Developmental Changes in the Coherence of Essentialist Beliefs About Psychological Characteristics

Susan A. Gelman  
University of Michigan

Gail D. Heyman  
University of California—San Diego

Cristine H. Legare  
University of Michigan

Essentialism is the belief that certain characteristics (of individuals or categories) may be relatively stable, unchanging, likely to be present at birth, and biologically based. The current studies examined how different essentialist beliefs interrelate. For example, does thinking that a property is innate imply that the property cannot be changed? Four studies were conducted, examining how children ($N=195$, grades 1–7; ages 7–13) and adults ($N=187$) reason about familiar and novel social characteristics. By 3rd grade (9 years), children showed some coherence of essentialist beliefs. In contrast, younger children expected less interrelatedness among dimensions than older children or adults. These findings suggest that essentialist attributions at first consist of separate strands that children eventually link together into a more coherent understanding.

This article examines how children and adults use essentialist reasoning when considering the psychological characteristics of others. Although essentialist reasoning has been defined in different ways in the literature (see Gelman, 2003; Medin & Ortony, 1989, for a review), for the present purposes we focus on several qualities that are often viewed as central to essentialist thought, including the beliefs that a trait (of a category or of an individual) may be relatively stable, unchanging, likely to be present at birth, or biologically based. For example, someone who is essentialist about "shyness" might believe that a child who is shy will be shy his whole life, regardless of environmental circumstances, and that scientists may one day uncover a gene for shyness. Essentialism is important because it has implications for how people reason about the world. On the positive side, essentialism encourages one to extend knowledge in new ways (e.g., inferring that dwarf hamsters in general have a gestation period of 3 weeks after learning that your pet dwarf hamster has a gestation period of 3 weeks). On the negative side, essentialism encourages stereotyping of social categories (e.g., inferring that girls are not good at math, after learning that Barbie thinks math is hard). Demoulin, Leyens, and Yzerbyt (2006, pp. 25–26) note that "Essentialist lay theories have been shown to affect group relations at a variety of levels." In particular, they note essentialist influences on prejudice, perceived differences between groups, dispositional attributions, and justifying social inequalities.

Past studies have examined each of several manifestations of essentialism separately, including innate origins, internal commonalities, immutability, and stability over time, and find that children show evidence for each by preschool age (see Gelman, 2003, for a review). To what extent are different essentialist beliefs interrelated? This question has important implications for understanding the coherence of children’s knowledge. More generally, it has been argued that children’s knowledge is organized into commonsense "theories" (e.g., theories of psychology, physics, and biology), but largely unexplored has been the question of whether children’s commonsense theories, like adult scientific theories, demonstrate coherence (Wellman & Gelman, 1998). "Coherence" itself is a multifaceted notion. Here, we use the term to mean that different essentialist components have implications for one another. For example, if essentialism is a coherent notion, then innateness should be correlated with immutability, which should be correlated with internal common-
alities. Likewise, if essentialism is coherent, then learning that a new property cannot change should lead one to infer that it has an innate basis. In contrast, if each of these essentialist strands is independent of one another, then there would be no evidence for coherence in essentialist reasoning.

Different patterns of results would have different implications for the concept of essentialism. For example, if the different strands of essentialism are coherent from the start, this would suggest that essentialism is a single concept with multiple outward manifestations. In contrast, if the different strands of essentialism are not correlated for young children, but are correlated for older children and adults, this would suggest that essentialism starts out not as a single conceptual entity, but rather as a variety of beliefs that over time become knitted together into a more cohesive framework. The issue of coherence therefore allows us to address whether essentialism is a monolithic concept that happens to be expressed in a variety of ways, or is instead a variety of different conceptual “pieces” that eventually come together over time. (For a discussion of related claims regarding the causal essentialism of living kind species, see Atran [1998], Barrett [2001], and Keil [1995] regarding essentialism as a single “mode of construal,” and Gelman [2003] regarding essentialism as an outgrowth of several distinct underlying capabilities.)

In order to address coherence, we first had to select a domain to focus on. Most prior developmental studies of essentialism have centered on the biological domain. For example, studies have examined children’s beliefs that animals cannot be transformed into members of another species (Keil, 1989), members of a species share “deep,” nonobvious similarities (Gelman & Markman, 1986), and species-typical behaviors and properties are passed down from the biological parents (Gelman & Wellman, 1991). However, more recently, a number of investigators have extended the study of childhood essentialism to social categories, including race (Hirschfeld, 1996), gender (Taylor, 1996), traits (Giles, 2003; Heyman & Diesendruck, 2002), and other personal characteristics (Gelman & Heyman, 1999). We have chosen to examine coherence with regard to traits and other psychological characteristics. We have done so for three primary reasons. First, psychological characteristics can be viewed in both essentialist and nonessentialist ways (Heyman & Gelman, 2000; Rothbart & Taylor, 1992). For example, intelligence can be seen as an inborn trait that is passed down from one’s parents, or as a malleable skill that improves with practice (in Dweck’s [1999] terms, an entity view or an incremental view). We can take advantage of this variability to study coherence. (It would be difficult to study coherence for characteristics or categories that are uniformly treated as essentialist in all respects.) Second, when looking at the development of beliefs about traits, researchers have typically focused on only one or two dimensions at a time (e.g., nature vs. nurture: Heyman & Gelman, 2000; stability over time or situations: Camhy & Ruble, 1994; Heyman & Dweck, 1998; Lockhart, Chang, & Story, 2002; Rholes, Jones, & Wade, 1988), so that currently we have little idea of how findings from one dimension might generalize to other dimensions. Examining the coherence of these beliefs would help provide a more complete understanding of the various findings in this research domain (see also Haslam, Bastian, & Bissett, 2004).

Finally, it is important to understand the development of children’s trait beliefs because such beliefs have important implications for children’s developing conceptions of themselves and others. Past research indicates that belief in trait stability is a strong predictor of how willing one is to attempt changes (e.g., believing that intelligence is fixed leads to less studying after doing poorly in an exam). For example, Heyman and Dweck (1998) found that a belief in trait stability was associated with a focus on categorizing people based on their abilities and personality attributes, rather than on understanding the relevant social and motivational processes. Levy and Dweck (1999) found that sixth graders who tended to endorse a fixed view of personality more readily formed stereotypes based on what they learned about the behavior of children at other schools. However, it is still not known whether other essentialist beliefs are related to stability judgments. For example, does thinking that a property is innate, or that it is internal, lead a child to think that the property cannot be changed?

One question that this approach raises is whether psychological characteristics are properly considered kinds (analogous to the animal species in prior studies of essentialism) or individuals. Put concretely: Is a “smart person” considered to be a member of the kind “smart people,” or as an individual with a particular characteristic? Although this question cannot be resolved with the present data, we are working with a definition of essentialism that can apply to individuals as well as categories. Therefore, we need not presume that the characteristics under study map onto subtypes of people (e.g., smart people, shy people). However, we return to the issue and its implications in the General Discussion section.
Prior work on this topic has focused almost exclusively on adults, and exclusively on familiar categories. For college students, essentialist beliefs about social categories (e.g., Blacks, women, gays, Jews, AIDS patients) cohere into two distinct dimensions: “naturalness” (how much a category is natural and real, including immutability, naturalness, discreteness, necessary features, and historical invariance) and “entitativity” (how homogeneous a category is, including informativeness, uniformity, and inherence) (Haslam, Rothschild, & Ernst, 2000, 2002). These factors are similar to the two factors theorized by Rothbart and Taylor (1992) to comprise essentialism: inalterability and inductive potential. Further research has continued to find coherence in adults’ essentialist beliefs, although the details of the structure vary depending on the constructs under study. For example, essentialist beliefs about homosexuality cluster into three dimensions (not two): (1) belief in a biological basis, immutability, and fixedness early in life, (2) belief in cross-cultural and historical invariance, and (3) belief in discreteness, entitativity, and defining features (Haslam & Levy, 2006). Furthermore, essentialist beliefs about personality (in contrast to social categories)—traits such as talkative, warm, anxious, polite, wise, aggressive, dramatic, contemptible—covary into a single factor (not two or three; Haslam et al., 2004).

Some research with children suggests that coherence among essentialist beliefs may be found as early as elementary school age. For example, Boxer and Tisak (2005) found in a sample of elementary school children that beliefs about continuity correlated with beliefs that aggressive behaviors were fixed and unchanging. However, this question has not been studied previously in a systematic or sustained way.

The goal of the present studies is to examine how whether essentialist dimensions interrelate at different points in development, with particular interest in childhood. Studies 1 and 2 focus on familiar personal characteristics; Studies 3 and 4 extend to novel social categories.

