The Development of Social Essentialism: The Case of Israeli Children's Inferences About Jews and Arabs

Dana Birnbaum, Inas Deeb, and Gili Segall
Bar-Ilan University

Adar Ben-Eliyahu
Duke University

Gil Diesendruck
Bar-Ilan University

Two studies examined the inductive potential of various social categories among 144 kindergarten, 2nd-, and 6th-grade Israeli children from 3 sectors: secular Jews, religious Jews, and Muslim Arabs. Study 1—wherein social categories were labeled—found that ethnic categories were the most inductively powerful, especially for religious Jewish children. Study 2—wherein no social category labels were provided—found no differences across sectors either in the inductive potential of ethnic categories or in children's capacity to visually recognize social categories. These results stress the importance of labels and cultural background in children's beliefs about social categories. The implications of these findings for accounts of the development of social essentialism are discussed.

For the past few years, a growing body of literature suggests that people across the world conceive of social categories in a way akin to how they conceive of living kinds. Specifically, people believe that, like categories of animals (Gelman, 2003), human categories (Gil-White, 2001) have unique properties that are innate and causal and that membership in these categories is absolute, permanent, and inductively powerful—a belief Medin and Ortony (1989) termed "psychological essentialism." In pioneering work among North American children, Hirschfeld (1996) found that kindergarteners believe racial identity to be absolute, established at birth, and thus independent of upbringing environment (see also Taylor, 1996, on gender). Analogous findings have been reported on other cultures. For instance, Mahalingam (2003) showed that Indian Brahmin adults believe that a person's caste is inherited from birth parents, rather than acquired from adoptive parents, and Gil-White (2001) found a similar nativist construal of ethnic differences among Mongolian Torguuds (see also Astuti, Solomon, & Carey, 2004).

This cross-cultural and developmental consistency in children and adults' tendency to essentialize human kinds on the one hand, and the cultural specificity of the particular categories essentialized on the other, has led researchers to ponder about the origins of such beliefs (see, e.g., Atran, 1990; Hirschfeld, 1996; Sperber, 1996). Manifest in these accounts is the need to understand how social essentialism emerges and to what extent it is susceptible to cultural and linguistic variations. The present studies attempt to do exactly that. In particular, we investigate how one component of essentialism is manifested in children from various subcultural groups, toward the same social categories. As briefly reviewed above, the cross-cultural developmental literature on social essentialism consists of studies assessing different components of essentialism (e.g., induction vs. innateness), conducted on different cultural groups (e.g., the Vezo...
of Madagascar vs. North Americans), and targeting different social categories (e.g., ethnicity vs. race). One of the main goals of the present studies was to attempt to minimize variance in these three dimensions, so as to more decisively assess the potential effect of cultural input on children’s beliefs.

The component we chose to focus on is the inductive power of social categories, that is, the extent to which knowledge of the category to which an object belongs allows one to infer properties of the object. Gelman and colleagues have documented how from a young age children are sensitive to the unique inductive power of animal categories (see Gelman, 2003, for a review). In the domain of human categories, Heyman and Gelman (2000) found that children draw inferences based on characters’ personality traits rather than their physical appearance, and Diesendruck and haLevi (2006) found that for Jewish Israeli adults and kindergarteners, ethnicity is the most inductively powerful social category. The present studies follow up on these latter findings.

For historical and political reasons, the category “Arab” for Jewish Israelis and the category “Jew” for Arab Israelis are extremely salient and important. Studies reveal that even 3-year-old Israeli Jewish children are somewhat familiar with the category “Arab” and that negative stereotypes regarding Arabs grow stronger from kindergarten to high school (Bar-Tal, 1996; Pitner, Astor, Benbenishty, Haj-Yahia, & Zeira, 2003; Slone, Tarrasch, & Hallis, 2000; Teichman, 2001). Bar-Tal and Teichman (2005) argue that a potential source for the development of negative stereotypes is the portrayal of Arabs and Jews in the media in general, and school books in particular (see also Al-Haj, 2002; Soen, 2002).

Taking into account the cultural and ideological heterogeneity of Israeli society, particularly with respect to issues pertaining to ethnicity, in the present studies we tested children from three different subcultural “sectors” within the Israeli society: secular Jews, modern religious (Zionist) Jews, and Muslim Arabs. Our selection of the three subgroups was motivated both by the diversity of ideological commitments vis-à-vis ethnicity typified by these groups, and by attention to the representativeness of these groups within the Israeli society. Specifically, secular Jews represent the majority group in Israel, modern religious Jews manifest different values for religious and often political reasons, and Muslim Arabs constitute the minority about whom the focal majority stereotypes are held (see the Participants section for further details). Evidence for cross-group differences would reinforce the notion that social essentialism is influenced by the type of input children are exposed to.

In order to trace the development of these potential cultural differences, we tested children from three age groups: kindergarteners, second graders, and sixth graders. Studies in North America found that these developmental points capture the transition from intuitive to more adult-like beliefs about human kinds (Hirschfeld, 1996; Taylor, 1996). Thus, in contrast to previous cross-cultural studies on social essentialism (Astuti et al., 2004; Gil-White, 2001; Mahalingam, 2003), the present design allowed us to assess how children’s concepts change as they immerse into the institutionalized views exposed in schools.

Study 1 employed a methodology similar to the one used by Diesendruck and haLevi (2006). The experimenter presented children with pictures of two test characters, told children the category membership of these characters vis-à-vis two social categories (e.g., their ethnicity and gender), and informed children about a novel psychological property each test character had (e.g., their game preferences). The experimenter then presented to children a target character that resembled one of the test characters on one of the categories but resembled the other test character on the other category. Children were asked which of the two properties attributed to the test characters the target character possessed. The rationale was that children’s pattern of responses across a number of trials would be indicative of the inductive power of the different social categories tested. To the extent that children’s essentialist beliefs about certain social categories—especially ethnicity—is indeed susceptible to cultural input, we should find differences in the inductive power of these social categories across children from the different sectors. The age factor would allow us to assess how early these potential cultural differences appear.

A second major goal of the present studies was to continue investigating systematically the factors contributing to the inductive power of social categories, and their potential interaction with children’s cultural group. A number of studies found that language (Bigler & Liben, 2007; Gelman, Taylor, & Nguyen, 2004; Heyman & Diesendruck, 2002), and perhaps especially labels (Cimpian, Arce, Markman, & Dweck, 2007; Gelman & Heyman, 1999; Patterson & Bigler, 2006), might be implicated in the development of social essentialism. The present studies investigate the role of language in children’s essentialist reasoning by comparing children’s pattern of induction when
provided with category labels versus when only provided with physical appearance information about category membership. This comparison allowed us to verify to what extent appearance information about social category membership is sufficient to bring about potential differences among social categories. This examination is relevant to the broader question regarding whether children’s essentialization of human kinds is a “bottom-up” process by which children simply represent categories they perceive in the world, or a “top-down” process by which the categories are first in the child’s mind and then looked for in the world. The comparison is also relevant for understanding potential differences between cultural groups. Specifically, if a particular cultural group essentializes a certain social category, then will this bias be manifested only when labeling is present? In order to address these two related questions, Study 2 replicated part of Study 1, but whereas in Study 1 all the social categories were verbally labeled, in Study 2 none were.

Study 1

The goal of Study 1 was to assess the inductive power of various social categories across children from different cultural and age groups. For this purpose, we designed the induction task so as to both maximize the salience of social categories and allow a comprehensive assessment of their inductive potential. To these ends, the induction task included two parts. In the first—so-called relative power part—the two dimensions children had to decide between were both social categories. For instance, children had to decide whether to draw inferences based on the ethnic similarity between two characters, or their gender similarity. The maximization of the salience of social categories was operationalized by having all social categories represented by stereotypical visual markers associated with each social category, as well as their familiar category labels. The second—so-called absolute power part—was a direct replication of Diesendruck and haLevi (2006), and in it the two dimensions children had to decide between were personality traits (e.g., niceness) and social categories (e.g., ethnicity). This design feature augmented the comprehensiveness of our assessment, by widening the range of inferential contexts presented to children, including conflicts between different types of human categories. Moreover, Diesendruck and haLevi found that while kindergarteners had a preference for drawing inferences based on social categories, adults had a preference for drawing inferences based on personality traits. Thus, the inclusion of the absolute power part in the current study also allowed us to track in more detail the developmental changes in children’s reliance on social categories versus personality traits as bases for inferences.

