Children's differential weighting of cues to social categories

Gil Diesendruck*, Eitan Weiss

Bar-Ilan University, Israel

ARTICLE INFO

Keywords:
Social categories
Categorization
Induction
Perceptual similarity
Labels

ABSTRACT

Two studies investigated the weights of physical similarity, labels, and internal properties in 5-year-olds' \((n = 64)\) categorization and inferences regarding three social categories: gender, race, and shirt-color. Participants saw exemplars of varying degrees of similarity to target categories and were asked to categorize the exemplars and draw inferences about them. Varied across studies was the kind of information pitted against visual similarity – labels (Study 1) or internal information (Study 2). Labels had the weakest effect on children's categorization of the most essentialized category – gender. (Essentialism was assessed independently.) Internal property information dominated physical similarity in determining children's categorization of all three categories. We conclude that essentialized social categories are defined as natural kinds, wherein appearances are indicative of intrinsic essences, and thus information about intrinsic properties – but not labels – can lead children to overlook physical dissimilarity.

© 2014 Elsevier Inc. All rights reserved.

A growing body of literature suggests that by age 5, children essentialize social categories. In particular, they treat certain categories as stable across time (Hirschfeld, 1996; Kinzler & Dautel, 2012), as inductively powerful (Diesendruck & haLevi, 2006), as objectively true (Rhodes & Gelman, 2009), and as passed on by inheritance (Taylor, 1996). One critical aspect of essentialism yet to be examined in regard to social categories is how category membership is assigned. It is of particular relevance because it touches on a core controversy in social sciences regarding the origins of social categories.

* Corresponding author at: Department of Psychology, Bar-Ilan University, Ramat-Gan 52900, Israel. Tel.: +972 3 531 8639. E-mail address: dieseng@mail.biu.ac.il (G. Diesendruck).

http://dx.doi.org/10.1016/j.cogdev.2014.06.001
0885-2014 © 2014 Elsevier Inc. All rights reserved.

Please cite this article in press as: Diesendruck, G., & Weiss, E. Children's differential weighting of cues to social categories. Cognitive Development (2014).
http://dx.doi.org/10.1016/j.cogdev.2014.06.001
(Haney Lopez, 2006; Omi & Winant, 1994; Rothbart & Taylor, 1992). Are social categories constructed via cultural factors such as language? Or are they inherent and manifested in objective partitions of social reality? The present studies ask an analogous developmental question: do children assign social category membership based primarily on how society refers to a person, on how a person looks, or on the internal properties of a person?

These different cues to category membership have a history in the developmental literature on categorization in other domains. Starting with Keil’s (1989) “transformation” tasks, studies have revealed that children weigh these types of cues differently for animal and artifact categories. For instance, Keil reported that making an animal of category X (e.g., raccoon) look like an exemplar of category Y (e.g., skunk) did not convince 8-year-olds that the animal had changed category membership, although changing its internal parts did. For artifacts, the appearance change sufficed for revision of category assignment (see also Gelman & Wellman, 1991; Newman & Keil, 2008, for evidence of such domain differences among younger children). Using different methods, 4-year-olds have been found to treat internal properties of animals (e.g., organs), but not artifacts (e.g., parts), as critical for category assignment (Diesendruck, Gelman, & Lebowitz, 1998). Within certain limits, labels lead 4-year-olds to overcome perceptual dissimilarity in their inferences based on animal categories (Davidson & Gelman, 1990; Jaswal & Markman, 2007).

Diesendruck and Peretz (2013) compared the relative weights of perceptual similarity, internal information, and labels in 3- and 5-year-olds’ determination of animal and artifact categories. Children were introduced to two novel categories and were shown test exemplars that varied systematically in degree of perceptual similarity to the categories. In one of their studies, children were either told that the test items had the same label as one of the categories or were not provided labels. For the crucial items, the label information conflicted with perceptual similarity, such that the test exemplar looked more similar to the differently labeled category than the similarly labeled one. They found that especially among 5-year-olds, labels had a stronger effect on children’s category assignment of artifacts than of animals. A second study had the same structure, but instead of being told labels for the target categories and test exemplars, children were told about their internal properties (biological for animals, mechanical for artifacts). Here it was found that internal properties had a stronger effect on the category assignment of animals than of artifacts.

These studies are consistent with the notion that animal categories are essentialized to a greater extent than artifact categories (Diesendruck, 2003; Gelman, 2003; Rhodes & Gelman, 2009). As such, children believe that animal categories reflect objective partitions of nature – partitions causally defined by internal hidden properties and manifested in the physical appearances of animals. Consequently, information about internal properties is weighted most heavily in the determination of category membership, but in the absence of such information, children rely on perceptual similarity. Artifact categories, in turn, are believed to reflect cultural conventions, rather than objective truths. Consequently, information about how cultures define artifacts – for instance, via labeling – is weighted heavily in category assignment. The goal of the present studies was to examine different social categories in these respects.

We focused on three social categories – gender, race, and shirt-color – selected for both theoretical and practical reasons. The theoretical reason is that there arguably are differences in the extent to which social categories are essentialized (Haslam, Rothschild, & Ernst, 2000). Specifically, in Israel – the country in which the present studies were conducted – 5-year-olds essentialize ethnicity more strongly than social status, gender, and religiosity (Diesendruck & haLevi, 2006), and gender more than race (Diesendruck, Goldfein-Elbaz, Rhodes, Gelman, & Neumark, 2013). The practical consideration is that systematically varying degree of perceptual similarity between exemplars and categories requires categories amenable to a morphing transformation. This criterion made it very hard to include ethnicity. Instead, we selected gender as a potentially highly essentialized category, race as a moderately essentialized one, and an artificial social category based on shirt color as a weakly essentialized category. Study 1 included a direct assessment of essentialism to confirm this order.

