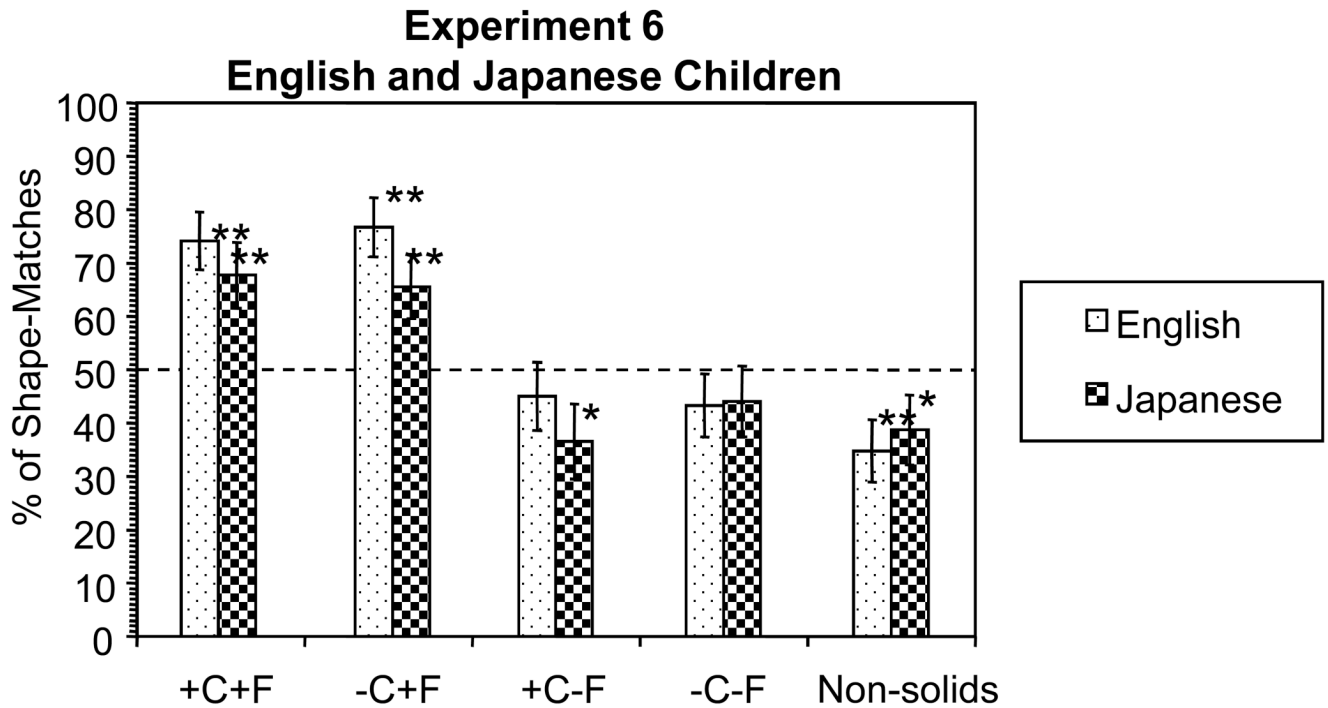


**Figure 8.** Effect of Entity type on English and Mandarin speakers using Experiment 2 stimulus set (Experiment 4).

Experiment 6



**Figure 9.**  
a-c. English and Mandarin By-item scatter plots (Experiment 6).



**Figure 10.**

Effect of Entity Type on English and Japanese speaking children. Percentages significantly above or below chance (i.e., 50% shape-match) are marked: \* for  $p < .05$ , and \*\* for  $p < .01$ .





**Table 1**

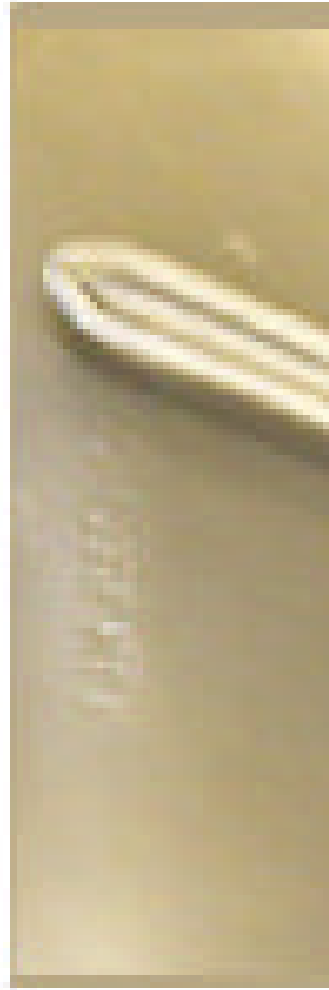
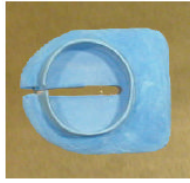
Stimuli and ratings for Experiments 1 and 4.