Method

Participants

A total of 144 children participated in this study. The sample consisted of 16 children from each of three sectors (secular Jews, religious Jews, and Muslim Arabs), and each of three age groups (kindergarten, second grade, and sixth grade). Only children with signed parental permission participated in the study. Table 1 presents age and gender information about the participants. The secular and religious Jewish children were recruited from secular and religious schools and kindergartens in the same cities surrounding the university. These cities have mixed populations of secular and religious Jews, though few Orthodox Jews, and practically no Arabs. The Arab children were recruited from an Arab and primarily Muslim city (Baka al Garbah) within Israel. No Jews live in this city.

In general, the Israeli Jewish population is composed of three main religiosity subgroups: secular Jews (71%), religious Jews (21%), and Orthodox religious Jews (8%; Central Bureau of Statistics, 2006). We targeted secular and religious Jews for a variety of reasons. First, they represent the majority of the Jewish population. Second, both groups are socially, economically, and politically involved in Israeli life to similar extents, and to a greater extent than the Orthodox Jewish population. Third, despite these similarities, these two groups are denominationally distinguished, have substantially different belief systems and norms of behavior in everyday life, and are differentially distributed across the political spectrum (Sikron, 2004). Denominationally, Hebrew has two separate words, which are frequently used either as adjectives or count nouns, to refer to secular (chiloni) and religious (dati) Jews. Secular and religious groups diverge substantially in beliefs regarding God, the Jewish commandments, and afterlife, and in their practices, such as keeping the Sabbath and its commandments, eating Kosher food, celebrating the Jewish religious holidays, praying, and attending synagogues (Levy, Levinson, & Katz, 2002). Reflecting this divide, while most children from these two
Table 1
Participants’ Age and Gender Distribution in Study 1 and Study 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Sector</th>
<th>Kindergarten</th>
<th>Second grade</th>
<th>Sixth grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td>Muslim Arabs</td>
<td>8 girls, 8 boys</td>
<td>8 girls, 8 boys</td>
<td>13 girls, 3 boys</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M_{age} = 5 \text{ y } 6 \text{ m}$</td>
<td>$M_{age} = 7 \text{ y } 6 \text{ m}$</td>
<td>$M_{age} = 11 \text{ y } 6 \text{ m}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4 m, 4 y 10 m–6 y)</td>
<td>(5 m, 6 y 9 m–8 y 2 m)</td>
<td>(5 m, 10 y 9 m–12 y)</td>
</tr>
<tr>
<td></td>
<td>Secular Jews</td>
<td>5 girls, 11 boys</td>
<td>10 girls, 6 boys</td>
<td>12 girls, 4 boys</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M_{age} = 5 \text{ y } 5 \text{ m}$</td>
<td>$M_{age} = 7 \text{ y } 5 \text{ m}$</td>
<td>$M_{age} = 11 \text{ y } 7 \text{ m}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4 m, 5 y 1 m–6 y)</td>
<td>(4 m, 7 y 1 m–8 y 6 m)</td>
<td>(6 m, 10 y 10 m–12 y 3 m)</td>
</tr>
<tr>
<td></td>
<td>Religious Jews</td>
<td>9 girls, 7 boys</td>
<td>9 girls, 7 boys</td>
<td>14 girls, 2 boys</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M_{age} = 5 \text{ y } 4 \text{ m}$</td>
<td>$M_{age} = 7 \text{ y } 6 \text{ m}$</td>
<td>$M_{age} = 11 \text{ y } 6 \text{ m}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4 m, 5 y 1 m–6 y)</td>
<td>(4 m, 7 y 1 m–8 y 6 m)</td>
<td>(5 m, 11 y 1 m–12 y 2 m)</td>
</tr>
<tr>
<td>Study 2</td>
<td>Muslim Arabs</td>
<td>10 girls, 6 boys</td>
<td>8 girls, 8 boys</td>
<td>7 girls, 9 boys</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M_{age} = 5 \text{ y } 5 \text{ m}$</td>
<td>$M_{age} = 7 \text{ y } 6 \text{ m}$</td>
<td>$M_{age} = 11 \text{ y } 7 \text{ m}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4 m, 4 y 11 m–6 y 2 m)</td>
<td>(5 m, 6 y 9 m–8 y 2 m)</td>
<td>(5 m, 10 y 9 m–12 y 4 m)</td>
</tr>
<tr>
<td></td>
<td>Secular Jews</td>
<td>6 girls, 10 boys</td>
<td>8 girls, 8 boys</td>
<td>13 girls, 3 boys</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M_{age} = 5 \text{ y } 6 \text{ m}$</td>
<td>$M_{age} = 7 \text{ y } 10 \text{ m}$</td>
<td>$M_{age} = 11 \text{ y } 8 \text{ m}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4 m, 4 y 11 m–6 y 2 m)</td>
<td>(4 m, 7 y 4 m–8 y 4 m)</td>
<td>(4 m, 10 y 10 m–12 y 3 m)</td>
</tr>
<tr>
<td></td>
<td>Religious Jews</td>
<td>7 girls, 9 boys</td>
<td>13 girls, 3 boys</td>
<td>9 girls, 7 boys</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$M_{age} = 5 \text{ y } 5 \text{ m}$</td>
<td>$M_{age} = 7 \text{ y } 6 \text{ m}$</td>
<td>$M_{age} = 11 \text{ y } 10 \text{ m}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(6 m, 4 y 6 m–6 y)</td>
<td>(7 m, 7 y–8 y 6 m)</td>
<td>(5 m, 11 y 2 m–12 y 5 m)</td>
</tr>
</tbody>
</table>

Note. $y = \text{ years; m = months (SD, range)}$.

$n = 16 \text{ per cell}$.

In contrast to Hebrew, Palestinian Arabic does not have separate count nouns for “religious” and “secular” Muslims—all are referred to simply as “Muslims.” While there is variation in the degree of observance of religious practices, the majority (approximately 70%) of the Muslim population identifies itself as “religious” and rigorously observes the basic Islamic commandments, such as the belief in Allah and the prophet Mohamed, providing for the needy, praying five times a day, and fasting during the month of Ramadan. Moreover, less observant Arab Muslims nonetheless respect and attempt to incorporate some of these traditions into their daily lives (Stendel, 1996). Smooha (1989) observes that while since the establishment of the State of Israel in 1948, Arab Muslims have been undergoing a process of “westernization,” they still maintain their traditions and values, especially in cities and villages populated exclusively by Arab Muslims. The children recruited for the present study reside in such a city. Baka al Garbiah is located in an area of Israel referred to as the “Triangle,” is one of the largest Arab cities in Israel (over 30,000 inhabitants), and is almost exclusively Muslim. Over 90% of the children in this city attend public schools, which in contrast to the Jewish sector, all belong to the same pedagogical branch. These schools are administered by the Israeli sectors attend the national public school system, they each attend the administratively and pedagogically separate branch that caters to each sector. In other words, the vast majority of students who attend secular schools come from secular families, and the vast majority of students who attend religious schools come from religious or traditional families. All major cities in Israel have public schools from each kind, and they are identifiable by their official names. One example of a major distinction between the two systems has to do with the extent of and approach toward teaching the Old Testament. In the secular system, the Old Testament is taught as a text recounting the myths and traditions of ancient Jews. In the religious system, the Old Testament is taught as a veridical account of the history of the world and the Jewish people, and the traditions are presented as normative rules to be followed presently as well. Lastly, these two sectors are characterized by distinct political orientations, especially toward the Israeli–Arab conflict. Namely, while the secular population endorses a variety of political opinions vis-à-vis the conflict, modern religious Zionist Jews tend toward opposing or stalling the creation of a Palestinian state in the West Bank and Gaza.