The present studies have the same structure as those conducted by Diesendruck and Peretz (2013). Children performed a categorization task that was a replication of theirs. They were shown typical exemplars of the three social categories: gender (White men/women), race (White/Black men), and shirt–color (men wearing red/yellow shirt), and four test exemplars for each category varying in degree...
of similarity to each social category. The main task was to assign each of the test exemplars to one of the two categories (e.g., men or women). In Study 1, we manipulated the presence/absence of labels for the categories and test exemplars. In Study 2, we manipulated the presence/absence of internal information about the categories and test exemplars. The crucial exemplars in each study were ones in which there was a conflict between either the label (Study 1), or the internal information (Study 2), and the perceptual similarity.

As an additional assessment of the weighting of the different cues, after completing the categorization task described, children also participated in an inference task. Its inclusion allowed a comparison to other work involving possible differences between categorization and induction tasks. Some researchers argue that labels are treated by children as features of objects, and as such are entered in an equivalent manner when computing the similarity between items in both categorization and induction tasks (Sloutsky & Fisher, 2004). Others argue that labels have special status for induction because they serve as proxies for category essences (Gelman, 2003). The present task was modeled after typical category-based induction tasks (Gelman & Markman, 1986). Children were taught novel properties regarding the two categories and were asked to infer which of the properties the test exemplars would have. As in the categorization task, in some trials children faced a conflict between the perceptual similarity of the test exemplar to the categories and either their labels (Study 1) or their internal properties (Study 2). We included this task after the categorization task so as to have a more direct and unconfounded comparison of children’s performance in a social categorization task to their performance regarding animals and artifacts documented by Diesendruck and Peretz (2013). Five-year-olds were the focus as this was the age group at which Diesendruck and Peretz (2013) found reliable domain differences in the weighting of cues and because this is an age at which clear signs of essentialist beliefs regarding social categories have been reported (Diesendruck et al., 2013; Rhodes & Gelman, 2009).

The general hypotheses were that to the extent that a social category is essentialized, the pattern of weighting of cues to category assignment (and induction) should resemble the one described by Diesendruck and Peretz (2013) for animals. Namely, in Study 1, social category assignment (and induction) should be relatively unaffected by labeling that conflicts with perceptual similarity, and in Study 2, assignment (and induction) should be swayed by internal property information even if it conflicts with perceptual similarity. The pattern of findings should be the opposite if a social category is conceptualized in a non-essentialist manner, as are artifacts.

1. Study 1: if it looks like a man but is called “woman”, what is it?

Numerous studies have documented the importance of labeling to children’s categorization decisions (Sloutsky & Fisher, 2011; Waxman & Braun, 2005), for the most part regarding the effect of labels as a domain-general phenomenon. As noted above, however, it has been hypothesized that given potential differences in the conceptual beliefs underlying certain categories, one might expect differences also in the beliefs about the susceptibility of various categories to the effect of labels.

In fact, there are two different hypotheses about how labels might interact with essentialist beliefs. According to one hypothesis, children (and adults) might believe in a “division of linguistic labor”, such that they defer to experts in a field to determine what things “really” are (Malt, 1990; Putnam, 1975). Following this hypothesis, it could be the case that for essentialized categories, children are more likely to believe that there might be a dissociation between how things look and what things really are, and thus an expert’s classification may be taken as authoritative. The effects of labels on children’s categorization and induction reviewed above are consistent with this possibility.

A second hypothesis is that there exist differences between domains in terms of how unconstrained children’s trust in a division of linguistic labor is. In particular, to the extent children view certain categories as capturing natural discontinuities in the environment, rather than arbitrary cultural conventions: (a) children will need firm evidence that the provider of the label is indeed an expert (see Braisby, 2001; Proctor & Keil, 2006, for relevant work on adults), and (b) children should be more willing to ignore cultural conventions such as labels, in favor of cues about the physical reality of the categories, as defining category membership (Diesendruck, 2003). In fact, Davidson and Gelman (1990) found that although children were generally swayed by labels for drawing inferences about

Please cite this article in press as: Diesendruck, G., & Weiss, E. Children’s differential weighting of cues to social categories. *Cognitive Development* (2014), http://dx.doi.org/10.1016/j.cogdev.2014.06.001
novel animals, this did not occur when the labels applied to exemplars that were physically highly dissimilar (Deng & Sloutsky, 2012; Noles & Gelman, 2012). The Diesendruck and Peretz (2013) study of animals vs. artifacts provides further evidence consistent with this hypothesis. Here we examine the effect of labeling on the categorization of social categories using the same methodology. We hypothesize that if children treat certain social categories as natural kinds, the effect of labels should resemble the one found in that study in regard to animals.

We created a set of test exemplars that varied systematically in degree of perceptual similarity to target categories. Through a calibration pre-test, categories were morphed to create exemplars rated by children as having very low to very high degrees of similarity to the alternative target categories. In the main experiment, critical trials were ones in which similarity information pointed to one category and label information to another. To address whether potential differences in the weights of labels versus physical similarity were related to degree of essentialism regarding the category, the three social categories selected were ones potentially varying in this respect, i.e., gender, race, and shirt-color. To ascertain these suspected differences, children were asked about their essentialist beliefs regarding these categories.

1.1. Method

1.1.1. Participants

Thirty-two 5-year-olds ($M_{\text{age}} = 65.72$ months, $SD = 3.41$; 15 boys) participated in the main experiment. An additional 20 5-year-olds participated in the calibration study. Children were Jewish monolingual Hebrew-speakers, recruited from Israeli kindergartens. All children received a sticker for their participation.

1.1.2. Design

A mixed design was employed, with Condition (Label vs. No-label) as a between-subjects factor, and Category type (gender, race, and shirt-color) and Similarity ratio (.12, .35, .6, .82) as within-subjects variables. Participants were assigned randomly to conditions. All participants completed the same three tasks, in a fixed order: Categorization, Inference, and Essentialism. Categorization was always presented first to allow a clear comparison to the findings on the same task regarding animals and artifacts, documented by Diesendruck and Peretz (2013).