Study 1

Although our primary interests are in children’s beliefs, Study 1 focuses on college students in order to enable comparison with the child studies, as well as to include a broader range of characteristics and essentialist dimensions than is possible in a developmental study (given the shorter tasks required when testing young children). The more detailed questionnaire used in Study 1 also allows us to conduct more in-depth analyses of the responses. We selected a sample of personal characteristics (i.e., “traits”) that span both social categories (e.g., homosexual) and personality characteristics (e.g., anxious).

Study 1 was designed to present a range of characteristics (e.g., intelligent, shy) and to ask a series of questions about each that would assess a variety of essentialist dimensions (e.g., Is intelligence inborn? Is shyness inborn?). Each of these dimensions has been linked to essentialism in past research: biological basis, as in brains or blood (Gelman, 2003; Gottfried, Gelman, & Schultz, 1999), naturalness (Haslam et al., 2000, 2002), heritability (Gelman & Wellman, 1991; Hirschfeld, 1996), inalterability (Keil, 1989; Rothbart & Taylor, 1992), and cross-cultural universality (Haslam & Levy, 2006). Our measure of cross-cultural universality involves prevalence rates in different subgroups of people, which have been linked to essentialist-like interpretations in prior research (e.g., Brüne, 2004, p. 41).

We sought to include a broad range of dimensions for inclusion in this research. There is some debate in the literature over the definition of essentialism, with some defining it broadly to include a wealth of factors (e.g., discreteness, biological basis, immutability, informativeness, consistency, and inheritance; Haslam et al., 2004), and others defining it much more specifically to include just belief in a non-obvious underlying reality with explanatory power (Demoulin et al., 2006). Essentialism has also been linked to related concepts such as genetic beliefs (Jayaratne et al., 2006; Keller, 2005). Our goal was to be inclusive in our definition and measure of essentialism, at the same time including dimensions that we thought would be accessible to children. Because no one had previously examined this issue developmentally (including multiple dimensions and multiple ages), we wished to include a broad sweep of relevant dimensions. At the same time, three of the elements that Haslam et al. (2004) included in their work were not included here: discreteness, informativeness, and inherence. Nonetheless, we were able to include a range of dimensions that are discussed in the literature and plausibly accessible to children.

Method

Participants. Participants were 98 college students enrolled in an Introductory Psychology course at a large Midwestern university (M age 19.25; range 18.0 – 23.25; 67 females, 31 males). The sample was predominantly White.
Items. The questionnaire focused on 12 characteristics (traits or social categories): anxious, conscientious, curious, homosexual, intelligent, lazy, musically skilled, nurturing, politically conservative, schizophrenic, shy, and superstitious. These represent a mixture of traits (e.g., intelligent, shy) and social categories (e.g., schizophrenic, homosexual). Although prior work has tended to focus on just one type of item or the other, we chose to include both types of items in order to obtain a broad sweep as a basis of comparison with the studies with children (Studies 2–4). For each characteristic, participants were asked about 9 essentialist dimensions: universal, inborn, in the brain, in the blood, consistent over time, detectable in infancy, similar to birth versus adoptive parents, change (reverse-coded), and environmentally affected (reverse-coded). For each characteristic, they were also asked about 3 control questions: talented in art, sex differences, and having friends. The control questions were included in order to include some diversity in the items, so that participants would not assume that all items were related. The only criteria for the control questions were that they be distinct from essentialist dimensions and applicable to judgments about people. Thus, altogether each participant answered 144 questions: 12 characteristics × 12 questions (9 essentialist dimensions + 3 control questions).

For each item, the response was indicated on a 6-point scale. The endpoints of the scales varied as a function of the question. The wording of the items and the endpoints are provided in Appendix A.

Procedure. Each participant received a questionnaire with all 144 items, grouped by dimension. The order of dimensions was randomly determined, with a separate random order for each participant. Within each dimension, the order of characteristics was randomly determined, separately for each participant and dimension.

Results

The data were examined to answer 4 questions: (1) Overall, how essentialist were the college students in this sample? (2) How consistent are individual participants across the different items on the questionnaire? For example, if a participant is highly essentialist on one dimension, does he or she tend to be highly essentialist on other dimensions as well? Similarly, if a participant is highly essentialist about one characteristic, does he or she tend to be highly essentialist about other characteristics as well? (3) How consistent are characteristics across the different dimensions on the questionnaire? Specifically, is the ordering of characteristics on one essentialist dimension (e.g., born) similar to the ordering of characteristics on another essentialist dimension (e.g., change)? (4) How many essentialist dimensions emerge from a factor analysis of the data? There are several possibilities: all the items could cluster together; the items could divide into two factors, naturalness and entitativity (Haslam et al., 2000); there could be some other set of clusterings; or there could be complete independence of the various measures. The analyses below address each of these four questions separately.

How essentialist were the college students in this sample? The means are presented in Table 1. For the most part, this set of characteristics was not highly essentialized. However, essentialism responses varied considerably as a function of the question. Certain characteristics received high essentialism scores overall (e.g., schizophrenia, intelligence, homosexuality, musical skill), as did certain dimensions (stability over age, universality, brain differences, born with predisposition). Moreover, each of the dimensions was essentialized for at least one characteristic (e.g., “in blood” was relatively little endorsed overall, although it was highly endorsed for schizophrenia), and each of the characteristics was essentialized on at least one dimension (e.g., “politically conservative” was the least essentialized characteristic, but it was viewed as highly stable over time). There were also considerable individual differences in each of the characteristics and dimensions, as indicated in Table 2. This variability in responding licenses us to proceed to examine the structure of the data (e.g., correlations; factor analysis).

In order to analyze the responses more systematically, we conducted a 2 (sex: male, female) × 9 (essentialist dimension) × 12 (characteristic) repeated measures analysis of variance (ANOVA). An unexpected though intriguing result was the main effect of sex, \( F(1, 94) = 5.75, p < .05 \), indicating that males generated higher essentialist scores than females, Ms (SDs) = 3.35 (0.54) and 3.12 (0.37), respectively. Across all the different characteristics and dimensions, males provided higher essentialist scores than females on 23 cells, whereas females provided higher essentialist scores than males on only 4 cells. We also obtained a main effect of dimension, \( F(8, 752) = 58.96, p < .001 \), and a main effect of characteristic, \( F(11, 1034) = 123.14, p < .001 \), indicating clear differences as a function of dimension and characteristic (as noted above). Finally, there were three significant interactions: Dimension × Sex, \( F(8, 752) = 3.05, p = .002 \); Dimension × Characteris-
tic, $F(88, 8272) = 16.83$, $p < .001$; and Dimension × Characteristic × Sex, $F(88, 8272) = 2.09$, $p < .001$. We conducted follow-up Bonferroni tests to locate the significant sex differences. Males were particularly more likely than females to endorse biological essentialism, blood: $M_{s} (S_{D}s) = 2.65 (1.20)$ and $1.99 (0.84)$; and brain: $M_{s} (S_{D}s) = 3.76 (0.94)$ and $3.30 (1.11)$.

How consistent are individual participants across the different items on the questionnaire? The question here is whether there are stable individual differences, with some individuals consistently displaying more essentialism than others (across dimensions and across characteristics). In order to examine this question, we conducted a series of Pearson correlations, displayed in Tables 3 and 4. As can be seen, there is considerable stability over items. This is particularly true across characteristics. For example, degree of essentialism regarding intelligence predicts degree of essentialism on all other characteristics except political conservatism. Out of 66 correlations, 65 were positive and 59 were statistically significant and positive. There is also some stability (although less) across dimensions. For example, degree of essentialism on the born question predicts degree of essentialism on the stability, brain, blood, birth-versus-adoptive parents, and detectability in infancy questions. Out of 36 correlations, 29 were positive and 15 were statistically significant and positive. Thus, on the whole, individual participants tend to be fairly consistent in the relative degree of essentialism, across items.

