Developmental Change in the Use of Base-Rates and Individuating Information

Samantha Gualtieri[1] and Stephanie Denison[2]
[1] Department of Psychology, University of Toronto
[2] Department of Psychology, University of Waterloo

Adults tend to make biased inferences when they are given base-rates (i.e., prior probabilities) that conflict with individuating information (i.e., a personality description), relying heavily on individuating information. Recent work has shown that six-year-olds do the same, whereas four-year-olds rely more on prior probabilities. In the present article, we revisit the argument that producing responses that align closely with base-rates should necessarily be seen as normative. We instead posit that rational inferences should be sensitive to all relevant information and should depend on its strength. In three experiments, we explored four-year-olds', six-year-olds' ($N = 200$), and adults' ($N = 196$) information use by manipulating the strength of individuating and base-rate information. Across base-rate manipulations, adults showed a bias for individuating information regardless of its strength. In contrast, six-year-olds appeared to use each type of information flexibly, depending on which was more informative. Four-year-olds' performance was less clear: Although they relied on base-rates when they were informative, they struggled to use the individuating information in their inferences and did not appreciate the manipulation of the strength of individuating information. Thus, six-year-olds appear to more flexibly use multiple sources of information than both younger children and adults, suggesting a period in development where children are able to weigh information before they are too biased toward individuating information.

Keywords: cognitive development, heuristics and biases, judgment and decision-making, probabilistic reasoning

Supplemental materials: https://doi.org/10.1037/xge0001121.supp

Adults often have trouble handling multiple, competing sources of information when making decisions. That is, in some cases, they make biased judgments because they weigh information incorrectly or even appear to ignore a relevant source (Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). For instance, in the classic lawyer-engineer problem, participants are told that a personality description was taken from a sample of lawyers and engineers. Notably, the personality description and sample information conflict: There are 70 lawyers and 30 engineers in the sample, yet the individual is described as enjoying math puzzles and carpentry. Participants are tasked with estimating the probability that the individual was one of the engineers in the sample. Thus, they should provide an estimate that reasonably considers the base-rate (i.e., number of lawyers and engineers in the sample) and the individuating information (i.e., the personality description). However, most participants rate the probability of the individual's occupation based almost exclusively on the similarity of the personality description to their representation of a typical member of each occupation, showing little sensitivity to the base-rate of occupations. This is often referred to as the *representativeness heuristic*.

Over the last 40 years, this bias for case-specific, individuating information has been extensively documented in adults. Due to the ease with which adult participants provide responses that are aligned with the individuating information, researchers have typically characterized these responses as resulting from reflexive, intuitive processing that pays little attention to the base-rates in the problem. Conversely, responses that consider the numerical information are regarded as resulting from more analytic and reflective processing (Evans, 2008; Kahneman, 2011; Kahneman & Frederick, 2002). However, this interpretation may oversimplify the patterns of findings that would characterize normative or rational

We thank Emily Stonehouse, Annika Voss, and Katarina Guillen for help with data collection and Tiffany Doan for help with analyses. Special thanks to parents and children for participating. We are grateful to the principals and teachers at the Waterloo Region District School Board and the Waterloo Catholic District School Board, and the staff at THEMUSEUM and the Ontario Science Center for their support. This research was supported by a Grant from the Natural Sciences and Engineering Research Council of Canada to Stephanie Denison. Data from this article appear in Chapter 3 of Samantha Gualtieri's dissertation, located here: http://hdl.handle.net/10012/14837. Portions of these data have been presented at the Society for Research in Child Development and have appeared in a Cognitive Science Society proceedings paper. There are no conflicts of interest to report. Data can be found here: https://osf.io/r8a6n/.

Correspondence concerning this article should be addressed to Samantha Gualtieri, Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada. Email: samantha.gualtieri@utoronto.ca

responding. In the lawyer-engineer problem, a mathematically or normatively correct response cannot be easily defined, as the appropriate judgment in these problems should likely take into account both individuating and base-rate information, and judgments should also be swayed by the strength, validity, and reliability of such information (Koehler, 1996).

In a related line of work, researchers have explored the emergence and strengthening of the tendency to ignore base-rates in favor of individuating information by examining children's ability to use such information in these problems. In these experiments, participants are given child-friendly adaptations of the adult problem (Davidson, 1995; De Neys & Feremans, 2013; De Neys & Vanderputte, 2011; Gualtieri & Denison, 2018; Jacobs & Potenza, 1991). For example, in one study, children were given a base-rate of characters (e.g., eight nice: two mean) and were asked to classify a random character from the group that they received individuating information about (Gualtieri & Denison, 2018). When the individuating information conflicted with the base-rates, four-year-olds appeared to take base-rate information into account more than older children and adults. By the age of 6, children overweighed the individuating information in their judgments at a similar rate as adults, indicating that they neglected the base-rate information in their decisions.

