Cognitive Psychology 130 (2021) 101422

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Evidence for multiple sources of inductive potential: Occupations and their relations to social institutions

Alexander Noyes [a,*], Yarrow Dunham [b], Frank C. Keil [b], Katherine Ritchie [c]

[a] New York University, United States
[b] Yale University, United States
[c] University of California, Irvine, United States

ARTICLE INFO

Keywords:
Concepts
Categories
Inductive potential
Social roles
Social cognition

ABSTRACT

Several current theories have essences as primary drivers of inductive potential: e.g., people infer dogs share properties because they share essences. We investigated the possibility that people take occupational roles as having robust inductive potential because of a different source: their position in stable social institutions. In Studies 1–4, participants learned a novel property about a target, and then decided whether two new individuals had the property (one with the same occupation, one without). Participants used occupational roles to robustly generalize rights and obligations, functional behaviors, personality traits, and skills. In Studies 5–6, we contrasted occupational roles (via label) with race/gender (via visual face cues). Participants reliably favored occupational roles over race/gender for generalizing rights and obligations, functional behaviors, personality traits, and skills (they favored race/gender for inferring leisure behaviors and physiological properties). Occupational roles supported inferences to the same extent as animal categories (Studies 4 and 6). In Study 7, we examined why members of occupational roles share properties. Participants did not attribute the inductive potential of occupational roles to essences, they attributed it to social institutions. In combination, these seven studies demonstrate that any theory of inductive potential must pluralistically allow for both essences and social institutions to form the basis of inductive potential.

1. Introduction

Why do certain categories support inductive inference more than others? For example, why does learning a particular property of some specific instances (e.g., that my dog produces its own vitamin C) seem to so easily extend to other members of the same category (dogs) whereas others do not (e.g., my chair is highly flammable)? This question is at the heart of how categories are mentally represented and used in daily life (Gelman, 1988). To answer it, we examined the inductive potential of occupational roles. These categories lie on the outskirts of current theories, many of which focus on essences as primary drivers of inductive potential (e.g., Gelman, 2003). For instance, people may infer dogs share indefinitely many properties because they share essences. Members of occupational roles lack a deeper nature that makes them what they are; for instance, there is no 'lawyer essence.' But these categories are ideal candidates for exploring the boundaries of current theories of inductive potential. If occupational categories have robust inductive potential, then essences cannot be the primary driving force. Instead, multiple sources or a pluralistic set of intuitions may underpin

* Corresponding author at: New York University, Department of Psychology, 6 Washington Pl, New York, NY 10003, United States.
E-mail address: NoyesAlexanderJ@gmail.com (A. Noyes).

https://doi.org/10.1016/j.cogpsych.2021.101422
Received 22 March 2021; Received in revised form 9 August 2021; Accepted 9 August 2021
Available online 4 September 2021
0010-0285/© 2021 Elsevier Inc. All rights reserved.

our inductive practices. The inductive potential of social roles may arise from an underexplored and completely different kind of basis: The position of occupational roles in stable social institutions.

1.1. What is inductive potential?

People can draw inductive inferences in all category domains, from food, office supplies, to furniture (e.g., Lafraire et al., 2020). But inductive potential differs in systematic ways across different domains of kinds. For example, paperclips share many correlated properties: they're composed of steel wire, they're bent in two loops, they're used to bind sheets of paper, and they measure an inch or two. This allows one to accurately guess the material, shape, function, and size of a new paperclip prior to seeing it. Therefore, the category *paperclip* is informative: Learning an unknown object is a paperclip carries considerable information. Further, members are homogenous: nearly all paperclips are highly similar to each other. Yet, the category *paperclip* may lack robust inductive potential by more stringent metrics.

Foremost, categories like *paperclip* fail to support new learning (Gelman, 1988; Gelman, 2003; Keil, 1989; Markman, 1989): When learning a novel property about a single paperclip (e.g., "contains silumin"), one would have little reason to suspect that other paperclips share the property, unless the novel property added detail to known functional properties (e.g., "inserting sheets of paper between the loops causes torsion force"). Likewise, much of the similarity between paperclips is perceptually observable: If you encountered an unusual looking paperclip, then most of the inferences about size, shape, and material would be lost; there is little unobservable structure.

Four criteria emerge from this discussion. If a category has robust inductive potential, then (1) participants should use the category to generalize novel or unknown properties to new category members; (2) participants should generalize unobservable properties; (3) participants should generalize properties to perceptually dissimilar category members; (4) participants should favor the category as a basis for inference over apparent similarity. A category that supports (1)-(4) has inductive potential by any commonly used account. Once established, one can then use categories with well-established inductive potential as benchmarks: Animal categories are the standard. Among categories of people, race, gender, ethnicity, and religion are the standard (depending on cultural context).

2. Social roles

Social roles are positions in social structures. For example, the role of President of the U.S. is embedded in the U.S. federal government. Individuals occupy a role when they perform the relevant functional part in an organization or are bound by norms related to the role. This distinguishes social roles from other human creations. Artifacts are created and become members of their category at the same time: For example, the Eiffel Tower came into existence at the same moment it became a tower. But, members of social roles can exist prior to entering a role and after they leave it. For example, Barack Obama existed prior to being the President and continues to exist after.

Occupational roles have clear connections to specific institutions, and so they are overtly institutional. This contrasts with entrenched social roles like race and gender (Haslanger, 2000; Mallon, 2016), which are covertly social, such that many people assume they are natural kinds (Rothbart & Taylor, 1992; Gelman, 2003). Occupational roles may suggest new ways to understand how we generalize properties across a wide range of social kinds. These discoveries could provide valuable insights into other social kinds—such as race and gender—as well.

3. Sources of inductive potential

When and why do people infer that a category has robust inductive potential? If people infer that category members share essences, category membership would enable generalization of many newly learned property types that are seen as outcomes of the essences. For example, people may anticipate tigers to share stripes, fur, and sharp claws if they assume those properties are inherited. And, people may not expect tigers to share eating raw chicken after seeing a tiger in a zoo eating raw chicken if they infer the property was merely incidental. Members of categories may share microstructural essences that explain their properties: e.g., atomic structure makes gold ductile (Kripke, 1972; Putnam, 1975). Whether this is a correct description of reality or not (Boyd, 1999; Leslie, 2013), people appear to infer that there are shared essences (Medin & Ortony, 1989; Gelman, 2003). Specifically, people infer natural objects (e.g., gold) have microstructural essences and artificial objects (e.g., paperclips) do not (Keil, 1989; Gelman, 1988). This assumption leads people to judge natural objects have greater inductive potential than simple artifacts. Thus, the benchmark of robust inductive potential are categories of animals, plants, and chemical substances, which people reliably infer have essences.

If occupational roles had robust inductive potential, could a belief in essences be the source? People can attribute essences to social-role categories. People attribute essences to gender early in development and across cultural contexts (Diesendruck, Goldfein-Elbaz, Rhodes, Gelman, & Neumark, 2013; Prentice & Miller, 2006; Taylor, Rhodes, & Gelman, 2009). People also attribute essences to categories associated with racial or ethnic differences, including religion and language (Diesendruck et al., 2013; Gil-White, 2001; Rhodes & Mandalaywala, 2017). Yet, people do not heedlessly attribute essences to the vast majority of social roles, including occupational roles and membership in clubs, teams, and organizations (e.g., Diesendruck et al., 2013; Noyes & Dunham, 2017, 2020; Noyes & Keil, 2020). Specifically, at baseline, children reason that groups are constituted by mutual intentions: An individual belongs to a group because she and other group members recognize and accept her membership; unlike gender, children reason that group membership can change (Noyes & Dunham, 2017, 2020). In another study, children and adults inferred, at baseline, that members of social categories shared properties because of their shared socialization; unlike gender, children reasoned individuals would not

express the same category-linked properties if they lived in a different social context (Noyes & Keil, 2019; 2020). Thus, if essences drove the inductive potential of occupational roles, it would likely have to be by proxy to another category. People might attribute essences to social composites that include occupational roles. For example, people might infer lawyers are upper-class White men and then import their stereotypes about these other categories into their thinking about lawyers. But such inferences would fail to support robust inductive potential of the kind *lawyers*, as the inferential strength would break down when we encounter conflicting perceptual features associated with race, gender, and age. Given the criteria outlined above, this suggests these social compositions lack robust inductive potential. A final possibility is that people have in mind a parallel natural category. There may be "natural lawyers" and "natural quarterbacks;" people may attribute essences to these parallel categories but not to their non-natural counterparts.

Social institutions, however, may support rich inductive inferences using a different but equally powerful inferential framework. Even if playing a role is transient, the role itself can be an extremely stable part of the institution: e.g., the quarterback role is a stable part of football. When individuals occupy the role, they inherit properties of the role: A quarterback acquires new privileges and obligations and performs new functional behaviors. When inferring properties from one quarterback to another, we could rely on their shared position in the institution and on the stability of the institution itself, rather than on the stability of the individual's internal structure.

Institutions can also have lasting effects on people. Even if occupational roles are impermanent and acquired, institutions shape people before, during, and after their membership. For example, all lawyers went to law school but none of them were lawyers at the time. Prolonged exposure to an institution systematically alters a person's experiences: They receive specialized training; individuals treat them differently; the job puts unique demands on their body and mind; the job alters their routines; they become socialized into unique organizational cultures; they imitate their career idols; they gain or lose access to resources. This onslaught of interacting causal forces can produce clusters of correlated properties, including new skills and personality traits. Individuals can carry these properties with them even as they move in and out of career-specific situations, and even as they retire. Therefore, institutions could shape deeper features of individuals.

A person's social role is unobservable. Lawyers vary in gender, race, ethnicity, age, and body composition. Though lawyers may dress alike and signal their social role in some contexts, they may dress differently in others. Social institutions, and positions in social institutions, are non-obvious. If members of occupational roles share properties, they would often be unobservable too, such as rights and obligations, functional behaviors, skills, and personality traits. If occupational roles support robust inferences at all, they would support invisible ones.

Even if social institutions could support inductive potential, they may not actually do so in our daily lives. We need empirical studies to determine whether institutions support inductive potential in ordinary reasoning. The evidence so far does suggest that people represent occupational roles as embedded in institutions (Noyes, Dunham, & Keil, 2018; 2020). People also infer social structures can sustain category-property associations (Noyes & Keil, 2019; 2020; Vasilyeva, Gopnik, & Lombrozo, 2018; Vasilyeva & Lombrozo, 2020). For example, people reasoned that "boys like trucks" could be true even absent any intrinsic property of boys; it could true because boys only have access to trucks (Vasilyeva et al., 2018). Likewise, when people heard a social category described with properties like "they value punctuality", they reasoned that socialization was the source of the similarity between members (Noyes & Keil, 2019; 2020). Thus, people could take categories to have inductive potential because of the way they systematically relate to social institutions. However, these studies were limited in several ways. Noyes and Keil (2019) did not systematically evaluate the four criteria of inductive potential enumerated above. As with the category paperclip, it may be that people merely took socialization as a source of homogeneity but did not infer robust inductive potential. Vasilyeva et al. (2018) demonstrated that people can explain isolated category-property links as social-structural when given evidence favoring that explanation. But people may not take social structures to support new learning beyond those known category-property links, and people may not infer social structures absent instruction. Without a systematic investigation of the inductive potential of occupational roles or other categories understood as

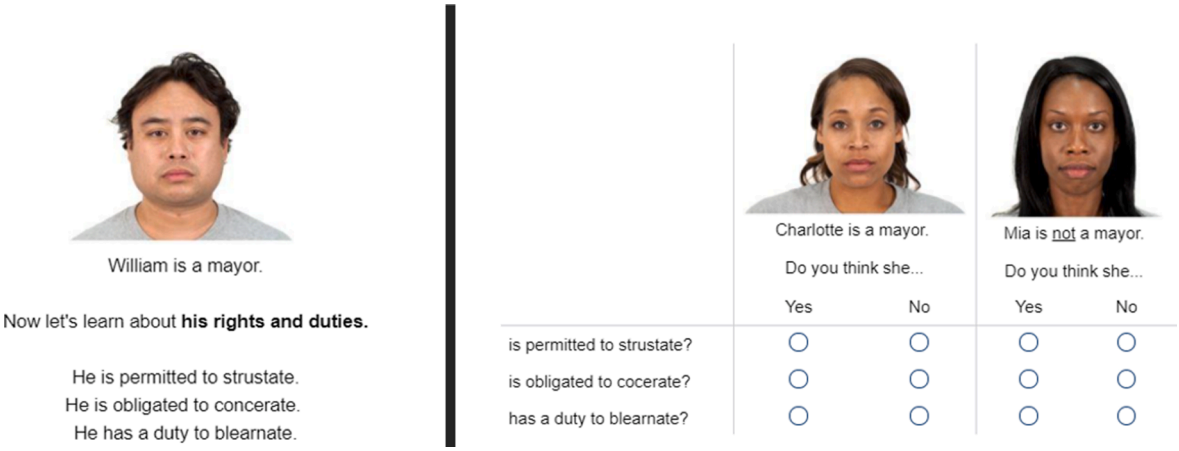

Fig. 1. The added-value paradigm used in Studies 1–4. Participants indicate whether they believe the test (right) individuals have the same properties as the target (left) individual.

institutional, we cannot know whether occupational-role categories promote robust inductive potential, and whether the source of that inductive potential is the position of individuals in social institutions.