**Table 1**

*Study 1 (Adults), Mean Essentialist Scores as a Function of Characteristic and Dimension*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Consistent</th>
<th>Universal</th>
<th>Brain</th>
<th>Born</th>
<th>Changea</th>
<th>Birth versus adoptive</th>
<th>Infancy</th>
<th>Environmenta</th>
<th>Blood</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenic</td>
<td>4.84</td>
<td>4.22</td>
<td>4.99</td>
<td>4.57</td>
<td>5.06</td>
<td>4.60</td>
<td>3.94</td>
<td>4.25</td>
<td>3.53</td>
<td>4.44</td>
<td>.80</td>
</tr>
<tr>
<td>Intelligent</td>
<td>5.18</td>
<td>3.96</td>
<td>4.62</td>
<td>4.31</td>
<td>3.46</td>
<td>4.10</td>
<td>3.50</td>
<td>2.67</td>
<td>2.69</td>
<td>3.83</td>
<td>.70</td>
</tr>
<tr>
<td>Homosexual</td>
<td>5.34</td>
<td>3.68</td>
<td>3.11</td>
<td>4.13</td>
<td>4.63</td>
<td>3.67</td>
<td>2.79</td>
<td>4.47</td>
<td>2.34</td>
<td>3.79</td>
<td>.91</td>
</tr>
<tr>
<td>Musically skilled</td>
<td>4.90</td>
<td>3.89</td>
<td>4.36</td>
<td>4.39</td>
<td>3.28</td>
<td>3.75</td>
<td>3.23</td>
<td>2.97</td>
<td>2.24</td>
<td>3.67</td>
<td>.64</td>
</tr>
<tr>
<td>Anxious</td>
<td>4.12</td>
<td>3.93</td>
<td>3.46</td>
<td>3.35</td>
<td>3.13</td>
<td>3.00</td>
<td>2.91</td>
<td>2.42</td>
<td>2.57</td>
<td>3.21</td>
<td>.67</td>
</tr>
<tr>
<td>Curious</td>
<td>4.21</td>
<td>4.49</td>
<td>3.31</td>
<td>3.35</td>
<td>3.04</td>
<td>2.62</td>
<td>2.77</td>
<td>2.47</td>
<td>1.91</td>
<td>3.13</td>
<td>.54</td>
</tr>
<tr>
<td>Shy</td>
<td>3.61</td>
<td>4.20</td>
<td>3.05</td>
<td>3.36</td>
<td>3.01</td>
<td>2.69</td>
<td>2.67</td>
<td>2.39</td>
<td>1.98</td>
<td>2.99</td>
<td>.56</td>
</tr>
<tr>
<td>Nurturing</td>
<td>4.79</td>
<td>4.07</td>
<td>3.10</td>
<td>3.46</td>
<td>2.59</td>
<td>2.08</td>
<td>2.80</td>
<td>1.77</td>
<td>1.96</td>
<td>2.96</td>
<td>.56</td>
</tr>
<tr>
<td>Conscientious</td>
<td>4.69</td>
<td>3.94</td>
<td>3.16</td>
<td>3.05</td>
<td>2.53</td>
<td>2.34</td>
<td>2.49</td>
<td>2.04</td>
<td>1.87</td>
<td>2.90</td>
<td>.54</td>
</tr>
<tr>
<td>Superstitious</td>
<td>4.30</td>
<td>3.15</td>
<td>2.86</td>
<td>2.33</td>
<td>2.66</td>
<td>2.15</td>
<td>1.96</td>
<td>2.29</td>
<td>1.77</td>
<td>2.61</td>
<td>.60</td>
</tr>
<tr>
<td>Lazy</td>
<td>3.70</td>
<td>3.35</td>
<td>2.92</td>
<td>2.51</td>
<td>1.80</td>
<td>2.15</td>
<td>2.17</td>
<td>2.10</td>
<td>1.93</td>
<td>2.51</td>
<td>.59</td>
</tr>
<tr>
<td>Conservative</td>
<td>4.66</td>
<td>2.79</td>
<td>2.44</td>
<td>2.11</td>
<td>2.33</td>
<td>1.79</td>
<td>1.91</td>
<td>1.74</td>
<td>1.57</td>
<td>2.37</td>
<td>.55</td>
</tr>
<tr>
<td>Mean</td>
<td>4.53</td>
<td>3.81</td>
<td>3.45</td>
<td>3.41</td>
<td>3.13</td>
<td>2.91</td>
<td>2.76</td>
<td>2.63</td>
<td>2.20</td>
<td>3.20</td>
<td>.32</td>
</tr>
<tr>
<td>SD</td>
<td>0.55</td>
<td>0.99</td>
<td>1.07</td>
<td>0.92</td>
<td>0.66</td>
<td>0.71</td>
<td>0.99</td>
<td>0.71</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Bold italics indicate high essentialist scores (i.e., $\geq 3.5$). aDimension that was reverse-coded.

**Table 2**

*Study 1 (Adults), Proportion of Essentialist Scores That Were Low (1–2), Medium (3–4), or High (5–6) on the 1–6 Scale, as a Function of Characteristic and Dimension*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Scores of 1–2</th>
<th>Scores of 3–4</th>
<th>Scores of 5–6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenic</td>
<td>.11</td>
<td>.32</td>
<td>.57</td>
</tr>
<tr>
<td>Intelligent</td>
<td>.24</td>
<td>.34</td>
<td>.41</td>
</tr>
<tr>
<td>Homosexual</td>
<td>.29</td>
<td>.27</td>
<td>.44</td>
</tr>
<tr>
<td>Musically skilled</td>
<td>.26</td>
<td>.38</td>
<td>.36</td>
</tr>
<tr>
<td>Anxious</td>
<td>.34</td>
<td>.46</td>
<td>.19</td>
</tr>
<tr>
<td>Curious</td>
<td>.39</td>
<td>.40</td>
<td>.21</td>
</tr>
<tr>
<td>Shy</td>
<td>.42</td>
<td>.42</td>
<td>.16</td>
</tr>
<tr>
<td>Nurturing</td>
<td>.47</td>
<td>.33</td>
<td>.20</td>
</tr>
<tr>
<td>Conscientious</td>
<td>.48</td>
<td>.34</td>
<td>.18</td>
</tr>
<tr>
<td>Superstitious</td>
<td>.57</td>
<td>.30</td>
<td>.13</td>
</tr>
<tr>
<td>Lazy</td>
<td>.60</td>
<td>.30</td>
<td>.11</td>
</tr>
<tr>
<td>Politically conservative</td>
<td>.64</td>
<td>.25</td>
<td>.11</td>
</tr>
</tbody>
</table>

Note. Scores were on a 1–6 Likert scale. Each participant contributed 9 scores per characteristic and 12 scores per dimension. These scores were tallied, and the proportions of scores falling into each of three ranges (1–2, 3–4, 5–6) are represented in the table.
How consistent are characteristics across the different dimensions on the questionnaire? To examine this question, we examined the ordering of the characteristics on each of the essentialist dimensions. For example, on the born question, schizophrenia received the highest score and politically conservative received the lowest score. How did these relative rankings compare with the question concerning change? The correlations are reported in Table 5. As can be seen, consistency and universality show little relationship to the other dimensions, perhaps because they tended to show a ceiling effect. However, the remaining 7 dimensions showed a high degree of consistency. This indicates that the characteristics that were high on one dimension tended to be high on other dimensions, and the characteristics that were low on one dimension tended to be low on other dimensions.

How many essentialist dimensions emerge from a factor analysis of the data? We conducted a principal components analysis of the items. From the rotated component matrix (using the Varimax rotation method; see Table 6), 3 factors emerged: One can be characterized as “biological”: in the brain, in the blood, inborn, and detectable in infancy. A second factor included 3 dimensions: consistent over time, universal, and changeable when desired. It is surprising that consistency and universality formed a factor with changeability (which would seem to be the inverse of consistency). Because this result was unexpected, it should be replicated before clear conclusions can be reached. A third factor was the environmental dimension: environmental effect and similarity to adoptive parents.

Analysis of control questions. The control questions permit a test of whether participants are sensitive to the content of the questions and engaged in the task. Although this should not be a serious concern with adult participants, it will become a more important issue when the task is extended to children (Studies

Table 3
Study 1 (Adults), Intercorrelations Among the Essentialist Dimensions, Collapsing Over Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Universal</th>
<th>Brain</th>
<th>Born</th>
<th>Change</th>
<th>Birth versus</th>
<th>Adoptive</th>
<th>Infant</th>
<th>Environment</th>
<th>Blood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent</td>
<td>.36**</td>
<td>—</td>
<td>.22*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Universal</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Brain</td>
<td>.28**</td>
<td>—</td>
<td>.29**</td>
<td>.49**</td>
<td>—</td>
<td>.37**</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Born</td>
<td>—</td>
<td>.31**</td>
<td>—</td>
<td>.42**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Change*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Birth versus adoptive</td>
<td>—</td>
<td>.29**</td>
<td>—</td>
<td>—</td>
<td>.37**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Infancy</td>
<td>—</td>
<td>.22*</td>
<td>—</td>
<td>.63**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Environment*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.28**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. *p<.05, two-tailed; **p<.01, two-tailed.
*Dimension that was reverse-coded.

Table 4
Study 1 (Adults), Intercorrelations Among Characteristics, Collapsing Over Essentialist Dimensions

<table>
<thead>
<tr>
<th></th>
<th>Intelligent</th>
<th>Homosexual</th>
<th>Musically skilled</th>
<th>Anxious</th>
<th>Curious</th>
<th>Shy</th>
<th>Nurturing</th>
<th>Conscientious</th>
<th>Superstitious</th>
<th>Lazy</th>
<th>Conservative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenic</td>
<td>.50**</td>
<td>.33**</td>
<td>.48**</td>
<td>.36**</td>
<td>.29**</td>
<td>.33**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intelligent</td>
<td>.42**</td>
<td>.66**</td>
<td>.46**</td>
<td>.47**</td>
<td>.47**</td>
<td>.47**</td>
<td>.43**</td>
<td>.46**</td>
<td>.27**</td>
<td>.28**</td>
<td>—</td>
</tr>
<tr>
<td>Homosexual</td>
<td>.50**</td>
<td>.47**</td>
<td>.45**</td>
<td>.40**</td>
<td>.40**</td>
<td>.40**</td>
<td>.35**</td>
<td>.35**</td>
<td>.33**</td>
<td>.28**</td>
<td>—</td>
</tr>
<tr>
<td>Musically skilled</td>
<td>.47**</td>
<td>.41**</td>
<td>.45**</td>
<td>.36**</td>
<td>.45**</td>
<td>.45**</td>
<td>.40**</td>
<td>.20**</td>
<td>.35**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Anxious</td>
<td>.68**</td>
<td>.66**</td>
<td>.57**</td>
<td>.62**</td>
<td>.62**</td>
<td>.72**</td>
<td>.61**</td>
<td>.66**</td>
<td>.40**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Curious</td>
<td>.63**</td>
<td>.67**</td>
<td>.70**</td>
<td>.69**</td>
<td>.69**</td>
<td>.52**</td>
<td>.61**</td>
<td>.66**</td>
<td>.40**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Shy</td>
<td>.62**</td>
<td>.72**</td>
<td>.70**</td>
<td>.70**</td>
<td>.70**</td>
<td>.52**</td>
<td>.70**</td>
<td>.64**</td>
<td>.64**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Nurturing</td>
<td>.64**</td>
<td>.64**</td>
<td>.63**</td>
<td>.64**</td>
<td>.64**</td>
<td>.63**</td>
<td>.64**</td>
<td>.63**</td>
<td>.63**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Conscientious</td>
<td>.61**</td>
<td>.62**</td>
<td>.63**</td>
<td>.61**</td>
<td>.62**</td>
<td>.62**</td>
<td>.62**</td>
<td>.61**</td>
<td>.62**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Superstitious</td>
<td>.62**</td>
<td>.62**</td>
<td>.63**</td>
<td>.62**</td>
<td>.62**</td>
<td>.62**</td>
<td>.62**</td>
<td>.62**</td>
<td>.62**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lazy</td>
<td>.60**</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. *p<.05, two-tailed; **p<.01, two-tailed.
Brain .85
Universal
Consistent

sion was most strongly associated with.

etitive characteristics more highly than negative char-
tistics, as predicted, the valence effect is greatest for the question

Table 5
Study 1 (Adults), Intercorrelations Among the Essentialist Dimensions, Collapsing Over Characteristics and Participants

<table>
<thead>
<tr>
<th></th>
<th>Universal</th>
<th>Brain</th>
<th>Born</th>
<th>Change</th>
<th>Birth versus adoptive</th>
<th>Infancy</th>
<th>Environment</th>
<th>Blood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Universal</td>
<td>—</td>
<td>—</td>
<td>.67*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Brain</td>
<td>.85**</td>
<td>.65*</td>
<td>.90**</td>
<td>.94**</td>
<td>.94**</td>
<td>.70*</td>
<td>—</td>
<td>.86**</td>
</tr>
<tr>
<td>Born</td>
<td>.80**</td>
<td>.92**</td>
<td>.94**</td>
<td>.73**</td>
<td>.78**</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Change</td>
<td>.86**</td>
<td>.76**</td>
<td>.91**</td>
<td>.82**</td>
<td>.89**</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Birth versus adoptive</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Infancy</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Environment</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note: *Dimension that was reverse-coded.
* p < .05, two-tailed; ** p < .01, two-tailed.

2–4). Therefore, as a basis of comparison, we examined performance on the control questions in this study as an indirect measure of engagement with the task.

Specifically, we examined responses to the control questions (regarding friendship, artistic talent, and sex differences) as a function of characteristic type. We first divided the characteristics into those that were positive in connotation (conscientious, curious, intelligent, musically skilled, and nurturing) and those that were negative in connotation (anxious, lazy, schizophrenic, shy, and superstitious). We refer to these as positive and negative characteristics, respectively. (Two further characteristics cannot be easily classified as clearly positive or negative [homosexual, politically conservative] and hence were dropped for the purposes of this analysis.)

We predicted that respondents would judge positive characteristics more highly than negative characteristics on the control dimensions, particularly for friendship and artistic talent, given an overall halo effect (i.e., one positive characteristic should predict another). To test this prediction, we computed a mean score for each cell of the design, per participant, and submitted these scores to a 3 (dimension: friendship, artistic talent, and sex differences) × 2 (valence: positive, negative) × 2 (sex of participant) ANOVA. The results indicated a main effect of valence, as predicted, F(1, 96) = 295.06, p < .001, with higher scores for positive than negative characteristics, Ms (SDs) = 3.99 (.50) and 3.14 (.55), respectively, out of 6. There was also a main effect of dimension, F(2, 192) = 40.10, p < .001, and a Dimension × Valence interaction, F(2, 192) = 75.34, p < .001. As predicted, the valence effect is greatest for the question about friends, Ms (SDs) = 4.69 (0.69) and 3.23 (0.76), respectively, and smallest for the question about sex differences, Ms (SDs) = 3.06 (0.97) and 2.81 (1.00), respectively. Finally, there was a Dimension × Sex interaction, F(2, 192) = 7.02, p = .001, indicating that women endorsed the friendship question more than men, Ms (SDs) = 4.05 (0.63) and 3.76 (0.62), respectively, but men endorsed the sex-difference question more than women, Ms (SDs) = 3.25 (0.98) and 2.79 (0.87), respectively.

Altogether, the results of the control questions support the notion that adults were highly and sensibly engaged with the task.

Discussion

The goal of Study 1 was to examine in some detail the extent to which essentialist beliefs in an adult (college) sample demonstrate coherence. Preliminary questions are whether we find essentialism of characteristics and whether we find individual differences. The answer to both preliminary questions is “yes.” Although the degree of essentialism varied
considerably as a function of the characteristics selected, a small number of characteristics (especially intelligence, schizophrenia, homosexuality, and musical skill) received essentialism scores above the midpoint of the scale. Variability in responses can be seen in the large range of scores, both within characteristics and within dimensions. We also unexpectedly found variability as a function of participants’ gender, with consistently greater essentialism among men than women. Because this result was unexpected and not reported in the published literature, it would need to be replicated. However, the finding of greater essentialism among males would be consistent with the gender stereotyping literature, which typically finds greater endorsement of gender stereotypes by males, and about males (Gelman, Taylor, & Nguyen, 2004; Hort, Fagot, & Leinbach, 1990).

The finding of essentialism, along with variability in the data, allowed us to look at the question of coherence. We found evidence of structure and coherence in the dataset. First, individual differences in one characteristic tend to be highly predictive of individual differences in other characteristics. For example, people who were more essentialist about nurturance tended to be more essentialist about laziness. Second, individual differences in one dimension tend to be fairly predictive of individual differences in other dimensions. For example, people who treat characteristics as present in infancy also tended to view characteristics as unchangeable. Third, the dimensions fall into three clusters: a biological cluster, a constancy cluster, and an environmental cluster. Thus, although it is not the case that these different ways of measuring essentialism fall tightly into a single, wholly coherent cluster, neither are these questions independent of one another.

Study 2

The purpose of Study 2 was to examine the coherence of essentialist beliefs across a range of ages. Participants ranged from first graders through college students. Essentialism has been found across a wide age span, from 2-year-olds to adults (Gelman, 2003; Mahalingam, 2003; Rehder & Hastie, 2001). We selected first graders as the youngest age group because we wanted to start with an age where children could clearly understand the questions being asked. In order to address the questions of interest, the items necessarily had rather high verbal demands, and thus we were unable to include children younger than first grade with these methods.

Method

Participants. Twenty-four first graders (M age 7.23; range 6.64–8.04; 8 girls and 16 boys), 20 third graders (M age 9.21; range 8.20–9.76; 6 girls and 14 boys), 23 fifth graders (M age 10.99; range 10.01–11.71; 14 girls and 9 boys), 20 seventh graders (M age 13.16; range 12.58–13.57; 7 girls and 13 boys), and 20 college students (M age 21.6; range 20.28–26.34; 10 females and 10 males) participated. Participants were 46% White, 31% Latino/a, 15% Asian, 5% African American, and 4% of unspecified race/ethnicity. The percentage of Whites in each age group ranged from 39% (third grade) to 65% (first grade). Response rates were comparable across the child samples, typically 80% or higher. The children were all drawn from the same school district, serving a lower-middle-class community.

An additional 21 adults (M age 40.15; range 29.15–57.24; 12 females and 9 males) participated as a community sample (reported in “Community Sample,” below). All were parents with at least one child between first grade and seventh grade. These participants were 38% White, 43% Latino/a, and 19% Asian, drawn from predominantly lower-middle-class and middle-class neighborhoods.

Items. The questions concerned 5 essentialist dimensions (born, brain, blood, change [reverse-coded], and environment [reverse-coded]) and 1 control question (concerning friends). Fewer dimensions were included in Study 2 than in Study 1 in order to make the task a manageable length for children, especially in the youngest age group. As in Study 1, the control was included to introduce diversity of items as well as an indirect measure of engagement with the task. For each dimension, 8 characteristics were included: curious, good at math, good at music, lazy, mean, shy, smart, and worried. These were selected in order to be familiar to young children, and primarily to overlap with those used in Study 1. Thus, altogether each participant answered 48 questions (8 characteristics × 6 questions [5 essentialist+1 control]). The wording of the items is provided in Appendix B.

Procedure. Participants were tested individually in an oral interview conducted by a trained researcher. Each participant was asked all 48 questions, grouped by characteristic. After introducing each characteristic, children were given a simple definition to assure understanding and control for their interpretation (e.g., “A boy named Daniel is very curious. He wants to know about lots of different things.”). They were then asked all questions related to that characteristic (e.g., all questions about curious,” below). All were parents with at least one child between first grade and seventh grade. These participants were 38% White, 43% Latino/a, and 19% Asian, drawn from predominantly lower-middle-class and middle-class neighborhoods.

Study 2

The purpose of Study 2 was to examine the coherence of essentialist beliefs across a range of ages. Participants ranged from first graders through college students. Essentialism has been found across a wide age span, from 2-year-olds to adults (Gelman, 2003; Mahalingam, 2003; Rehder & Hastie, 2001). We selected first graders as the youngest age group because we wanted to start with an age where children could clearly understand the questions being asked. In order to address the questions of interest, the items necessarily had rather high verbal demands, and thus we were unable to include children younger than first grade with these methods.
ous), one at a time. Response choices were “yes,” “no,” or “maybe.” This 3-point scale was used, rather than the 6-point scale in Study 1, to simplify the task for the child participants (especially those in the youngest age group). The order of dimensions was randomly determined, with a separate random order for each participant. Within each dimension, the order of characteristics was randomly determined, separately for each participant and dimension. The researcher wrote down all responses.

Results and Discussion

For each item, the response was yes (coded as 1), no (coded as 0), or maybe (coded as 0.5). We coded “maybe” responses as 0.5 because they indicated a reluctance to commit to either “yes” or “no,” and therefore are intermediate. Furthermore, this coding scheme is unbiased, weighted toward neither extreme. Overall, most responses to the noncontrol items (82.5% for children, 62.5% for college students) were either “yes” or “no,” with “maybe” used for only 17.5% of children’s responses and 37.5% of college students’ responses. When summed over characteristics, scores could range from 0 to 8. We did not separately analyze by characteristic, as in Study 1, because scores for individual characteristics could only vary from 0 to 1 and therefore were not amenable to parametric statistical analyses.

Preliminary analyses revealed no significant effects of ethnicity (whether divided into five ethnic categories or just two [White, non-White]), and hence this factor was dropped from further analyses. We conducted a 5 (dimension: born, brain, blood, change [reverse-coded], environment [reverse-coded]) × 5 (age group: first grade, third grade, fifth grade, seventh grade, college) ANOVA, with number of essentialist responses (ranging from 0 to 8) as the dependent measure. Preliminary analyses revealed no effects of gender, and hence gender was excluded from this analysis. The results indicated a main effect of age group, $F(4, 102) = 10.88, p < .001$. The Ms (SDs) are 4.16 (0.93) in grade 1, 3.04 (0.93) in grade 3, 2.72 (0.83) in grade 5, 2.32 (1.09) in grade 7, and 3.12 (1.03) in the college sample. Post hoc Bonferroni tests revealed that the youngest age group gave higher essentialism scores than all other age groups, $ps < .01$. The greater rate of essentialism in the youngest age group is consistent with earlier findings showing that younger children tend to be more essentialist than older children and adults when reasoning about gender (Taylor, 1996), traits (Heyman & Gelman, 2000), and language (Hirschfeld & Gelman, 1997).

There was also a main effect for dimension, $F(4, 408) = 45.83, p < .001$, and a Dimension × Age Group interaction, $F(16, 408) = 2.51, p = .001$. None of the other results were statistically significant. The main effect for dimension indicated a consistent ordering overall: brain responses were most essentialized ($M = 4.92, SD = 2.76$), followed by environment responses ($M = 4.10, SD = 1.97$), followed by responses to the questions about born ($M = 2.66, SD = 2.10$), and finally change ($M = 2.00, SD = 1.72$) and blood ($M = 1.85, SD = 2.10$) (which did not differ from one another). However, the size and details of the effect vary as a function of age group. For all ages, the “brain” question elicits the most essentialist responses. However, the least essentialized dimension changes with age: “change” was least essentialized among first and third graders (i.e., they often reported that change was possible), whereas “blood” was least essentialized among the older three age groups.

Most important for current purposes was the question of how the different dimensions intercorrelated. The sample as a whole shows significant intercorrelations among several factors: born and blood (.39), born and brain (.25), born and change (.22), and brain and change (.19), all $ps < .05$. Being born with a characteristic most consistently predicts other essentialist factors. When we examined individual age groups, however, only college students and first graders showed significant effects. College students showed positive correlations between born and brain (.46) and between brain and change (.65), $ps < .05$. First graders showed a positive correlation between born and blood (.57) as well as negative correlations between born and environment (−.56) and between blood and environment (−.70), $ps < .05$. The first graders’ correlations appeared to be spurious, reflecting a bias of some children to tend toward “yes” responses overall and other children to tend toward “no” responses overall, as can be seen by the negative correlations involving dimensions that were reverse-coded. (Despite this effect, it does not reflect a lack of sensitivity to the questions on the part of the children. None of them answered “yes” on every item or “no” on every item.)

Community sample. The fact that only the adults demonstrated coherence raises the question of whether the differences between children and adults are due to age or whether it is because the college students are a more select sample. To examine this question, we conducted the same study with a community sample of adults. As mentioned earlier, all were parents with at least one child between first
grade and seventh grade, which were the ages of the child participants.

First, we compared responses from this sample with responses from the community sample, with a 2 (sample: college, community) × 5 (dimension: born, brain, blood, change, environment) ANOVA. The results indicated a main effect of dimension, $F(4, 156) = 9.02, p < .001$, but no effects of sample. The mean scores were 3.15 (out of 8) for the college sample and 3.10 for the community sample, $SD = 1.03$ and 1.63, respectively. Furthermore, when the community sample was analyzed separately, we found significant correlations between born and change ($r = .45, p < .05$), brain ($r = .83, p < .001$), and blood ($r = .47, p < .05$), as well as a significant correlation between brain and blood ($r = .61, p < .01$). These results demonstrate that the degree of essentialism and coherence shown in the college sample generalizes beyond a population of elite students.

**Analysis of control questions.** As in Study 1, the control questions were analyzed to assess whether participants were sensitive to the content of the questions and engaged in the task. Specifically, we examined responses to the control questions regarding friendship as a function of characteristic type. We divided the characteristics into those that were positive (curious, math, music, smart) and those that were negative (lazy, mean, shy, worried). We computed two summed scores for each participant (one for positive characteristics and one for negative characteristics) and submitted them to a 2 (valence: positive, negative) × 5 (age group) × 2 (sex of participant) ANOVA.

The results indicate a main effect of valence, as predicted, $F(1, 96) = 150.03, p < .001$, indicating higher “friendship” scores for positive than negative characteristics ($M_s = 2.92$ and 1.45, respectively, out of 4; $SDs = .89$ and .92, respectively). Furthermore, there was a main effect of age group, $F(4, 96) = 3.27, p < .05$, and a Valence × Age Group interaction, $F(4, 96) = 2.75, p < .05$. The interaction indicates that the valence is significant within each of the age groups considered separately, all $p_s < .001$, by Bonferroni’s test. The main effect of age group indicates that the only significant age difference on the control question is between first graders and college students ($M_s = 2.58$ and 1.94; $SDs = .82$ and .43), $p < .05$. Altogether, these data demonstrate that even the youngest participants were engaged in the task and answered sensibly.

**Summary.** Study 2 investigated adults’ and children’s real-world social reasoning, with the goal of determining whether beliefs about familiar characteristics would reveal an implicit degree of coherence in essentialism. Although the data indicate that the participants as a whole demonstrate coherence among various essentialist dimensions, the adults (both college sample and community sample) were the only age group individually to demonstrate coherence. In contrast, children do not seem to expect much coherence among different essentialist dimensions, with familiar characteristics. However, we do not know whether this result implies that children fail to expect essentialist dimensions to interrelate. There are undoubtedly a rich set of influences (besides essentialism per se) that affect children’s beliefs about real-world characteristics, and these other influences may have overshadowed any sensitivity to coherence. We therefore complemented Study 2 with an additional pair of studies that would directly manipulate essentialist dimensions and study their effects on children’s beliefs. We designed Studies 3 and 4 with this goal in mind: to examine coherence using an experimental approach, with novel properties.

**Study 3**

Studies 1 and 2 reveal that for adults, various dimensions of essentialism intercorrelate. Yet for children in Study 2, they do not. There are different possible explanations for children’s lack of coherence. One possibility is that the results reflect true developmental change, with children considering the various components as distinct from one another, and only adults being capable of linking them into a more coherent framework. However, another possibility is that the approach used in Studies 1 and 2 may have underestimated the scope of children’s coherence beliefs. In particular, Studies 1 and 2 focused on familiar characteristics, about which people (even children) have a wealth of prior knowledge and beliefs. Therefore, other, irrelevant factors may have influenced participants’ responses and masked any coherence beliefs. For example, a child might expect in general that being born a certain way implies a relative lack of changeability for that characteristic. Yet on any particular item, a child might bring to mind salient examples that could contradict that association (e.g., when asked about “born,” the child might think about her younger sibling who was shy from birth, but when asked about continuity, the child might think about her own capacity to overcome shyness over time). Thus, correlating beliefs about familiar characteristics may not be a very sensitive test of whether children hold systematic expectations about how dimensions of essentialism interrelate.
In order to provide a more sensitive test of coherence, in Studies 3 and 4 we opted for a more direct approach. That is, we experimentally manipulated essentialist dimensions of novel characteristics (e.g., explaining that a person with a particular characteristic either was born that way or was not born that way). Then we examined children’s and adults’ beliefs about other essentialist characteristics (e.g., would this characteristic change over time?). In this way, Studies 3 and 4 manipulated one essentialist factor at a time and assessed its effects (if any) on a range of other essentialist factors. We hoped that this method would be more suited to the capabilities of young children.

**Method**

**Participants.** Twenty first graders (M age 6.81; range 6.12–7.35; 9 girls and 11 boys), 20 third graders (M age 8.90; range 8.26–9.73; 7 girls and 13 boys), 20 fifth graders (M age 10.78; range 10.4–11.28; 12 girls and 8 boys), 21 seventh graders (M age 12.71; range 12.27–13.24; 16 girls and 5 boys), and 48 college students (M age 22.18; range 18.5–31.25; 25 females and 23 males) participated. Child participants were approximately 60% White, 15% Latino/a, 15% Asian, 5% African American, and 5% unspecified/other. (Unfortunately, ethnicity data were not kept for the college sample.) The percentage of Whites among the children was highest in first grade and seventh grade (80% and 70%, respectively) and lowest in third grade and fifth grade (35% and 55%, respectively). None of the participants had participated in Studies 1 or 2. However, the elementary school children (grades 1, 3, and 5) were from the same schools as those in Study 2, and the seventh graders were from the same school district. Response rates were comparable across the child samples, typically 80% or higher.

**Items.** Each participant received 12 vignettes describing a character with a novel property (e.g., “A girl named Hannah is very gogi”). Each character was paired with one fact, either essentialist (e.g., “She was born that way. She was born gogi.”) or antiessentialist (e.g., “She was not born that way. She was not born gogi.”). The facts involved 5 essentialist dimensions (born, change, brain, environment, blood) and 1 control dimension (friend, which was the control in Studies 1 and 2). After learning the new fact, the participant was asked whether the fact generalized to each of the other dimensions (e.g., “Do you think Hannah could change whether or not she’s gogi, if she wants to?” “Why is Hannah gogi—is it because of things that people around her did?”). The wording of the items is provided in Appendix C.

**Procedure.** Participants were tested individually by a trained researcher. The researcher introduced the task for the children by asking, “What are some ways that people are different from each other?” After the child answered, the researcher replied, “Good, that’s right. We’re all different in some ways. For instance, some people are nice, other people are mean. Some people are good at sports, other people aren’t so good at sports. Some people are good at drawing, other people aren’t so good at drawing. Some people like to go to parties, other people would rather stay home and read. People are different in lots of different ways. I’m going to tell you about some ways that people are different from each other, but you probably never heard about these before. So listen carefully, and you’ll learn about these new ways.”

Altogether, each participant received 60 questions in an oral interview (6 essentialist vignettes, 6 antiessentialist vignettes, each with 5 questions). The responses in each case were yes (scored as 1), no (scored as 0), or maybe (scored as 0.5). As before, the 3-point scale was adopted for use with children in order to simplify the task.

Adults received the same items in written questionnaire format, with responses indicated on a 6-point Likert scale. There was no general warm-up, but each characteristic was introduced (e.g., “Imagine that you have learned about a new personality characteristic, vooperness. Some people are vooper, other people are not vooper.”) before the initial information was provided.

**Results**

Because the adults and the children answered on different response scales (yes/no for children; 6-point scale for adults), we converted the adult responses to a 0–1 scale (subtracting 1 from the score and then multiplying it by 0.20).

We conducted a 5 (age group: first grade, third grade, fifth grade, seventh grade, and adult) × 5 (dimension: born, change, brain, environment, blood) × 2 (item type: +essentialist, −essentialist) ANOVA. As predicted, there was a main effect of item type, F(1, 124) = 91.73, p < .001, indicating that participants were more likely to endorse essentialist responses when the initial teaching fact was essentialist than when it was antiessentialist. However, this effect was modulated by age group, as indicated by a significant Item Type × Age Group interaction, F(4, 124) = 24.06, p < .001. Means relevant to this
interaction are given in Table 7. Post hoc Bonferroni tests revealed that the item type effect was significant among third graders (p < .05), fifth graders (p = .001), seventh graders (p < .001), and adults (p < .001), but not first graders. There was also a main effect of dimension, F(4, 496) = 9.34, p < .01, and a Dimension × Item Type interaction, F(4, 496) = 3.41, p < .01, and a three-way interaction among dimension, item type, and age group, F(16, 496) = 1.87, p < .05. Post hoc Bonferroni tests showed that the item type effect was significant for different dimensions at different ages. (All significant item types were in the predicted direction: +essential responses were greater than −essential responses.) For third graders, item type was significant for change only (M50.61, SD50.46, M50.52, SD50.47, p < .05). For fifth graders, the item type effect was significant for born (M50.59, SD50.41, M50.62, SD50.47, p < .01), change (M50.62, SD50.47, p < .05), and environment (M50.61, SD50.47, SD50.47, p < .01). For seventh graders, the item type effect was significant for born (M50.68, SD50.46, SD50.25, SD50.28, p < .001), change (M50.69, SD50.39, SD50.25, SD50.28, p < .001), and environment (M50.58, SD50.24, SD50.35, SD50.21, p < .001), and blood (M50.68, SD50.46, SD50.28, SD50.30, p < .01). For adults, the item type effect was significant for each of the dimensions tested, all ps < .001: born (M50.63, SD50.27, SD50.18, SD50.17), change (M50.70, SD50.26, SD50.20, SD50.11), environment (M50.65, SD50.31, SD50.18, SD50.13), blood (M50.66, SD50.33, SD50.17, SD50.13), and brain (M50.70, SD50.30, SD50.15, SD50.14).