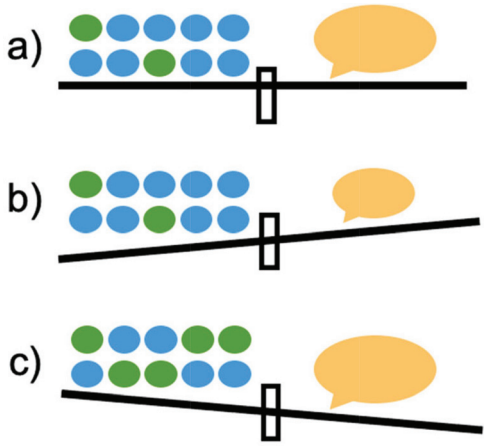
These results are quite surprising: Six-year-olds produced responses that overweighed the individuating information, whereas four-year-olds provided responses that were more aligned with the base-rate. This raises significant questions regarding the nature of the processing required to produce these patterns of performance, notably because responses that consider the base-rate are typically characterized as the result of more analytic and reflective decision-making (Stanovich et al., 2011; Toplak, 2018). Indeed, it would be quite difficult to argue that four-year-olds' base-rate use results from *better* analytical processing than six-year-olds, given the gains children make over these years in general cognitive functioning (Diamond, 2013; Garon et al., 2008) and the amount of knowledge amassed over a two-year span in early childhood. That said, it is not necessarily unusual for younger children to use presented data more rationally than older children, and even adults, in other reasoning problems (e.g., Lucas et al., 2014; Seiver et al., 2013; Walker et al., 2016). However, more work is required to determine whether younger children are in fact behaving more rational in these judgment and decision-making tasks or if they may instead be relying on a relatively quick, computationally efficient base-rate shortcut, which only appears more normative on the surface.

Further, a growing body of work in the adult literature has called into question the notion that base-rate use *requires* analytic processing. Indeed, it has been suggested that base-rate responding can also result from relatively efficient processing (De Neys, 2007). Evidence in support of this view comes from studies showing that response times remain similar when base-rates are included in problems compared with when they are not (Pennycook & Thompson, 2012). Additionally, base-rate information can create a similar amount of interference to individuating information when participants solve problems where such information conflicts, suggesting that individuating information and base-rate information are both processed efficiently (Pennycook et al., 2014). Thus, while many experiments have shown very high levels of individuating information use in adults, it appears that base-rates are also impacting their reasoning.

Together, these findings and interpretations call for a closer look into these response patterns. We suggest that analytic reasoning should be characterized by judgments that are sensitive to manipulations of information strength, with flexible use of either type of information based on whether one source should be weighed more strongly (see Figure 1 for a schematic of this interpretation). Further, manipulating the strength of base-rates and individuating information and presenting them to children and adults should be informative about how each piece of information is being used at different ages. To our knowledge, neither the developmental nor adult cognitive literatures have examined how manipulations of information strength influence judgments.

In the current experiments, we presented four-year-olds, six-year-olds, and adults with problems that varied the strength of the base-rate and individuating information provided. Unlike in previous work, we examine both sides of the coin, considering the possibility that both base-rate and individuating information use can result from heuristic processing that simply favors different types of information at different points in development: reflexive use of individuating information at older ages (i.e., the representativeness heuristic) and reflexive use of the base-rates (a base-rate heuristic) at younger ages. This contrasts with much of the literature, in that we examine whether base-rate use may also be characterized as an intuitive response, particularly for young children who have not yet built up the same strong reliance on trait information.

Figure 1
Schematic of Analytic Reasoning



Note. In analytic reasoning, information should be weighed flexibly depending on the strength of both pieces of information in the problem. Circles on the left depict base-rates, speech bubbles on the right depict individuating information. (a) When given a skewed base-rate with a lengthy piece of individuating information (as in the classic work), the base-rate and individuating information should be weighed together (though adults often neglect the base-rate in favor of individuating information). (b) When given a skewed base-rate with a shorter piece of individuating information, the base-rate should be weighed more heavily. (c) When given a lengthy piece of individuating information and an uninformative (i.e., equal) base-rate, the individuating information should be weighed more heavily. See the online article for the color version of this figure.

Across three experiments, participants were presented with child-friendly versions of the lawyer-engineer problem, in which background statistical data in the form of base-rates is provided and individuating information about a specific case is then presented. This problem was similar to those presented to children in prior examinations; however, it was modified to achieve three main goals. In Experiment 1, we manipulated the individuating information to determine if its use is influenced by the strength of the information provided. The length of the descriptions were varied such that if participants are only reflexively using this information, they should apply it equally across manipulations of strength. This manipulation provides insight into the reasoning of six-year-old children and adults from Western cultures, who have shown a strong tendency to rely on personality information in previous work (Gualtieri & Denison, 2018; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974; Seiver et al., 2013). In most previous work, the individuating information presented has been strong, detailed descriptions that are overwhelmingly representative of one group over the other (Davidson, 1995; Gualtieri & Denison, 2018; Jacobs & Potenza, 1991; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). Because previous experiments have used these heavy-handed descriptions, it is unclear if six-year-olds rely on the individuating information indiscriminately, or if they weigh the strength of the information in their judgment. Regardless of its strength, six-year-olds may apply a very simple heuristic where they determine whether the individuating information sounds more like one category or the other, and then choose that category reflexively. However, if children do consider the strength of the information, this may influence their decision about whether to rely on the individuating information. In cases where the individuating information is weaker and conflicts with the base-rate of the group, more sensitivity to those base-rates may emerge.

To this end, we manipulated the base-rate information to examine children's weighing of numerical information in Experiment 2a. This manipulation is particularly informative when interpreting the behavior of four-year-olds, who produced responses that were more in line with base-rates than individuating information in prior examinations (Gualtieri & Denison, 2018). Four-year-olds have also shown a general tendency to rely on observed statistical data in other paradigms (Lucas et al., 2014; Seiver et al., 2013). Similar to how the individuating information manipulation provides insight into six-year-olds' overreliance on personality descriptions, manipulating the strength of the base-rate information will provide insight into whether four-year-olds use a *base-rate heuristic*. Because children use base-rate information to make other social inferences as early as the second year (Kushnir et al., 2010; Ma & Xu, 2011), it is possible that four-year-olds might be using a simple shortcut in which they match their response to the base-rates, rather than considering the utility of the base-rate information (and other useful information, such as individuating information, when base-rates are less informative). Thus, when base-rates become much less informative (i.e., they are equal), responses produced via a base-rate heuristic might not be different from chance, whereas more effortful, reflective responding might show signs of considering the other available information, such as personality descriptions.

Our second goal involved pursuing a conceptual replication of recent findings by presenting alternative types of individuating information. In previous work, the individuating information has involved quite ingrained traits and stereotypes for Western children. In the present experiments, our aim was to examine this type of social learning more generally to see how weighing of different types of information influences everyday judgments that do not involve deeply rooted stereotypes (such as those based on gender, as examined in Gualtieri & Denison, 2018). To achieve this, we created two problems that contained age-appropriate group information that signified a preference for an activity, rather than a distinct social category. In one problem, participants were told a story about children in a class that could play baseball or make crafts during free time. In a second problem, participants were told a story about children in a class that could learn about space or wild animals during a trip to a museum.

Our final goal was methodological and exploratory in nature. In adult versions of these tasks, participants are asked to indicate the likelihood that the individual belongs to an occupation (e.g., they are asked to provide an estimate using a percent value that the individual is an engineer). Because young children cannot provide estimates in percentage values, previous child examinations have employed a forced-choice procedure (e.g., De Neys & Vanderputte, 2011; Gualtieri & Denison, 2018). Although suitable for young children, a binary dependent measure is also limiting and may be masking nuances in children's weighing of base-rates and individuating information. To examine the feasibility of using a more continuous measure, we asked children to choose between two options, but we also asked them to estimate their confidence in their response using a 3-point scale. We created inference scores by combining these measures to explore whether additional insight on children's sensitivity to the strength of the information can be gained (Thevenow-Harrison & Kalish, 2016).

Experiment 1

In Experiment 1, we examined four- and six-year-old children's use of individuating information when it conflicted with base-rate information. To explore the effects of individuating information strength on children's performance, we included two between-subjects conditions that varied the length of individuating information provided.

Method

Participants

In all experiments, children were tested individually at their schools or at a local museum. Formal demographic data were not collected, but the Kitchener-Waterloo region is predominantly middle-class, and approximately 81% of residents are White (Statistics Canada, 2017). Eighty children were included in the final analyses. Twenty four-year-olds (M_{age} = 55.33 months; females = 9) and twenty six-year-olds (M_{age} = 77.51 months; females = 12) participated in the more individuating condition. Twenty four-year-olds (M_{age} = 54.24 months; females = 11) and twenty six-year-olds (M_{age} = 78.12 months; females = 12) participated in the less individuating condition. An additional five children were tested and excluded for not finishing the task (n = 3), refusing to agree that there were more kids in the majority group on both trials (despite correction, n = 1), and parental report of atypical development (n = 1). This sample size was decided in advance based on the lab stopping rule at the time of data collection.

Materials and Procedure

Participants heard two stories about children in a class at a school, narrated live by the experimenter using a PowerPoint

presentation (see Figure 2 for an overview of the procedure). In the activity story, participants were told that children in a class could make a craft or play baseball during free time. Participants were shown a base-rate of children who completed each activity. The base-rate for each problem consisted of eight children completing one activity and two completing the other activity (e.g., eight children making a craft and two children playing baseball). After the experimenter counted the base-rate aloud, participants were asked which activity was completed by more children. Depending on the participant's response, the experimenter agreed or disagreed with the child's answer and stated that there were more children who made a craft and less who played baseball. Following this, participants were told that one child in the class went home for lunch, and they were given case-specific individuating information about the child's traits and preferences that were more stereotypical of a child who, for example, is more likely to enjoy sports versus arts and crafts. The individuating information was representative of the minority group and thus conflicted with the base-rate information. More specifically, if a participant was shown a base-rate of eight children making a craft and two playing baseball, they were given information that corresponded to an interest in sports. After they were told the individuating information, the experimenter asked the participant to indicate which activity they thought the child completed earlier in the day. Once the participant made their choice, the experimenter asked them to rate their confidence in their response ("for sure, maybe, or just guessing," modified from Thevenow-Harrison & Kalish, 2016). We computed scaled scores based on the dichotomous choice and the confidence judgment, giving us a more sensitive estimate of children's weighing of the information.

Children in the more individuating condition were given a lengthier, more informative description that was typical of those used in the classic studies with adults and recent child investigations (Gualtieri & Denison, 2018; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). Children in the less individuating condition were given a shorter, less informative description of the child's preferences, which was still representative of one of the groups (see Table 1 for individuating information for each condition).

Figure 2

Overview of Procedure

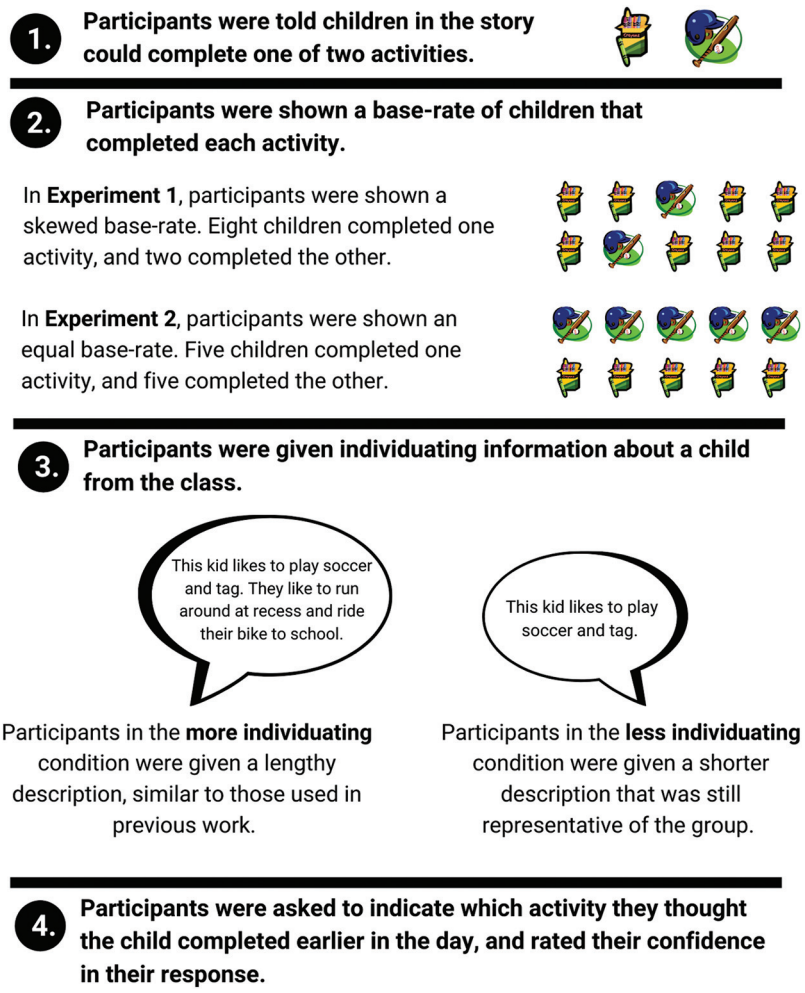


Table 1
Individuating Information Used in Each Condition

Condition	More individuating	Less individuating
Craft	This kid likes to paint and play with playdough. They like to make cool things and use their imagination.	This kid likes to paint and play with playdough.
Baseball	This kid likes to play soccer and tag. They like to run around at recess and ride their bike to school.	This kid likes to play soccer and tag.
Wild animal	This kid wanted a cat on their cake. They want to take care of animals when they grow up, and they read books about sharks.	This kid wanted a cat on their cake.
Space	This kid wanted a rocket on their cake. They want to be an astronaut when they grow up, and they read books about aliens.	This kid wanted a rocket on their cake.

In the museum story, participants were told that children on a field trip chose to visit either a wild animal exhibit or a space exhibit, with eight children choosing one exhibit and two choosing the other. Similar to the activity story, participants were given individuating information that corresponded with preferences more representative of the minority group. The order of the stories presented (i.e., activity story first, museum story first), the majority group (and, thus, the individuating information used), the placement of the pictures in the base-rate array, and the group introduced first were counterbalanced across participants.

Results

Data for all experiments are available at https://osf.io/r8a6n/.

In all experiments, participants received a score of 1 on each test trial if they selected the group that corresponded to the individuating information and received a score of 0 if they selected the opposite group. Though selecting the majority group in line with the base-rate would be a closer approximation to a normative response (as defined in past research), we coded the data in this way because we manipulated the strength of the individuating information. This also facilitated comparison across experiments, because a 50/50 base-rate is used in Experiment 2a. Turning to the confidence scores, participants received a score of 1 if they indicated they were guessing, 2 if they said maybe, and 3 if they said they for sure thought the child completed the activity. To gain further insight on children's ability to weigh base-rate and individuating information, we created inference scores from their responses to the test question and their confidence scores. We derived these scores by adjusting responses on the binary measure in light of children's confidence: Participants who used the individuating information and were confident received a score of 6, whereas those who said they were guessing received a score of 4. Conversely, those who used the base-rate and were confident received a score of 1, whereas those who said they were guessing received a score of 3. This created a 6-point scale (see Table 2 for a detailed breakdown of the scores).

One hundred sixty trials were included in the final analyses, as each child completed two trials (see Figure 3 for overall means). Preliminary analyses indicated there were no effects of counterbalancing or story type (i.e., problem order: first, second; story: activity, museum) on children's responses. We ran a Generalized Estimating Equation (GEE) binary logistic regression including children's age (four-year-olds, six-year-olds), condition (more individuating, less individuating) and the interaction between age and condition to examine their impact on children's use of individuating information. This revealed a significant main effect of condition, Wald's $\chi^2(df = 1) = 5.51$, $p = .019$, and age, Wald's $\chi^2(df = 1) = 7.12$, $p = .008$. This also revealed a significant Age $\times$ Condition interaction, Wald's $\chi^2(df = 1) = 4.04$, $p = .044$.

Pairwise comparisons (Bonferroni corrected) provided additional insight on the Age $\times$ Condition interaction. There was a significant difference between six-year-olds in the more condition ($M = .70$, $SD = .46$) and four-year-olds in the more condition ($M = .30$, $SD = .46$; $M_{Difference} = -.40$, $p < .001$) and four-year-olds in the less condition ($M = .28$, $SD = .45$; $M_{Difference} = -.43$, $p < .001$). There was also a significant difference between the six-year-olds in the more condition and those in the less condition ($M = .33$, $SD = .47$; $M_{Difference} = -.38$, $p = .001$).

To determine whether children's responses were different from chance, we examined performance at each age separately using intercept-only GEE binary logistic regressions. Because four-year-olds did not show an effect of condition, we collapsed their performance across more and less conditions. Four-year-olds provided the individuating response 29% of the time, Wald's $\chi^2(df = 1) = 12.44$, $p < .001$, suggesting that they responded according to the individuating information at rates significantly *below* what would be expected by chance. Thus, children provided the base-rate response 71% of the time, which is not statistically different from the base-rate value of 80% (see the online supplemental materials for this analysis[1]). Six-year-olds used individuating information at rates significantly above chance in the more condition, Wald's $\chi^2(df = 1) = 7.45$, $p = .006$, but not in the less condition, Wald's $\chi^2(df = 1) = 3.28$, $p = .07$, which did not differ from chance. Indeed, six-year-olds provided the base-rate response 67% of the time in the less condition, which is also not statistically different from the base-rate value of 80% (see the online supplemental materials for this analysis).

We next examined children's inference scores. One hundred fifty-eight trials were included in the final analyses, because one child did not provide confidence ratings (see Figure 4 for overall means). A GEE linear regression with age (four-year-olds, six-year-olds), condition (more individuating, less individuating), and their interaction revealed significant main effects of child's age, Wald's $\chi^2(df = 1) = 4.97$, $p = .026$, and condition, Wald's $\chi^2(df = 1) = 11.20$, $p = .001$. The interaction between age and condition

[1] Some analyses are presented in the online supplemental materials because they were not planned a priori and required slightly different coding.

Table 2
Breakdown of the Inference Scores

Inference score	Participant's responses	
	Binary choice	Confidence scale
1	Base-rate	For sure/Very confident
2	Base-rate	Maybe/Slightly confident
3	Base-rate	Guessing/Not very confident
4	Individuating	Guessing/Not very confident
5	Individuating	Maybe/Slightly confident
6	Individuating	For sure/Very confident

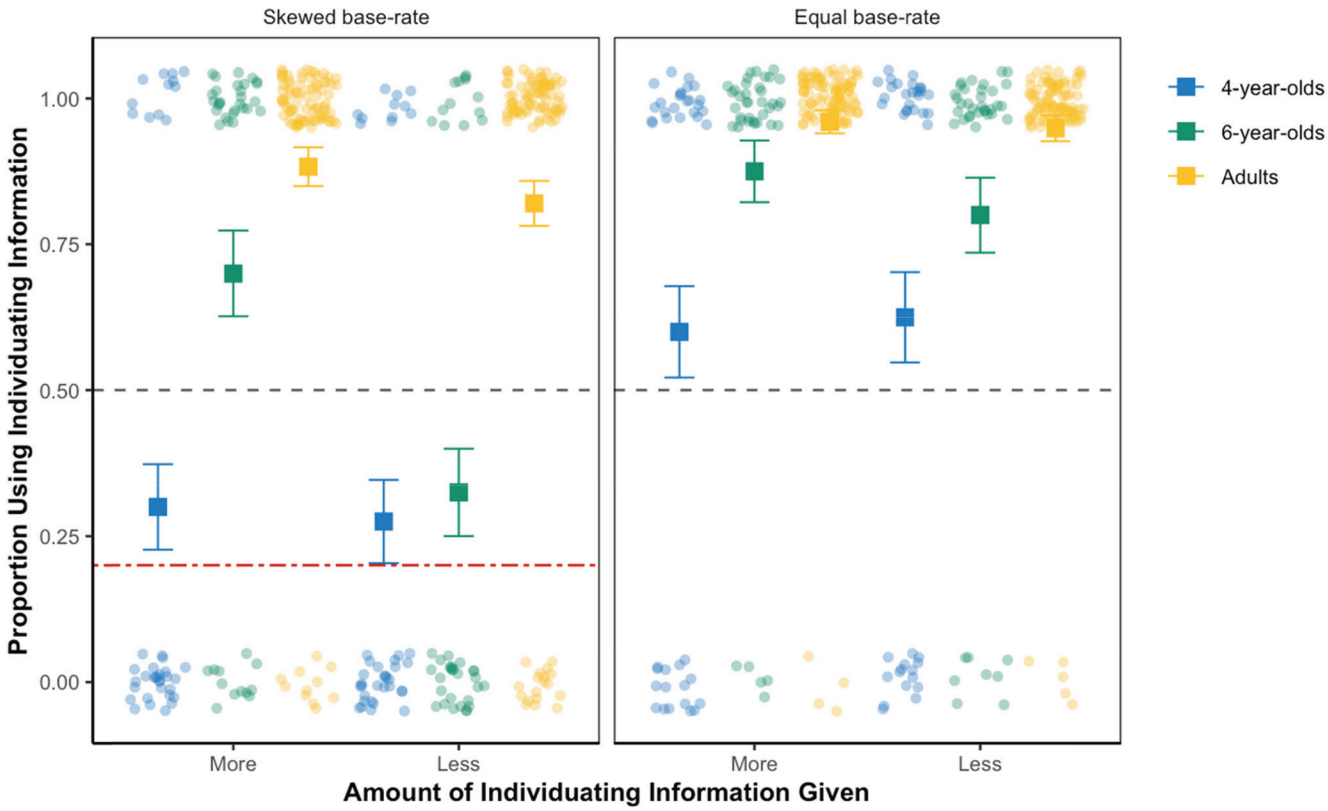
Note. The confidence scale for children (*for sure*, *maybe*, or *guessing*) was modified for adults (*very confident*, *slightly confident*, or *not very confident*).

was also significant, Wald's $\chi^2(df = 1) = 7.47$, $p = .006$. The general trends found with the binary measure held with the inference scores. Pairwise comparisons (Bonferroni corrected) indicated there was a significant difference between six-year-olds in the more condition ($M = 4.15$, $SD = 1.44$) and four-year-olds in the more condition ($M = 3.15$, $SD = .83$; $M_{Difference} = 1$, $p < .001$), and four-year-olds in the less condition ($M = 3.03$, $SD = 1.15$; $M_{Difference} = 1.12$, $p < .001$). There was also a significant difference between the six-year-olds in the more condition and those in the less condition ($M = 2.93$, $SD = 1.16$; $M_{Difference} = 1.23$, $p < .001$). Although this measure did not provide any additional nuance in children's weighing of the information, it validates that the trends found with a binary measure hold when confidence is taken into account.

Discussion

In Experiment 1, four- and six-year-old children were presented with adapted versions of the lawyer-engineer problem. When children were given more descriptive pieces of individuating information, six-year-olds relied on the individuating information in their inferences while four-year-olds gave responses that were not statistically different from the base-rate value. This conceptually replicated previous findings in which four-year-olds' inferences indicated base-rate use, while six-year-olds more heavily relied on the individuating information (Gualtieri & Denison, 2018). Notably, six-year-olds continued to prefer the individuating information in these problems even though it described personal preference information that is less culturally ingrained than the trait and gender stereotypes used in prior work. However, six-year-olds did not indiscriminately rely on the individuating information. When presented with the less descriptive individuating

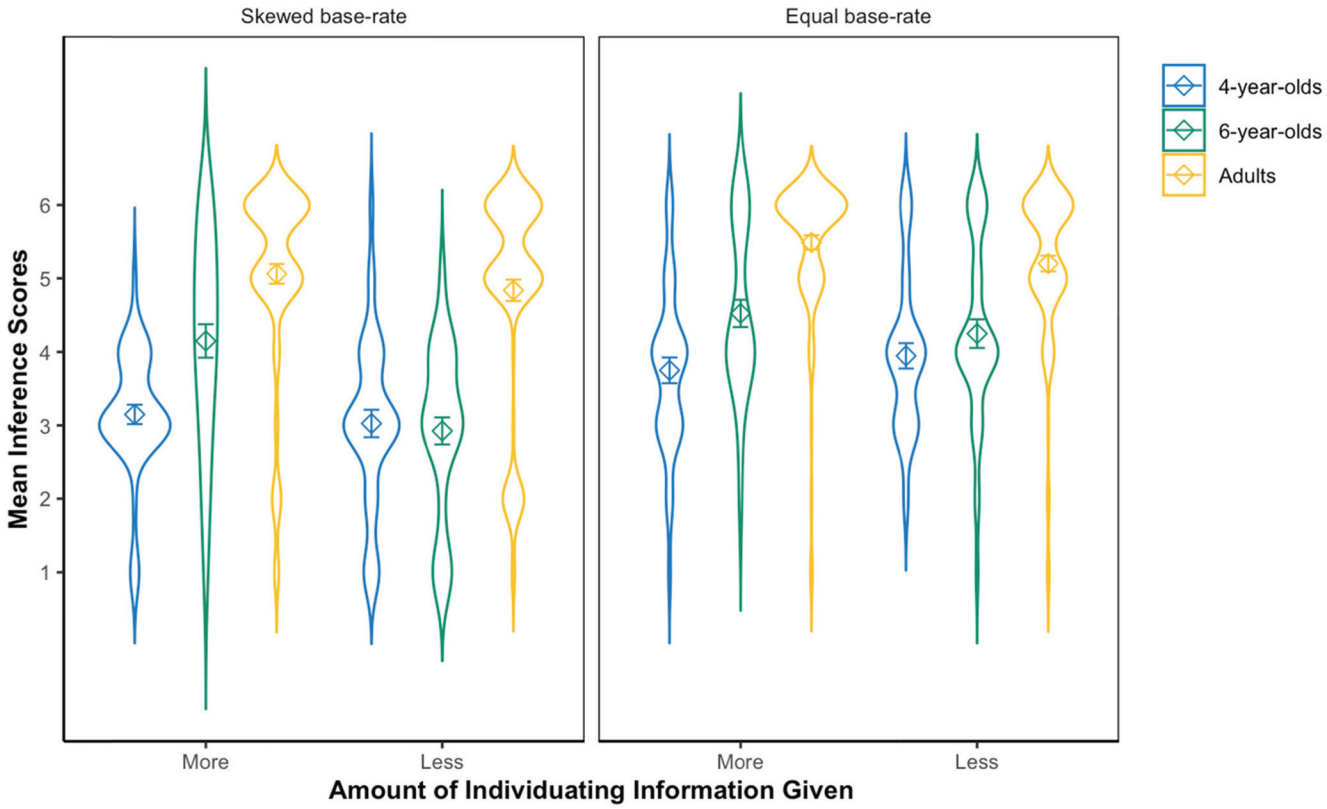
Figure 3
Proportion Using Individuating Information, on Average, Per Trial



Note. Means and standard error bars are plotted, with individual data points jittered to avoid overplotting. Skewed base-rate data are from child participants in Experiment 1 and adult participants in Experiment 3. Equal base-rate data are from child participants in Experiment 2a and adult participants in Experiment 3. The red dashed line indicates the skewed base-rate. See the online article for the color version of this figure.

Figure 4

Average Inference Scores



Note. Means and standard error bars are plotted. Skewed base-rate data are from child participants in Experiment 1 and adult participants in Experiment 3. Equal base-rate data are from child participants in Experiment 2a and adult participants in Experiment 3. See the online article for the color version of this figure.

information, both four- and six-year-olds used the base-rate information in their decisions. Thus, although six-year-olds show a general tendency to rely on individuating information when it conflicts with base-rates, they seem to monitor the strength of the information provided and opt to rely more on base-rates when the individuating information is less informative.

The results from Experiment 1 suggest that by age 6, children evaluate the strength of the individuating information in their inferences and weigh this information with the base-rate, though they show a general preference toward individuating information. Conversely, four-year-olds seem to align their responses with the base-rate, regardless of the informativeness of the individuating information. From these findings, it is unclear whether six-year-olds are attending to the base-rate information or if their responses are simply based on the strength of the individuating information. That is, six-year-olds could select the opposite of the individuating information in the less individuating condition because they find it uninformative or unrepresentative, rather than the richer interpretation that they perhaps weighed that information less heavily than the base-rates. It is also unclear as to whether the four-year-olds are *rationally* relying on base-rates or if they are simply tracking base-rates in their responses with little sensitivity to any other available information in the problem (i.e., the individuating information). In Experiment 2a, we presented children with an equal base-rate to examine their use of individuating information. If children are attending to both base-rate and individuating information, they should rely on the individuating information in their inferences when base-rates are neutral.

Experiment 2a

We examined children's use of individuating information with an equal base-rate (i.e., 5:5). As in Experiment 1, we varied the length of the individuating information across two between-subjects conditions.

Method

Participants

Eighty children were included in the final analyses. Twenty four-year-olds (M_{age} = 56.16 months; females = 8) and twenty six-year-olds (M_{age} = 76.58 months; females = 9) participated in the more individuating condition. Twenty four-year-olds (M_{age} = 56.21 months; females = 5) and twenty six-year-olds (M_{age} = 77.39 months; females = 10) participated in the less individuating condition. An additional five children were tested and excluded for not finishing the task (n = 2) or refusing to agree that there were an equal number of kids in each activity on both trials (despite correction, n = 3).

Materials and Procedure

The procedure was identical to Experiment 1; however, children were presented with a 5:5 base-rate in both problems (see Figure 2 for an overview of the procedure). Thus, the base-rate information indicated a 50/50 chance that the child completed either activity. Accordingly, after the experimenter counted the base-rate aloud, she asked participants if more children completed one of the activities or if the same number of children completed both activities. Depending on the child's response, the experimenter agreed or disagreed with their answer and stated that the same number of children completed both activities. As in Experiment 1, the individuating information was varied between conditions, with half of the participants completing the more individuating condition and half completing the less individuating condition. The order of the stories presented (i.e., activity story first, museum story first), the individuating information used, the placement of the pictures in the base-rate array, and the group introduced first were counterbalanced across participants.

Results

One hundred sixty trials were included in the final analyses, as each child completed two trials (see Figure 3 for overall means). Preliminary analyses indicated there were no effects of counterbalancing (i.e., problem order: first, second; story: activity, museum) on children's responses. We conducted a GEE binary logistic regression with age (four-year-olds, six-year-olds), condition (more individuating, less individuating), and their interaction to examine their impact on children's use of individuating information. This revealed a significant main effect of age, Wald's $\chi^2(df = 1) = 8.50$, $p = .004$. Condition, Wald's $\chi^2(df = 1) = .30$, $p = .58$, and the Age $\times$ Condition interaction, Wald's $\chi^2(df = 1) = .64$, $p = .42$, did not have a significant effect on children's scores. Across conditions, six-year-olds (more individuating $M = .87$, $SD = .34$; less individuating $M = .80$, $SD = .40$) tended to rely more on the individuating information than four-year-olds (more individuating $M = .60$, $SD = .49$; less individuating $M = .62$, $SD = .49$).

To further understand performance, we examined children's responses at each age separately using GEE binary logistic regressions. Because there were no effects of individuating information at either age, we examined the individuating conditions together. Four-year-olds' use of individuating information was not significantly different from chance, Wald's $\chi^2(df = 1) = 3.03$, $p = .08$, although six-year-olds used individuating information at significantly greater than chance rates, Wald's $\chi^2(df = 1) = 24.95$, $p < .001$.