1.1.3. Materials

The main stimuli consisted of 11 cm by 11 cm pictures of people, presented on a 13-inch laptop screen. The pictures were collected from internet images and consisted of typical exemplars of polar values in each category: for gender, a White man and a White woman; for race, a White man and a Black man; and for shirt-color, a White man wearing a red t-shirt, and another White man wearing a yellow t-shirt. For each of the three pairs of typical exemplars, four test exemplars were created, each with a different degree of similarity to the typical exemplars. Thus each category set included two typical exemplars and four test exemplars (see Fig. 1).

The test exemplars were created, and eventually selected, following the procedures used by Sloutsky and Fisher (2004) and Diesendruck and Peretz (2013). First, the pair of typical exemplars from each category (e.g., gender) was used to create a 16-step sequence in which one member of the pair was morphed into another member (e.g., the man into a woman). These morphing sequences were created using Squirrel Morph software. Of the 16 pictures in a sequence, we selected every second picture to serve as a "candidate" test exemplar in the actual experiments (total of 8 pictures per set). We repeated this procedure for all three categories, creating 24 candidate test exemplars. A calibration study was conducted to quantify the similarity of each candidate test exemplar to each of the typical category exemplars.

A group of 20 5-year-olds were presented the 24 candidate test exemplars (8 per category), one at a time, and asked whether the candidate test exemplar looked more like the typical exemplar from Category A (e.g., the woman) or more like the typical exemplar from Category B (e.g., the man). For each trial, the number of children selecting Category A divided by 20 (the number of participants), was the measure of similarity between the candidate test exemplar and Category A. For instance, if
12 of the 20 children judged that a given candidate-test exemplar was more similar to the exemplar from Category A than to that from Category B, that candidate test exemplar received a similarity ratio of .6 to Category A. In this manner, 12 test exemplars were selected (4 per category), representing four “similarity ratios” of test exemplars to Category A: .12, .35, .6, and .82 (see Fig. 1).

1.1.4. Procedure
Each child was interviewed individually in a quiet room in his or her kindergarten. The entire procedure was conducted using a laptop computer for presentation of the pictures.

1.1.4.1. Categorization task. In both conditions, participants first saw a picture of a couple the experimenter [E] said were named Yossi and Rinat. E said that Yossi and Rinat, “like travelling around the world and like taking pictures of different kinds of people from far away and unfamiliar places, and who often times do things that we don’t do. After they take pictures, Yossi and Rinat like organizing the pictures in albums, putting people from each kind together into an album. There are different albums with pictures of the different kinds of people. One day, Yossi and Rinat found out that someone had mixed up all the pictures. Now they want to organize the pictures again. Let’s help them return the pictures to the albums, each picture to the kind of people it belongs with.”

![Gender](image1)
![Race](image2)
![Shirt-color](image3)

**Fig. 1.** Typical and test exemplars used for the three categories.
The first trial was then presented. E revealed the two categories, represented by three typical exemplars of the category, saying to the participant, for instance: “Here are the men” [pointing to Category A, which appeared on one side of the screen], “and here are the women” [pointing to Category B]. E then revealed one of the test exemplars for the set, shown at the bottom-center of the screen, equidistant from the two categories.

In the Label condition, E pointed to the test exemplar and labeled it with the same label previously applied to Category A. For instance: “Look at the picture of this man. Can you return it to the group of the same kind? Should it go with the men or with the women?” In all trials, the test exemplars were given the same label as Category A. This means that in trials of low-similarity ratios (e.g., .12, .35), label and similarity provided conflicting cues about the category membership of the test exemplar, as the label referred to the category least similar to the test exemplar. In trials of high-similarity ratios (.6, and especially .82), label and similarity cohered, as the label was applied to the category most similar to the test exemplar. In the No-label condition, E pointed to the test exemplar and simply said: “Look at this picture. Can you return it to the group of the same kind? Should it go with the men or with the women?” Thus, in the No-label condition, there were no conflicting cues, as children only had similarity as a cue to category membership. Given that Category A was always the category defining the label, we refer to it as the Target Category, and the other category as the Alternative Category. Our main dependent measures was whether or not children selected the Target Category.

After recording the participant’s dichotomous selection (either verbal or pointing), E introduced the Inference trial related to that test exemplar.

1.1.4.2. Inference task. The inference task started with E reminding children that the pictured people come from far away countries where they do something that we don’t. E then described two novel activities as stereotypical of each of the two categories. For instance, E said: “Men in those countries like playing with a donga. See, this is a donga [pointing to a picture of a novel object]. Women in those countries like playing with a flerit. See, this is a flerit [pointing to a picture of a different novel object].”

E made sure children recognized which activities went with which category by questioning them. E then presented again the test exemplar and asked, while pointing to pictures of the objects that varied in size: “What do you think this [man/person] most likes to play with? Play only donga [pointing to a picture of a large donga]? Play donga a lot and flerit a bit [pointing to a picture of a large donga and a small flerit]? Play both donga and flerit [pointing to a picture of mid- and equally sized donga and flerit]? Play donga a bit and flerit a lot [pointing to a picture of a small donga and a large flerit]? Play only flerit [pointing to a picture of a large flerit]?” Children in the Label condition were told the label of the test exemplar (e.g., “this man”), and children in the No-label condition were not told a label (e.g., “this person”). Children’s responses were coded on a scale from 1 to 5, with 1 signifying a choice exclusively consistent with the Target Category (e.g., selection of “play only donga” in the example above), and 5 a choice exclusively consistent with the Alternative Category (e.g., selection of “play only flerit” in the example above).

The next test exemplar pertinent to the category was then shown and participants went through a cycle of Categorization-then-Inference tasks. After all four test exemplars for a given category were shown, the next category was presented. Participants completed 12 Categorization-then-Inference trials, four for each category type. Both order of categories and test exemplars within each category were randomized for each participant, as was the introduction order and left-right placement of the categories on the screen. In the Inference task, the pairing of activities to the different category types was also randomized across participants. No feedback was given to participants on their performance.

For gender, we used the labels “men” and “women”; for race, “white men” and “black men”; and for shirt-color we used made-up names that we associated with the shirt-color so as to mark them as categories, i.e., “bargovans who wear yellow”, and “tirpals who wear red”. The activities included in the Inference task, in addition to the one described above, were: “like playing an instrument called targon/grempel”, “like building a premin/cobert”.

Upon completion of all 12 cycles of Categorization-then-Inference trials, participants answered a series of questions from the Essentialism Components Questionnaire (ECQ: Deeb, Segall, Birnbaum, Ben-Eliyahu, & Diesendruck, 2011). The ECQ was left for last so as to not bias responses in the other two tasks.
1.1.4.3. Essentialism questions. These questions from the ECQ were included so as to assess the hypothesized differences across category types in their degree of essentialism. The first six questions regarded differences between members of the two categories. For the race category, for instance, they were: “To what extent do Black and White people differ in the way they think/what they like/in the way they behave/in the way they look/in what they have inside their body/kind of blood they have?” The response options were: (1) not at all different; (2) differ a little; (3) very different; (4) totally different. The 7th and 8th questions regarded the possibility of changing one’s category membership. For instance: “Let’s say that a Black/White man wants to become a White/Black man, is it possible for him to do some things and then become a White/Black man?” The 9th and 10th questions regarded the effect of environment on changing category membership: “If a baby is born and raised in a White/Black family, is it possible that he will be Black/White?” Finally, the 11th and 12th questions regarded the inheritance of category membership: “Is it possible that a White/Black mother will give birth to a Black/White baby?” Response options for questions 7–12 were: possible, maybe possible, or impossible, assigned scores of 1, 2.5, and 4, respectively. Thus the higher the score, the more strongly essentialist the child’s answer was. Children completed these questions for each of the three categories, with slight adjustments. A single essentialism score was calculated as the average of all questions.

1.2. Results

1.2.1. Essentialism questions

We first assessed whether there were differences in the extent to which children held essentialist beliefs about the categories. We conducted a repeated-measures analysis of variance (ANOVA) with Condition (Label, No-label) as a between-subjects factor, and Category Type (gender, race, and shirt-color) as a within-subjects factor, using the mean essentialism scores (1–4) of each category as the dependent variables. The analysis revealed only a significant effect of Category type, $F(2, 60) = 24.17, p < .001, \eta^2 = .45$. Post hoc Bonferroni comparisons revealed differences across all three categories, with gender the most essentialized ($M = 3.52, SD = .48$), followed by race ($M = 3.34, SD = .55$), and shirt color ($M = 2.89, SD = .67$) ($p < .05$, one-tailed for the difference between gender and race).

1.2.2. Categorization task

We conducted a repeated-measures ANOVA on the likelihood of children selecting the Target Category, including children’s gender and Condition (Label, No-label) as between-subjects factors, and Similarity ratio (.12, .35, .6, and .82) and Category type (gender, race, and shirt-color) as within-subjects factors. There were effects of Condition, $F(1, 28) = 8.54, p < .01, \eta^2 = .23$, Similarity ratio, $F(3, 84) = 33.84, p < .001, \eta^2 = .55$, the two-way interaction between Category type and Similarity ratio, $F(6, 168) = 2.21, p < .05, \eta^2 = .07$, and the theoretically central three-way interaction among Condition, Category type, and Similarity ratio, $F(6, 168) = 2.96, p < .01, \eta^2 = .10$. There were no other significant effects.

We examined the three-way interaction by analyzing each Category type separately and conducting repeated-measures ANOVAs with Condition as a between-subjects factor and Similarity ratio as a within-subjects factor (see Fig. 2 for the pertinent data). For the gender category, there was only an effect of Similarity ratio, $F(3, 90) = 23.44, p < .001, \eta^2 = .44$, with children gradually more likely to place test items with the most similar category across both conditions. Condition did not have a significant effect ($p = .3$), indicating that children were unaffected by the label in categorizing mixed-gender test exemplars. For the category race, there was an effect of Similarity ratio, $F(3, 90) = 17.69, p < .001, \eta^2 = .37$, and although the effect of condition did not reach significance ($p = .09$), it was more marked than the one for gender (Fig. 2). Finally, for the category shirt-color, all effects were significant: Similarity ratio, $F(3, 90) = 6.20, p < .005, \eta^2 = .17$, Condition, $F(1, 30) = 11.37, p < .005, \eta^2 = .28$, and the interaction between the two, $F(3, 90) = 3.80, p < .05, \eta^2 = .11$. Independent samples t-tests on each ratio of similarity revealed that Condition had a significant effect only on the lower two ratios, with children more likely to assign a .12 and .35 test exemplar to the Target Category in the Label than in the No-label condition – $t(30) = 5.21, p < .001$, and $t(30) = 2.24, p < .05$, respectively for .12 and .35 ratios.

As a final analysis, we conducted one-sample t-tests comparing performance to chance selection (.5), for each Category type by Similarity ratio. The results appear in Fig. 2 (marked by asterisks).
Perhaps the most telling pattern is the one regarding test exemplars at the .12 ratio of similarity – the ratio with the starkest conflict between the two categorization cues in the Label condition. Not surprisingly, given that in the No-label condition there was no conflict between cues, .12 exemplars from all three categories were systematically assigned to the Alternative Category (i.e., were assigned to the Target Category significantly less often than expected by chance). The interesting differences among categories were manifested in the Label condition. In the gender category, .12 exemplars...
were – as in the No-label condition – systematically categorized with the Alternative Category. In other words, children ignored the label, and categorized by perceptual similarity. For race, however, assignment of the .12 exemplars was at chance. In other words, the presence of labels moderated the pull of perceptual similarity. Finally, .12 exemplars in the shirt-color category were systematically assigned to the Target Category (i.e., were assigned to the Target Category significantly more often than expected by chance). In other words, the labels led children to switch their default categorization decision.

1.2.3. Inference task

Participants’ responses were scored according to the degree to which they thought the test exemplar would prefer the activity of the Target Category (range: 1 = exclusively consistent with Target Category, 5 = exclusively consistent with Alternative Category). Thus, a lower score represented assignment according to the Target Category. A repeated-measures ANOVA including gender and Condition (Label, No-label) as between-subjects factors, and Similarity ratio (.12, .35, .6, and .82) and Category type (gender, race, and shirt-color) as within-subjects factors was conducted on these scores.

The analysis revealed an effect of Similarity ratio, \( F(3, 84) = 11.28, p < .001, \eta^2 = .29 \), such that the more similar the test exemplar was to the Target Category, the more likely children were to infer that it would have the same preference as the Target Category (\( M_{.12} = 3.06, SE = .22; M_{.35} = 2.66, SE = .23; M_{.6} = 2.37, SE = .21; M_{.82} = 1.90, SE = .19 \)). There was also an effect of Condition, \( F(1, 28) = 5.38, p < .05, \eta^2 = .16 \), with children in the Label condition (\( M = 2.05, SD = 1.06 \)) more likely than those in the No-label condition (\( M = 2.90, SD = .79 \)) to infer that test exemplars would have the same preference as the Target Category. In other words, across all categories, a common label persuaded children that people would have the same preferences. There were no other significant effects.

1.3. Discussion

The findings on essentialism confirmed results from previous studies showing that children hold stronger essentialist beliefs about gender than they do about race. Although expected, this finding is noteworthy given that the measures used in the various studies differ substantially. Here children were asked a series of direct questions about aspects of essentialism, and in other studies children were asked to determine whether gender or race categories reflect universal objective truths (Diesendruck et al., 2013; Rhodes & Gelman, 2009). The finding of consistency, both within and across cultures, suggests that the measures indeed cohere in assessing children’s beliefs about the categories.

A further finding was that in the absence of conflicting information, children made category assignments and drew inferences, based systematically on the perceptual similarity between test exemplars and the target categories. In this sense, children responded to social categories analogously to how they have been shown to respond to categories in other domains (Sloutsky & Fisher, 2004; Smith, 2005).

Results for the inference task revealed that, different from the categorization task, labeling had a similar effect across all three social categories. Children were more likely to infer that test exemplars would share the preferences of a category if the category had the same label as the test exemplars than if it did not. This finding is consistent with previous work on category-based induction showing the effect of labels both in social (Diesendruck & HaLevi, 2006; Heyman & Gelman, 2000), and other domains (Gelman & Markman, 1986). In the General Discussion, we return to possible explanations for the lack of category type effect in this task.

The most central findings, however, pertain to the categorization task. This task, a replication of the one used in the domains of animals and artifacts by Diesendruck and Peretz (2013), allowed us to compare social categorization patterns to the patterns in these domains. We found that the weighting of perceptual and linguistic cues to category assignment was not uniform but instead varied corresponding to the degree of essentialism of the categories. The least essentialized category – shirt-color – was treated as the least grounded on physical information and thus the most open to linguistic revision. In fact, children completely disregarded perceptual dissimilarity, “blindly” following the labeling cue. At the intermediate level of essentialism of the category race, exemplars were moderately subject to the effect of labeling. Finally, the most essentialized category – gender – was treated as the most
grounded on physical cues and thus the least tolerant of perceptual dissimilarity in favor of a linguistic criterion. If someone looked like a man but was referred to as a “woman”, children categorized the person as a man.

It could be argued that this pattern reflects not different degrees of essentialism, but rather different degrees of confidence. Namely, perhaps children were more confident about their intuitive categorization of gender than of shirt-color, and thus were more likely to ignore the conflicting labels for the former than for the latter category type. We doubt this possibility for two main reasons. First, if this had been the case, we would expect children to behave similarly in the inference task. There, however, we found that children’s inferences regarding gender were as strongly determined by labels as their inferences about shirt-color. Second, Diesendruck and Peretz’s (2013) study showed that even though animal and artifact categories were equally unfamiliar, children’s categorization in these two domains was differently influenced by labeling – thus supporting the idea that it was only the degree of essentialism that moderated their responses.

The results thus show that children construed gender categories akin to the way in which they construe animal categories, and construed shirt-color categories akin to how they construe artifact categories. The former are essentialized, and as such are believed to capture natural and physical discontinuities in the social continuum. The latter are not as essentialized and are thus believed to be culturally constructed and indexed by conventional labels. For Israeli 5-year-olds, racial categories fall between the two.

2. Study 2: if it looks like a man but thinks like a woman, what is it?

Study 1 assessed children’s essentialist beliefs about social categories by evaluating the extent to which a cultural (extrinsic) factor – labels – could alter children’s appearance-based categorization. Study 2 assessed the same question, but by evaluating the complementary process, namely, the extent to which an arguably essential (intrinsic) factor could alter appearance-based categorization.

A number of studies using various methodologies have shown that children treat internal properties of animals – more so than of artifacts – as definitional of category membership (Diesendruck et al., 1998; Gelman & Gottfried, 1996; Gelman & Wellman, 1991; Keil, 1989). Diesendruck and Peretz (2013) found that internal properties had a stronger effect on 5-year-olds’ assignment of category membership for animals than for artifacts. The conclusion drawn from these studies is that internal properties tap children’s essentialist beliefs about animals. Study 2 asks whether internal properties have a similar effect on social categories.