The Arab population in Israel is primarily Muslim (about 90%; Central Bureau of Statistics, 2006).
Ministry of Education, and they incorporate Islamic studies in their curricula. For instance, the schools celebrate all Muslim holidays, and there are at least 2 hr a week of classes on Islamic religion and history. Even though there are different Arab groups in Israel (e.g., Christians), for the sake of brevity, we refer to our Muslim Arab sample as “Arabs.” In fact, while young Jewish Israeli children may be familiar with the term “Arab,” it is unclear to what extent they differentiate between Muslims, Christians, Bedouins, or Druze.

**Design**

Study 1 consisted of two parts. In the first—“relative power”—part, children had to complete a series of 12 trials in which they had to draw inferences based on one of two possible social categories. This was named the relative power part because it assessed the inductive power of each social category relative to the other social categories. In the second—“absolute power”—part, children had to complete a series of 8 trials in which they had to draw inferences based either on a social category or a personality trait. This was named the absolute power part because in it, children did not have to choose between competing social categories but rather between each social category and a personality trait. The relative power part was always presented first, so as to allow an unbiased comparison between children’s responses to this part across the two studies (to be described later).

The social categories used in both parts were: ethnicity (Jew/Arab), gender (boy/girl), social status (rich/poor), and religiosity (religious/secular). The four personality traits used in the absolute power part were “niceness” (nice/not nice), “shyness” (shy/friendly), “creativity” (creative/not creative), and “activeness” (active/guile). These categories and traits are salient in the Israeli society, and have been found to be familiar to Israeli kindergarten children (Diesendruck & haLevi, 2006).

**Materials**

In each trial, children were shown a triad of line-drawn black and white human characters: a target character and two test characters. Each character was professionally drawn with physical markers that are typical of specific social categories. In the relative power part, all three characters within a triad belonged to the same two social categories, but each was assigned a specific value of these categories (e.g., Arab or Jew for the category ethnicity).

All possible combinations of social categories were presented to children, resulting in six types of triads: ethnicity/gender, ethnicity/social status, ethnicity/religiosity, gender/social status, gender/religiosity, and social status/religiosity. In each triad, each of the test characters resembled the target character on one of the social categories but was different from the target character on the other social category. For example, in a triad contrasting the social categories ethnicity and gender (as in Figure 1), one of the test characters was drawn as an Arabic boy, the other test character as a Jewish girl, and the target character as an Arabic girl. In this example, the test character “Arabic boy” resembled the target character (“Arabic girl”) with regard to ethnicity (i.e., it was an “ethnicity match”) but differed from her with regard to gender. In turn, the test character “Jewish girl” resembled the target character with regard to gender, but differed from her with regard to ethnicity (i.e., it was a “gender-match”).

The visual cues consisted primarily of dressing attire and accessories typical of the social category. All Jewish characters had on their shirt or as a necklace the Star of David sign and a letter in Hebrew. Secular Arab characters wore on their shoulders Kaffiyah (a Palestinian scarf) and had a letter in Arabic on their shirt. Although there are no consistent visual cues that differentiate secular Arabs from secular Jews within Israeli society, the visual cues described above are typical markers of each ethnicity. In turn, there are distinct markers for both religious/practicing Jews and Muslims, and these were represented in the pictures. Specifically, male religious Jewish characters wore a yarmulke, fringes that come out of the shirt (tzitziot), and side-locks (peyot). Female religious Jewish characters wore long dresses and had a head cover. Male religious Muslims wore a long robe and a skullcap. Female religious Muslims wore a head cover (hijab) and a long dress (jilbah). Rich characters wore fancy clothes and jewelry, and held expensive accessories. Poor characters wore ragged clothes and torn shoes or sandals. When the character was “religious” without a specific ethnicity, for example, a “religious boy” or “religious rich,” it was depicted as a religious Jew for participants in the Jewish sector and as a religious Muslim for participants in the Arab sector. Figure 2 provides examples of the pictures used here.

In order to emphasize the relevant social categories in each triad in either part of the study, the physical characteristics of the characters not indicative of their social category membership were made...
irrelevant, either by equating them (e.g., height) or by randomizing them across all characters (e.g., hair color). Moreover, in all triads in which gender was not a target social category, the gender of all characters was the same.

Each triad (e.g., ethnicity/gender) was presented to each participant on two separate trials. Each time, the target character was of a different value (e.g., for one child, on one of the two trials with this triad, the target character was an Arabic/girl, and on the other trial, the target character was a Jewish/boy). All possible target characters composed of the crossing of all social categories were presented an equal number of times across children (e.g., another child would get the target characters Arabic/boy and Jewish/girl, for the ethnicity/gender triad). The pairings of characters were randomly assigned to participants, and they appeared the same number of times across children.

The absolute power part of the study resembled the relative power part, except that each character was described by a social category and a personality trait rather than by two social categories. One of the test characters was presented to children as being of the same social category as the target but as having a different personality trait (social category match), while the other test character was described as having the same personality trait as the target character but as belonging to a different social category (personality trait match). Given that personality traits cannot be depicted visually in a conventional and exclusive manner, they were presented only via verbal labels. The visual cues indicating the social categories were the same as in the relative power part. All possible pairings of social category and personality trait were used for each child. In other words, for each child, each social category (e.g., social status) was paired with two different personality traits, such that each social category value (e.g., rich and poor) was paired with a different personality trait. The pairings were randomly assigned to participants, and they appeared the same number of times across children.

For both parts, all the properties attributed to the characters were novel, so that children could not base their answers on stigmas tying particular properties to categories. The properties had to do with preferences, patterns of behaviors, aspirations,
and physiological properties, and were framed within sentences that allowed children to understand their general meaning. There were 20 different novel properties (12 for the relative power part of the study and 8 for the absolute power part), so that on each trial children were asked to draw inferences about a different property. The novel words were Hebrew-sounding words for the Jewish children (see the list below) and Arabic-sounding words for the Arab children. The full list of properties was: eats Dabo/Fifo for lunch, plays Zigo/Zabo at the school yard, pretends that s/he is a Talan/Falkan, listens a lot to Rendit/Mishnit music, in her/his school s/he is best in Salmon/Takton, wants to be a Mashitz/Nagim when s/he grows up, believes that in the moon there is Tugit/Pushrit, thinks it is important to be Manil/Gakish, thinks that Vitot/Bidot are the nicest animals on earth, believes that Dushon/Pitzug is the most beautiful place on earth, enjoys to read books about Tushib/Masrib, loves to watch TV programs about Dulon/Sulon, feels happy when s/he is in Nunga/Smuda, is afraid of Namtagim/Natgarim, gets angry when someone says s/he is Taluf/Mitlon, has blood type Papuch/Tirel, has Lidon/Kiron in her/his stomach, feels bad when s/her eats Hamgal/Lamdash, the smell of Miker/Lakanf makes her/him sneeze, has too little Vitamin G/R in her/his body.

Procedure

A female experimenter tested participants individually in a quiet area of their kindergarten or school. All Jewish children were tested in Hebrew by Jewish research assistants, and all Arab children were tested in Arabic by Muslim Arab research assistants from the same area that children were from. The relative and absolute power parts were conducted one after the other, with a brief intermission between them when testing kindergarteners. The procedure lasted between 15 and 20 min.

Relative power part. The procedure started with the experimenter placing one of the test pictures of the first trial in front and to the right of the child, telling the child the character’s two social categories, and then informing the child about a novel property this test character had. In the same manner, the experimenter placed the second test character to the left of the child, assigning to this character his or her corresponding social categories and a different value for the novel property. The experimenter then placed the target character in between the two test characters, and slightly closer to the child. The experimenter told the child the target character’s two social categories, explicitly noting the similarities to each of the test characters. Finally, the experimenter asked the child to infer whether the target character had the same property as the first test character or as the second test character. Figure 1 depicts a schematic example of the procedure. Once children responded by pointing to one of the test characters, the experimenter removed all the pictures, and brought out the next set of pictures. The experimenter avoided giving any feedback to the child that could be interpreted in terms of accuracy in their performance.

There were 12 trials in this part: 2 trials with each triad type. The order of presentation of the triads for each participant was random and was counterbalanced across participants within a sample (i.e., there...
were 16 different presentation orders). The right-left placement of the test characters, and consequently the order of presentation of the test characters, was also counterbalanced within and between subjects. For instance, for the ethnicity/gender triad, on one trial the gender match was presented first and was placed to the right of the child, and on the other trial, the gender match was presented second and placed to the left of the child. Finally, the order in which the social category labels were applied to each character was also counterbalanced within and between subjects. For instance, on one trial of the ethnicity/gender triad, the experimenter first noted the ethnic membership of the character and then the gender, and on the other trial she did the opposite (e.g., “He is a boy, he is Jewish” vs. “He is Jewish, he is a boy”). It is important to note that in the sentences we used in Hebrew and Arabic, all the social category labels were morphologically and syntactically ambiguous as to whether they were adjectives or nouns—a fact that is representative of Hebrew and Arabic in general. Here, we translated the labels into the form that seemed most appropriate for the English construction.

Participants’ responses in the first part were scored according to the number of trials in which children drew inferences based on a particular social category. Given that each triad of social categories appeared twice for each participant, scores for each triad could vary from 0 to 2 (e.g., on ethnicity/gender triads, children could select the ethnicity match on 0, 1, or both of the trials). These Triad Relative Inference Scores were used in the analyses of children’s preferences for specific social categories. In order to evaluate potential differences across sectors and age groups, we summed all such scores for any given social category, creating Total Relative Inference Scores. Given that each social category appeared on three triads, with each triad appearing twice, the Total Relative Inference Score for each category could range from 0 (no inferences based on this category) to 2 (inferences based on this category in both trials involving this category).

### Results

#### Comparisons Across Social Categories

In order to compare the inductive power of social categories, we compared the frequency of children’s selections of the different social categories to chance, conducting separate analyses for each of the two parts of the study.

On the relative power part, to avoid violating an assumption of independence of measurements, and given that the Triad Relative Inference Scores ranged from 0 to 2, we conducted nonparametric tests assessing the distribution of children on each triad type. For instance, we counted the number of children, out of the total 144, who on the two ethnicity/gender triads picked the ethnicity match on both, only one, or none of the trials. The expected (chance) distribution was 25% (36 children), 50% (72 children), and 25% (36 children), respectively. These analyses revealed that the only triad types in which the distribution significantly deviated from chance were ones involving ethnicity. In particular, on the ethnicity/gender triads, significantly more children than were expected by chance selected ethnicity on both trials (distribution for both, one, and none, was 57, 62, and 25, respectively), $\chi^2(2, N = 144) = 17.00, p < .001$. A similar pattern was found on ethnicity/religiosity triads (distribution for both, one, and none, was 48, 75, and 21, respectively), $\chi^2(2, N = 144) = 10.38, p < .01$. On ethnicity/social status triads the pattern was in a similar direction, but not significantly different from chance (distribution for both, one, and none, was 46, 68, and 30, respectively), $\chi^2(2, N = 144) = 4.00, p = .13$. On all other triad types the distributions did not differ from chance, $ps > .2$.

Similar tests were conducted on the absolute power part, in which each social category appeared twice, each time paired against a different personality trait. The analyses showed that only ethnicity was selected more often than expected by chance (distribution for both, one, and none, was 37, 84, and 23, respectively), $\chi^2(2, N = 144) = 6.72, p < .05$.

Altogether, these findings reveal that across all sectors and ages, ethnicity trumped alternative
Comparisons Across Sectors and Age Groups

The second main goal of Study 1 was to compare the inductive power of various social categories across Israeli children from different ages and sectors. In order to maximize the representativeness of this assessment, we calculated a Total Inference Score for each category by adding each category’s Total Relative Inference Score (0–6) and Absolute Inference Score (0–2). These scores could thus range from 0 (no inferences based on the category) to 8 (all inferences based on the category).

We entered the Total Inference Scores for each social category into a multivariate analysis of variance (MANOVA), with age (kindergarten, second grade, and sixth grade) and sector (Arabs, secular Jews, and religious Jews) as between-subjects factors (see Figure 3 for means). The overall MANOVA revealed a significant effect of sector, \( F(8, 266) = 2.19, p < .05, \eta^2 = .06 \), but neither the effect of age nor the interaction between sector and age were significant, \( ps > .09 \). Looking at each social category separately, we found that sector had a significant effect on ethnicity, \( F(2, 135) = 4.21, p < .05, \eta^2 = .06 \). Tukey’s honestly significant difference (HSD) post hoc tests revealed that religious Jews made ethnicity-based inferences more than Arabs, \( p < .05 \). Sector also had a significant effect on social status, \( F(2, 135) = 3.39, p < .05, \eta^2 = .05 \), with Arab children drawing the most inferences based on social status and secular Jewish children the least (Tukey’s HSD, \( p < .05 \)). We also found that age had a significant effect on gender, \( F(2, 135) = 3.97, p < .05, \eta^2 = .05 \), with kindergarteners drawing the most inferences based on gender, and sixth graders the least (Tukey’s HSD, \( p < .05 \)).

In order to further evaluate children’s inferential tendencies, we compared the Total Inference Score of each social category, for each age group and sector, against chance (chance = 4). As can be seen in Figure 3, these analyses revealed distinct sectorial and developmental patterns of responses. Among Arab children, the only significant finding was that sixth graders drew fewer inferences based on gender than would be expected by chance. Among secular Jews, none of the Total Inference Scores, at any age, were significantly different from chance. Finally, among religious Jews, the pattern was quite different, as children at all ages made ethnicity-based inferences—and only ethnicity-based inferences—significantly more than would be expected by chance.

![Figure 3](image-url)

**Figure 3.** Mean Total Inference Scores for each social category, age group, and sector, in Study 1. 
*Note.* The horizontal bold line indicates chance value (4). Asterisks denote difference from chance, on t tests (15), \( p < .05 \).
Analyses of Group Membership

Given the structure of the stimulus set, the study also allowed us to explore an additional set of issues having to do with group membership. In particular, we analyzed whether there were differences in the pattern of participants’ inferences, based on whether the target of the inferences belonged to the same group as the participant. For this purpose, we conducted two sets of analyses, one looking at the ethnic membership of the participants (Jew or Arab), and another looking at the gender membership (boy or girl).

For the analyses on ethnicity, we created two new variables: (a) the number of trials in which children drew inferences based on ethnicity when the target character was a Jew and (b) the number of trials in which children drew inferences based on ethnicity when the target character was an Arab. Given the counterbalancing precautions noted before, for each participant, half of all eight trials involving ethnicity across the two parts of the study had a Jewish character as target, and half had an Arab character as target. Thus, these scores ranged from 0 to 4. For the analyses, we divided the participants into two sectors according to their ethnicity: Arabs and Jews (including both secular and religious). (Separate analyses on religious and secular participants generated the same results as the ones reported below.)

The above scores were entered into a repeated measures analysis of variance (ANOVA) in which target type (Jew, Arab) constituted a within-subjects factor, and ethnic sector (Jews, Arabs) a between-subjects factor. This analysis revealed only a significant interaction between target type and ethnic sector, F(2, 138) = 5.27, p < .05, $\eta^2 = .04$. To uncover the source of this interaction, we conducted separate paired t tests within each ethnic sector, comparing between the inference scores for each target type. These analyses revealed that while among Jewish participants, there was no difference between their tendency to draw inferences based on ethnicity when the target character was a Jew ($M = 2.49, SD = 1.19$) versus when the target character was an Arab ($M = 2.32, SD = 1.14$), $p > .3$, Arab participants were less likely to draw inferences based on ethnicity when the target character was a Jew ($M = 1.89, SD = 0.95$) than when the target character was an Arab ($M = 2.35, SD = 0.93$), $t(47) = 2.44, p < .05$. As can be seen by the means, the exceptional pattern of inferences stemmed from the Arab participants’ responses to Jewish target characters.

Analogous analyses were conducted with regard to gender. For these analyses, we created two new scores: (a) the number of trials in which children drew inferences based on gender when the target character was a boy (out of 4) and (b) the number of trials in which children drew inferences based on gender when the target character was a girl (out of 4).

These scores were entered into a repeated measures ANOVA in which target type (Boy, Girl) constituted a within-subjects factor, and participants’ gender (boys, girls) a between-subjects factor. This analysis revealed a significant effect of target type, $F(1, 138) = 8.88, p < .005, \eta^2 = .06$, such that children were overall more likely to draw inferences based on gender when the target character was a girl ($M = 2.11, SD = 1.19$), than when the target character was a boy ($M = 1.67, SD = 1.15$). This main effect was subsumed by a close to significant interaction between target type and participants’ gender, $F(1, 138) = 3.14, p = .079, \eta^2 = .02$. Given the exploratory nature of this analysis, and the analogy with the results on ethnicity, we followed up this analysis by again conducting separate paired t tests within each participants’ gender group, comparing between the inference scores for each target type. These analyses revealed that while among male participants, there was no difference between their tendency to draw inferences based on gender when the target character was a boy ($M = 1.89, SD = 1.15$) versus when the target character was a girl ($M = 2.00, SD = 1.25$), $p > .6$, female participants were less likely to draw inferences based on gender when the target character was a boy ($M = 1.53, SD = 1.9135$) than when the target character was a girl ($M = 2.18, SD = 1.16$), $t(47) = 3.93, p < .001$. As can be seen by the means, the exceptional pattern of inferences stemmed from the female participants’ responses to male target characters.

Comparison of Social Categories Against Personality Traits

A minor goal of Study 1 was to assess the development of children’s overall social category-based induction. Earlier studies conducted among secular Jews in Israel had found that while kindergarteners were more likely to draw inferences based on social category membership than based on personality traits, adults showed the reverse pattern (Diesendruck & haLevi, 2006). The absolute power part of Study 1 allowed us to assess
this developmental trend in more detail. For this purpose, we computed a Social Categories Inference Score by simply adding the Absolute Inference Scores across the four social categories. This score could range from 0 (all inferences based on personality traits) to 8 (all inferences based on social categories). We then conducted an ANOVA on these Social Categories Inference Scores, with age and sector as between-subjects factors. The only close to significant effect was of age, \( F(2, 135) = 2.81, p = .06, \eta^2 = .04 \). Given our interest in the development of this tendency, we compared the Social Categories Inference Score for each age group separately to chance (chance = 4). These analyses revealed that both kindergarteners \( (M = 4.67, SD = 1.50), t(47) = 3.07, p < .005, \) and second graders \( (M = 4.67, SD = 1.62), t(47) = 2.86, p < .01 \), drew inferences based on social categories significantly more than would be expected by chance. In turn, sixth graders did so at a level no different than chance \( (M = 4.02, SD = 1.56), p = .9 \). In other words, it seems that by sixth grade, children are midway toward the adult pattern of reliance on personality traits.

**Discussion**

The results of Study 1 reveal that, overall, ethnicity was the most powerful source of induction. Importantly, it appears that ethnicity was especially salient for religious Jewish children, and it became a powerful source of induction earlier for them than for children from the other two sectors. In other words, while there are some indications that by sixth grade, children from other sectors (e.g., Arabs) start construing ethnicity as the most powerful source of induction among a variety of human categories, religious Jewish children manifest this bias already by kindergarten.

The above conclusions raise questions about why is it that ethnicity is such a powerful source of induction for children in Israel, and why is it especially important for religious Jewish children. Some of the plausible responses to these questions are beyond the scope of the present studies, and will be raised in the General Discussion. Nonetheless, a direct question has to do with whether ethnicity's inductive power derived from the salience of its visual markers. That is, was the presence of the unique attire and visual symbols associated with the Jewish and Arab characters in our materials, a more visually compelling cue to children than the other social category markers, making that dimension of similarity stand out above all others? Were these cues especially salient for religious Jewish children? Or did the inductive power of ethnic categories derive from their linguistic markers (i.e., labels), and thus exerted their effect independently of children's capacity to visually recognize the categories? Study 2 addressed these alternatives.

**Study 2**

In order to assess the relative contributions of visual versus linguistic markers to the effects found in Study 1, Study 2 adopted the same procedure as in the relative power part of Study 1, except that the experimenter did not provide children any social category labels. Thus, in Study 2, the only information available to children about the social category membership of the characters displayed was their physical markers. The rationale was that if the inductive power of ethnicity—or any other social category—derives from bottom-up perceptual processes, then sheer exposure to stereotypical visual markers of ethnic categories should lead children to draw inferences based on them. Moreover, we should find that religious Jewish children will show the same higher sensitivity to ethnicity as found in Study 1. Alternatively, if the inductive power of ethnicity—as revealed in Study 1—is attributable primarily to the presence of cultural markers such as labels, then children in Study 2—including religious Jewish children—should not draw inferences based on ethnicity when only physical information about category membership is available. A comparison between children’s responses in the two studies would allow us to estimate directly the impact of labeling on children's inferences.

In a sense, this comparison complements one reported in Diesendruck and haLevi (2006). In that study, children were tested in one of two conditions: a category labels + visual markers condition—similar to the present Study 1—and a category labels, no visual markers condition. That study thus assessed the relative contribution of visual markers to the inductive potential of various social categories. In fact, Diesendruck and haLevi found that visual markers did not contribute significantly to the inductive power of social categories. The present Study 2 assessed the complementary effect by adding the missing cell, that is, a visual markers-only condition. It thus allowed us to evaluate how much labels contribute to the inductive potential of social categories.
In Study 2, we did not replicate the absolute power part of Study 1, because that part contrasted social categories with personality traits. While one can visually hint at the personality trait of a person by, for instance, illustrating his or her behavior, we believe these visual cues differ from those associated with most of the social categories examined here insofar as they are not conventional and stereotypical physical markers that exclusively identify the category membership of an individual.

In addition, in order to estimate potential differences in the salience or recognizability of the social categories, we conducted a “category recognition” task at the end of the triad induction task. Children’s accuracy would allow us to evaluate to what extent differences in the inductive power of social categories and across sectors could be attributed to the ease of category recognition.

Method

Participants

A total of 144 children participated in this study (see Table 1 for details). The samples and recruitment methods were identical to those in Study 1. None of the children had participated in Study 1.

Design

Study 2 consisted of two parts: an induction part and a category recognition part. The induction part was nearly identical to the relative power part of Study 1, except that the experimenter did not provide category labels for any of the characters. The same 12 trials presented in Study 1 were presented in this part of Study 2. In the category recognition part, children were shown a series of six triads that had appeared in the induction part and, on separate trials, were asked to describe and identify the characters corresponding to the different social category labels mentioned by the experimenter.

Materials

The materials were the same as those of the relative power part of Study 1.

Procedure

The procedure was identical to the one in Study 1. The category recognition part followed the induction part after a short intermission between them.

Induction part. The procedure in the induction part was identical to the procedure in the relative power part of Study 1, except that the experimenter did not mention the social categories of each character. Children’s responses were scored in the same way as in Study 1: namely, according to the number of trials (0–2) within each triad type in which they selected a particular social category (Triad Relative Inference Scores) and according to the total number of trials (0–6) across triad types in which children drew inferences based on a particular social category (Total Relative Inference Scores).

Category recognition part. Category recognition was assessed in both an unprompted and a prompted manner. For the two unprompted trials, the experimenter showed children the six pictures from the first two trials of characters that children had seen in the induction part. Children were simply asked to describe each character in their own words, and the experimenter wrote down these descriptions. After these two unprompted trials, the experimenter moved on to the prompted trials, which tested children’s capacity to recognize the characters as exemplars of the target categories. For this purpose, the experimenter laid down the three pictures of the third trial each child had seen in the induction part and asked the child to identify the referent characters of the different social categories depicted in those pictures. For example, regarding the triad displayed in Figure 1, the experimenter asked: “Is there an Arab among these characters? Who is an Arab?”. The experimenter wrote which characters the child recognized as Arabs and then asked about the next categories depicted (e.g., “Is there a Jew among these pictures? Is there a girl? A boy?”). The same procedure was repeated for the next three triads in the child’s set, thus resulting in the presentation of a total of 12 pictures.

Children’s responses in the unprompted trials were scored according to the number of times in which children explicitly mentioned the target category value when describing the character (e.g., said ‘Jew’ of a Jewish character). We refer to these as the Unprompted Category-Recognition Scores. Children’s responses in the prompted trials were scored according to the number of characters identified correctly for each social category. Given that in this part, each social category appeared on six pictures, children’s Prompted Category-Recognition Scores for each social category varied from 0 (no category exemplar correctly identified) to 6 (all category exemplars correctly identified).
Results

Induction Part

As in Study 1, we first compared the inductive power of the various social categories and then analyzed the effects of sector and age group on children’s choices.

Comparisons across social categories. As in Study 1, in order to evaluate the inductive power of each social category, we analyzed the distribution of children according to their Triad Relative Inference Scores on each triad type, comparing it against the expected chance distribution (36, 72, and 36, respectively, for selecting a particular dimension on both, one, or none of the trials of a given type). These analyses revealed a pattern of results different from the one found in Study 1. In particular, on ethnicity/gender triads, again significantly more children than were expected by chance selected ethnicity on both trials (distribution for both, one, and none, was 53, 58, and 33, respectively), $\chi^2(2, N = 144) = 11.00, p < .005$. However, on ethnicity/social status triads, the pattern was the reverse. Namely, significantly more children selected ethnicity on neither one of the two trials (i.e., selected social status on both) than was expected by chance selected ethnicity on both trials (distribution for both, one, and none, was 22, 70, and 52, respectively), $\chi^2(2, N = 144) = 12.61, p < .005$. Social status was also frequently selected on both trials in social status/gender triads (distribution for both social status, one, and none, was 65, 46, and 33, respectively), $\chi^2(2, N = 144) = 33.00, p < .001$. On all other triad types the distributions did not differ from chance, $p > .2$. In sum, when no labels were offered, children did not consistently privilege ethnicity over all other social categories and, in fact, seemed to privilege another social category—social status—instead.

Comparisons across sectors and age groups. In order to evaluate the effects of sector and age group on children’s choices, we entered the Total Relative Inference Score of each social category, for each age group and sector, against chance (chance = 3). As can be seen in Figure 4, these analyses revealed sectorial and developmental patterns of responses that were quite different from the ones found in Study 1 (compare asterisks in Figures 3 and 4). Among Arab children, in kindergarten, none of the scores were significantly different from chance; in second grade, religiosity was selected significantly less often than expected by chance, and in sixth grade, ethnicity was selected significantly more than expected by chance, and gender significantly less than expected by chance. Among secular Jews, the only score that differed significantly from chance was for gender in sixth grade, with children drawing gender-based inferences less often than expected by chance. Finally, among religious Jewish children, only by sixth grade, children were more likely to draw inferences based on social status than would be expected by chance and were less likely to draw inferences based on gender than would be expected by chance.

Comparisons Between Studies 1 and 2

In order to further evaluate children’s inferential tendencies, we compared the Total Relative Inference Score of each social category, for each age group and sector, against chance (chance = 3). As can be seen in Figure 4, these analyses revealed sectorial and developmental patterns of responses that were quite different from the ones found in Study 1 (compare asterisks in Figures 3 and 4). Among Arab children, in kindergarten, none of the scores were significantly different from chance; in second grade, religiosity was selected significantly less often than expected by chance, and in sixth grade, ethnicity was selected significantly more than expected by chance, and gender significantly less than expected by chance. Among secular Jews, the only score that differed significantly from chance was for gender in sixth grade, with children drawing gender-based inferences less often than expected by chance. Finally, among religious Jewish children, only by sixth grade, children were more likely to draw inferences based on social status than would be expected by chance and were less likely to draw inferences based on gender than would be expected by chance.
The overall MANOVA revealed a main effect of study, $F(4, 267) = 3.85$, $p < .01$, $\eta^2 = .06$ (see Figure 5). The separate ANOVAs showed that this effect was significant for ethnicity, $F(1, 270) = 10.03$, $p < .005$, $\eta^2 = .04$, with children more likely to draw ethnicity-based inferences in Study 1 ($M = 3.52$, $SD = 1.39$) than in Study 2 ($M = 3.03$, $SD = 1.28$). The effect of study was also significant with regard to social status, $F(1, 270) = 8.11$, $p < .01$, $\eta^2 = .03$, with children less likely to draw social-status-based inferences in Study 1 ($M = 2.99$, $SD = 1.43$) than in Study 2 ($M = 3.47$, $SD = 1.49$). Finally, the interaction between age and study was significant with regard to gender, $F(2, 270) = 3.40$, $p < .05$, $\eta^2 = .03$, and the interaction between sector and study was significant with respect to ethnicity, $F(2, 270) = 3.18$, $p < .05$, $\eta^2 = .02$.

In order to identify the source of the theoretically central Sector x Study interaction on ethnicity, we conducted separate ANOVAs on each of the three sectors. In these analyses, the Total Relative Inference Scores for ethnicity were used as the dependent measure, and study as a between-subjects factor. These ANOVAs revealed that while study did not have a significant effect on either Arab or secular Jewish children's tendency to draw inferences based on ethnicity ($ps > .3$), study did have a significant effect on religious Jewish children's tendency to do so, $F(1, 94) = 14.19$, $p < .001$, $\eta^2 = .13$. In particular, these children were more likely to rely on ethnicity when it was labeled (Study 1, $M = 3.90$, $SD = 1.48$) than when it was not (Study 2, $M = 2.85$, $SD = 1.22$).

**Category Recognition Part**

The main goals of the category recognition part of Study 2 were to estimate potential differences in the visual salience of the different social categories, and to evaluate whether there are differences in this respect between ages and sectors.

Preliminary analyses of the unprompted trials revealed no significant differences between the values within each category (e.g., no differences between children's spontaneous recognition of Jews vs. Arabs). We thus analyzed the data using as the dependent measures children's Unprompted Category-Recognition Scores for each of the three social categories. Given that gender is pervasively marked in both Hebrew and Arabic, it would have been pragmatically odd for children to mention a character's gender category. For this reason, we excluded the category gender from the analyses.

We entered these scores into a MANOVA with age and sector as between-subjects variables. This analysis revealed only a significant effect of age, $F(6, 244) = 11.89$, $p < .001$, $\eta^2 = .23$. Neither the effect of sector nor the interaction between age and sector was significant ($ps > .2$). The individual ANOVAs on each category revealed that age had a
significant effect on all categories (ps < .001), with sixth graders more likely to mention the category labels than children from the other two age groups (Tukey’s HSD, ps < .001). Additionally, we conducted paired t tests comparing the frequency with which children mentioned the category labels across the different categories. These tests revealed that the proportion of trials in which children mentioned the social status label of the characters (M = 0.27, SD = 0.38) was significantly higher than the proportion of trials in which children mentioned either a religiosity label (M = 0.16, SD = 0.26), t(134) = 3.76, p < .001, or an ethnicity label (M = 0.18, SD = 0.34), t(132) = 2.90, p < .005.

For the analyses of the prompted trials, we compared the Prompted Category-Recognition Score for each social category (range = 0–6), and in each age group and sector, against chance (chance = 3 pictures recognized correctly). The Gender Recognition Scores were close to ceiling (above 5) in all ages and sectors, and were therefore omitted from the analyses. As depicted in Figure 6, these analyses showed that in kindergarten, Arab children identified ethnicity at a rate significantly lower than expected by chance, and religiosity and social status at chance. Secular Jewish children identified religiosity at a rate lower than expected by chance, and ethnicity and social status at chance. Finally, religious Jewish children identified all categories at chance. By second grade, and even more so at sixth grade, all children identified all categories at a rate better than expected by chance.

Last, to directly address potential associations between children’s capacity to recognize members of a social category, and children’s tendency to draw inferences based on a social category, we assessed the correlations between children’s Prompted Category-Recognition Score on any given social category, and their Total Relative Inference Score on these same categories. None of these correlations was significant (ps > .16).

Discussion

One of the main goals of Study 2 was to assess whether the presence of visual markers of social category membership was sufficient to generate the
pattern of induction observed in Study 1. The simple answer is "no." The most consistent findings from Study 2 were that with development, children's tendency to draw inferences based on social status and ethnicity increased, and their tendency to draw inferences based on gender decreased. Crucially for the present purposes, however, Study 2 revealed neither a consistent effect of sector on children's inferences nor a distinctive sensitivity to ethnic categories. In particular, there was no evidence that religious Jewish children were more likely to draw ethnicity-based inferences than children from the other sectors.

The comparison between the two studies revealed a number of important patterns. First, they confirmed that with development, children's reliance on gender as a source for induction declines, but their reliance on ethnicity rises. These results on gender are consistent with Taylor's (1996) finding of a decrease, from kindergarten to fifth grade, in children's beliefs about the innateness of gender-typical behavior. Second, the analyses showed that labeling has a distinctive effect on children's ethnicity-based inferences. In particular, labeling significantly increased the frequency of ethnicity-based inferences but had no effect on either religiosity- or gender-based inferences. There was an effect of labeling on social-status-based inferences but in a negative direction; that is, labeling decreased social status-based inferences. Given that the task required children to choose between social categories, it is possible that the decrease in social-status-based inferences in Study 1 was partly on account of the increase in ethnicity-based inferences. It is nonetheless revealing that the increase was only in terms of ethnicity. Third and finally, the analyses by sector showed that labeling had a significant effect only among religious Jewish children. In other words, labeling in itself does not seem to create an inferential bias in a vacuum. Rather, it is only within a population already disposed toward such a bias that labeling carries an effect. We will return to these issues in the General Discussion.

The findings from the category recognition part of Study 2 further confirmed the overall conclusions noted above. In particular, first, we found that there were no differences across children from the different sectors in their capacity to recognize social category members in the pictures we used as stimuli. In other words, any differences that exist between children from the different sectors in terms of the inductive power of a given social category—especially ethnicity—do not seem to reflect differences across sectors in the ability to recognize the visual cues associated with a category, or in the visual salience of a category to a particular sector. This was further confirmed at an individual child level, by the lack of significant correlations between children's recognition and induction scores. Second, we did find a consistent increase with age in children's capacity to recognize social category members, but again, this developmental effect held across all three sectors. Thus, even though in Study 1, religious Jewish children—and only them—favored ethnicity as a source of induction already by kindergarten and second grade, it does not seem that this pattern derived from the stronger perceptual salience of ethnicity to children in this sector at those ages. In fact, by second grade, children from all sectors recognized ethnic categories in the prompted trials at above chance rates, and yet this increased recognition capacity was not accompanied by an increase in the rate of ethnicity-based inferences.

General Discussion

A number of studies conducted with both children and adults have found that in different societies, different social categories seem to be essentialized (e.g., Astuti et al., 2004; Hirschfeld, 1996; Mahalingam, 2003). One of the main goals of the present studies was to assess potential differences among sectors within a given society in their tendency to essentialize the same social categories and to examine this matter from a developmental perspective. The three sectors of the Israeli society evaluated here—secular Jews, Zionist modern religious Jews, and Muslim Arabs—were selected for the presumed diversity in their perspectives regarding ethnicity. The hypotheses were that to the extent that essentialist beliefs about a given social category are influenced by the cultural input to which children are exposed, we should find differences in a measure of essentialism among children from these three sectors. Moreover, to the extent that essentialism is a bias available to children from early on, then we should find evidence for these hypothesized differences among sectors from the earliest age assessed.

The findings from Study 1 were consistent with the above sectorial and developmental hypotheses. In that study, we found that across all ages and sectors, children were most likely to draw inferences based on ethnicity. That is, overall, when characters were described as Jews or Arabs, children were more likely to focus on that categorical similarity
when drawing novel inferences about these characters than on any dimension of similarity between the characters, be the dimension another social category (e.g., gender) or a personality trait (e.g., niceness). This privileged status of ethnicity seems to reflect the salience of this category in public discourse in Israel (Bar-Tal & Teichman, 2005) and is generally consistent with previous findings among Jewish secular kindergarten children (Diesendruck & haLevi, 2006). In fact, this focus on ethnicity has been found in other societies as well (see, e.g., Astuti et al., 2004; Gil-White, 2001; Kanovsky, 2007).

A further important finding from Study 1 is that there were nonetheless differences in the emphasis on ethnicity among sectors. In particular, compared to secular Jews and Muslim Arabs, we found that religious Jewish children were the most likely to draw inferences based on ethnicity. In fact, among kindergarteners and second graders, only the religious Jewish children drew inferences based on ethnicity significantly more than would be expected by chance. This differs from the findings of Diesendruck and haLevi (2006) who found distinctive attention to ethnicity also among secular Jewish kindergarteners. One possible explanation for this apparent disparity in the responses of secular children regarding ethnicity is that while in Diesendruck & haLevi, ethnicity was contrasted exclusively against personality traits, in the present study ethnicity was first contrasted against other social categories. This exposure to other social categories may have weakened the “appeal” of ethnicity to secular Jewish kindergarten children in the present study. In other words, it is possible that for secular Jewish children, ethnicity is a powerful enough category that it trumps personality traits, but it is not powerful enough to trump other social categories. Ethnicity does have this latter kind of power for religious Jewish children, intimating that this group may be the most essentialist about ethnicity among the three sectors who participated in these studies.

The above conclusion raises the question of what might give rise to these differences among the sociocultural sectors in Israeli society, or more generally, what causes differential essentialist beliefs toward a particular social category to emerge? There are a number of possibilities raised in the literature. One possibility derived from the social psychological literature (Allport, 1954) is that the degree of contact among members of different social categories may be inversely related to the degree to which the social categories are essentialized. It seems to us, however, that this factor cannot fully account for the differences we found in the present study. For instance, the religious and secular Jewish samples were recruited from the same cities, and so it does not seem to be the case that the religious Jewish children would have less contact with Arabs than the secular Jewish children. A second possibility is that essentialism may be particularly enforced within groups that wish to maintain a power differential toward a separate group. Mahalingam (2003) raised this possibility as an account of the essentialist beliefs of Brahmin Indian adults toward other castes. While this notion may indeed be applicable to the context of Israeli society, it is unlikely that already by kindergarten, religious Jewish children are aware of this ideology of power. In other words, this ideology would have to be somehow transmitted to children by means they could understand and enforce.

A third possibility that addresses this latter matter is that due to ideological values, religious Jewish adults may emphasize distinctions between ethnic groups, which children may then pick up and reproduce (see Bigler & Liben, 2007, for a discussion of such processes). One way in which adults might emphasize ethnicity is via linguistic markers, such as chronic labeling of ethnic categories, or use of generics to describe ethnic categories. Gelman and colleagues found that such cues may be related to essentialist beliefs about gender (Gelman et al., 2004), and thus a similar process may occur in the context of ethnic categories. We are currently examining this possibility. A final possibility we would like to raise is that religious beliefs per se may endorse essentialist beliefs. In fact, in a recent study among secular and religious Jewish children, Diesendruck and Haber (2009) found that children’s beliefs about God as the creator of social categories was positively related to children’s essentialist construal of the social categories.

Clearly, there are a number of research routes to undertake in order to uncover why a particular group within a society comes to essentialize a particular social category. The present findings not only establish the need to pursue these routes but also provide some initial guidelines for doing so. One such guideline mentioned above that carries important theoretical implications has to do with the age at which differences across sectors start emerging. In Study 1, we found that the differences between religious Jewish children and the other two groups of children in terms of privileging ethnicity as a basis for induction were evident already in the youngest age group tested, that is,
kindergarteners. This finding suggests that whatever effect cultural input might have on children's essentialist beliefs, it takes place early on in development and prior to children's intensive immersion into their sector's unique educational institutions. In fact, recent studies reveal that even under conditions of arbitrary social classification, preschoolers manifest ingroup favoritism, thus corroborating the notion that processes of social categorization have significant implications very early on in development (Patterson & Bigler, 2006). This conclusion seems most consistent with the claim that rather than having to acquire essentialist beliefs per se (Fodor, 1998), children might start off with an intuitive essentialist bias, and cultural input orients this bias to specific social categories (Hirschfeld, 1996).