3.1. Empirical approach

We aimed to assess the inductive potential of occupational roles and identify its sources. We focused on social roles that are unambiguously embedded in social institutions: lawyer, judge, mayor, quarterback, and police officer. This embedding allowed us to focus on the possibility that social institutions are a source of inductive potential. We excluded social roles that are consistently essentialized, such as gender. We excluded occupational roles, like farmers, which may not be institutional, as one might classify a person as a farmer solely because they grow crops, even if they do so entirely on their own without any recognition by others. We also excluded categories that previous research has found to be dual character (a category, like *scientist*, is dual character when people think an individual can be a member by realizing certain values, such as the pursuit for objective truth; Knobe, Prasada, & Newman, 2013; we return to this topic at length in the General Discussion). All of our studies were pre-registered: osf.io/ec7q3/. Note, however, that we are presenting our studies out of chronological order, and have re-grouped by method and question to aid comprehension.

In the first set of studies (Studies 1–4), we used an 'added value' paradigm (Fig. 1). Participants learned about a target individual's occupation and properties. Participants then encountered two new individuals. Both test individuals were the same race, gender, and age as each other, and both differed from the initial target individual in race and gender. One of the test individuals had the same occupation as the target individual and one did not. Participants indicated whether they thought each individual had the properties of the target individual. This allowed us to measure the added value of knowing occupational category membership, as there was no perceptually-guided reason to generalize properties and the test individuals differed only in occupational role.

An important component of our project was distinguishing between different types of properties. For example, we were interested in whether occupational roles support inferences about 'internal' properties like personality traits and skills, or only 'external' properties like rights and obligations and functional behaviors. In all cases, we focused on novel properties. Participants had just enough information to identify the type of property: e.g., "She dances the scubblebet" or "he conducts tingforms." We used novel properties to guarantee that participants were not relying on information they already had about the category or category members (Gelman, 2003). In Studies 1–3, we compared occupational roles to region of origin (e.g., "Northeast resident") and religion (e.g., "Christian"). We selected region of origin as a low inductive potential control. By selecting regions like "Northeast" we could avoid any municipal, state, or federal institutions. These categories also do not map onto essences, especially as we aimed for less conventional cuts and avoided lexicalized noun phrases ("resident of the Northeast" instead of "New Englander;" "resident of the Southeast" instead of "Southerner"). We also compared occupational roles with religion; unlike region, religion categories have high inductive potential (Diesendruck & HaLevi, 2006; Haslam, Rothschild, & Ernst, 2000; Smyth, Feeney, Eidson, & Coley, 2017). In Study 2, we removed faces to test whether seeing diverse faces reduced the inductive potential of any of the categories. In Study 3, we removed property information, to estimate the effect of knowing the type of property. In Study 4, we compared occupational roles to animal categories, as animal categories are the benchmark for robust inductive potential.

In Studies 5 and 6, we used an 'appearance foil' paradigm (Fig. 7; Gelman & Markman, 1986; 1987). The target individual was the same race and gender as one test individual and a different race and gender from another. The test individual who differed in race and gender belonged to the same occupation. Participants were then given a forced choice: They had to either select the occupational-role match or the appearance (race/gender) match. This manipulation allowed us to test whether participants use occupational roles when category membership conflicts with perceptual similarity – a hallmark of robust inductive potential (Gelman & Markman, 1986, 1987). It also allowed us to pit the inductive potential of occupational roles against race and gender. As noted above, race and gender are often used as benchmarks for robust inductive potential among social categories in the U.S. context (Gelman, 2003; Prentice & Miller, 2007). In Study 6, we tested animal categories in this paradigm, to examine whether occupational roles support perceptually unguided inferences as well as animal categories.

In Study 7, we examined the source of the inductive potential of occupational roles. Social institutions may select individuals from the population who already possess certain properties. For example, the competitiveness of Olympic gymnastics may select people with short, thin, muscular builds. This type of causal process straddles the institutional and the essential: There is no Olympic gymnast essence, and individuals only become Olympic gymnasts in virtue playing a role in an institution, but we might want to say that the institution selects for a naturally-occurring kind of person, such that the correlations between properties are largely sustained by natural processes. Alternatively, social institutions may operate through entirely social processes: Individuals acquire properties in virtue of their position within the institution (social constitution), and social institutions can also shape people in deep and lasting ways by systematically altering their experiences (social production). To disambiguate these sources, we told participants a property generalized to members of an occupation. We then asked them to predict whether various non-prototypical individuals would possess the property: e.g., a person who was randomly assigned to be a lawyer or a retired quarterback.

In short, across these seven studies we investigate the possibility that people take occupational roles to have robust inductive potential because of their position in stable social institutions. In assessing this possibility, we can determine whether our inductive practices are guided by a pluralistic set of sources (essences in some cases, institutions in others) or whether current theories based in essences adequately predict the boundaries of inductive potential.

4. Study 1

We examined the inductive potential of occupational roles using the added-value paradigm and compared their added value with

religion and region-of-origin categories. We considered 6 types of properties: cultural activities, deontic properties, functional behaviors, personality traits, physiological traits, and skills. Though their type was specified, they were novel: e.g., "She conducts tingforms." Given that inductive potential involves more than just already knowing some information about that category, we needed to appeal to novel properties. The added-value paradigm was designed to specifically address inductive potential criteria (1)-(3), namely whether occupational roles support extending novel (1) and unobservable (2) properties between people who are visually dissimilar (3). We thought they would support these inferences. We expected them to strongly support deontic, functional, and skill-based inferences, followed by personality, and to weakly or negligibly support physiological and cultural inferences.

4.1. Method

4.1.1. Participants

We recruited 150 participants based on a power analysis described in our pre-registration (https://osf.io/hxzem). To ensure data quality, we had an initial, brief comprehension check before the consent form, which helps reduce the number of automated (non-human) respondents that can access the study. We also had two text entry questions at the end of the survey. We use these text entries only to exclude automated respondents. After exclusions, we had 118 participants. We did not record participant demographic information. Because mTurk participants have reliable demographic characteristics that researchers regularly assess (Berinsky, Huber, & Lenz, 2012; Burnham, Le, & Piedmont, 2018; Huff & Tingley, 2015), and because we did not plan any demographic-based analyses, we decided to reduce the burden placed on participants by removing the additional work of filling out demographic surveys. Research reliably finds that mTurk participants are close to representative on race and gender but skew younger, liberal, and non-religious. Research reliably finds that mTurk participants are close to representative on race and gender but skew younger, liberal, and non-religious. This study and all subsequent studies were conducted under Protocol Number 1311013027 approved by Yale's IRB Board.

4.1.2. Design & procedure

Participants were told they would make judgments about others. They were instructed to treat any novel words as expressing unfamiliar true information. Participants were randomly assigned between-subject to reason about occupation, religion, or region of origin. Participants reasoned about one category at a time in randomized order. For each category, they received all 6 property types in randomized order, but they reasoned about properties of only one type at a time (e.g., just deontic properties) (see Fig. 1). For each property type, participants made 3 inductions for a total of 18 inductions. For each category, participants learned about three individuals: The target individual was a member of the category and the individual whose properties participants learned about. The other two individuals were the test individuals, who differed in race and gender from the target individual, and were the same race, gender, and age as each other. Participants answered yes/no for each property for each test individual; thus, it was possible for participants to extend the property to neither test individual, to one, or both.

4.1.3. Stimuli

The five occupational roles were lawyer, mayor, judge, quarterback, and police officer. The five religions were Mormon, Catholic, Jew, Protestant, and Muslim. The five regions of origin were: resident of the Southwest, Midwest, Northeast, Southeast, and West (specified as regions of the United States).

To examine how inductive potential might vary across types of properties, properties were divided into six types. Cultural properties were described as "[his/her] favorite activities," and were always about music, dancing, and games: "She listens to grynn music," "she dances the scubblebet," "enjoys playing breeween." Deontic properties were described as "[his/her] rights and duties" and included properties like "she is permitted to tuate," "she is obligated to gaspiralyze," and "she has a duty to imperfy." Functional behaviors were described as "ways [he/she] helps others." They included properties like "she helps others apprefix," "she conducts tingforms," and "she works to complete rousurs." Personality traits were described as "[his/her] personality and character," and included properties like "she is sificatible," "she is efulen," and "she is izzilent." Physiological properties were described as "[his/her] body," and included properties like "she has digisol in her blood," "she has internara inside her body," and "she is prone to ociram." Skills were described as [his/her] skills and abilities," and included properties like "is skilled at draptionism," "has the ability to pebby," and "is really knowledgeable about mulbering."

We used 5 face triads: A White male face (target) and two East Asian female faces (test); a Black female face (target) and two White male faces (test); an East Asian male face and two Black female faces; a Black male face and two White female faces; an East Asian female face a and two Black male faces. With 3 faces each, the 5 triads involved 15 unique faces. Faces always had the same role (target versus test) in their respective triad. Faces were from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015).

4.2. Results & discussion

4.2.1. Pre-registered analyses

We pre-registered effect-size interpretations. We determined that percentage point differences of +10 were negligible, +20 were weak, +30 were moderate, and +40 were strong. We included these effect sizes to make our predictions quantitatively robust and clarify a priori the effect sizes we anticipate based on our theorizing. Briefly, we considered 10% negligible based on the assumption that mTurk participants tend to respond at a rate of 10% on dichotomous questions that should elicit 0% responding. We derived 40% from a hypothetical 70% (same-category target)/30% (different-category target) response pattern, which is the type of response pattern we a priori anticipate for strong inferences based on our experience with these types of measures.In all reported results, we

used generalized estimation equations (GEE) to account for within-subject data. In all reported results, we used generalized estimation equations (GEE) to account for within-subject data. Specifically, we used the 'gee' function, with default settings (Gaussian, independent correlation structure), from the 'gee' package in R. P-values were derived from the robust Z-scores provided by the function's robust sandwich estimator. For all studies, data files, R code files, and pre-registration files can be accessed at osf.io/ec7q3/.

On average, participants were 41 percentage points more likely to extend a property when individuals belonged to the same occupation than when they did not, $b = 0.41$, $SE = 0.03$, $p < .001$. Participants were 35 percentage points more likely with religious categories, $b = 0.35$, $SE = 0.03$, $p < .001$, and 27 percentage points more likely with region-of-origin categories, $b = 0.27$, $SE = 0.03$, $p < .001$. Occupational roles added significantly more value than religion, though this was difference was small: $b = -0.06$, $SE = 0.03$, $p = .031$. Occupational roles added significantly more value than region-of-origin categories, $b = -0.14$, $SE = 0.03$, $p < .001$. Next, we looked at each property (see Fig. 2; see /osf.io/j6m9k/ and osf.io/tesvy/ for supplemental figures for religion and region of origin, respectively).

4.2.1.1. Cultural activities. Occupational roles (+17 percentage points) added less value than both religion (+36 points), $b = 0.18$, $SE = 0.05$, $p < .001$, and region of origin (+29 points), $b = 0.11$, $SE = 0.04$, $p = .005$. Consistent with our expectations, occupational roles weakly supported inferences about cultural properties, and did so less well than religion and region of origin. The cultural activities at question here, such as styles of dancing, likely map onto social groups more than roles within groups. Therefore, occupational roles may be distinctive in *not* supporting inferences about these types of cultural activities.

4.2.1.2. Deontic. Occupational roles (+61 percentage points) added more value than both religion (+48 points), $b = -0.13$, $SE = 0.05$, $p = .005$, and region of origin (+32 points), $b = -0.29$, $SE = 0.04$, $p < .001$. This was consistent with our expectation that occupational roles would strongly support inferences about deontic properties. This estimate actually outpaces our expectations in magnitude and in being higher than religion. Assigning and altering deontic properties may be one of the primary functions of social institutions (Searle, 1995; 2010) and the roles within them (Kalish & Lawson, 2008).