One important modification of the work on animal concepts incorporated in Study 2 has to do with the type of internal properties. It has been argued that given that children conceptualize animals as biological kinds, the type of internal properties that may tap their essential qualities are biological properties, having to do with physiology and anatomical structure (Keil, 1995). Although there is not sufficient research assessing this aspect of children’s conceptualization of social categories, there are some indications that the most central internal properties are ones having to do with psychological characteristics, namely behavioral roles (Kalish & Lawson, 2008) or beliefs and preferences (Diesendruck & Eldror, 2011). Thus in Study 2, we provided children with internal psychological characteristics of the target categories and test exemplars. Moreover, in Study 2, we did not provide any labels for the target categories or test exemplars.

2.1. Method

2.1.1. Participants

Thirty-two 5-year-olds (Mage = 65.81 months, SD = 3.67; 19 boys) participated. None had participated in Study 1. Participants’ demographics and recruitment were similar to those in Study 1.

2.1.2. Design

As in Study 1, the design was mixed with Condition (Internal-information vs. No-information) as a between-subjects factor and Category type (gender, race, and shirt-color) and Similarity ratio (.12, .35, .6, .82) as within-subjects factors. Participants were assigned randomly to one of the conditions.

Please cite this article in press as: Diesendruck, G., & Weiss, E. Children’s differential weighting of cues to social categories. Cognitive Development (2014), http://dx.doi.org/10.1016/j.cogdev.2014.06.001
Study 2 included only the Categorization and Inference tasks as assessments of essentialist beliefs were available from Study 1.

2.1.3. Materials

The stimuli were the same as those used in Study 1.

2.1.4. Procedure

The procedure for both tasks was the same as in Study 1, with the exception that instead of being told the labels of the target categories and the test exemplars, children were told about internal properties they had. As in Study 1, every participant completed 12 Categorization-then-Inference trials, four for each category type. In the Categorization task, E presented the two categories, saying, for instance: “See these people [pointing to the Target Category]; they think it’s important to eat food from the forest. And see these people [pointing to the Alternative Category]; they think it’s important to eat food from the sea.” Then revealed one of the test exemplars for the set, and in the Internal-information condition said: “Look at a picture of someone who thinks it’s important to eat food from the forest. Can you return it to the group of the same kind? Should it go with the people who think it’s important to eat food from the forest or with the people who think it’s important to eat food from the sea?” As in Study 1, the test exemplar was given the same internal information as the Target Category. Thus, in trials of low-similarity ratios (e.g., .12, .35), internal information and similarity provided conflicting cues about the category membership of the test exemplar, as the information matched the category least similar to the test exemplar. In trials of high-similarity (.6, and especially .82), information and similarity cohered. In the No-information condition, E pointed to the test exemplar and said: “Look at this picture. Can you return it to the group of the same kind? Should it go with the people who think it’s important to eat food from the forest or with the people who think it’s important to eat food from the sea?” Thus, in the No-information condition, there were no conflicting cues, as children only had similarity as a cue to category membership.

After recording the participant’s dichotomous selection (either verbal or pointing), E started the Inference task related to that test exemplar. For instance, E said: “People in those countries who think it’s important to eat food from the forest, like playing with a donga. See, this is a donga [pointing to a picture of a novel object]. People in those countries who think it’s important to eat food from the sea, like playing with a flerit. See, this is a flerit [pointing to a picture of a different novel object].” The experimenter then presented again the test exemplar and asked, “What do you think this [person who thinks it’s important to eat food from the forest/person] most likes to play with? Play only donga? Play donga a lot and flerit a bit? Play both donga and flerit? Play donga a bit and flerit a lot? Play only flerit?” Children in the Internal-information condition were told that information about the test exemplar, and children in the No-information condition were not. Children’s responses were coded as in Study 1.

Ordering and counterbalancing were as described in Study 1. The other two pairs of internal properties used were “people who think the moon/sun is important for our health” and “people who think it’s important to read books/do sports.”

2.2. Results

2.2.1. Categorization task

As in Study 1, participants’ responses were scored according to whether or not they selected the Target Category. A repeated-measures ANOVA including children’s gender and Condition (Internal information, No-information) as between-subjects factors, and Similarity ratio (.12, .35,.6, and .82) and Category type (gender, race, and shirt-color) as within-subjects factors, revealed effects of Condition, $F(1, 28) = 21.81, p < .001, \eta^2 = .44$, Similarity ratio, $F(3, 84) = 17.19, p < .001, \eta^2 = .38$, and the two-way interaction between Condition and Similarity ratio, $F(3, 84) = 7.63, p < .001, \eta^2 = .21$. There were no other significant effects.

We followed-up on the two-way interaction by analyzing each Condition separately, conducting repeated-measures ANOVAs with Similarity ratio as the within-subjects factor, and selection of the Target Category across Category types as the dependent variable. In the No-information condition,
there was an effect of Similarity ratio, \(F(3, 45) = 17.78, p < .001, \eta^2 = .54\). As expected, children were gradually more likely to place test items with the most similar category \((M_{12} = .5, SE = .22; M_{35} = 1.13, SE = .20; M_{6} = 1.75, SE = .25; M_{82} = 2.31, SE = .24)\). In contrast, in the Internal-information condition, the effect of Similarity ratio was not significant. In other words, children were as likely to assign to the Target Category a test exemplar with a .12 similarity ratio to that category as they were a test exemplar with a .82 similarity ratio.

As a further analysis of the above interaction, and in order to allow for a direct comparison with the results of Study 1, we compared the performance of children in each of the conditions to chance selection (chance = .5), for each Category type by Similarity ratio. The results appear in Fig. 3 (marked with asterisks). As seen there, in the No-information condition, children assigned exemplars at the .12 ratio of similarity to the Target Category at rates significantly lower than that expected by chance – a pattern that held for all three Category types. In the Internal-information condition, children assigned exemplars at all ratios of similarity to the Target Category at rates significantly higher than that expected by chance. Interestingly, the only case in which children in the Internal-information condition performed at chance levels was for the test exemplar at .12 similarity ratio in the category of shirt-color. In other words, for almost all test exemplars, children completely switched their baseline tendency and determined category membership based on the internal information.

### 2.2.2. Inference task

Participants’ responses were scored as in Study 1. A repeated-measures ANOVA including children’s gender and Condition (Internal-information, No-information) as between-subjects factors, and Similarity ratio (.12, .35, .6, and .82) and Category type (gender, race, and shirt-color) as within-subjects factors was conducted. Effects of Similarity ratio, \(F(3, 84) = 11.78, p < .001, \eta^2 = .30\), and of the interaction between Condition and Similarity ratio, \(F(3, 84) = 3.32, p < .05, \eta^2 = .11\), were significant. There were no other significant effects.

We followed-up on the two-way interaction by analyzing each condition separately and conducting repeated-measures ANOVAs with Similarity ratio as the within-subjects factor and inference score across Category types as the dependent variable. Fig. 4 portrays this interaction. Similar to the findings from the Categorization task, in the No-information condition, there was an effect of Similarity ratio, \(F(3, 45) = 15.54, p < .001, \eta^2 = .51\), such that the more similar the test exemplar was to the Target Category, the more likely children were to infer that it would have the same preference as the Target Category \((M_{12} = 3.44, SE = .22; M_{35} = 3.00, SE = .30; M_{6} = 2.42, SE = .26; M_{82} = 1.83, SE = .15)\). In contrast, and again analogous to the findings from the Categorization task, in the Internal-information condition, the effect of Similarity ratio was not significant. In other words, children were as likely to infer that a test exemplar with a .12 similarity ratio would have the preference attributed to the Target Category, as they were a test exemplar with a .82 similarity ratio.

### 2.3. Discussion

As in Study 1, in the absence of conflicting information, children in Study 2 systematically categorized and drew inferences based on the perceptual similarity of test exemplars to the target categories. This consistency notwithstanding, in Study 2 we found slightly different effects of the conflicting cue – namely, internal property information – on children’s categorization and induction decisions, than the ones found in Study 1.

First, internal property information had an identical effect in both tasks. In both, internal property information significantly swayed children away from relying exclusively on perceptual similarity, irrespective of the category type. It might be that by defining the categories in terms of internal property information – rather than by familiar labels as done in Study 1 – we more directly tapped children’s abstract concepts of social categories, thus neutralizing potential knowledge-based differences. For instance, whereas in Study 1 we made the concepts of gender and race readily available and thus accessible to children’s decisions, in Study 2 we did not explicitly mark these concepts, emphasizing instead the novel properties.
A second and related difference between studies is that in Study 2 internal property information completely toppled the effect of perceptual similarity in almost all categories. Both in categorization and in induction, responses to exemplars of very low similarity to a target category were statistically indistinguishable from those of very high similarity. One could argue that children responded in such a way to the internal property information for pragmatic reasons. Namely, they might have misunderstood the task as implying that when internal information is provided, they should
categorize test exemplars based on that information. We dispute this possibility for two main reasons. First, the present study was modeled precisely after the study by Diesendruck and Peretz (2013), with children of similar ages and backgrounds. There it was found that different types of information affected children’s categorization differently depending on the type of category. In particular, whereas internal information had a stronger effect on the categorization of animals than of artifacts, intentional information had the opposite effect. In fact, intentional information had no effect on 5-year-olds’ categorization of animals, indicating that simply because a test exemplar was described as having the same property as a target category was not enough to induce children to categorize according to such a criterion. Second, in the present study we found a case in which internal information did not decisively lead children to ignore perceptual similarity – namely, in very low similarity exemplars in the shirt-color category. In other words, children relied the least on the internal information when it applied to the least essentialized categories, indicating that they weighed the internal and perceptual information differently depending on the type of category.

Taken together, these results reproduce those reported by Diesendruck and Peretz (2013) for the domain of animals. Children here took internal property information as definitional of social categories – especially gender and race – suggesting that these categories are essentialized by children. For essentialized categories, physical appearances are manifestations of causally inherent properties. Thus, although in the absence of information about inherent properties, physical appearances serve as valid cues to category membership – as in fact demonstrated for gender in Study 1 – when present, information about inherent properties takes precedence over physical appearance.

3. General discussion

The present studies were designed to assess the relative weights of physical appearance, labels, and information about internal properties in 5-year-olds’ categorization and inferences in the social domain. Drawing from work on these questions in the domains of animal and artifact categories, the hypotheses were that to the extent a social category is essentialized, (a) labels should not overtake perceptual similarity in determining category assignment, and (b) internal property information should overtake perceptual similarity in determining category assignment. The findings from the present two studies mostly confirmed these hypotheses.

In Study 1, the degree to which social categories were essentialized modulated the extent to which category assignment was affected by labeling. The most essentialized category – gender – was quite impervious, with children making practically the same kinds of assignments irrespective of whether

Fig. 4. Mean scores in the Inference task of children in the no-information and internal-information conditions at each ratio of similarity, across Category types, in Study 2.

Please cite this article in press as: Diesendruck, G., & Weiss, E. Children’s differential weighting of cues to social categories. *Cognitive Development* (2014), http://dx.doi.org/10.1016/j.cogdev.2014.06.001
test exemplars had the same label as the target category. In turn, the least essentialized category—shirt-color—was highly susceptible to the effect of labeling. Children in the Label condition completely switched from perceptually based category assignments—displayed by their peers in the No-label condition (Fig. 2). This difference between types of categories is analogous to the difference reported by Diesendruck and Peretz (2013) regarding animals and artifacts. There, whereas animal category assignment was resistant to the effect of labels, being determined instead primarily by the degree of perceptual similarity between test exemplars and target categories, artifact category assignment was susceptible to the effect of labeling.

Explanation of the difference between animals and artifacts was couched in terms of how categories in these domains are conceptualized (Gelman, 2003; Keil, 1995). Specifically, the claim is that animal categories are essentialized, believed to represent natural kinds and defined by intrinsic properties. As such, the physical appearance of an animal is believed to be a direct index of its essential nature, and extrinsic factors such as labels can only minimally affect categorization. Artifacts in turn, are believed to represent conventional kinds, defined by relational properties such as how people use, refer, or intend to construe them. In this case, labels are one of the features that in fact define artifact kinds. This general argument about the relation between essentialism and the weight of labels in categorization received substantive support in the present studies. The direct assessment of children’s essentialist beliefs about the different social categories corresponded to how susceptible to the effect of labeling the categories were.

This interaction between degree of social essentialism and the effect of labels also relates to the social sciences literature on the origins of social categories (Haney Lopez, 2006; Omri & Winant, 1994; Rothbart & Taylor, 1992). For children, certain social categories (e.g., shirt–color) are indeed cultural constructions, arbitrary to the point that labels can completely violate perceptual similarity. Others (e.g., gender), however, are highly ingrained, reaching the conceptual status of natural kinds. In fact, this analysis intimates that whereas novel social categories are treated like “artifacts”, highly familiar ones are treated like “animals”. In this regard, it is interesting that for the Israeli children in the present study, the culturally less salient category of race fell somewhere between these two extremes. Given differences between Israel and the U.S. in the conceptual status of race (Diesendruck et al., 2013), it would be interesting to examine how North American children would respond in our task.

Findings for the induction task in Study 1 nonetheless did not show an interaction between category type and the effect of labels. Instead, the main finding was that labels affected children’s inferences uniformly across social categories. A possible explanation for this lack of consistency may have to do with the different demands of categorization versus induction (Gelman, 2003). In categorization, participants are provided with a variety of cues and they have to decide how to weigh them in their decision about category assignment. In induction, the category assignment is given and participants have to decide what to infer based on it. Thus, whereas labels served as one of a few potential cues to category assignment in the categorization task, labels set the category based on which children were asked to draw inferences, thus leaving less room for the assignment to be disputed.

This analysis resonates with the notion that children rely on a division of linguistic labor when it comes to understanding novel exemplars (Malt, 1990; Putnam, 1975). Work with adults shows that deference to linguistic assignment is influenced by the task and the categories involved (Braisby, 2001; Proctor & Keil, 2006). Along these lines, future work could examine whether manipulating how a labeler is described—e.g., as an expert or novice in a field—affects children’s deference to labeling. An additional suggestion for future work is to vary the order of presentation of the tasks.

The inconsistency between categorization and induction has also been found in the domain of animals, with some induction studies finding labels to guide children’s inferences (Jaswal & Markman, 2007), and categorization studies finding only a slight effect of labels (Diesendruck & Peretz, 2013; Sloutsky & Fisher, 2004). Most pertinent to the present study, using a somewhat different task, Gelman, Collman, and Maccoby (1986) found that labels had a more substantial impact on children’s gender-based inferences than on gender categorization—thus replicating the present findings on gender.

The results of Study 2 presented a more straightforward pattern. Information about internal psychological properties of exemplars and categories determined responses in the two tasks and across all three social categories, to the point that children completely ignored information about perceptual similarity. As is strikingly apparent in Fig. 3, in the absence of internal property information, children
categorized only 17% of the exemplars highly dissimilar to a target category (i.e., at the .12 ratio) as members of that category. In contrast, when given internal property information, children did so for over 73% of these very exemplars.

Previous studies show that children expect members of the same social category to share internal psychological properties (Diesendruck & Eldror, 2011; Diesendruck & haLevi, 2006). Here we show that, like for animal categories, internal properties are definition of social category membership (Diesendruck & Peretz. 2013; Diesendruck et al., 1998; Gelman & Wellman, 1991; Keil, 1989) – so much so, in fact, that even if a category is novel and in principle open to revision – e.g., shirt-color – defining its internal psychological properties seems to have triggered children’s essentialization of that category. Thus, in a twist on the classic “color-blind” educational policy, when children essentialize social categories based on beliefs about their psychological distinctiveness, children become blind to the categories’ physical distinctiveness.

Taken together, the present two studies add further evidence on children’s essentialist construal of social categories. A highly essentialized social category – gender – was overall categorized based on the same “weights” typically applied for the categorization of natural kinds – animals. This is particularly salient when we compare results from the two studies (Fig. 2a and 3a). When a gender test exemplar was highly dissimilar from the Target Category (the .12 exemplar), the shared label did not drive children to move their assignment of the exemplar to that category at all. In turn, when that test exemplar was described as having the same internal properties as the Target Category, children moved from assigning that exemplar to the Target Category around 20% of the time to doing so over 90% of the time.

In sum, the present studies show that for 5-year-olds, if someone looks like a member of essentialized social category A, but is named with the label of social category B, that person is still a member of social category A. But if that person thinks like someone from category B, he or she indeed belongs to category B. This hierarchy in the weights of various cues to category membership is consistent with an essentialist construal of social categories. These categories are treated by children as representing objective and natural partitions of the social world, partitions grounded on inherent properties causally responsible for people’s physical appearance and thus impervious to cultural revisions.

Acknowledgments

This research was funded by grant# 672/00 awarded by the Israel Science Foundation to Gil Diesendruck.

References


G Model
http://dx.doi.org/10.1016/j.cogdev.2014.06.001


Please cite this article in press as: Diesendruck, G., & Weiss, E. Children’s differential weighting of cues to social categories. *Cognitive Development* (2014). http://dx.doi.org/10.1016/j.cogdev.2014.06.001