One important general caveat with regard to the above conclusions is that, while we have been discussing the sectorial and developmental findings in terms of "essentialism," all the results come from induction tasks. Although these tasks have been used in the past as one of the components of essentialist reasoning, it is possible that not all components of essentialism pattern in the same way; that is, that they do not form a coherent construct (Gelman, Heyman, & Legare, 2007). In fact, comparing the same populations as the ones included here but on different components of essentialism, Birnbaum (2009) found that Muslim Arab children, rather than religious Jewish children, manifested the most essentialist-like response patterns. At a minimum, nonetheless, the present findings on induction invite the same kinds of questions about the origins of the status of ethnicity as the ones raised above about essentialism in general.

The second major goal of the present studies was to evaluate the contribution of labeling to the inductive potential of social categories. This was accomplished by comparing children's responses in a task where all categories were labeled (Study 1) to a task where categories were not labeled (Study 2). There have been claims about the importance of language in the generation of essentialism in general (Carey, 1995), and social essentialism in particular (Sperber, 1996). In fact, a number of studies highlight the importance of labeling to the inductive potential of categories in general (Gelman & Markman, 1987; Sloutsky, Lo, & Fisher, 2001), and social categories in particular (Diesendruck & haLevi, 2006; Gelman & Heyman, 1999).

The findings from these two studies converged onto the same conclusion regarding this issue. Namely, that labeling plays an important role in the inductive power of social categories, and that visual cues—or perceptual similarity—are not sufficient to account for the inductive potential of social categories. First, we found that different from Study 1, in Study 2 ethnicity was not the most powerful source of induction. Second, in Study 2 we found that religious Jewish children were no more likely than secular Jewish children or Muslim Arab children to rely on ethnicity for drawing inferences. Third and finally, we found that there were no systematic differences across sectors in children's capacity to recognize the social categories depicted in the stimuli, and there were no significant correlations between children's capacity to recognize a social category and children's tendency to rely on that category as a basis for induction. In fact, the finding that religious Jewish kindergarteners privileged ethnicity as a basis for induction when ethnicity was labeled, even though they were incapable of systematically recognizing Jews and Arabs in pictures, suggests that children's concept of ethnicity predates their perceptual classification. Altogether, social categories—in the present case, ethnicity—earn their inductive potential not from their visual distinctiveness, but rather from their labels (see also Hirschfeld, 1993).

There is considerable debate regarding the mechanisms by which labeling affects induction and essentialism. Some researchers argue that young children rely on labels for induction because of the privileged status of the auditory modality (Sloutsky et al., 2001). Others claim that labels invite children to search for the commonalities among instances of a category, thus fostering the acquisition of a notion of kind (Waxman, 1999; Xu, 2002). Finally, there are suggestions about the possibility of labels causing children to believe that deep commonalities between category members exist (Carey, 1995; Gelman & Heyman, 1999). The comparison between Studies 1 and 2 provides important new insights regarding this issue. That comparison revealed that labeling significantly increased the inductive potential only of ethnic categories, and only for religious Jewish children.

The category specificity of the effect of labeling is inconsistent with any generalist account of such an effect, perhaps especially one based exclusively on low-level perceptual capacities. This specificity intimates that children are relying on more than what is made available to them in the task; that is, children seem to be relying also on their underlying concepts about the categories. As a consequence, not all categories are equally susceptible to this effect. The sector specificity of the
effect of labeling reinforces the above conclusion, by emphasizing that not all children are equally susceptible to such an effect. Labeling does not seem to create essentialism out of vacuous concepts but rather from already fertile ones. The fact that this was true for the religious Jewish children across all ages also undermines the argument that labeling effects in young children are due to the early maturation of the auditory system (Sloutsky et al., 2001).

These findings about the specificity of labeling are also relevant to a recent theory expounded by Bigler and Liben (2007) about the development of intergroup concepts. According to Bigler and Liben, the psychological salience of a particular category derives primarily from four factors: perceptual discriminability, labeling, differential distribution, and implicit use. We would add that prior concepts about real-life categories may modulate the weighting of these factors. Using the effect of labeling as an example, it is only within a particular population that a certain labeled category stood out. One caveat to this conclusion is that it is possible that while we have manipulated—and thus equated—children’s exposure to social category labels in the task, children from different sectors may nonetheless be differentially exposed to social category labels in their daily lives. In other words, consistent with Bigler and Liben, ethnicity may be salient to religious Jewish children because this category is more frequently labeled in their regular input. As mentioned earlier, we are currently examining this possibility.

The conclusion that seems most consistent with all these findings is that labels are particularly effective triggers of an essentialist bias that, while intuitive to all children, may be differentially latent with regard to particular social categories depending on a child’s sociocultural background. In other words, the argument is that religious Jewish children in Israel may have a stronger essentialist belief about ethnicity than children from the other two sectors tested here. This belief does not derive from children’s perception of the categories or from their experiences with members of the categories. Rather, as discussed earlier, aspects of these children’s sociocultural input may endorse this belief. Interestingly, it is only by sixth grade that children have sufficiently fine-tuned visual schemas for the different categories, that they do not “need” labels to activate their rich concepts about the categories. Up until then, labels are one of the primary means by which this function is achieved. In other words, rather than construing ethnic categories “from the bottom up,” children are doing so by first associating abstract conceptual richness with labels, and only then recognizing the referents of these labels in their social environment.

One final finding that may have important implications for the development of social essentialism came from our analyses of the effect of group membership on children’s inferences. In Study 1 we found that the experimental context least conducive to ethnicity-based inferences was one in which the participant was an Arab and the target of the inference was described as a Jew. Similarly, the context least conducive to gender-based inferences was one in which the participant was a girl and the target of the inference a boy.

A line of research that may aid in the interpretation of the above pattern of results are adult social psychological studies on intergroup perceptions. Following this literature, we may describe the present pattern of findings regarding category-based inferences in terms of the perceived homogeneity of categories. The more one relies on a particular category as the basis for inference, the more one can be said to conceive of the members of the category as being similar. Thus, for instance, the asymmetry summarized above can be described as Arab children viewing “Jewishness” as more heterogeneous than “Arabness.” The casting of the findings in this light makes them directly relevant to a well-documented phenomenon on adults’ intergroup perceptions—the so-called outgroup homogeneity effect. The effect refers to the finding that adults tend to perceive members of an outgroup as more homogenous than members of their ingroup (see Ostrom & Sedikides, 1992, for a review). At face value, if children were susceptible to this effect, then we would have expected the exact opposite pattern of findings from the one reported here; for example, Arab children should have perceived Jewishness as more homogeneous than Arabness. Interestingly, nonetheless, studies in the adult social psychological literature report that under certain circumstances, there may occur a reversal of the outgroup homogeneity effect, analogous to the one found in the present study. In particular, the effect has been found to reverse when the participants belong to a minority group with whom they identify highly, and whom they positively value (Doosje, Ellemers, & Spears, 1995). To the extent that Arab children (and girls) correspond to these characteristics, the present findings may be taken as evidence for the early manifestation of this adult social psychological phenomenon. In general, these findings invite the
incorporation of assessments of self-identification with groups, to the study of the development of social essentialism. 

To conclude, the present studies shed light on a number of aspects regarding social essentialism. First, at a descriptive level, the studies confirm that within the Israeli society, ethnicity stands out as the most inductively powerful human category, above a number of other social categories or personality traits. Second, the studies reveal that ethnicity has a privileged status particularly within a sector of that society, namely, modern religious Zionist Jewish children. This finding underscores the role of cultural input in the emergence of social essentialism. Third, the early emergence of these sectorial differences intimates that cultural input is of a form capable of affecting children’s concepts from early on. Fourth, it becomes evident that the distinctive status of ethnicity is only manifest when labels are used, but again this effect is both category and sector specific. Finally, the results implicate the role of self-identification in children’s concepts of social categories.

References