4.2.1.3. Functional. Consistent with predictions, occupational roles (+50 points) added similar value as religion (+45 points), $b = -0.05$, $SE = 0.05$, $p = .222$, and more value than region of origin (+27 points), $b = -0.24$, $SE = 0.05$, $p < .001$.The basic purpose of occupational roles is to divide labor and so it follows that members of occupational roles would engage in particular functional behaviors consistent with their duties.

4.2.1.4. Personality. Occupational roles (+43 percentage points) added more value than religion (+33 points), $b = -0.10$, $SE = 0.05$, p

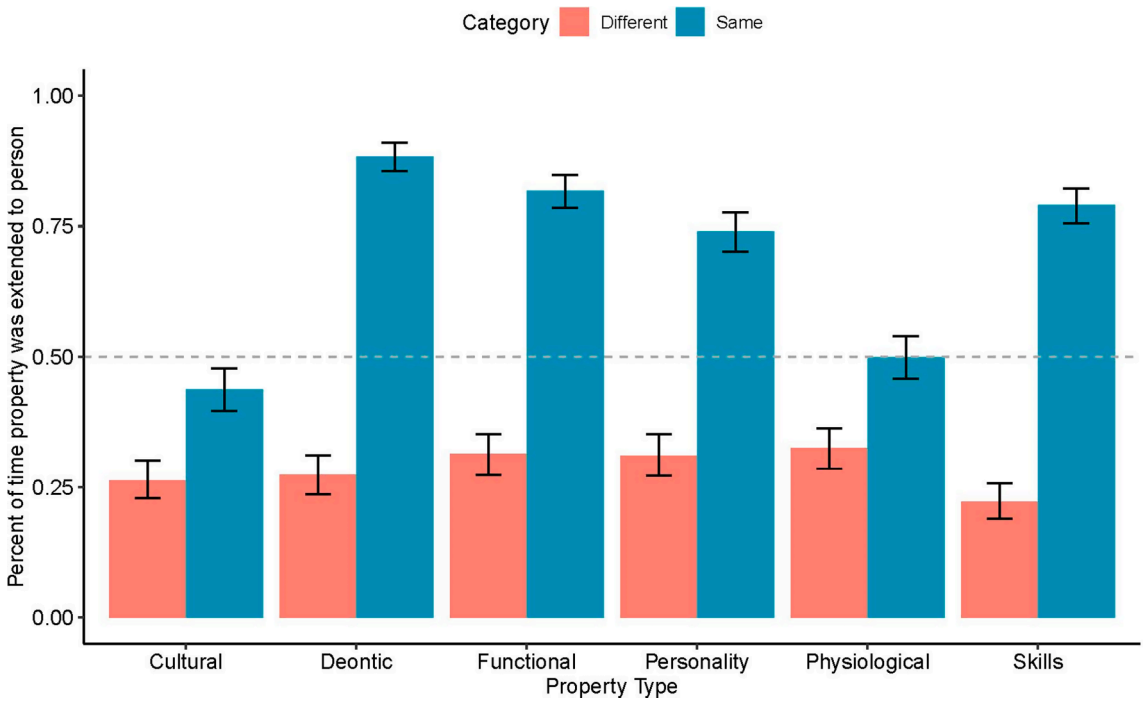

Fig. 2. Full results for occupational roles, Study 1. Participants learned about an individual's properties distributed across 6 types enumerated on the X-axis. Participants then indicated whether a new individual with the same (blue) or different (red) occupational role also shared those properties. Error bars are 95% bootstrapped confidence intervals.

= .038, and region of origin (+26 points), b = -0.17, SE = 0.05, p < .001. This result was surprising. We expected that occupational roles would weakly (but not negligibly) support inferences about personality traits. This was motivated by prior findings that people can be essentialist about personality in and of itself (Gelman, Heyman, & Legare, 2007; Haslam, Bastian, & Bissett, 2004), suggesting they would not necessarily expect institutions to shape personality traits. Indeed, we thought that occupational roles might not do much better than region of origin. Nevertheless, occupational roles strongly supported inferences about personality traits, and exceeded even religion (though only slightly).

4.2.1.5. Physiological. Occupation (+17 points), religion (+16 points), and region of origin (+19 points) negligibly supported inferences about physiological properties, and they were not significantly different from each other (ps > 0.50). This is consistent with our prediction that occupational roles would negligibly support inferences about physiological properties. We also suspected there might be floor effects here, consistent with the lack of statistical differences.

4.2.1.6. Skills. Occupational roles (+57 percentage points) added far more value than religion (+32 points), b = 0.25, SE = 0.05, p < .001, and region of origin (+27 points), b = 0.30, SE = 0.04, p < .001. This is consistent with our predictions. We expected occupational roles to strongly support inferences about skills and distinctively so. A key social function of occupational roles is to distribute labor, so that individuals can specialize. The difference in added value was greatest for skills, so occupational roles may be distinctive in this regard.

4.2.2. Exploratory analyses

We used K-means clustering to explore how the different categories clustered (R function stats::kmeans). Average silhouette width suggested 2 clusters was the optimal number of factors (using the R function 'fviz_nbclust' from package 'factoextra'). The 5 occupational roles clustered together, and all the non-occupational roles clustered together. Repeating K-means clustering with 3 clusters revealed all categories clustered in their respective domains (occupations, religion, region of origin). Together, this provides some evidence that all the occupational roles we explored patterned similarly, without any obvious outliers.

4.2.3. Summary

Occupational roles support perceptually unguided inferences about novel properties, and they do so at a magnitude similar to at least one social category with robust inductive potential: religion (Diesendruck & HaLevi, 2006). As expected, occupational roles strongly promote inferences about rights and obligations, functional behaviors, and skills and abilities. Occupational roles supported inferences about personality traits to a surprisingly strong degree, suggesting that participants are open to the possibility that personality traits are shaped by social institutions (see Study 7).

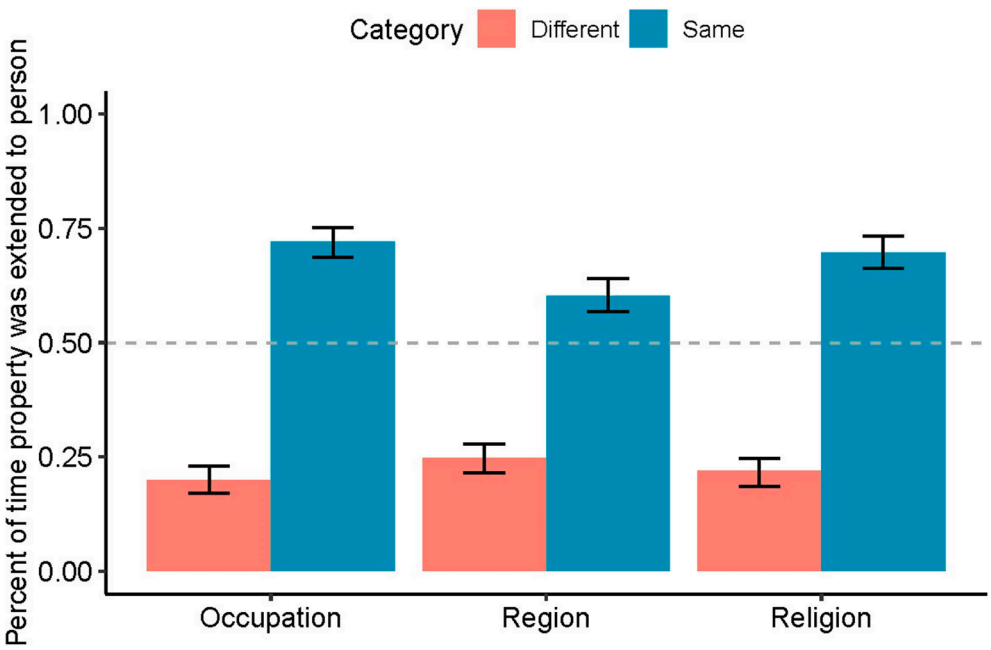

Fig. 3. Study 2: Comparison of occupation, region, and religion categories when 'blank' properties were used.

5. Study 2

The design is identical to Study 1 except now the type of property is not indicated: e.g., "She has property X." This allowed us to test the effect that property type information (e.g., skills, personality) had on our inductive potential estimate. We recruited 150 participants from mTurk. There were 137 participants after excluding automated responses.

5.1. Results & discussion

Participants were 52 percentage points more likely to extend a property when individuals belonged to the same occupation than when they did not, $b = 0.52$, $SE = 0.04$, $p < .001$. Participants were 48 percentage points more likely with religion, $b = 0.48$, $SE = 0.04$, $p < .001$, and 36 percentage points more likely with region of origin, $b = 0.36$, $SE = 0.03$, $p < .001$. Occupational roles were not significantly different from religions, $b = -0.04$, $SE = 0.04$, $p = .317$. Occupational roles added significantly more value than region-of-origin categories, $b = -0.17$, $SE = 0.04$, $p < .001$. See Fig. 3. Therefore, participants used occupational roles to make robust inductive inferences even when property information was completely unknown. Overall, the magnitude of inferences was similar to Study 1, suggesting that the inclusion of property information (and the specific set of properties chosen) was not greatly affecting the overall magnitude estimates.

6. Study 3

The design is identical to Study 1 except we removed face information. The diversity of faces in Study 1 could have altered inductive potential estimates. People may represent social categories as bundles: e.g., White Anglo-Saxon Protestants and Arab Muslims, and so depicting people that differ in race and gender could undermine the inductive potential of social categories, making them look inductively weaker than they may be in practice. We recruited 150 participants; there were 123 participants after exclusions.

6.1. Results & discussion

We found a qualitatively similar pattern of results as in Study 1. Occupational roles (+40) supported similar levels of inductive inference as religions (+37), $b = -0.03$, $SE = 0.03$, $p = 0.381$. Occupational roles supported greater inductive potential than regions of origin (+30), $b = -0.10$, $SE = 0.03$, $p = .001$. We also found a qualitatively similar pattern for the 6 properties (see Fig. 4; see https://osf.io/ufrjy/ and osf.io/z2tsq/ for supplemental figures for religion and region of origin, respectively).

Compared to religion, occupational roles supported inferences about cultural properties significantly less ($b = 0.23$, $SE = 0.04$, $p < .001$), deontic properties significantly more ($b = -0.10$, $SE = 0.05$, $p = .03$), functional behaviors non-significantly more ($b = 0.03$, $SE = 0.05$, $p = .560$), personality non-significantly more ($b = -0.07$, $SE = 0.04$, $p = .178$), physiological properties non-significantly less ($b = 0.01$, $SE = 0.04$, $p = .862$), and skills significantly more ($b = -0.21$, $SE = 0.05$, $p < .001$).

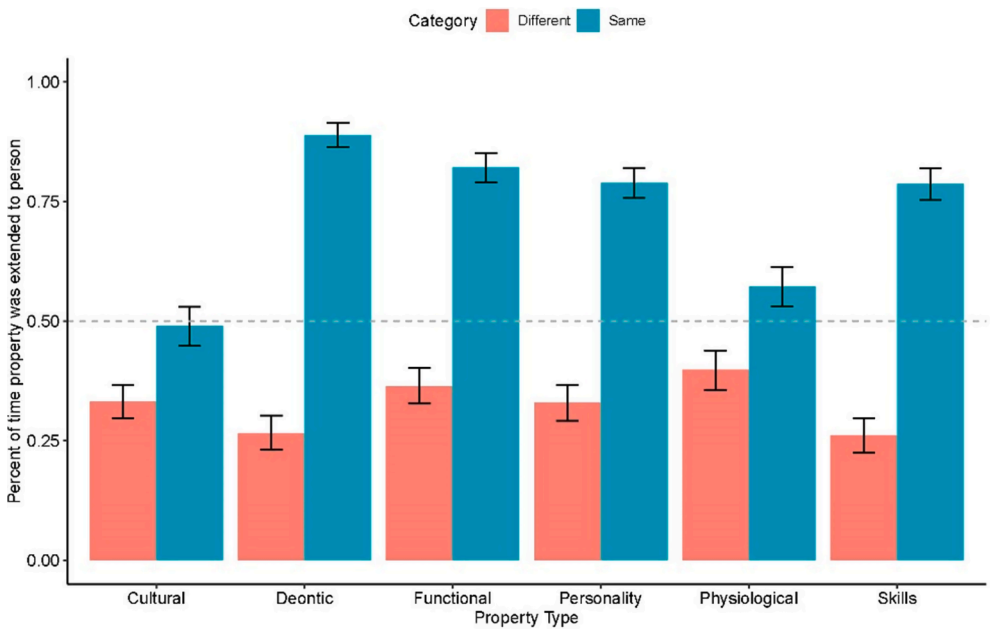

Fig. 4. Study 3: People's inferences about occupational roles when there was no face information.

