Evaluating the evidence of shape bias in children cross-culturally

Samah Abdelrahim[1] & Michael C Frank[1]

[1] Stanford University

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. Samah Abdelrahim: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Michael C Frank: Writing - Review & Editing, Supervision.

Correspondence concerning this article should be addressed to Samah Abdelrahim, Postal address. E-mail: samahabd@stanford.edu

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words "**here we show**" or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: keywords

Word count: X

Evaluating the evidence of shape bias in children cross-culturally

Introduction

How children acquire meanings of words rapidly from interactions with the people around them in the absence of negative feedback and sparse data represents a very hard induction problem which they seem to navigate through without going through the many possible hypotheses. Children are thought to use some biases and constraints to eliminate unlikely hypotheses. For example, the whole object constraint suggests that a child assumes the word labels the entire object, not parts or characteristics of the object, while according to the taxonomic assumption a child extends a word to new objects based on taxonomic relations such that they would extend a soup with pizza rather than a spoon. Additionally, a child tends to assign only one label to an object as per the mutual exclusivity constraint by. While these constraints are considered conceptual biases the child could be using to facilitate their learning, a huge body of literature showed that children show a perceptual bias towards shape when applying lexical categories. The shape bias refers to the tendency of children to attend to the shape of objects more than any other attributes like color, texture, or material when mapping a newly learned word to a referent, in other words, when forming categories that govern generalization of words to new instances (Imai et al., 1994); Landau et al., 1988). In (Graham & Diesendruck, 2010) they showed 15 months old children an exemplar that have a specific non perceptual property. Then children were asked which of the test objects that matched the exemplar in either color, or shape, or material. Children extended the non-perceptual property to the test objects that resembled the exemplar in shape. WHY IS THE SHAPE BIAS IMPORTANT? MAYBE USE SMITH et al. 2002 TRAINING STUDY AS AN EXAMPLE? The role shape bias could be playing in acquiring meanings of nouns was demonstrated by (Smith et al., 2002) in which they found evidence that attention to the specific relevant properties for naming could be tuned by training using the least amount of data. In that study, they trained 17-month-old children throughout 7 weeks of repeatedly playing with and hearing names of unfamiliar objects that are well organized by shape. Children who received the training formed a generalization that only objects with similar shape have the same name, and showed more rapid acquisition of object names masured by parental reports during and after the intervention. The degree of shape bias observed in young children's word learning in lab experiments varies across cultures and languages. English speaking children were found to show a salient bias towards categorizing labels by shape specially around the age of three/two. On the other hand, Eastern Asian language speakers like Japanese and Mandarin were found to be less prone to attend to shape for word extension (more details here?). For example, Imai and Gentner 1997 tested Japanese and English speaking children on the word generalization task, and found that English speaking children are generally more likely to exhibit shape bias than Japanese children. (Gathercole & Min, 1997; Imai & Gentner, 1997; Jara-Ettinger et al., 2022; Samuelson & Smith, 1999; Soja, 1992; Subrahmanyam & Chen, 2006; Yoshida & Smith, 2003). The shape bias observed with English speaking children in the US was shown to be elicited only in linguistic tasks and not in non-linguistic similarity judgment tasks, and is thus, unlikely to be a general perceptual bias (Landau et al., 1988; Smith et al., 2002) . The cross-linguistic differences between English speakers and other language's speakers were attributed to a variety of factors. One of the explanations for the East Asian languages' speaker's weak bias towards shape is the lack of marking of the ontological caegorization of count and mass nouns in their grammar. Those cross-linguistic differences were attributed to the East Asian languages' lack of marking for the ontological categorization of count and mass in its grammar. In English, mass nouns are characterized by the impossibility of being directly modified by a numeral or being combined with an indefinite article (a or an). Extending a word to a referent needs prior knowledge about the physical nature of the referent because the principles that govern word meaning extensions for solid objects seem to be fundamentally different than the ones governing substances. So children must approach this problem having prior knowledge about the distinction. (Quine & Van, 1960) views young children as lacking ontological knowledge about anything in the world and bears no commitment to any ontological definition before the syntax. However, this view was challenged by some studies that claim that the ontology underlying natural language is induced in the course of language learning, rather than constraining learning from the beginning (Soja et al 1991, 1992). Additionally, evidence by (Barner et al., 2009) showed that those cross-lingusitc disparity reported in studies like (Gathercole & Min, 1997; Imai & Gentner, 1997) could be inferable from the lexical statistics of the language and that both speakers of languages that either have count-mass grammar or lack it construe solid and non-solid referents similarly. Which leads to the second explanation that was argued for by examining the early vocabulary of English speaking children in the US which revealed a distribution that is dominated by count nouns and nouns that refer to solid objects and hence, shape bias might be a product of the statistical regularities of early language rather than the syntactic structure (Gershkoff-Stowe & Smith, 2004; Samuelson, 2002; Samuelson & Smith, 1999). A study that was carried out with the Tsimane' group in Bolivia displayed a similar pattern to Mandarin and Japanese speakers Jara ettinger et al 2022. Authors argued for a partial explanation of the phenomenon by the statistical regularities of the environment that is dominated by shape-based solid artifacts in the case of English-speaking children in the US which was suggested to partially drive the shape bias. Nevertheless, the absence of information on the structure of the early vocabulary of the Tsimane' language, as well as the East Asian languages, leaves us with a gap in teasing apart the effect of environmental statistics from the lexcial one and in understanding how entagled they are. The same cross-linguistic effect was observed in studies that incorporated the solidity of objects as a continuum of complex solid objects, simple solid objects, and substances, but this time the effect was only observed in simple solid objects for which English speakers manifested a shape bias, while Speakers of Mandarin and Japanese did not. This picture is further complicated by developmental changes in the magnitude of the shape bias. To illustrate, Landau et al 1988 tested 2, 3 yrs and adults and found that adults are at ceiling while there is no effect of age between the 2 and 3 yr olds. In a later study landau et al 1989 showed that 3 yr olds showed a strong bias towards shape regardless of the presence of functionality, while 5 yrs old and adults showed a weaker but existing preference for shape that disppeared when functionality was highlited in the test question instead of only the label. Adding the cross-lingustic dimension on top of the developmental question adds another layer of complexity. In a study with English and Mandarin 3 and 4 yrs old children along with adults (Subrahmanyam & Chen, 2006) found that 3 yrs old from both language groups prefered shape when generalizing the label of an object or a substance while Mandarin 4 yrs old and adults prefered material for both exemplars while English 4 yrs old and adults showed a preference for shape but only in the case of an object while no preference was displayed in the case of substance. In addition, when testing for word extension, most of the studies use a novel object for which the child has no prior conceptual knowledge so they don't relate it to any previously existing category in order to elicit their spontaneous labeling mechanism (Samuelson & Smith, 2005). However, the real life instances in which a child learns a new label to object mapping is probably full of a repertoire of perceptual, conceptual and social information that is linked to the label-object pair. Many studies have also attempted to capture the effect of providing more contextual information related to the object on children's lexical categorisation, by varying many dimensions of (functionality), (animacy), (lingusitc cues), (priming) and so on… Therefore, comparing the magnitude of the shape bias across studies, ages, environments, and languages is difficult because these studies use different types of stimuli, tasks, measures, and analyses. In light of all that, a meta-analysis of the literature on shape bias becomes necessary and beneficial on many levels. Given the seemingly contradictory findings, the range of possible explanations, as well the huge disparity between studies on many dimensions such like the presence and length of training, the contrasted dimensions in test trials (shape against size, color, material, function ..etc), the stimuli (complext against simple objects, solid against non-soild), and the measure used (forced choice, endorsment task), this study uses meta-analysis to estimate an overall effect size. Additionally, this meta-analysis (MA) study leverages date points from studies conducted with different age groups to describe the developmental trajectory of shape bias using the state of the evidence in the extant empirical literature. Lastly, this MA study attempts to test for moderation by coding the between-study differences. In particular, we were interested in testing four hypotheses about moderating factors of the effect size, that are derived from the prior literature. Firstly, Shape bias will increase with age, such that a greater shape bias will be observed in older children. There will be a consistent association between population (as coded by language and location) and effect size such that shape bias effect will be gradient, with the highest effect in English speaking populations, and the lowest in populations with the least industrialized environments. Additionally, Eastern countries with comparable life style to English speaking countries wil lie in the middle and the strongest effects will be observed with simple uniform solid objects. Lastly, There will be no effect of how informative the syntactic structure with respect to count and mass nouns in older children, but an effect will be observed for younger children.

Methods

All hypotheses, literature search criteria, and statistical analyses were preregistered on the OSF (link) except where noted.

Literature search

Effect sizes of shape bias were extracted from papers sourced from google scholar based on the Key words “shape bias”, “word generalization”, “word learning” and citations of two papers (Landau et al., 1988) and (Imai & Gentner, 1997). PRISMA CHART HERE One hundred sixty one papers were found. We filtered these based on eligibility criteria including being an experiment (e.g., with random assignment to at least two conditions), children participants from 2-5 years old, and using a word extension task that contrasts shape with other properties of the referent. Applying these criteria resulted in 32 papers that both satisfied our criteria and reported enough information to calculate the effect size (es), either by directly reporting the es, reporting the proportion of choosing shape in text, tables or graphs, or reporting a statistical test against chance like a one sample t-test.

Coding of effect size and moderators

After coding the paper's metadata that includes the citation and a unique identifier, for each effect size, many moderating variables are coded. Those moderators include the age, participants number, type of syntax used (informative or neutral, count or mass, common or proper, animate or inanimate for east asian languages), the matching and alternative properties (shape, color, material, ..etc) , the type of stimuli in terms of solidity, animacy and dimensionality, type of exposure before testing (presence or absence of training), vocabulary size, country and spoken language, response mode (picking or pointing or verbalizing), measure (behavioural, eye tracking, gaze-video coding, neurophysiological), population type (typically developing or not), and number of trials. The measures used in most of the literature are typically either a forced choice task in which the child is forced to make a choice from the test objects or an endorsement task in which the child is asked to accept or reject a test object as a member of the exemplar category by saying yes or no. one study used an open choice task in which the child have the chance to retract from making a choice by saying i don't know (cimpian et al 2005). Because studies can include many age groups, different types of stimuli, and various manipulations, one experiment usually yields more than one effect size. Estimation of those effect sizes was carried out following this order: If the paper reports cohen's d, we used it as the effect size in our MA (N=?). if the cohen's d was not directly revealed, we looked for one sample t-test values to calculate the effect size (see supplementary materials for equations used). The third step in the absence of one sample t-tests was to use proportions reported either in text, tables, or graphs along with the standard deviation SD or the standard error SE (in the case of the absence of SD or SE, equations to calculate them were used, supplementary materials) The final sample was 43 papers and 286 effect sizes.

Analytic approach

All analyses were performed using the R metafor package. We pooled effect sizes using a multivariate meta analysis model with experiment number within a paper as a random effect nested grouping variable. We performed confirmatory analysis of the hypotheses via a multi-level meta-regressions that include random effects that control for non-independence between effect sizes based on grouping by paper and grouping by experiment. The moderators that were included as fixed effects were "age", "linguistic group", and "solidity of the stimuli". Because data points contributed from individual languages strikingly varied such that we have 160 effect size for English language while Japanese was the closest language with only 17 data points, we grouped language into the linguistic groups of Indo and Non-Indo European.

Results

The meta-analytic model revealed a pooled positive overall effect size of 0.44 [0.34,0.55] ($P<0.01$) with a between study heterogeneity of 0.73 standard deviation (SD). To test for the theoretically relevant possibility of moderation by stimuli, we conducting a separate meta-analysis on the data points that only involve using a solid object as a stimuli. This yielded an overall effect size for solid objects of 0.52 with a heterogeneity of 0.69 SD. Since the heterogeneity remains very high, we subsequently incorporated age in the multi-level model to determine whether part of the heterogeneity could be explained by the developmental change.

To determine the best functional curve that fits the development of shape bias, we compared four types of functional forms in our multi-level model using the corrected

Akaike information criterion AICc for model selection. The functional forms to be compared were: constant, linear, logarithmic, and quadratic. A difference of more than 4 between the minimal AICc and any other AICc was used as a meaningful difference. Based on this convention, the difference between linear fit and quadratic fit wasn't meaningful so we decided to fit a linear developmental model for age. ## Moderating factors Language was hypothesized to be moderating the tendency to generalize by shape. Nonetheless, testing for this hypothesis using individual languages was not possible because of the scarcity of data from languages other than English (data on this ?). So languages were grouped into Indo-European (included English, German, and Spanish) and Non-IndoEurpean (included). Both age and language group displayed significant effects when the language group was added as a moderator in the multi-level model. It showed a significant effect size of 0.44 for the average age of the IndoEuropean group, but this effect size signitifcantly decreases with age and it decreases more for speakers of the Non IndoEuropean language group. No significant interaction between age and language was found. Because both language and location are interchangeable in our data, for example English will mostly correspond to the US and Japanese to Japan, it is yet to be uncovered whether the effect is due to the syntax of the language or other factors that correspond to the other hypothesese about the environment or the lexical statistics. Directly testing for the statistical regularities hypotheses was not feasible due to the sparcity of data on the vocabulary size and structure from both language groups. Likewise, we lack data on the distribution of shape based objects in the environment, and the only evidence we have is the comparison study between English speakers in the US and the Tsimane'population in Bolivia by Jara-ettinger et al 2022, which leaves the same questions unanswered. Figure 3: The developmental curve in Figure [3] shows that shape bias decreases with age for both language groups which poses the question of what type of a hypothesis shift children might be undertaking around the age of 5 making them less reliant on the perceptual cues like shape. An abundance of evidence in the literature suggests that ## Publication Bias

Discussion

References