Analysis of control questions. As in Studies 1 and 2, we analyzed the control questions (regarding “friends”) as a comparison with the experimental items. However, in contrast to Studies 1 and 2, the novel properties used in this study were all neutral in connotation, and therefore could not be divided into ones with positive versus ones with negative connotations. Thus, the “friends” control does not serve as a test of whether participants are engaged with the task, but does enable a comparison with the experimental items. In particular, we predicted that, in contrast to the experimental items, the question about friendship would be unrelated to whether the participant had heard +essentialist or −essentialist information about the characteristic.

To examine this issue, we conducted a 5 (age group) × 2 (item type: +essentialist, −essentialist) ANOVA, with the dependent measure being respondents’ score on the friendship question. We obtained a main effect of age group, F(1, 124) = 4.19, p < .01, indicating greater endorsement of friendship in the younger age groups (M50.55, 50.53, 50.44, and 50.40 at first grade, third grade, fifth grade, seventh grade, and adults; SD50.14, 50.17, 50.17, and 50.12, respectively). There was also a main effect of item type, F(1, 124) = 6.47, p < .05, and an Item Type × Age Group interaction, F(4, 124) = 3.92, p < .01. These results indicate that, overall, participants judged the character to be more likely to have friends when he or she was described in nonessentialist terms (M50.50; SD50.21) than in essentialist terms (M50.45; SD50.20). However, the effect was carried entirely by only two of the age groups: fifth graders (M50.52 and 50.35, SD50.29 and 50.27, p < .001) and seventh graders (M50.49 and 50.38, SD50.25 and 50.21, p < .05). This pattern differs from that found with the experimental items, which showed increasing effects of item type with age, beginning in third grade.

Discussion

Study 3 provided a more direct test of coherence by experimentally manipulating various essentialist dimensions overtly and one at a time. On this new task, in contrast to Study 2, children do show evidence for some coherence starting in third grade, although not before. Learning that a new property is essentialist on one dimension leads children to make more essentialist projections on other dimensions. However, there were also clear developmental changes, with greater interrelatedness among dimensions with age.

Study 4

Study 3 showed some interrelatedness of essentialist characteristics by third grade but not before. The purpose of Study 4 was to provide a more sensitive test to determine whether younger children also
show any coherence among essentialist dimensions. Study 4 was similar in design to Study 3, but with three changes designed to reduce the information-processing demands: (1) all the novel characteristics were worded so as to have meaningful content (e.g., “banana-hater” rather than “vooper”), (2) fewer vignettes were included, and (3) we restricted our focus to two of the dimensions that showed the greatest effect in Study 3 (born, change). Because our focus in this study was on the younger children, basically a replication with a simplified method, we included only the younger two age groups in this study.

Method

Participants. Twenty-one first graders (M age 7.06; range 5.6–7.66; 7 girls and 14 boys) and 20 third graders (M age 8.85; range 7.84–9.4; 8 girls and 12 boys) participated. Participants included 44% Latino/a, 41% White, 7% African Americans, and 7% other. None had participated in Studies 2 and 3. However, they were from the same schools as children in Studies 2 and 3.

Items. Vignettes concerned four novel characteristics: banana-hater, early-waker, guinea-pig-lover, and easy-laugher. Each vignette introduced a different character who was labeled with a novel trait word. Children then heard one novel fact, either essentialist (+born, −change) or antiessentialist (−born, +change). For each vignette, participants were tested on five properties: blood, brain, environment, friends (control), and either born or change (depending on which one they had not also been taught on). For example, a sample item is as follows:

“How Rachel thinks bananas taste yucky. She is a banana-hater. She was born that way; she was born a banana-hater. Do you think Rachel could change whether or not she’s a banana-hater, if she wants to? Someday, will scientists be able to figure out who is a banana-hater, by looking at their blood under an x-ray or microscope? (etc.)”

Procedure. Participants were tested individually, by a trained researcher. Each participant received 20 questions (2 essentialist vignettes, 2 antiessentialist vignettes, each with 5 questions). The responses in each case were yes (scored as 1), no (scored as 0), or maybe (scored as 0.5).

Results

We conducted a 2 (age group: first grade, third grade) × 2 (dimension: born, change) × 2 (item type: +essentialist, −essentialist) ANOVA. The dependent measure was the number of essentialist responses (ranging from 0 to 4, across the 4 relevant questions presented per vignette). The results indicated a main effect of age group, F(1, 39) = 4.56, p < .05, indicating that the younger children provided more essentialist responses overall than the older children (Ms = 2.48 and 2.14, respectively; SDs = .54 and .47). Furthermore, there was a significant Item Type × Age Group interaction, F(1, 39) = 5.07, p < .05. Post hoc Bonferroni tests indicated that only the third graders showed a significant item type effect, p < .05. Whereas first graders gave essentialist responses equally often when the initial information was +essentialist as when it was −essentialist (Ms = 2.43 and 2.52, respectively; SDs = .57 and .68), third graders gave significantly more essentialist responses on +essentialist items compared with −essentialist items (Ms = 2.31 and 1.96, respectively; SDs = .62 and .51). No other results in this ANOVA were statistically significant.

Analysis of control questions. As in the earlier studies, we analyzed the control questions (regarding “friends”) as a comparison with the experimental items. As in Study 3, we predicted that the question about friendship would be unrelated to whether the participant had heard +essentialist or −essentialist information about the characteristic. To examine this issue, we conducted a 2 (age group) × 2 (item type: +essentialist, −essentialist) × 2 (dimension: born, change) ANOVA, with the dependent measure being respondents’ score on the friendship question. The results indicated no significant effects for any of these factors, confirming that the control questions elicited a different pattern of responses than the essentialist questions.

We also conducted an analysis, similar to that of Studies 1 and 2, where we examined “friend” responses to characteristics that we judged likely to be seen as positive in valence (guinea-pig-lover, easy-laugher) versus negative in valence (banana-hater). (“Early-waker” was judged to be neither clearly positive nor clearly negative.) For each participant, we computed two scores: the mean response on the friendship question to the two positive valence items, and the response on the friendship question to the one negative valence item. These scores (each of which could range from 0 to 1) were entered into a 2 (age group) × 2 (valence) × 2 (dimension) ANOVA. The results indicated a main effect for valence, F(1, 39) = 4.19, p < .05, with friendship scores higher on positive than negative valence items (Ms = .65 and .51, respectively; SDs = .33 and .44). There was no significant interaction with age. This finding
indicates that children in this study were attentive to the task and answering sensibly.

Discussion

Study 4, using a modified task, replicated the results of Study 3: Third graders show interrelatedness among essentialist characteristics whereas first graders do not. Specifically, for third graders, learning that a characteristic is inborn or cannot change leads children to expect that it will be unchanging or inborn, detectable in the blood and brain, and not due to environmental factors. The fact that first graders did not show the effect does not mean that children of this age cannot demonstrate coherence. It is important to acknowledge that, in the future, some other means of tapping into young children’s beliefs may reveal greater sensitivity.

General Discussion

The primary goal of these studies was to examine the structure of essentialist beliefs, in particular the degree of coherence among various essentialist dimensions. Taken together, these four studies demonstrate coherence among essentialist characteristics starting by about third grade. Coherence steadily increases over childhood, with greatest coherence among adults. For adults, biological and environmental aspects of essentialism form an interrelated set of beliefs. We examined coherence in adults’ beliefs from several different perspectives (correlations, factor analysis, and experimental manipulations of essentialist factors), and all approaches reveal consistent interrelatedness of adults’ beliefs. Adults did not display a single essentialist dimension, but rather three different essentialist strands (biological essentialism, stability, and environmental essentialism). However, even these factors are not wholly independent. For example, college students in Study 2 report beliefs about the brain that correlate with beliefs about changeability.

In contrast, first graders show no evidence for coherence in any of the studies. This result could indicate that different essentialist beliefs are initially independent, or that the measures we have used are not sensitive enough to tap into the beliefs of such young children. It is still possible that first grade or even younger children have some degree of coherence in essentialist beliefs for some characteristics. For example, Giles and Heyman (2004) compared essentialist beliefs about aggression and withdrawal among a sample of preschool children and 7- to 8-year-olds. They found that preschool children who endorsed the notion of stability over time were more likely than other children to make attributions to internal causes, for both aggressive and withdrawn characters (Study 1). It may be that children’s understanding is more piecemeal and tied to particular contents early on, and only later expanded to a more general “theory” of traits or personal characteristics.

However, it is important to emphasize that these results do not imply that children’s biological knowledge more generally lacks coherence. Indeed, even preschool-aged children have coherent concepts of life and death. Slaughter and her colleagues provide compelling evidence that concepts of vitalism, human body functioning, and death are all interrelated from an early age (Slaughter, Jaakkola, & Carey, 1999; Slaughter & Lyons, 2003). The current studies do not examine this type of coherence, instead addressing coherence of essentialism per se. Relatedly, one might ask how these developmental findings relate to children’s developing folk biological concepts. To what extent are the developmental
differences due to cognitive developmental factors, and to what extent do they reflect effects of schooling? To what extent do other sorts of theories (e.g., religious beliefs) contribute to the patterns we have obtained? These are all important questions that await further research.

Within the study of essentialism, it would be interesting to know how these findings would generalize beyond the particular characteristics and dimensions that were chosen for study. For example, when studying personal characteristics (e.g., smart, shy), one can ask whether they are properly considered to refer to individuals rather than kinds. Prior research supports the idea that people with a distinctive characteristic can be construed as forming a kind, even by young children and even when the characteristic is novel (e.g., “carrot-eaters,” Gelman & Heyman, 1999). The wording of the novel items in Study 4 is particularly designed to highlight such an interpretation. Nonetheless, there are also difficulties with this formulation. For example, for “shy people” to be an essentialized category, members of the category would presumably need to have a cluster of characteristics in common beyond the single property of shyness. It is also unclear how different cross-classifications of people would relate to one another (e.g., “shy,” “intelligent,” and “gay” could all apply to the same individual; would that mean that the individual has multiple categorical essences?). In contrast, for biological kinds, the basic level label seems to be privileged. It is possible that one would find evidence for the coherence of essentialist beliefs at an earlier age on a task that asked children to consider kinds. A related unresolved issue is whether the same results would obtain if we were to examine coherence of essentialism in a different domain. The present set of studies focused exclusively on personal characteristics. Yet the clearest evidence for essentialism in the literature comes from biological kinds (Astuti, Solomon, & Carey, 2004; Gelman & Wellman, 1991). It would therefore be informative to conduct an analogous set of studies within the biological domain (e.g., regarding animal species). See Gelman (1992), Giles (2003), Rothbart and Taylor (1992), and Yuill (1992) for a discussion of how essentialism of animal kinds might relate to essentializing of human characteristics.

A further issue that should be noted is that the dimensions under investigation may vary in terms of the extent to which they reflect basic, foundational aspects of essentialism versus more contingent, empirical instantiations of these basic beliefs. To give an extreme example of this distinction: Believing boys and girls to have different kinds of insides would be a basic aspect of essentialism that we would expect any proponent of essentialism to endorse, whereas believing boys and girls to have different chromosomes would be a specific, learned instantiation that only people with biologically sophisticated knowledge would hold. Because psychological essentialism is often characterized as a placeholder notion (Gelman, 2003; Medin, 1989), foundational aspects would provide a better test. We intended the dimensions we tested to tap into more foundational aspects of essentialism. However, one could argue that the dimensions regarding brains and blood pertain more to specific, learned associations, and therefore may underestimate the coherence of children’s essentialist responding. Yet we suggest that children need not have detailed biological knowledge in order to appreciate the significance of brains and blood for essentialism. Merely knowing that blood is internal and runs through the body, and that brains are internal and linked to mental states (Johnson & Wellman, 1982) could be sufficient to link them to the essentialist assumption that insides have critical significance. In any case, this issue would benefit from further empirical investigation.

In conclusion, these findings suggest that essentialist attributions do not at first constitute a single coherent system, but rather consist of separate strands that children link together over time into a more powerful theory. In this way, although essentialism is pervasive in the United States, it changes substantially over development. One unanswered question is why coherence increases with age. There may be domain-general changes in children’s capacity to integrate information across contexts (e.g., Kuhn, 1989). Alternatively, there may be particular beliefs that children encounter as they get older (e.g., understanding that biological factors are relatively unchanged), which contribute to coherence. It would also be of great interest to study what input children receive from parents, teachers, and other sources that might influence their beliefs in this regard.

Finally, we might ask what implications these changes have for children’s behavior and beliefs more generally. By and large, the more coherence we find, the more likely it is that an essentialist belief in one domain will have implications for other beliefs and behaviors. For example, once children come to expect that different essentialist beliefs interrelate, then children with an entity view of intelligence (Dweck, 1999) may be more likely to embrace a biological explanation for individual differences in intelligence, and may be more likely to assume that environmental interventions will have little effect.
Our results suggest that beginning by at least third grade, messages about the extent to which a particular essentialist characteristic is essentialized (e.g., “He inherited his musical talent from his dad”) are likely to have broader implications for children’s understanding of that characteristic.

References


Appendix A. Wording of Questionnaire in Study 1, Using “shy” as Sample Characteristic

(Change)—To what extent can people change whether they are shy, if they want to? (1 = not at all; 6 = a lot)

(Brain)—Do you think that there are consistent differences between the brains of people who are shy and the brains of people who are not shy? (1 = no differences at all; 6 = definitely are differences)

(Blood)—In the future, will scientists be able to determine which people are shy by analyzing samples of their blood? (1 = definitely no; 6 = definitely yes)

(Consistent)—Do people who are shy at age 20 tend to be shy at age 45? (1 = definitely no; 6 = definitely yes)

(Environment)—How much does the environment a person grows up in affect whether or not he or she will become shy? (1 = not at all; 6 = a lot)

(Infancy)—Do you think scientists will ever be able to tell which babies will grow up to be shy? (1 = definitely no; 6 = definitely yes)

(Birth vs. adoptive parents)—Will a baby who is adopted at birth grow up to be as shy as the birth parents, or as shy as the adoptive parents? (1 = adoptive parents; 6 = birth parents)

(Control—talented in art)—To what extent are people who are shy likely to be talented in art? (1 = not at all likely; 6 = very likely)

(Control—having friends)—To what extent are people who are shy likely to have a lot of friends? (1 = not at all likely; 6 = very likely)

(Control—sex differences)—As a group, do men tend to differ from women in their tendency to be shy? (1 = no difference; 6 = big difference)

Appendix B. Wording of Study 2, Using “A boy named George is very shy” as Sample Characteristic

(Born)—“Do you think that George was born shy?”

(Brain)—“A boy named Kyle is NOT shy. Do you think that George’s brain is different from Kyle’s brain?”

(Blood)—“In the future, will scientists be able to figure out who is shy, by looking at their blood under an x-ray or microscope?”

(Change)—“Do you think George can change whether or not he’s shy, if he wants to?” [antiessentialist; reverse-coded]

(Environment)—“Why is George shy—is it because of things that people around him did?” [antiessentialist; reverse-coded]

(Control—having friends)—“Do you think George has a lot of friends?”

Appendix C. Wording of Study 3

Note: + indicates that the property was essentialist (or, in the case of the control property “friends”, was positive); – indicates that the property was antiessentialist (or, in the case of the control property “friends”, was negative).
Teaching

[change+]: A boy named Anthony is very erdel. He CANNOT change, even if he wants to change. He will always be erdel, even if he wants to change.

[change−]: A boy named Jason is very bliff. He can change if he wants to change. He could be NOT bliff if he wants to.

[blood+]: A boy named Eric is very noof. Someday scientists will be able to figure out who is noof, by looking at their blood under an x-ray or a microscope.

[blood−]: A boy named Zachary is very gogi. Scientists will NEVER be able to figure out who is gogi, just by looking at their blood under an x-ray or a microscope.

[brain+]: A boy named Kyle is very tibby. A boy named John is NOT tibby. Their brains are very different from each other.

[brain−]: A boy named Andrew is very banny. A boy named Samuel is NOT banny. Their brains are almost the same.

[born+]: A boy named Dylan is very tukal. He was born that way—he was born tukal.

[born−]: A boy named Thomas is very mozet. He was NOT born that way—he was NOT born mozet.

[environment+]: A boy named Robert is very prebber. His family and teachers didn’t make him prebber—they did NOT do anything to make him prebber.

[environment−]: A boy named Hunter is very zep. His family and teachers made him zep—they did things that made him zep.

[CONTROL: friends+]: A boy named Austin is very darrow. He has a lot of friends.

[CONTROL: friends−]: A boy named Ryan is very vooper. He does NOT have a lot of friends.

Testing (With Essentialist Prediction in Parentheses Following the Question)

- In the future, will scientists be able to figure out who is X, by looking at their blood under an x-ray or microscope? (yes)
  - Why is Anthony X—is it because of things that people around him did? (no)
  - A boy named Nathan is NOT X. Do you think that Anthony’s brain is different from Nathan’s brain? (yes)
    - Do you think that Anthony was born X? (yes)
    - Do you think Anthony can change whether or not he’s gogi, if he wants to? (no)
    - [CONTROL] Do you think Anthony has a lot of friends?