Compared to regions of origin, occupational roles supported inferences about cultural properties significantly less ($b = 0.19$, $SE = 0.04$, $p < .001$), deontic properties significantly more ($b = -0.28$, $SE = 0.05$, $p < .001$), functional behaviors significantly more ($b = -0.19$, $SE = 0.04$, $p < .001$), personality significantly more ($b = -0.17$, $SE = 0.04$, $p < .001$), physiological properties non-significantly less ($b = 0.04$, $SE = 0.04$, $p = .256$), and skills significantly more ($b = -0.19$, $SE = 0.05$, $p < .001$). Thus, it does not appear that the visual cues presented in Study 1, which provided information about race and gender, had a major effect on reported results.

6.1.1. Exploratory analyses

We repeated the cluster analysis described in Study 1 and again found that 2 clusters were optimal, and that the clusters were occupational roles and non-occupational roles; as before, extracting 3 clusters clearly returned the 3 expected domains.

7. Study 4

Next, we compared occupational roles and animal species. Animal categories have reliably high inductive potential and provide a standard of robust inductive potential (Gelman, 2003). We wanted to see how occupational categories compared. To make the animal triads, we took two animals that looked similar but belonged to different taxonomic groups: e.g., mice and pika resemble each other, but mice are rodents and pika are lagomorphs. Then, we selected a member of the same taxonomic group as one of the animals: e.g., capybara and mice look different but both are rodents (see Fig. 5).

We excluded deontic and cultural properties, which are plausibly unique to social categories. Instead, we focused on functional behaviors, physiological properties, "personality" traits (referred to here, simply as "traits and characteristics"), and skills. We slightly tweaked the language to make them animal-appropriate: To introduce functional behaviors, we said "Now let's learn about what this one does" instead of "Now let's learn about ways she helps others." To introduce traits, we said "Not let's learn about this one's traits and characteristics" instead of "Now let's learn about her personality and character." Otherwise, the design was the same as Study 1 (minus religion and region of origin). We recruited 170 participants and had 167 after exclusions.

By excluding properties that were unlikely to apply to animals, we designed the study to maximize the inductive potential of animal categories, so that we could have a robust standard against which to measure the inductive potential of occupational roles. But we can only compare the inductive potential of animal categories and occupational roles on a property-by-property basis. That is, it makes little sense to collapse across property type and compare their 'total' inductive potential. Animals presumably strongly guide inductive inferences about physiological properties, whereas we already know (and would not expect) occupational roles to support these inferences. In contrast, occupational roles support inferences strongly about deontic properties but we excluded those because we did not expect them to be at all relevant to animal categories. Of these four properties, we expected functional behaviors and skills to have the greatest possibility of animal-level inductive potential. We expected animals to support physiological properties dramatically more than occupational roles, reflecting their different ontological status as natural and social kinds, respectively.

7.1. Results & discussion

See Fig. 6. Occupational roles (+0.76) supported function-based inferences significantly more than animal categories (+0.57), $b = 0.19$, $SE = 0.05$, $p < .001$. Occupational roles (+0.53) supported personality-based inferences significantly less than animal categories (0.70), $b = -0.16$, $SE = 0.05$, $p = .002$. As expected, and consistent with their different causal structures, occupational roles (+0.15) supported physiological properties dramatically less than animal categories (+0.66), $b = -0.52$, $SE = 0.04$, $p < .001$. Finally, occupational roles (+0.69) supported skill-based inferences significantly more than animal categories (+0.58), $b = 0.11$, $SE = 0.05$, $p = .037$ – though this difference was barely significant.

Animal categories provide the benchmark for robust inductive potential (Gelman, 1988; Gelman, 2003; Brandone & Gelman, 2009). Occupational roles only had substantially less inductive potential on the one property that we would not expect occupational

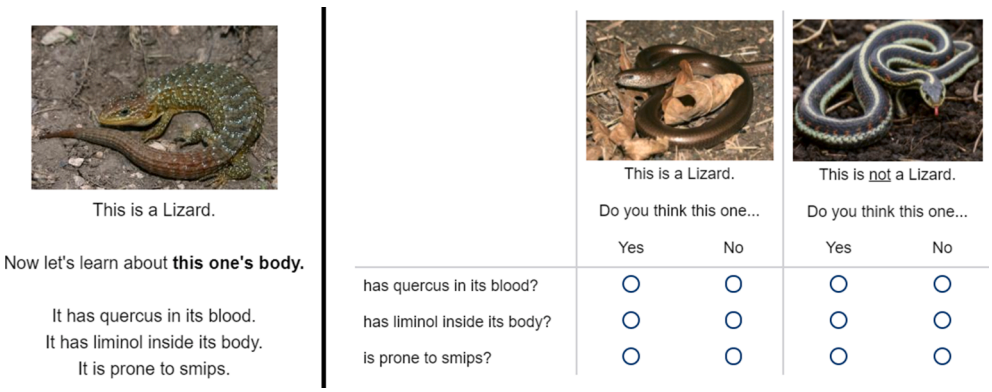

Fig. 5. In a between-subject design, half of participants responded to animals (depicted above) and half of participants responded to social roles (as in Study 1).

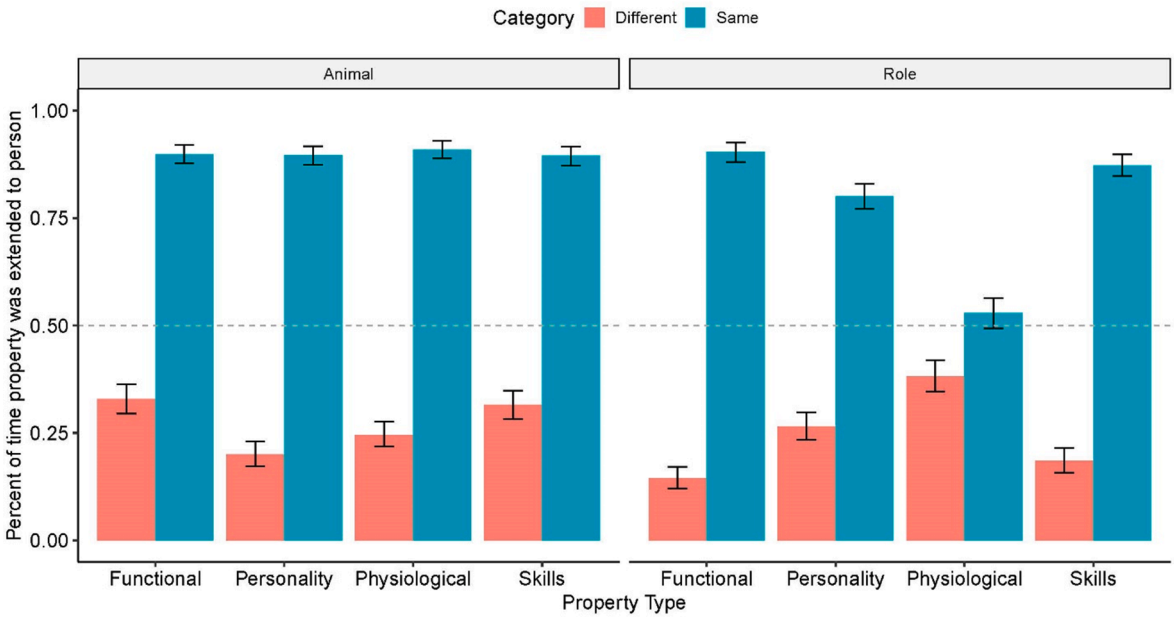

Fig. 6. Results for Study 4, which compares occupational and animal categories on four properties: functional behaviors, 'personality' traits, physiological properties, and skills.

roles to promote: physiological properties. And, again, we set up the study aiming to both replicate the strong inductive potential of animal categories, while innovating on the design by using distinct property types, so we did not include property types that we expected animal categories to fail to support (e.g., deontic properties). Of the three property types that are appropriate for both domains—function, personality, and skill—the two domains come out as on par in terms of inductive potential, with no significant difference, $b = 0.05$, $SE = 0.05$, $p = .321$. Occupational roles were moderately higher on function-based inferences, moderately lower on 'personality,' and comparable on skill-based inferences. Occupational roles apparently have the same magnitude of inductive potential as animal categories.

One source of doubt is that the 'added value' paradigm does not provide participants another basis for drawing inferences. In the absence of any other information, occupational roles may be sufficient to make inferences, but participants may disregard occupational roles as soon they have an alternative. For this reason, we turned to an appearance-foil paradigm (Gelman, 1988; Gelman & Markman, 1986, 1987).

8. Study 5

In a new 'appearance foil' paradigm, we set up a conflict between generalizing on the basis of occupations and on the basis of visual

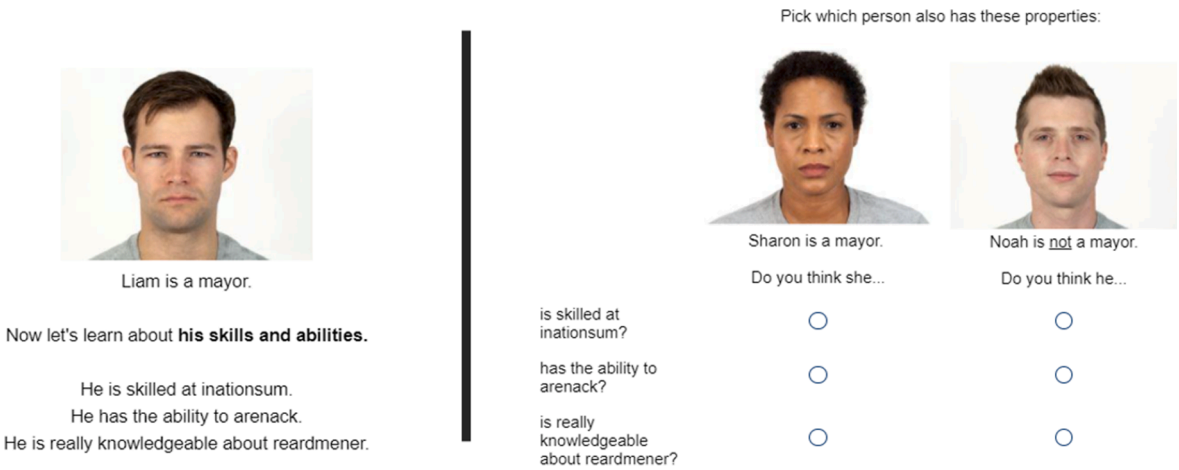

Fig. 7. The appearance-foil method used in Study 5 and 6.

similarity (including race and gender-related face cues). This type of paradigm provides a robust test of inductive potential because participants are not merely using the category as a basis of induction but favoring it over salient, competing cues. We focused on both race and gender because they provide the benchmark for robust inductive potential among social categories (Gelman, 2003; Rothbart & Taylor, 1992). As the task makes them visually salient, and race and gender were presented simultaneously, this provides a rigorous test of the inductive potential of occupational roles. This addresses criterion (4): If a category has robust inductive potential, participants should favor the category as a basis for inference over apparent similarity.

8.1. Method

8.1.1. Participants

We recruited 125 participants from Amazon Mechanical Turk. We had 118 participants after exclusions.

8.1.2. Design & procedure

Participants reasoned about 2 of the 5 occupational roles in randomized order. For each occupation, they reasoned about all 6 types of properties. Participants learned about a target individual's occupational category and his/her properties. Then, participants saw two test individuals: One had the same race, gender, and age as the target individual but was not the same occupation, and the other was a different race and gender (and sometimes also different age) but had the same occupation. Participants then selected (forced choice) which individual they thought also had the same properties (Fig. 7).

8.2. Results & discussion

To analyze the results, we compared the proportion of times participants selected occupation instead of race and gender to make inferences by property type. By comparing this proportion to 0.50, we can examine whether participants favored occupational roles, race/gender, or neither. Participants significantly favored occupational roles for making inferences about deontic properties, $b = 0.34$, $SE = 0.02$, $p < .001$, functional behaviors, $b = 0.34$, $SE = 0.02$, $p < .001$, personality traits, $b = 0.15$, $SE = 0.03$, $p < .001$, and skills, $b = 0.31$, $SE = 0.02$, $p < .001$ (Fig. 8). Participants significantly favored gender/race for making inferences about cultural activities, $b = -0.14$, $SE = 0.03$, $p < .001$, and physiological properties, $b = -0.28$, $SE = 0.03$, $p < .001$.

Overall, this pattern of results is consistent with the rank-order of effect sizes observed in Study 1 and 3: Occupational roles strongly supported inferences about deontic, function, personality, and skill-based properties and negligibly supported inferences about cultural activities and physiological properties. People frequently use gender and race to make inferences about personality and ability (Prentice & Miller, 2007), making participants' use of occupational roles over gender and race combined especially striking. However, the task did pragmatically emphasize occupational roles by labelling them. But given that we also emphasized race and gender by organizing all triads around race and gender and using face photographs that isolated facial cues (instead of other cues like clothing) and because the verbal labels were provided in plain text, whereas race and gender were provided in salient photographs, this seems an unlikely limitation. Indeed, participants did readily rely on race and gender for properties that occupational roles are not strong indicators of (i.e., cultural activities and physiological properties). Nevertheless, this limitation can be further reduced by contrasting the inductive potential of occupational roles with animal categories in this paradigm, where both would receive a similar treatment.

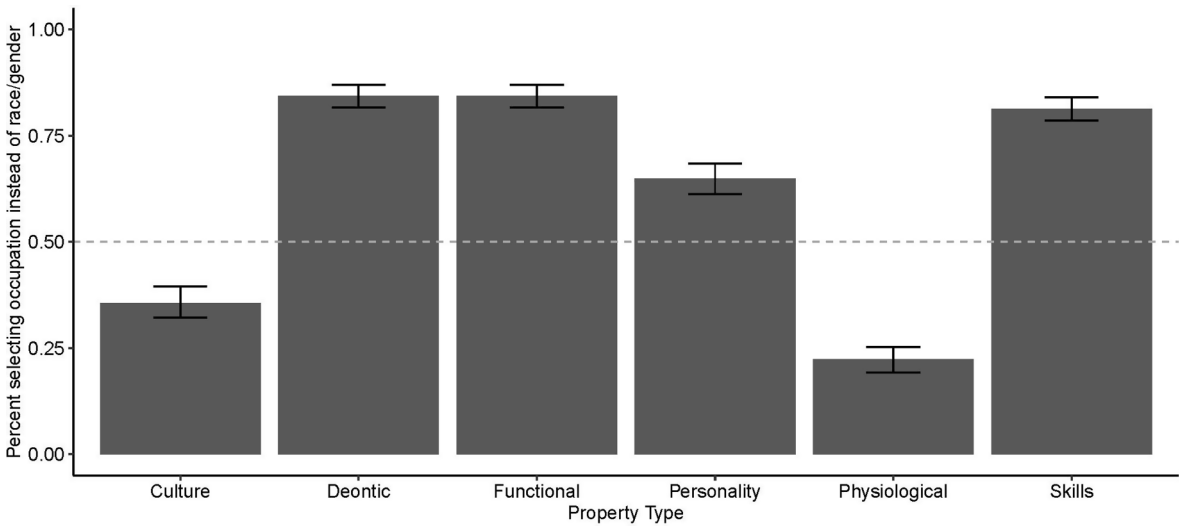

Fig. 8. Study 5: Proportion of times participants selected occupation rather than race and gender for making inferences.

9. Study 6

We conducted another appearance foil paradigm with occupational roles, and we added an additional between-subject condition: We converted Study 4 into the appearance foil method for animal categories. That is, we used the same animal triads but switched the target animal and the same-category test animal, such that now the target animal resembled a different-category test animal but looked dissimilar to a same-category test animal. And, the task was now forced choice. The only other change from Study 5 is that we excluded deontic and cultural properties given that these were, as noted above, inappropriate in the animal domain.

9.1. Results & discussion

As with Study 5, participants favored occupational roles over race/gender for making inferences about functional behaviors, $b = 0.39$, $SE = 0.02$, $p < .001$, personality, $b = 0.19$, $SE = 0.03$, $p < .001$, and skills, $b = 0.36$, $SE = 0.02$, $p < .001$; participants favored gender/race for making inferences about physiological properties, $b = -0.34$, $SE = 0.03$, $p < .001$. Animal categories were favored over visual similarity for all property types ($ps < 0.001$; see Fig. 9). The magnitude of this preference was similar in the case of occupational roles and animals for functional behaviors, $b = 0.06$, $SE = 0.03$, $p = .090$, and skills, $b = 0.04$, $SE = 0.03$, $p = .269$, less for personality traits, $b = -0.15$, $SE = 0.04$, $p = .001$, and dramatically less for physiological properties, $b = -0.71$, $SE = 0.04$, $p < .001$. Therefore, occupational roles were preferred over visual similarity just as strongly as animals for two of the properties, even though occupational roles were contrasted with a far more potent foil (gender/race).

10. Study 7

Studies 1–6 demonstrated robust inductive potential in occupational roles: Occupational roles support inferences about novel properties that are not guided by perceptual similarity, occupational roles support inferences about several types of properties at a greater magnitude than race and gender, and to a similar extent as animal categories. Their inductive potential was strongest for four properties: deontic properties, functional behaviors, personality traits, and skills. What is the source of this inductive potential?

Social institutions involve a huge variety of interacting forces and dependence relations—some of these are causal, others are constitutive or normative. These processes seem to unfold across three timescales: First, there are pre-membership processes, which include the steps individuals take to attempt to gain membership in an occupational category (e.g., taking pre-law classes in undergrad to become a lawyer) and the filtration process by which only certain individuals can eventually enter the role. Filtration processes create a possibility for quasi-essentialist correlations to emerge: e.g., to become an Olympic athlete, one may require an innate build and metabolic profile. Second, there are properties 'built into' or constitutive of roles. That is, there are properties members have in virtue of their position within a social institution. It is not that being a lawyer causes one to have an obligation to loyally respect the interests of one's clients. Rather, that is part of what it *is* to be a lawyer. We expected deontic properties to be represented as a constitutive component of occupational role membership. Functional behaviors may also partially constitute what occupational roles are; each involves performing a role. Third, there are causal processes that occur over prolonged periods of membership in occupational roles: Occupying a role over time can shape one's experiences, which can cause more lasting changes that persist after membership ends. These processes shape skills and could even shape personality traits. In sum, we expected participants to indicate that properties generalize to members of social roles primarily because of social constitution and social production and not because of

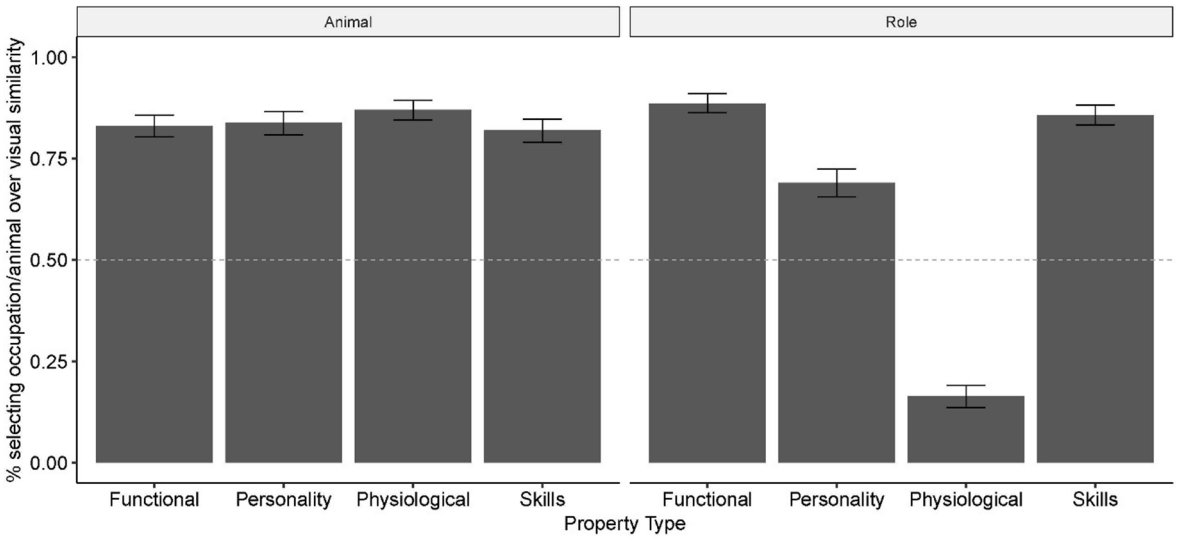

Fig. 9. Study 6: Proportion of times participants selected occupation/animal category rather than visual similarity for making inferences.

quasi-essentialist reasons like pre-existing dispositions.

10.1. Method

10.1.1. Participants

We recruited 150 participants from Amazon Mechanical Turk. 122 remained after exclusions.

10.1.2. Design & procedure

We included the 5 occupations used in prior studies: lawyer, judge, quarterback, police officer, mayor. Participants read 8 vignettes in randomized order. These vignettes represented the full factorial of 3 variables: Selection processes that happen prior to membership, membership itself, and length of membership. These three factors varied within subject. To vary selection processes, we varied whether an individual became a member intentionally (i.e., through actively and intentionally pursuing the career and being hired in a way that fit with employer intentions) or randomly. To vary membership, we varied whether an individual had officially started or had not yet started, or whether an individual had officially retired or had not yet retired after holding the occupation for a prolonged period of time. To vary length of membership, we varied whether an individual had been on the job for 30 years or whether they had not yet started or only recently started. Here is an example of the vignette where selection was intentional, the individual was currently a member, and they had prolonged exposure:

> "In Hacking City, people pursue the job they want and companies hire who they want. 30 years ago, Alex was hired as a lawyer. Alex has been a lawyer for 30 years. Alex is still a lawyer."

Here is an example where selection was random, the individual had not officially started, and so they did not have prolonged exposure:

> "In Haslanger City, people are randomly assigned a job. Hayden was randomly assigned lawyer. Hayden will officially be a lawyer January 1st, 2021. Hayden is not a lawyer yet."

Participants then learned a new property of the category. We focused on deontic, function, personality, and skill-based properties, as participants used occupational roles to make robust inferences about these properties in the preceding studies. The facts were taken from previous studies: "Here is a fact about lawyers' ways of helping others: lawyers help others apprefix." Participants were then asked "Do you think Hayden helps others apprefix?" and they responded on a 1–6 "definitely no" to "definitely yes" scale.

10.2. Results and discussion

First, we examined a model collapsing across property type, with selection processes (i.e., intentional or random), membership (member or not member), and length of membership (30 years in role or 0–1 days in role) entered as fixed effects without interactions. Being a member of the role currently, $b = 0.97$, $SE = 0.08$, $p < .001$, and being in the role for 30 years, $b = 1.02$, $SE = 0.08$, had large, significant effects; being intentionally selected had a significant but relatively smaller effect: $b = 0.37$, $SE = 0.08$, $p < .001$. With the interactions added, only one higher-order effect was significant: The interaction between intentional selection and length of membership: $b = -0.28$, $SE = 0.10$, $p = .007$. Specifically, being intentionally selected had a smaller effect when a person had been in the role for 30 years, $b = 0.23$, $SE = 0.04$, $p < .001$, than when they had not been, $p = .50$, $SE = 0.07$, $p < .001$. As described below, this occurred because intentional selection and length of membership affected the same kind of properties, and so their marginal effect was diminished when the other was already present. Membership affected different properties than the other two, so these interactions (and the three-way interaction) were not significant ($ps > 0.10$). Next, we examined each of the four property types individually:

10.2.1. Deontic properties

By far the largest effect was membership, $b = 1.87$, $SE = 0.16$, $p < .001$. Individuals had the typical rights and obligations of the role when they were currently in the role and not before they started or after they retired. Being in the role for 30 years had a small but significant effect: $b = 0.57$, $SE = 0.09$, $p < .001$. Intentional selection had a negligible effect: $b = 0.11$, $SE = 0.05$, $p = .039$. In a model with interactions, none of the interactions were significant ($ps > 0.4$). As anticipated, deontic properties are strongly tied to membership: By occupying a position within a social institution, we are assigned certain rights, duties, and obligations. However, people did think length of membership was relevant. Participants may have had in mind various entitlements like pensions that individuals earn after being in a role.

Next, we examined the three conditions where only one of the factors were present (Figure 11): (*Intentional selection only*) A person was intentionally selected for the role but because they had not started yet they were neither currently a member nor did they have prolonged experience in the role; (*Membership only*) A person was a current a member of the role but they were randomly selected and just started; (*30 years of experience only*) A person has worked in the role for 30 years but they were randomly selected in the beginning and just retired. Being intentionally selected was not sufficient to confer deontic properties if an individual was not yet in the role: Participants rated this condition below the midpoint, $b = -0.46$, $SE = 0.14$, $p = .001$. Participants had middling intuitions about a person that had retired after 30 years of experience, as they were not significantly different from the midpoint: $b = 0.004$, $SE = 0.16$, $p = 0.959$. Being in the role, even on the first day after random assignment, was sufficient to confer deontic properties: Participants were significantly above the midpoint, $b = 1.32$, $SE = 0.12$, $p < .001$.

10.2.2. Functional behaviors

Functional properties were most influenced by membership, $b = 1.13$, $SE = 0.13$, $p < .001$, followed by length of membership, $b = 0.87$, $SE = 0.09$, $p < .001$, followed by intentional selection, $b = 0.32$, $SE = 0.05$, $p < .001$. The interactions were not significant ($ps > 0.4$). Being in the role, even on the first day after random assignment, was sufficient for functional behaviors: $b = 0.80$, $SE = 0.11$, $p < .001$. Thirty years of experience in the role, even after leaving it, was sufficient for functional behaviors too: $b = 0.55$, $SE = 0.14$, $p < .001$. Being intentionally selected for the role without starting was not sufficient, $b = -0.03$, $SE = 0.13$, $p = 0.812$. So, the source of functional behaviors is mixed: On the one hand, an individual takes on certain functional behaviors merely by entering the role. On the other hand, the capacity to engage in certain functional behaviors may be restricted by experience: For example, it may take years being a lawyer to master oral speaking, and a lawyer would carry this property with them after leaving. Therefore, functional behaviors may be derived both from deontic properties and skills, which explains why they visually look intermediate between them (see Fig. 10).

10.2.3. Personality

Personality traits were most strongly affected by length of membership, $b = 1.05$, $SE = 0.09$, $p < .001$, followed by intentional selection, $b = 0.49$, $SE = 0.06$, $p < .001$, and mere membership, $b = 0.45$, $SE = 0.07$, $p < .001$. One of the interactions was significant: Length of membership interacted with intentional selection, $b = -0.53$, $SE = 0.15$, $p < .001$. Being intentionally selected had a smaller effect when a person had been in the role for 30 years, $b = 0.24$, $SE = 0.06$, $p < .001$, than when they had not been, $b = 0.75$, $SE = 0.09$, $p < .001$. Participants may reason that being intentionally selected for the role implies experience prior to the role, such as attending law school, which is made up for by having a lot of experience in the role. Or, participants may infer that although people intentionally selected for roles are more likely to have typical traits of members of the role, those that do not will gradually acquire them anyways. Of the three solo conditions, thirty years of experience in the role had the largest effect by itself, $b = 1.33$, $SE = 0.09$, $p < .001$. Being intentionally selected for the role, $b = 0.59$, $SE = 0.12$, $p < .001$, and currently occupying the role, $b = 0.41$, $SE = 0.11$, $p < .001$, were also significantly above the midpoint. Therefore, participants understood personality traits as resulting from multiple sources. Nevertheless, having had years of experience in the role was by far the most important source of personality traits.

10.2.4. Skills

People reasoned about skills similarly to how they reasoned about personality traits. Length of membership had the largest effect, $b = 1.57$, $SE = 0.11$, $p < .001$, but intentional selection, $b = 0.54$, $SE = 0.07$, $p < .001$, and membership, $b = 0.41$, $SE = 0.06$, $p < .001$, also had significant, moderate effects. As with personality traits, intentional selection interacted with length of membership (the only significant interaction): $b = -0.41$, $SE = 0.16$, $p = .009$. Participants reasoned that being in the role for 30 years was sufficient to have the typical skills of the role: $b = 1.36$, SE = 0.11, p < .001. Being intentionally selected for the role was also sufficient to have the typical skills (though this effect was markedly smaller): $b = 0.30$, $SE = 0.13$, $p = .023$. Finally, being a member of the role was by itself not sufficient, b = 0.09, SE = 0.12, p = .425.

Taken together, membership and length of membership were the most important factors. Deontic properties are tied to being a member of the role: One acquires rights, obligations, and duties when one enters the role, and loses them when one leaves. Therefore, these properties are *socially constituted*, as they are not caused by exposure to the role over time but are constituted by being in it.

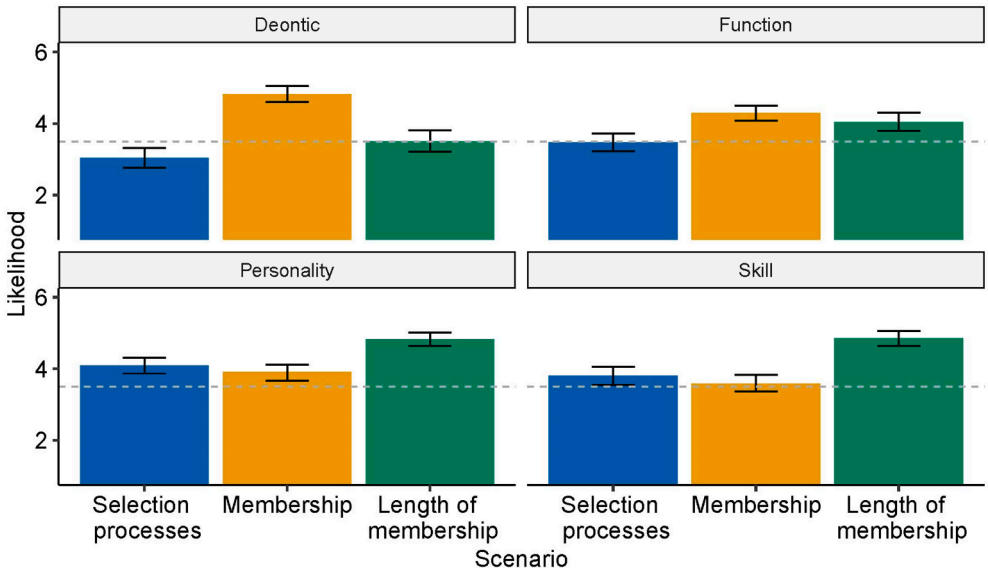

Fig. 10. Results for Study 8. This figure depicts 3 of the 8 total vignettes: The three singleton vignettes that isolated the effects of intentional selection, membership, and length of membership.

Personality traits and skills are accumulated over time when one has years of experience in the role and persist after one has left the role. Therefore, social institutions can exert causal forces that affect people in deeper and more lasting ways. Being intentionally selected into a role, with all that entails, has a much smaller role in people's inferences. People rarely believed members of occupational roles share properties because there are 'natural lawyers' or people who are 'born mayors' who already had those properties before the entered the role. If they had, then an individual being intentionally or randomly selected for the role would have been critical. Even when participants did rely on intentional selection, participants may have considered exposure to the institution prior to the role such as attending law school or playing football in high school. That would explain why intentional selection had similar, but smaller, effects as length of membership. Presumably there are roles where people place greater importance on natural ability (e.g., Olympic gymnasts and swimmers); but assumptions about natural ability are not important for all occupational roles and not necessary for inductive potential. We think that both intentional selection and length of membership tap into *social production*, which encompasses the way institutions induce new property correlations by systematically altering people's experiences.

11. General discussion

In combination, our studies demonstrate that occupational roles possess robust inductive potential. Upon learning that one person with an occupation had a property, participants were willing to generalize that property to another individual with that occupation. Participants used occupational roles to draw inferences about diverse property types; participants drew particularly robust inferences about deontic properties, functional behaviors, personality traits, and skills. This occurred even if the specific property was novel and unobservable, even if membership was not visually apparent, and even if there were other bases for inferences (i.e., visual similarity related to race and gender). Indeed, the inductive potential of occupational roles rivalled animal categories and surpassed race and gender. Therefore, occupational roles did not merely surpass some minimum standard; they supported inductive inference as well as paradigmatic cases of robust inductive potential. This makes it necessary to include occupational roles in any theory of ordinary cognition. Broadening the purview of our theories, by encompassing a greater range of categories with robust inductive potential, helps to reveal the foundations of our inferential practices.

Our results indicate that deontic properties, functional behaviors, personality traits, and skills can be divided into two classes of properties based on their presumed source. Deontic properties and functional behaviors are 'external' to individuals and internal to roles. When roles are created, part of what that involves is putting in place functions and norms (e.g., rights and obligations) that people have in virtue of playing the role (Ásta, 2018; Haslanger, 1995; Ritchie, 2020; Searle, 1995; 2010; Thomasson, 2019). These are put in place via various social forces—some of them intentional, some unintentional, some related to explicit rules (e.g., like those outlined in an employee manual), some more implicit. Personality traits and skills are properties 'internal' to individuals. Playing a role or even setting out to play a role can shape people in these ways. So, while internal, we take these properties to also be, at least in part, socially produced. An institution can shape more lasting features of individuals through prolonged shaping of their experiences. Social institutions can create new roles and change people as they set out to and then actually do play these roles. Thus, social institutions provide a source of inductive potential that is fundamentally extrinsic rather than inherent.

11.1. Domain-differences in inductive potential

Gelman (1988) aptly observed that when one studies artifacts, one is usually noting regularities in the environments that artifacts are embedded in, rather than regularities of the artifacts *per se*. For example, an archaeologist might make important observations about the distribution of Chinese porcelain bowls along early modern trade routes. The archaeologist is not making a discovery about Chinese porcelain bowls in and of themselves but making a discovery about ancient trade routes and the tastes of consumers. We think this qualification of the inductive potential of simple artifacts is well observed and correct. This observation connects to the difference between principled and statistical properties (Prasada & Dillingham, 2009). "Cows have four-part digestive tracts" expresses a principled property of cows. A cow has this property in virtue of being a cow. "Cows wear bells around their necks" expresses a statistical property of cows. It generalizes incidentally: For convenience and due to custom, people often put bells on their cows. Simple artifacts and natural objects may both have many statistical features that generalize across category members. But natural objects have more principled properties than simple artifacts.

If the sorts of inductive potential we found for occupational roles were just statistical, we ought to be concerned that their inductive potential is more like artifactual categories than animal categories or race and gender categories. However, upon examination, we see that properties that generalize to members of occupational roles are often principled rather than statistical. Even though occupational roles involve properties that are not inherent in individuals, they can be principled. The permission to defend clients is not inherent to an individual; an individual acquires this right (and the obligations that accompany it). But the reason the individual acquired the property is because they entered into an occupational role: They can defend clients in virtue of being a lawyer. "Lawyers are permitted to defend clients" expresses a generalization involving a principled property. One could even say this property is inherent to the role. The role 'lawyer' is a stable part of the legal system, and an inherent feature of that role is the permission to defend clients. So, this property is not merely a contingent property of the environment but a stable feature of an institutional system.

Occupational roles can also shape people's internal properties and these properties can be principled too, e.g., "Lawyers know how to compose convincing arguments." As with external properties, this skill does not need to be based in biology nor do we think people assume it is. But once acquired, this property is internal, such that an individual will carry it with them even after they exit the role (e.g., by retiring).

This is not to deny that there are differences between natural objects and occupational roles. There certainly are. Instead, we argue

that occupational roles manage to have the robust inductive potential that is on a par with natural objects despite involving profound differences in the sorts of structures that underpin their existence and natures.

To describe our view more clearly, we first consider a view that has some initial plausibility, but which our studies suggest ought to be rejected. One might have thought that occupational roles are a species of cultural artifact and so lack inductive potential. On this view, artifactual categories fall on the negative side of important *natural/non-natural* and *essentialist/non-essentialist* distinctions that run through our category representations. According to this view, natural categories are taken to have essences and this is the source of their strong inductive potential. In contrast, non-natural categories are not taken to have essences and, so, they have low inductive potential.

We saw, however, that occupational roles have significant inductive potential. A category not being thought to be "found in nature" or not being essentialized, is not sufficient for it having low inductive potential. We acknowledge that simple artifacts could possess greater inductive potential than past theories have granted (Chaigneau & Barsalou, 2008). But, if they really do lack inductive potential, it is not only because they lack essences but also because they are not shaped and supported by social institutions. To be clear then, we are not arguing that all the categories we might call "socially constructed" or "conventional" have robust inductive potential. Many of these categories are conventional in the trivial sense: e.g., *white things*. Others might be informative and homogenous, but not have the four features we took to be characteristic of inductive potential, e.g., *paperclips*. Rather, we are arguing that social processes, when sufficiently stable and complex, can promote robust inductive potential.