<i>Complex Solids</i>		
Complexity Rating M=4.31, SD=1.33, Meaningfulness M=5.53, SD=1.5		
<b>Standard</b>	<b>Shape alternative</b>	<b>Material alternative</b>
Clear plastic clip	Metal clip	Clear plastic piece
Copper T	Ivory plastic T	Copper piece
Wooden reamer	Rubber reamer	Wooden pieces
Metal whisk	Nylon whisk	Metal piece
<i>Simple Solids</i>		
Complexity Rating M=2.57, SD=.74, Meaningfulness M=3.45, SD=.83		
<b>Standard</b>	<b>Shape alternative</b>	<b>Material alternative</b>
Cork stopper	Plaster stopper	Cork piece
Terracotta wedge	Marble wedge	Terracotta pieces
Red Sculpey ½ egg	Styrofoam ½ egg	Sculpey pieces
Pink wax kidney	Yellow plaster kidney	Pink wax piece
<i>Non-solids</i>		
Complexity Rating M=2.78, SD=1.03, Meaningfulness M=2.23, SD=.92		
<b>Standard</b>	<b>Shape alternative</b>	<b>Material alternative</b>
Nivea Spiral	Sparkle gel Spiral	Nivea piles
Orange sand Trapezoid	Blue Bakit Crystal Trapezoid	Orange sand piles
Shaving foam Gamma	Brown sugar Gamma	Shaving foam pile
Sawdust Omega	Flour Omega	Sawdust pile

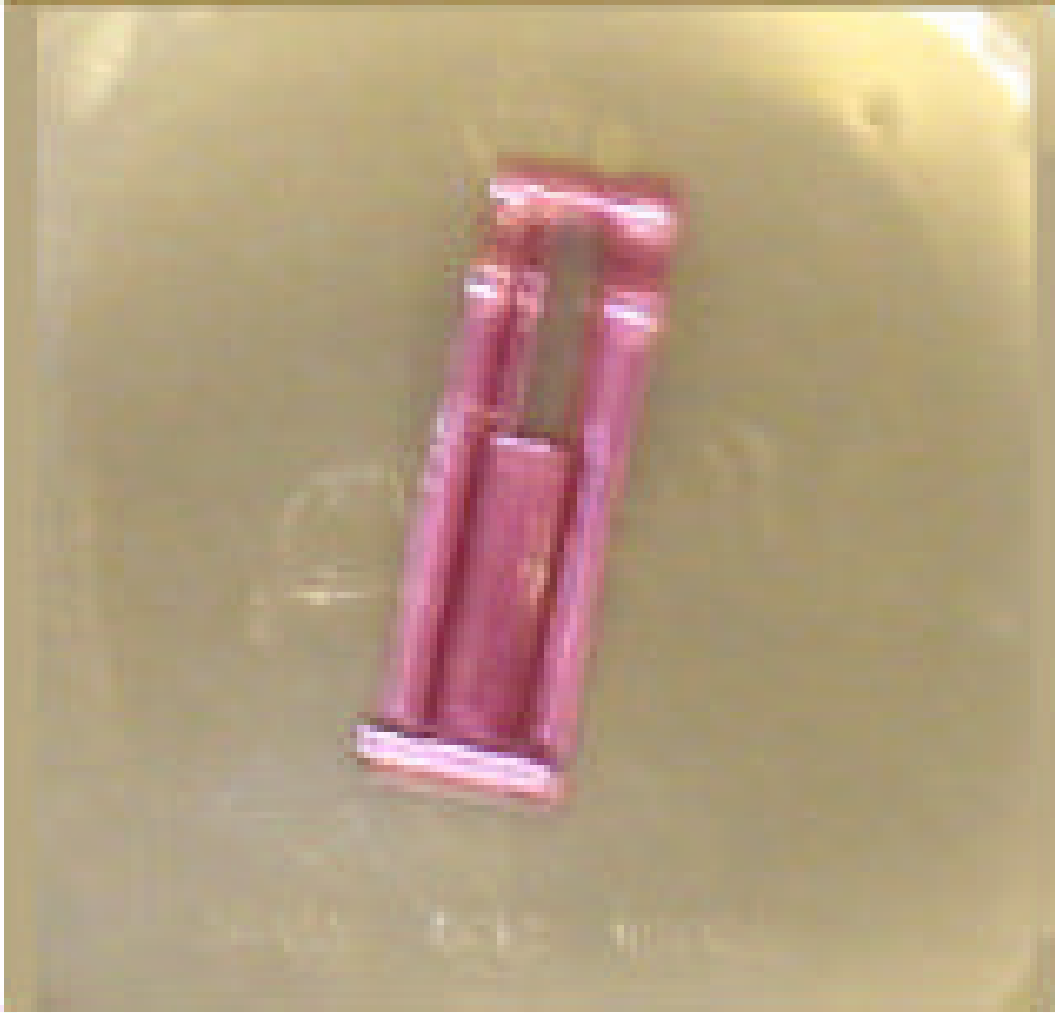
**Table 2**  
Stimuli and ratings for Experiments 2, 3, 5, and 6.

<b>+C+F Complex and Functional</b>	
	M=4.08 SD=.92
	M=6.42 SD=.69
	

**+C+F Complex and Functional**



**+C+F Complex and Functional**



NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript