The Development of Structural Thinking about Social Categories

Nadya Vasilyeva (vasilyeva@berkeley.edu),
Alison Gopnik (gopnik@berkeley.edu), Tania Lombrozo (lombrozo@berkeley.edu)
Department of Psychology, University of California, Berkeley, 3210 Tolman Hall, Berkeley, CA 94720 USA

Abstract

Representations of social categories help us make sense of the social world, supporting predictions and explanations about groups and individuals. Here we explore whether children and adults are able to understanding category-property associations in *structural* terms, locating an object of explanation within a larger structure and identifying *structural constraints* that act on elements of the structure. We show that children as young 3-4 years of age show signs of structural thinking, but that this capacity does not fully develop until after 7 years of age. These findings introduce a viable alternative to internalist accounts of social categories, such as psychological essentialism.

Keywords: structural explanation, structural factors, social categories, essentialism, category representation

Imagine that a school introduces a dress code stating that all items of a student's clothing must match in color. When school begins, most boys show up wearing blue, and most girls show up wearing pink. What explains the correlation between gender and color? One explanation is that boys naturally prefer blue, and girls pink. But a quick glance at history shows that in the 19th century, pink was considered the vigorous, masculine color, whereas girls wore "delicate and dainty" blue (Fausto-Sterling, 2012). If an explanation that appeals to intrinsic preferences is inadequate, an alternative might be to appeal to a structural feature of the environment: department stores reliably stock more pink options for girls than for boys. In this case, availability could be a sufficient explanation for the observed correlation.

This example illustrates what we call "structural thinking." A hallmark of structural thinking is locating an object of explanation within a larger structure and identifying *structural constraints* that act on components of the structure to shape the distribution of outcomes for each component. In our example, girls occupy a position within larger social and institutional structures that make them more likely than boys to choose pink over blue. A structural approach to social categories differs from *internalist* approaches, which focus on essential or inherent properties of the category itself. In the current paper, we ask whether and when children develop the ability to think about social categories in structural terms.

Internalist approaches to category representation. One prominent approach to theorizing about the representation of social categories (such as "girl") is based on the notion of *psychological essentialism*, which refers to the tendency to represent (some) categories in terms of an underlying essence that is constitutive of category membership and/or causally responsible for key category features (Gelman, 2003). Psychological essentialism can support efficient generalizations about natural kinds, but can also lead to unwarranted normative expectations about categories, stereotypical generalizations, and prejudice (Leslie, 2015).

A related internalist approach comes from Cimpian and Salomon (2014), who argue for the *inherence heuristic*, defined as the tendency to explain observed patterns in terms of the inherent properties of the objects that instantiate them. If girls wear pink, people might infer that it must be due to something inherent about pink ("it is delicate") and/or girls ("they are attracted to delicate colors"), rather than considering a broader range of external, historical factors. Cimpian and Salomon argue that the inherence heuristic is distinct from, but potentially a precursor to, essentialized representations of social categories.

A final approach, the *aspect hypothesis*, comes from Prasada and Dilllingham (2006, 2009), who offer a non-essentialist account of categorical representation. On this view, some features of a category are viewed as *aspects* of the kind. For example, "fighting crime" is an aspect of being a police officer (in contrast to merely statistical associations, such as between police officers and "eating donuts").

While psychological essentialism, the inherence heuristic, and the aspect hypothesis are importantly distinct in their commitments regarding categorical representations, they all support internalist explanations for associations between a category and a feature (e.g., "she chose pink because girls like warm colors"), as well as formal explanations that appeal to category membership (e.g., "she chose pink because she is a girl"). By contrast, they lack mechanisms for differentiating *kinds* (i.e., "girls") from the *structures* in which they are embedded (i.e., the social position occupied by girls). As a result, they cannot readily accommodate the kind of structural thinking supported by a structural approach.

A structural approach to category representation. Our study explores an alternative to internalist accounts. According to a structural view of categorical representation, reliable connections between properties and categories can be represented as a consequence of stable structural constraints acting on categories from the outside.

This approach is based on the notion of structural explanation developed in philosophy of social sciences, where it is defined by situating the object of explanation in a network of relationships within a larger, organized whole (a structure), and identifying how relationships to other parts of the whole modify the probability distribution over possible states of the part whose behavior is explained (relative to a hypothetical case outside a structure, relative to other nodes within the structure, or relative to different structures; Haslanger, 2015). For example, an internalist explanation for why many women in heterosexual relationships leave their jobs after having a child might appeal to women's priorities or abilities, whereas a structural explanation would identify constraints that affect women in virtue of their position within the social structure (e.g., lack of paid parental leave, a

gender wage gap, etc.). These structural constraints shift the probability distribution across different outcomes for women versus men. For another position subjected to different structural constraints (e.g., "men," "women in a different culture"), the same event (having a child) need not trigger the same outcomes. Rather than pinpoint triggering causes (e.g., the baby's arrival), structural explanations identify constraints that shape the causal relationships between triggering causes and their effects (Dretske, 1988).

The structural view capitalizes on the distinction between *nodes* (positions within social structures) and *node-occupiers* (categories that occupy those positions, and come to possess particular properties by virtue of their location within the structure; Haslanger, 2015). This distinction brings to light a potential ambiguity in formal explanations (e.g. "Smith quit her job after having a baby because she is a woman"), given that the term "woman" can refer to either the node or the node-occupier. Such explanations could attribute properties directly to the node (i.e., women's location in a structure), without necessarily tying them to its inherent nature (i.e., to women themselves). In other words, a formal explanation could support both structural and internalist interpretations, a prediction that our experiments test.

Structural vs. other externalist approaches. One way to appreciate what constitutes a structural explanation is to consider what it is not. Structural explanations are not merely "situation" explanations from the traditional person-situation dichotomy (Ross & Nisbett, 2011), nor "causal history of reasons" explanations from Malle's (2004) taxonomy, which are narrower in their restriction to intentional behavior, yet broader in allowing for *non-structural* antecedents to reasons. Structural explanations are a sub-type of externalist explanations that invoke *stable* constraints acting on a *category* in virtue of its *position in a structure*.

It's useful to think of structural explanations in terms of the ANOVA or "cube model" (Kelley, 1973), in which a behavior is attributed to co-varying factors (person, situation, or stimulus). However, the cube model assumes that the data (behaviors) come from an "unconfounded" factorial design, where person and external factors vary independently. Structural thinking is instead sensitive to *confounds* between people and situations; within a social structure, categories are often constrained by their nodes. The category "women" can only occupy the "women" node, which constrains the range of properties the occupier can display.

The notion of a confound between a category and its social location also helps to position the structural view of categories relative to role-based categories, such as *guest*, which specify a role in a relational structure (Asmuth & Gentner, 2016; Markman & Stilwell, 2001). Role-based categories involve relational structure, but structural thinking about social categories critically applies to cases in which a relational position is *confounded* with membership in a (perceived) taxonomic category.

Cross-cultural research on independent vs. interdependent (object vs. field) construals (Nisbett, 2003) suggests that the reasoning style associated with structural thinking is not as "unnatural" as it may seem. Research on analogy (Gentner, 1983; 2005) and recent work on role-based concepts (Goldwater, Bainbridge, & Murphy, 2016) offer additional indications that people have the representational capacities to reason about structures. If people possess the requisite resources for engaging in structural reasoning, the question is: do they? And if so, when does this capacity develop? These are the questions our study addresses.

The development of structural thinking. Our study evaluates two competing hypotheses. *Hypothesis 1* is that young children lack the conceptual prerequisites and/or knowledge to engage in structural thinking. *Hypothesis* 2 is that young children can successfully engage in structural thinking from an early age.

Each hypothesis receives some support from existing research. In favor of the first hypothesis, prior work demonstrates that children view some social categories (such as gender) as essentialized natural kinds from an early age (Rhodes & Gelman, 2009; Taylor, 1996), even when cultural input suggests otherwise (Astuti et al., 2004). There is also evidence that young children have trouble endorsing environmental mechanisms that could produce category features (Rhodes & Taylor, 2009), although the "environmental factors" that were examined were primarily non-structural in nature. Finally, as young as 4-5 years of age, children tend to generate and endorse "inherent" explanations of categorical patterns over "extrinsic" ones (Cimpian & Markman, 2011; Cimpian & Steinberg, 2014).

Beyond evidence of early essentialist and inherence-based reasoning, there is evidence that children lack capacities involved in structural thinking. Structural explanation could rely on structure-wide counterfactual alternatives, which do not fully emerge until age 7-8 (Beck et al., 2006; Rafetseder, Cristi-Vargas, & Perner, 2010). Structural reasoning also relies on representing relations, and research on relational reasoning suggests a developmental shift in relevant capacities throughout and beyond the preschool years (e.g., Gentner, 1988; Richland, Morrison, & Holyoak, 2006).

On the other hand, there is evidence that potentially favors Hypothesis 2. Several findings suggest that young children appreciate external constraints on social categories. Seiver, Gopnik, and Goodman (2013) demonstrated that children as young as 4 can use *situational* information in explanation and prediction when appropriate covariation evidence is available. Four-year-olds also recognize moral constraints on their own behavior (Chernyak & Kushnir, 2014) and acknowledge that the behavior of members of a social category can be driven by common norms (Kalish, 2011; see also Kalish & Shiverick, 2004; Rakoczy, Warneken, & Tomasello, 2008; Smetana, 1981; Turiel, 1983).[1]

[1] Translating research on norms into predictions about structural reasoning is not straightforward. First, moral norms carry deontic content, which distinguishes them from other kinds of structural constraints (such as a wage gap) that do not. Second, category-specific norms can be interpreted in either essentialist or structural terms (e.g., if girls are not allowed to go out after 9 pm, this could stem from inherent characteristics of girls, or structural forces). Existing studies about norms have not made these distinctions, complicating their interpretation with regard to structural reasoning.

A final and more intriguing possibility is that young children may be *more* open to structural reasoning than older children and adults. Rhodes and Gelman (2009) showed that young children are more flexible than older children about some social categories, such as race. In Seiver et al. (2013), it was older children, not younger children, who showed an overall bias for person over situation explanations. Moreover, young children may be less biased by prior assumptions than adults, and thus open to learning a broad range of causal relationships (Lucas, Bridgers, Griffiths, & Gopnik, 2013). This body of work suggests that relative to older children and adults, young children could be more open to integrating external constraints in their representations of social categories and relying on structural relations in reasoning.

Experiment

This study had three goals: to determine whether and when children can successfully engage in structural thinking in explaining the association between a category and a property; to determine whether a structural construal can be experimentally induced; and to evaluate the prediction that structural thinking can support formal explanations under a structural interpretation of the category.

To address these three goals, we adopted an approach mirroring Prasada and Dillingham (2006, 2009), who developed a set of tasks that can be used to identify whether people construe the connection between a feature and a category as principled (such as between "fighting crime" and being a police officer) or statistical (such as between "eating donuts" and being a police officer). They showed that only the principled connections between kinds and features supported judgments of feature *immutability* (a person who does not fight crime is not really a police officer), *partial definitions* (a police officer is a person who fights crime), and *formal explanations* ("this person fights crime because she is a police officer"). With the aim of detecting structural thinking and differentiating it from internalist thinking, we modified these three measures (described below). Vasilyeva and Lombrozo (in prep) found that with adults, responses across these judgments can successfully be used as a "profiling tool" to detect structural thinking, which generates a unique signature: relatively high mutability ratings, low partial definition ratings, and high formal explanation ratings. In contrast, the pattern for an internalist construal should be low – high – high. To further validate the profiling tool, we additionally included an open-ended explanation prompt and close-ended causal explanation evaluations.

Method

Participants We recruited 41 3-4-year-olds (mean age 4.3 years, range 3.0-4.9; 23 females, 18 males) and 48 5-6-year-olds (mean age 5.6 years, range 5.0-6.9; 23 females, 25 males). Additionally, 67 adults (mean age 33 years, range 19-71; 33 females, 64 males) were recruited via Amazon Mechanical Turk; participation was restricted to users with an IP address within the US and an approval rating of at least 95% based on at least 50 previous tasks. Children were tested in person using an illustrated storybook presented on a laptop; adults were tested online.

Materials, Design, and Procedure Participants were first introduced to a school where girls and boys study in separate classrooms, and presented with fictitious data about students playing different games during recess: girls predominantly played Yellow-Ball while boys predominantly played Green-Ball. Participants were told that the game each child played was determined by tossing a pebble towards two buckets standing side by side: if the pebble fell into the yellow bucket, the child played Yellow-Ball that day, and if the pebble fell into the green bucket, that child played Green-Ball that day (Figure 1a). The critical manipulation concerned the sizes of the buckets. In the internalist condition, both buckets were of the same size (Figure 1b), inviting participants to infer that the correlation between category membership and game choice was the product of inherent preferences (see Kushnir, Xu, & Wellman, 2010, for evidence that even younger children can use statistical evidence to infer a preference). In the structural condition, one bucket was instead much larger than the other: in the girls' classroom the yellow bucket was larger than the green bucket, with the reverse in the boys' classroom (Figure 1c). The size difference imposed a stable structural constraint on the probability distribution over options available to members of each category, inviting a structural interpretation of the category-property connection.

After comprehension checks, all participants completed a series of measures designed to differentiate an internalist from a structural construal of the property. First, in the *open-ended explanation* task, participants were asked why girls in the girls' classroom play Yellow-Ball a lot at their school. Second, participants completed a causal explanation evaluation task and the three profiling tools measures: a mutability judgment, a partial definition, and formal explanation ratings.

In the *causal explanation evaluation* task, children evaluated three kinds of causal explanations offered by puppets that "sometimes say things that are smart, and sometimes say things that are silly." The puppets explained that girls tend to play Yellow-Ball "because girls like playing Yellow-Ball" (internalist); "because in the girls' classroom, it's easier to throw a pebble in the yellow bucket" (structural); or "because they got sprinkled with water" (an incidental explanation invoking an irrelevant fact from the cover story). Participants evaluated each explanation using a two-step, four-

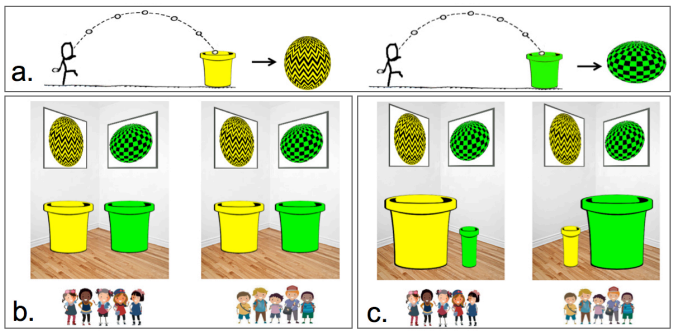

Figure 1: Illustrations of the procedure determining which game each student played in the story (a) and of the different constraints on the probability of outcomes in the internalist (b) and structural conditions (c).

point thumb scale: they first chose one of two thumbs representing "good explanation" (up) and "bad explanation" (down), and they then chose between two subsequent options based on their choice: "kind of good/bad" (small thumb) or "really good/bad" (big thumb) – a scale previously shown to work well to measure children's agreement with explanations (Cimpian & Steinberg, 2014; Hussak & Cimpian, 2015).

For the *mutability judgment*, participants were told that after a change in the school's rules allowing children to attend any classroom, Suzy's parents transferred her to the boys' classroom. Participants were asked to guess which game Suzy would play after transferring, responding on a two-step, four-point scale ranging from "for sure Yellow-Ball" to "for sure Green-Ball." This mutability judgment mirrors more familiar "switched at birth" tasks in the essentialism literature (Gelman & Wellman, 1991), in which children are asked, e.g., whether a cow raised by pigs will moo or go oink. Similarly, our mutability judgment involves a change in environment (structural constraints), and participants are asked to infer whether a property will match the exemplar's category (the node occupier) or the new environment (the node). A shift in predictions from Yellow-Ball to Green-Ball should track the causal influence of the node, and indicate structural thinking (as well as show that structural positions are seen as influencing behavior, rather than merely reflecting existing internal preferences).

For the *partial definition* task, participants rated whether an alien did a good job telling what a girl is to another alien who had never heard about girls: "A girl is a person who plays Yellow-Ball a lot." Participants used a two-step, four-point scale ("really bad job" - "really good job").

In the *formal explanation* task, participants met Suzy who "plays Yellow-Ball a lot at her school" and were asked to evaluate a formal explanation offered by a puppet - "Because Suzy is a girl" - using the two-step, four-point thumb scale ranging from "really bad" to "really good."

Results and Discussion

Due to differing test formats and sample sizes, data from children and adults were analyzed separately. For the open-ended explanation task (see Figure 2), participants' explanations were coded as internalist ("maybe the girls just like it better, so they always aim to get their pebbles into the yellow ball bucket"), structural ("because the pebble went into the yellow bin, because the yellow one is bigger"), or miscellaneous, comprised of "I don't know," question restatements, and unclassifiable responses ("the yellow ball is brighter"). The distribution of explanations was affected by condition for each age group (3-4-year-olds: $\chi^2(N=41)=6.19$, $p=.045$; 5-6-year-olds: $\chi^2(N=48)=16.80$, $p<.001$; adults: $\chi^2(N=67)=42.86$, $p<.001$). Critically, in the structural condition some proportion of participants in each age group produced structural explanations (Figure 2, right panel, black bars). There was also evidence of developmental change: in the structural condition, the percentage of internalist explanations dropped as the percentage of structural explanations increased, so that the overall preference for internalist explanations in the younger age group flipped to a preference for structural explanations for older children.

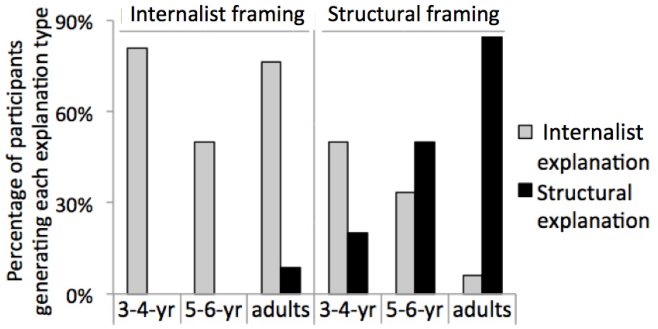

Figure 2: Distribution of internalist and structural explanations generated in response to question about why girls play Yellow-Ball, as a function of condition and age group.

Children's evaluations of *causal explanations* (see Figure 3) were analyzed as a function of explanation type (internalist, structural, incidental), condition (internalist, structural), and age group (3-4, 5-6 year-olds) in a mixed ANOVA, with the key prediction concerning an *interaction* between explanation type and condition. The analysis revealed a main effect of explanation type, $F(2,170)=9.87$, $p<.001$, $\eta_p^2=.104$, which was qualified by a significant interaction between explanation type and condition, $F(2,170)=6.00$, $p=.003$, $\eta_p^2=.066$: only the structural explanation ratings were boosted by the structural framing. Most importantly, we observed the target three-way interaction: $F(2,170)=3.73$, $p=.026$, $\eta_p^2=.042$, driven by the selective effect of condition on 5-6-year-olds' evaluations of the structural explanation: older children, but not younger children, rated structural explanations higher in the structural condition than in the internalist condition ($p_{older}<.001$, $p_{younger}=.390$). The interaction remained significant when restricting the analysis to internalist and structural explanations, $p=.012$. For adults, an explanation type (essentialist, structural, incidental) by condition (essentialist, structural) mixed ANOVA revealed a significant effect of explanation type, $F(2,126)=171.15$, $p<.001$, $\eta_p^2=.731$, and a marginal effect of condition, $F(1,63)=3.74$, $p=.058$, $\eta_p^2=.056$, qualified by a significant interaction, $F(2,126)=117.83$, $p<.001$, $\eta_p^2=.652$: adults favored the internalist explanation over the structural in the internalist condition, with the reverse in the structural condition (p's$<.001$, see Figure 3).

Having succeeded in finding evidence of structural thinking in our open- and close-ended causal explanation tasks, we next turn to the profiling tool measures to see whether they reveal developmental differences mirroring these patterns. For adults, who exhibited high levels of structural thinking, we would predict the following for the structural condition relative to the internalist condition: more frequent predictions

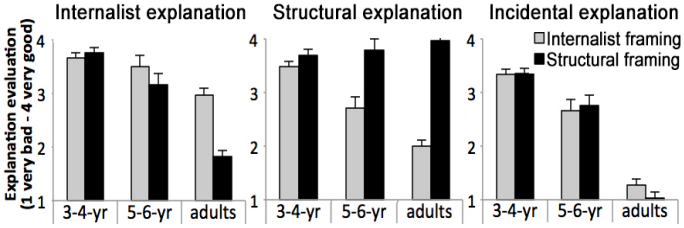

Figure 3: Explanation evaluation as a function of explanation type, framing condition, and age group.

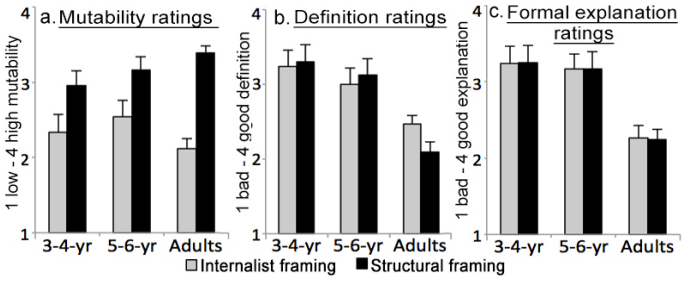

Figure 4: Mutability (a), partial definition (b), and formal explanation ratings (c) as a function of framing condition and age group.

that Suzy will play Green-Ball when switched to the boys' classroom, lower endorsement of the partial definition, and no difference in endorsement of the formal explanation. For children, we would predict the same patterns, with smaller effects for the younger children. This is what we found.

For the *mutability judgment* task (see Figure 4a), the predicted main effect of condition was marginal for the youngest group, $t(39)=1.96$, $p=.057$, $d=.42$, but significant for the older children, $t(46)=2.29$, $p=.027$, $d=.63$, and for adults, $t(65)=8.04$, $p<.001$, $d=2.00$. The age by condition interaction for children was not significant, $F(1.85)<.01$, p=.984, but it appears that the property (playing Yellow-Ball for girls) was seen as more mutable in the structural condition than in the essentialist condition by age 5-6.

For the *partial definition* task (Figure 4b), we predicted that properties construed as internalist should support definitions better than properties construed as structural. Neither younger nor older children displayed such a pattern (p's$\geq$.687), but adults did, $t(65)=2.11$, $p=.039$, $d=.52$.

Finally, as predicted, *formal explanation* ratings did not significantly differ across the essentialist and structural conditions for any age group, all p's$\geq$.915 (see Figure 4c), suggesting that these explanations support both internalist and structural construals.

These results show that even young children are capable of structural thinking, as reflected in their open-ended explanations. They also provide the first demonstration that across all age groups formal (categorical) explanations support two interpretations: essentialist and structural. Beyond these age-general effects, they reveal developmental changes in structural thinking, with older children and adults more readily engaged in structural thinking. (Notably, we have reasons to believe that the observed pattern of developmental change is not due to younger children simply not understanding the task or explanations: in the explanation generation task younger children produced predominantly internalist explanations regardless of the framing, and when asked to break ties in the explanation evaluation task, they ranked internalist explanations higher under the internalist framing.) Moreover, these results suggest that internalist versus structural construals can be effectively induced, though in reality, they likely coexist, and are triggered by different cues. Finally, our results show that the profiling tool can effectively track internalist versus structural thinking across development.

General Discussion

Using novel tasks designed to assess structural thinking, we find evidence that even young children are able to reason about social categories in structural terms. By age 5-6, children preferentially generated and accepted structural explanations for a category-property association when a structural constraint was presented, with hints of an emerging sensitivity by ages 3-4.

Recognizing structural reasoning as a distinct cognitive phenomenon invites us to rethink some of the findings in the literature on psychological essentialism. For example, many discussions of essentialism emphasize its capacity to support predictions and promote generalizations across category members (Gelman, 2003). In fact, generalization tasks are often used to *measure* the extent to which a category representation is essence-based. However, a structural representation of a category can likewise support such generalizations: structural forces shape properties of the nodes within the structure, and the occupiers of the nodes, being subject to these forces qua occupiers, are likely to obtain the properties in spite of idiosyncrasies in their individual histories and predispositions. Haslanger (2015) correspondingly praises structural explanations for their stability and identification of broad patterns that hold across "inessential perturbations," suggesting that such explanations may be particularly good in supporting generalizations within stable structures. These features of structural thinking challenge the widespread assumption that the stability and generalizability of category properties imply internalist (essentialist) representations. More generally, our findings lay the groundwork for refining internalist claims and the evidence that is taken to support them, and for making more fine-grained distinctions when it comes to externalist alternatives.

We have also demonstrated that formal explanations support both structural and internalist interpretations. Introducing structural connections as a new type of non-accidental relationship between a property and a category raises new questions about generics (e.g., "Girls prefer pink"), which are implicated in perpetuating stereotypes. On most accounts, generics are interpreted as expressing something about the underlying nature of the category, reinforcing essentialist beliefs (Cimpian & Markman, 2011; Leslie, 2014; Prasada & Dillingham, 2009). For example, Leslie argues that generics are by default interpreted as expressing "generalizations that hold because of common, inherent features of the members of the kind" (p. 217), where the only alternative available to people is interpreting generics as describing statistical connections, along the lines of "police officers eat donuts," on the basis of "specific worldly knowledge." But if people can interpret generics structurally, by construing features of category members as products of structural constraints rather than inherent attributes of the kind, this potentially opens up a new way to mitigate harmful side-effects of generic language without purging it from everyday speech (or, equally implausibly, convincing people that many associations between properties and social categories are merely "accidental").

By introducing a structural alternative into the dichotomy of internal vs. vaguely and variably defined external (situational) factors in explanations of behavior, we have unmasked a gap in our understanding of categorical reasoning, and opened up new directions of study that may help account for some of the mixed evidence in research on the development of relational reasoning, essentialist beliefs about social categories, and reasoning about moral and conventional norms. The reported work already calls for revision of current accounts of generic language and formal explanation, and highlights the need to study categories embedded in relational structures. But of course, a lot more remains to be done.

Acknowledgments

This work was supported by the Varieties of Understanding Project, funded by the John Templeton Foundation, and by the McDonnell Foundation.

References

Asmuth, J., & Gentner, D. (2016). Relational categories are more mutable than entity categories. *Quarterly Journal of Experimental Psychology, 0*, 1-19.

Astuti, R., Solomon, G. E., Carey, S., Ingold, T., & Miller, P. H. (2004). Constraints on conceptual development: A case study of the acquisition of folkbiological and folksociological knowledge in Madagascar. *Monographs of the society for Research in Child Development*, i-161.

Beck, S. R., Robinson, E. J., Carroll, D. J.,& Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77*(2), 413- 426.

Chernyak, N., & Kushnir, T. (2014). The self as a moral agent: Preschoolers behave morally but believe in the freedom to do otherwise. *Journal of Cognition and Development*, *15*(3), 453-464.

Cimpian, A.,& Markman, E.M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development, 82*(2), 471–492.

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences, 37*(5), 461–480.

Cimpian, A., & Steinberg, O. D. (2014). The inherence heuristic across development: Systematic differences between children's and adults' explanations for everyday facts. *Cognitive Psychology, 75*, 130–154.

Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. Cambridge, MA: MIT Press.

Fausto-Sterling, A. (2012). Sex/Gender: Biology in a social world. Routledge.

Gelman, S.A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.

Gelman, S.A., & Wellman, H.M. (1991). Insides and essences: Early understandings of the nonobvious. Cognition, 23, 183-209.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science*, *7*(2), 155-170.

Gentner, D. (1988). Metaphor as structure-mapping: The relational shift. *Child Development, 59,* 47-59.

Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories in developmental time* (pp. 245–275). Mahwah, NJ: Erlbaum.

Goldwater, M. B., Bainbridge, R., & Murphy, G. L. (2016). Learning of role-governed and thematic categories. *Acta Psychologica*, *164*, 112-126.

Haslanger, S. (2015). What is a (social) structural explanation? *Philosophical Studies*, 1–18.

Hussak, L. J., & Cimpian, A. (2015). An early-emerging explanatory heuristic promotes support for the status quo. *Journal of Personality and Socisal Psychology, 109*(5), 739-752.

Kalish, C.W. (2011). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology*, *48*(4), 1133-1143.

Kalish, C.W., & Shiverick, S.M. (2004). Children's reasoning about norms and traits as motives for behavior. *Cognitive Development*, *19*(3), 401–416.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologists*, *28*(2), 107-128.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*(8), 1134-1140.

Leslie, S.-J. (2015). The original sin of cognition: Fear, prejudice and generalization. *The Journal of Philosophy, 3*.

Leslie, S.-J. (2014). Carving up the social world with generics. In T. Lombrozo, J. Knobe, & S. Nichols (Eds.), *Oxford Studies in Exp. Philosophy, (pp.* 208-232).

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284–299.

Malle, B. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction.* Cambridge, MA: The MIT Press.

Markman, A.B., & Stilwell, C.H. (2001). Role-governed categories. *Journal of Experimental & Theoretical Artificial Intelligence*, *13*(4), 329-358.

Nisbett, R.E. (2003). *The geography of thought: How Asians and Westerners think differently… And why.* New York, NY: The Free Press.

Prasada, S., & Dillingham, E. M. (2006). Principled and statistical connections in common sense conception. *Cognition*, *99*, 73–112.

Prasada, S.,& Dillingham, E. M.(2009). Representation of principled connections: a window onto the formal aspect of common sense conception. *Cogitive Science*, *33*(3), 401–48.

Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of "nearest possible world". *Child development*, *81*(1), 376-389.

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Developmental Psychology*, *44*(3), 875–881.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*, 249–273.

Rhodes, M., & Gelman, S. (2009). A developmental examination of the conceptual structure. *Cognitive Psychology*, *59*(3), 244–274.

Ross, L., & Nisbett, R. E. (2011). *The person and the situation: Perspectives of social psychology*. Pinter & Martin Publishers.

Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development, 84,* 443–454.

Smetana, J.G. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52,* 1333–1336.

Taylor, M.G. (1996). The development of children's beliefs about social and biological aspects of gender differences. Child Development, 67, 1555-1571.

Thibaut, J. P., French, R., & Vezneva, M. (2010). The development of analogy making in children: Cognitive load and executive function. *Journal of Experimental Child Psychology, 106*(1), 1-19.

Turiel, E. (1983). *The development of social knowledge: Morality and convention.* Cambridge, UK: Cambridge University Press.

Vasilyeva, N. & Lombrozo, T. (in prep.) Structural thinking.