11.2. Structural explanation

A recent theoretical framework called 'structural explanation' relates to our arguments (Haslanger, 2016; Vasilyeva et al., 2018; Vasilyeva & Lombrozo, 2020). Structural explanations distinguish between "nodes" and "node occupiers." A node is a stable position in social structure. A node occupier is what currently occupies that position, which could be a person (e.g., Obama) or whole categories of people (e.g., women) and objects (e.g., dollar bills).

Proponents of the structural-explanation framework in the psychological literature argue that generic statements, such as "girls play with dolls," are not always explained by inherent or essential features of the category; sometimes they are explained by the position members of the category stably occupy in society (Vasilyeva et al., 2018; Vasilyeva & Lombrozo, 2020). To consider a case where a category might be thought to have an essence and to figure in social structures, consider Jewish people living in Nazi Germany. A generic statement like "Jews wear yellow star badges" held true there. People may also believe that being Jewish reflects common ancestry or a commitment to shared values, both of which can be essential and inherent (Bailey, Knobe, & Newman, 2020). But clearly wearing stars, as well as other regularities relating to Nazi-era policies, reflected the position Jewish people occupied in Nazi Germany not their inherent preferences or values. People recognize this difference, and can evaluate generic statements as true either because of inherent, essential reasons or because of social-structural ones (Vasilyeva et al., 2018; Vasilyeva & Lombrozo, 2020).

Vasilyeva & Lombrozo (2020)'s framework motivates a classification of generic properties as follows: Internal and inherent features are tied to kinds and are connected to essences, whereas external and extrinsic properties are tied to social positions and are connected to social structures. However, a distinction between kinds and positions in social structures cannot always be drawn; some social kinds *just are* positions in social structure (Noyes & Dunham, 2017; 2020; Noyes & Keil, 2020; Noyes et al., 2018; 2020; Ritchie, 2020; Shapiro, 1997). Kinds like "lawyer" and "money" are socially constituted: Members belong to the category because of social facts and processes, such as agreed-upon rules and reoccurring behaviors. Deontic and functional properties are at least partially constitutive of what it means to be a member of an occupational kind. For example, what is a mayor without the power to shape policy, determine budgets, or make pronouncements on city plans? Unlike the case of Jews in Nazi Germany, on which people plausibly do have a representation of the kind Jewish people and of the social position they were forced into, it is not always possible to dissociate kinds from their social positions. For some social kinds even their most indispensable and defining features (e.g., "Mayors shape municipal policy") are explained by non-inherent, non-essential processes—they are social through and through.

An important advance of our work is demonstrating the distinction between social constitution and social production in ordinary concepts. Social institutions not only shape 'nodes,' they also shape 'node occupiers.' Social institutions can set in motion a barrage of causal forces that shape a person's personality traits, skills, and preferences. Individuals in occupational roles do not merely express the same preferences because of constraints in the 'choice architecture' of their society (Vasilyeva & Lombrozo, 2020); individual members of occupational roles may be shaped in lasting ways that carry with them even as they leave one position in society and enter a new one, as happens during retirement. We observed this distinction clearly in how participants reasoned about the properties that members of occupational roles share: Participants reasoned deontic properties and functional behaviors were socially constituted, and that skills and personality traits were socially produced. This is a striking demonstration. The distinction philosophers make between constitutive and causal social processes (e.g., Haslanger, 1995) might well have had no connection to how ordinary people think. Yet, in our studies these are not just metaphysical distinctions. Participants represent deontic properties and functional behaviors as socially constituted, and personality traits and skills as socially produced by causal processes. Whereas causation is a cross-temporal relation, constitution occurs at a single moment. This fits with the dynamics we observed in participants' judgments. Participants judged that personality traits and skills accumulated over time persisted after a person exited a role (consistent with causation), whereas participants judged that people gained deontic properties and functional behaviors immediately upon entering a role and lost them immediately upon leaving the role (consistent with constitution).

Therefore, our findings motivate three modifications of the structural-explanation framework in psychology (Vasilyeva et al., 2018; Vasilyeva & Lombrozo, 2020): (1) The framework needs to recognize the distinction between social constitution and social production, as these are both features of how people understand social kinds. (2) By being socially constituted, the defining and indispensable

properties of some social kinds depend on the position individuals occupy in social structures; in this way they are closely connected and might even be identified with positions in social structures. (3) Even the internal properties of individuals can be socially produced, such that social institutions can cause individuals to share personality traits and skills (and so can become more stereotypical members of categories). Therefore, representing properties as internal to individuals (such that an individual carries the property with them when they leave a social position) is not equivalent to representing the property as inherent or essential.

Studying occupational roles might seem an odd way to gain insight into social categories connected to discrimination, prejudice, and oppression. But, by focusing on people's concepts of categories that are not taken to be natural in any way, we gain insights into social explanation as well. These insights may help us understand psychological representations of race, gender, sexuality, and other important social categories. Greater engagement with a broad range of social categories may reveal new ways people represent the relationship between social structures and social categories, opening additional pathways for exploring how race and gender are represented, and how stereotypes originate.

11.3. Role-governed categories

A related theoretical framework proposes that role-governed categories, including *guest*, *predator*, *home*, *pet*, and *game* differ representationally from feature-based ones, such as *tiger*, *chair*, and *telescope* (Markman & Stilwell, 2001; Goldwater, Markman, & Stilwell, 2011; Rein, Goldwater, & Markman, 2010; Goldwater & Markman, 2011). Relations are often absent from formal models of categorization built to handle feature-based categories (e.g., Nosofsky, 2011); yet, representations of role-governed categories require the explicit representation of relations. We have approached the topic of roles from a different vantage point, focusing not directly on how these categories are represented, but showing that relations between roles are sufficient for robust inductive potential. Nevertheless, our perspectives are complementary and offer similar general insights: Relations are significant in higher-order cognition and models of higher-order cognitive processes (i.e., category representation, formation, and inductive potential) need to explicitly include relations between roles to successfully capture the scope of ordinary reasoning.

Our studies offer several insights for the literature on role-governed categories. First, we have undertaken one of the most substantive explorations of the inductive potential of relational structures. One might have assumed that relational structures fail to support inductive potential. Relational categories, such as *guest* and *leader*, often involve high-level abstraction, such that the roles cut across disparate contexts: e.g., *guest* could include foreign diplomats or a friend staying the night, and *leader* could include the President of the United States or the shift leader at a retail store. As is typical of analogy and abstraction (Gentner, 1983), these categories sacrifice shared features to isolate common relations. However, not all relational structures fit this mold. Social institutions are highly specific, content-rich relational structures. As we found in our studies, the specificity and richness of these relations can support the densely correlated property clusters that support inductive potential in other domains (e.g., animals). Thus, our studies help reveal when and why relational structures may support content-full inductions.

The specificity and richness of institutional roles also opens up new avenues for future research. The literature on role-governed categories has focused on category formation (Goldwater, Bainbridge, & Murphy, 2016); categories like *guest* and *leader* require participants to identify relational similarities across diverse contexts and suppress alternative feature-based representations. Therefore, the hard part is engaging in the abstraction or analogical reasoning needed to isolate the roles; there is less need for protracted learning about the relational structures per se. Institutional roles raise questions about how people acquire content knowledge. For example, Americans acquire substantial domain-specific knowledge about the U.S. Federal government; we may anticipate this process to be lifelong. Thus, institutional roles could provide an important test case for investigating the knowledge enrichment processes needed to build up content-rich representations of relational structures.

To conclude, we have developed our theorizing about institutional roles by contrasting institutional roles with essentialized categories. A related literature has contrasted a broad class of role-governed categories with feature-based ones (Markman & Stilwell, 2001; Goldwater, et al., 2011; Rein, et al., 2010; Goldwater & Markman, 2011). These complementary perspectives help highlight the importance of relations in higher-order cognition and the need to include relations in models of concepts and categorization.

11.4. Dual-character concepts

Dual-character concepts have also been linked to occupational roles and their possible inductive potential (Bailey, Knobe, & Newman, 2020; Del Pinal & Reuter, 2017; Foster-Hanson & Rhodes, 2019; Knobe, et al., 2013; Leslie, 2015). In this framework, certain concepts possess two sets of criteria for category membership: (1) A set of concrete features, and (2) A set of abstract values. These can be dissociated. For example, 'minister' has a dual character: A jaded, callous individual that leads Protestant religious services is a minister according to (1) but not (2). A devout, compassionate individual that provides spiritual guidance to people but has never led religious services in any official capacity may be taken to be a true minister given (2) but not (1).

Recently, researchers have argued that (2) involves a value-based essence and that this supports inductive potential (Newman & Knobe, 2019; Bailey et al., 2020). On this view, the true nature of a category member can be evaluative rather than biological. Both shared values and shared biology can support robust inductive inference. In the dual-character framework, social roles without dual character lack functional and normative structure. For instance, 'second cousins' is not a dual character concept. No significant social function or values are realized in second cousins. Is "dual characterhood" required to be functionally and normatively significant? Our investigation suggests not: Social institutions themselves can be a source of functional and normative structure.

In our studies, we selected occupational roles embedded in rich institutional structures but without dual character: e.g., 'mayor,' 'lawyer,' and 'quarterback.' In Knobe, et al. (2013), 'mayor' was one of the roles rated lowest in dual character. Yet, mayors are

embedded in stable municipal institutions. While these sorts of occupational roles and dual character categories are both related to functions and norms, they differ in how they are related. Dual character categories relate to broader social-cultural value systems tied to morality and how to live a good and virtuous life. For example, a 'true leader' will include morally-laden features like integrity, honesty, humility, and courage – none of which relate to any specific occupation but rather reflect general virtues that all people could embody. This creates the possibility for a person to realize 'true leadership' independently of occupying any particular social position. And, we may even think individuals are inherently disposed to 'true leadership,' whether because of innate characteristics or a deep commitment to certain values. In contrast, other occupational roles relate to norms tied to specific social institutions or social groups (e.g., football teams, city government, judicial system); these norms relate to the part an individual plays in an institutional whole. It does not make sense for such individuals to be a member of the category outside of the institution: There is little sense in thinking of a quarterback with no connection to a football team or a mayor who is not a mayor of some city or other. Indeed, it seems the evaluative component of a concept's dual character can disappear by focusing on a more specific but related role in an institution, as in the pairs 'leader' and 'mayor' or 'scientist' and 'university professor.' There is nothing inherent in individuals that make them members of institutional occupational roles; individuals acquire membership by virtue of their place within an institution.

In summary, our argument differs from the literature on dual character concepts in at least two major ways. First, occupational roles embedded in social institutions reflect different kinds of functions and norms than occupational roles with dual characters. Second, social institutions are sufficient to support robust inductive potential. Occupational roles do not need to be connected to deeper values or value systems to have robust inductive potential.

11.5. Enumerations of essentialist beliefs

Many theories of psychological essentialism encompass several beliefs understood as 'components' of essentialist reasoning. The exact enumeration of these beliefs differs by author. However, a recent, prominent description of these beliefs is provided by Rhodes & Mandalaywala (2017). In Rhodes & Mandalaywala (2017), homogeneity encompasses the belief that 'category members will share properties with one another, even if they have other dissimilarities.' Social institutions also support this inference. Despite considerable differences in individual lawyers (mayor, etc.), participants infer they share many properties in virtue of their shared position in social institutions. Social constitution, in particular, can support powerful homogeneity with respect to deontic properties and functional behaviors. The belief that 'category memberships cause the development of category-typical properties' is accomplished by social production. Being a member of the category (i.e., being in the role) causes individuals to acquire category-typical properties; specifically, shared skills and personality traits. According to Rhodes & Mandalaywala (2017), "All of the inter-related beliefs just described…can be understood as a consequence of the intuition that category memberships are conferred via intrinsic, stable, and causally powerful 'essences.'" But, as we have shown, these two beliefs are not diagnostic of a belief in essence because they can be supported by inferences to different structures, including social institutions. The presence of these beliefs is not a unique consequence of essences but instead a consequence of a more general intuition that category membership reflects real-world (causal and constitutive) structures. We think a belief that a category is a natural kind (in the sense of being naturally constituted and not socially constituted), that category membership is intrinsic to individuals, and that category-linked properties are specifically naturally produced, are unambiguous symptoms of a belief in essences; however, we argue that other purported components lack a privileged connection to essences and may need to be revised in, or removed from, future enumerations of essentialist beliefs (Noyes & Keil, 2019; 2020; Noyes & Dunham, pre-print).

11.6. The sources of inductive potential

We began with the question of why people take certain categories to have robust inductive potential. Sometimes people attribute essences to categories. Underlying essences shared by each kind member explain why kind members share properties. Recently, researchers have posited more views of what essences might involve: In addition to the internal, biological essences that people might attribute to animal kinds and race and gender kinds (Gelman, 2003; Rothbart & Taylor, 1992), people might also attribute value-based essences (Newman & Knobe, 2019; Bailey et al., 2020) and teleological essences (Rose & Nichols, 2019, 2020). What unites these accounts is the presence of a true nature shared by category members that is responsible for the similarities among them. Although an essence-based framework informs why people make some generalizations, essences do not seem to be the primary source of inductive potential. In our studies, similarities among members of categories occur because of external social institutions that have nothing to do with the true or inherent nature of individuals, whether their deeper values, their telos, or their biology. It is completely acceptable to say that someone might not have been a lawyer or they used to be a lawyer and are no longer. It is not essential to an individual that they be a lawyer and so there is nothing about the 'true nature' of an individual that makes them a member of the category. Therefore, people easily attribute the source of inductive potential to social institutions rather than just to essences shared by category members.

More abstract similarities do hold between social institutions and essences and these may reveal deeper cognitive principles. People may take correlations between properties to reflect real-world structures, whether these involve social institutions, genes, ideals, or purposes. That is, people may often infer property co-occurrence as arising from principled regularities in the environment rather than from incidental ones. Kind-indicating language like generic language and noun labels may then act as a means of transmitting these beliefs, highlighting co-occurrences really are non-accidental (Gelman, 2003; Gelman, Ware, & Kleinberg, 2014; Ritchie & Knobe, 2020; Ritchie, 2019; Rhodes, Leslie, & Tworek, 2012). In contrast to psychological essentialism though, our work raises doubts about a pervasive tendency to infer essences as the source of inductive potential. Any theory of kind representations must pluralistically allow for not only essences but also for external structures like social institutions to form the basis for kind membership and to support the

judgments people make. Institutional categories are a part of general cognition just as animal and gender categories are. So, although psychological essentialism may be one mode of representing categories, it is not be the only, or perhaps not the most common, or even the default case (Noyes & Keil, 2020).

11.7. Conclusion

Ordinary cognition adeptly uses categories to draw robust inductive inferences. These categories allow us to track, explain, and predict regularities in the environment so that we can act on that knowledge. And they allow individuals to construct simpler and more general representations of the vast complexity we encounter in the world. Indeed, categories with robust inductive potential allow entire communities to coordinate around stable real-world structures. Consequently, categories with robust inductive potential are fundamental to community-wide knowledge-building projects like science. They also figure in pernicious community-wide projects and agendas, like those involving racist, sexist, and homophobic prejudice and oppression.

Understanding how people represent natural kinds has provided deep insight into the ways in which we use inductive inference and has given specific insight into how people engage with the natural sciences (see Gelman & Rhodes, 2012; Shtulman & Schulz, 2008; Coley, Arenson, Xu & Tanner, 2017). Studying how people represent kinds that are obviously social promises another avenue for understanding the *multiple* sources people use to guide their judgments about the inductive potential of kinds. It also promises unique insights into how people engage with the social sciences.

Psychological theories of inductive inference need to accommodate the diversity of real-world structures that people take to underlie inductive potential. In recognizing this diversity, deep and fundamental commonalities also emerge, such as the importance of inferring non-accidental structure as a source of property correlations. This accommodation will also require recognizing that our judgments about the sources of inductive potential may be important in a particular domain without being fundamentally and intrinsically connected to all forms of inductive potential. For instance, internal or inherent structure might be central to why we take certain natural categories to be rich with inductive potential involving perhaps inexhaustibly many shared features. Yet, as our studies here suggest, not all inductive potential is underpinned by inherent natural features. Our results enjoin previous work in demonstrating that people approach the world as composed of stable categories with robust, non-accidental inductive potential, and that this assumption is at the heart of higher-order cognition. This assumption extends to people's representation of the very parts of the world they help to create: People understand institutions as shaping us in deep and possibly unexpected ways, shifting not only what actions are possible or expected but also the nature of the actors themselves.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ásta. (2018). *Categories we live by: The construction of sex, gender, race, and other social categories*. Oxford University Press.

Bailey, A., Knobe, J., & Newman, G. (2020). Value-based Essentialism: Essentialist Beliefs About Non-biological Social Groups. *Journal of Experimental Psychology: General.*

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*, 351–368.

Boyd, R. (1999). Homeostasis, species, and higher taxa. *Species: New interdisciplinary essays, 141*, 185.

Brandone, A. C., & Gelman, S. A. (2009). Differences in preschoolers' and adults' use of generics about novel animals and artifacts: A window onto a conceptual divide. *Cognition, 110*, 1–22.

Burnham, M. J., Le, Y. K., & Piedmont, R. L. (2018). Who is Mturk? Personal characteristics and sample consistency of these online workers. *Mental Health, Religion & Culture, 21*, 934–944.

Chaigneau, S. E., & Barsalou, L. W. (2008). The role of function in categories. *Theoria et Historia Scientiarum, 8*, 33. https://doi.org/10.12775/ths.2008.003.

Coley, J. D., Arenson, M., Xu, Y., & Tanner, K. D. (2017). Intuitive biological thought: Developmental changes and effects of biology education in late adolescence. *Cognitive Psychology, 92*, 1–21.

Del Pinal, G., & Reuter, K. (2017). Dual character concepts in social cognition: Commitments and the normative dimension of conceptual representation. *Cognitive Science, 41*, 477–501.

Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development, 77*, 539–553.

Diesendruck, G., Goldfein-Elbaz, R., Rhodes, M., Gelman, S., & Neumark, N. (2013). Cross-cultural differences in children's beliefs about the objectivity of social categories. *Child Development, 84*, 1906–1917.

Foster-Hanson, E., & Rhodes, M. (2019). Normative social role concepts in early childhood. *Cognitive Science, 43*, Article e12782.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65–95.

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

Gelman, S. A., Heyman, G. D., & Legare, C. H. (2007). Developmental changes in the coherence of essentialist beliefs about psychological characteristics. *Child development, 78*, 757–774.

Gelman, S. A., Ware, E. A., Kleinberg, F., Manczak, E. M., & Stilwell, S. M. (2014). Individual differences in children's and parents' generic language. *Child development, 85*(3), 924–940.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183–209.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532. https://doi.org/10.2307/1130693.

Gelman, S. A., & Rhodes, M. (2012). Two-thousand years of stasis. *Evolution challenges: Integrating research and practice in teaching and learning about evolution*, 200-207.

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

Gil-White, F. (2001). Are ethnic groups biological species to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology, 42*, 515–553.

Goldwater, M. B., & Markman, A. B. (2011). Categorizing entities by common role. *Psychonomic Bulletin & Review, 18*, 406–413.
Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition, 118*, 359–376.
Goldwater, M. B., Bainbridge, R., & Murphy, G. L. (2016). Learning of role-governed and thematic categories. *Acta Psychologica, 164*, 112–126.
Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*, 113–127.
Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin, 30*, 1661–1673.
Haslanger, S. (2000). Gender and race:(What) are they?(What) do we want them to be? *Noûs, 34*, 31–55.
Haslanger, S. (2016). What is a (social) structural explanation? *Philosophical Studies, 173*, 113–130.
Haslanger, S. (1995). Ontology and social construction. *Philosophical Topics, 23*, 95–125.
Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics, 2*.
Kalish, C. W., & Lawson, C. A. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development, 79*(3), 577–593.
Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
Knobe, J., Prasada, S., & Newman, G. E. (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition, 127*, 242–257.
Kripke, S. (1972). Naming and necessity. In D. Davidson, & G. Harman (Eds.), *Semantics of natural language* (pp. 253–355). Dordrecht, Netherlands: D. Reidel Publishing.
Lafraire, J., Rioux, C., Hamaoui, J., Girgis, H., Nguyen, S., & Thibaut, J. P. (2020). Food as a borderline domain of knowledge: The development of domain-specific inductive reasoning strategies in young children. *Cognitive Development, 56*, Article 100946.
Leslie, S. J. (2013). Essence and natural kinds: When science meets preschooler intuition. *Oxford studies in epistemology, 4*, 108–165.
Leslie, S. J. (2015). "Hillary Clinton is the only man in the Obama administration": Dual character concepts, generics, and gender. *Analytic Philosophy, 56*, 111–141.
Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122–1135.
Mallon, R. (2016). *The Construction of Human Kinds*. New York: Oxford University Press.
Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.
Markman, A. B., & Stilwell, C. H. (2001). Role-governed categories. *Journal of Experimental & Theoretical Artificial Intelligence, 13*, 329–358.
Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 21–59). Cambridge: Cambridge Univ Press.
Newman, G. E., & Knobe, J. (2019). The essence of essentialism. *Mind & Language, 34*, 585–605.
Noyes, A., & Dunham, Y. (2017). Mutual intentions as a causal framework for social groups. *Cognition, 162*, 133–142.
Noyes, A., & Dunham, Y. (2020). Groups as institutions: The use of constitutive rules to attribute group membership. *Cognition, 196*, 104–143.
Noyes, A., & Keil, F. C. (2019). Generics designate kinds but not always essences. *Proceedings of the National Academy of Sciences, 116*, 20354–20359.
Noyes, A., & Keil, F. C. (2020). There is no privileged link between kinds and essences early in development. *Proceedings of the National Academy of Sciences, 117*, 10633–10635.
Noyes, A., Keil, F. C., & Dunham, Y. (2020). Institutional actors: Children's emerging beliefs about the causal structure of social roles. *Developmental Psychology., 56*, 70–80.
Noyes, A., Keil, F. C., & Dunham, Y. (2018). The emerging causal understanding of institutional objects. *Cognition, 170*, 83–87.
Prasada, S., & Dillingham, E. M. (2009). Representation of principled connections: A window onto the formal aspect of common-sense conception. *Cognitive Science, 33*, 401–448.
Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science, 17*, 129–135.
Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science, 16*, 202–206.
Putnam, H. (Ed.). (1975). *The meaning of 'meaning'. In Mind, language, and reality: Philosophical papers,* (Vol. 2). Cambridge: Cambridge University Press.
Rein, J. R., Goldwater, M. B., & Markman, A. B. (2010). What is typical about the typicality effect in category-based induction? *Memory & Cognition, 38*, 377–388.
Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*, Article e1437.
Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences, 109*(34), 13526–13531.
Ritchie, K. (2019). Should we use racial and gender generics? *Thought: A. Journal of Philosophy, 8*, 33–41.
Ritchie, K. (2020). Social structures and the ontology of social groups. *Philosophy and Phenomenological Research, 100*, 402–424.
Ritchie, K., & Knobe, J. (2020). Kindhood and essentialism: Evidence from language. *The Development of Social Essentialism, 59*, 133.
Rose, D., & Nichols, S. (2019). Teleological essentialism. *Cognitive Science, 43*, e12725. https://doi.org/10.1111/cogs.2019.43.issue-410.1111/cogs.12725.
Rose, D., & Nichols, S. (2020). Teleological essentialism: generalized. *Cognitive science, 44*, Article e12818.
Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin, & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 11–36). Thousand Oaks, CA, US: Sage Publications Inc.
Searle, J. (2010). *Making the social world: The structure of human civilization*. Oxford: Oxford University Press.
Searle, J. R. (1995). *The construction of social reality*. New York, NY: The Free Press.
Shapiro, S. (1997). *Philosophy of mathematics: Structure and ontology*. Oxford University Press.
Shtulman, A., & Schulz, L. (2008). The relation between essentialist beliefs and evolutionary reasoning. *Cognitive Science, 32*, 1049–1062.
Smyth, K., Feeney, A., Eidson, R. C., & Coley, J. D. (2017). Development of essentialist thinking about religion categories in Northern Ireland (and the United States). *Developmental Psychology, 53*, 475–496.
Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development, 80*, 461–481.
Thomasson, A. (2019). The Ontology of Groups. *Synthese., 196*, 4829–4845.
Thomasson, A. L. (2019). The ontology of social groups. *Synthese, 196*, 4829–4845.
Vasilyeva, N., & Lombrozo, T. (2020). Structural thinking about social categories: Evidence from formal explanations, generics, and generalization. *Cognition, 204*, 104383. https://doi.org/10.1016/j.cognition.2020.104383.
Vasilyeva, N., Gopnik, A., & Lombrozo, T. (2018). The development of structural thinking about social categories. *Developmental Psychology, 54*, 1735–1744.