We next examined children's inference scores. One hundred fifty-eight trials were included in the final analyses, as one child did not provide confidence ratings (see Figure 4 for overall means). A GEE linear regression with age (four-year-olds, six-year-olds), condition (more individuating, less individuating), and their interaction revealed a significant main effect of child's age, Wald's $\chi^2(df = 1) = 7.14$, $p = .008$. Condition, Wald's $\chi^2(df = 1) = .04$, $p = .85$, and the Age $\times$ Condition interaction, Wald's $\chi^2(df = 1) = .1.37$, $p = .24$, did not have a significant effect on children's scores. As with the binary dependent measure, six-year-olds' inference scores (more individuating $M = 4.53$, $SD = 1.18$; less individuating $M = 4.25$, $SD = 1.24$) were significantly higher than four-year-olds' inference scores (more individuating $M = 3.75$, $SD = 1.1$; less individuating $M = 3.95$, $SD = 1.06$).

Discussion

We presented children with problems that contained an equal base-rate and varied the strength of the individuating information. In both conditions, six-year-olds appropriately used the individuating information in their responses. Notably, they used the individuating information in the less individuating condition, indicating that their use of base-rate information in the less individuating condition of Experiment 1 was not because they found the less descriptive information unrepresentative. In both conditions, four-year-olds' mean use of the individuating information was not statistically different from 50%.

Four-year-olds in Experiment 2a used the individuating information at rates around chance. This result is difficult to interpret for two reasons: first, because the base-rate is situated at 50%, it is unclear if their current performance indicates use of the 50% base-rate or a tenuous ability to use individuating information at all, which has been documented in other experiments (see Kalish & Shiverick, 2004, for findings suggesting that four-year-olds do not predict behavioral consistency; but see Gualtieri & Denison, 2018, for evidence that they do use such information in a nearly identical paradigm to the one presented here, but with different and possibly more well-practiced information). Second, four-year-olds' responses, while not statistically above chance, are leaning in that direction. Given that there is a possibility that the design lacks the statistical power to reveal a difference from chance, this leaves open the possibility that a small effect of individuating information might be present. Thus, we tested an additional 40 children in a baseline experiment that only contained individuating information to examine whether four-year-olds can use the individuating information on its own and to see whether their information use at baseline would differ from Experiment 2a.

Experiment 2b

Method

Participants

Forty children were included in the final analyses. Twenty four-year-olds (M_{age} = 54.66 months; females = 9) participated in the more individuating condition and twenty four-year-olds (M_{age} = 54.95 months; females = 11) participated in the less individuating condition. An additional three children were tested and excluded due to experimenter error ($n = 1$) or for not finishing the task ($n = 2$).

Materials and Procedure

Participants were presented with the same stories as in Experiments 1 and 2a, except the base-rate information was omitted. For instance, participants were told that children in a class could make a craft or play baseball. Instead of seeing the base-rate information on the next slide, participants were given the individuating information about one child from the class and were asked to indicate which activity they thought that child completed earlier in the day.

Results and Discussion

Eighty trials were included in the final analyses, as each child completed two trials. Preliminary analyses indicated there were no effects of counterbalancing (i.e., problem order: first, second; story: activity, museum) on children's responses. We conducted a GEE binary logistic regression with condition (more individuating, less individuating) as a factor to examine its impact on four-year-olds' use of individuating information, which did not have a significant effect on children's responses, Wald's $\chi^2(df = 1) = 1.12$ $p = .29$. Collapsing across conditions, an intercept-only GEE binary logistic regression revealed that four-year-olds used individuating information at significantly higher than chance rates ($M = .66$, $SD = .48$), Wald's $\chi^2(df = 1) = 6.39$, $p = .01$. Similar to prior investigations (Gualtieri & Denison, 2018), these data suggest that four-year-olds use individuating information when it is the only information in the problem. That said, 66% is not particularly strong individuating information use. We also examined whether four-year-olds' responses in Experiment 2a and 2b differed using difference bootstraps and found that they used individuating information at similar rates across these experiments (see the online supplemental materials for this analysis). This suggests that, although four-year-olds used the individuating information at above chance levels in Experiment 2b, their use of this information did not differ statistically from children's performance in Experiment 2a.

Experiment 3

In Experiment 3, we presented our problems to adults. Investigations of adults' base-rate neglect have not varied the length of the individuating information (although see Pennycook et al., 2012, for experiments with briefer descriptions, but not a systematic manipulation of length or strength). Thus, it is unclear whether adults' preference for individuating information is limited to the typical lengthy descriptions, or if they consider the strength of the information in their decision. We crossed the manipulations from Experiments 1 and 2a to create a 2 × 2 design for adults.

Method

Participants

We recruited adults from Canada and the United States via MTurk, with the aim of having approximately 50 participants per condition after necessary exclusions (M_{age} = 36 years, *range* = 18–73 years; females = 76, males = 116, nonbinary = 1; three declined to report gender). After exclusions there were 47 participants in the skewed base-rate more individuating condition, 50 in the skewed base-rate less individuating condition, 50 in the equal base-rate more individuating condition, and 49 in the equal base-rate less individuating condition (total $N = 196$). An additional 96 adults were tested and excluded for failing at least one of the comprehension or memory checks (see below), and/or providing an unsatisfactory response to a question aimed to identify bots rather than human responders (see below). We required that participants had 100 HITs and an approval rating of 97. Participants were paid $.50 USD.

Materials and Procedure

Adults were given the same two stories as children, although they read through the materials themselves and provided responses via Qualtrics. Adults participated in one of four between-subjects conditions that crossed base-rate composition (skewed 8:2 base-rates that conflicted with the individuating information vs. equal 5:5 base-rates) and amount of individuating information (more vs. less). We included the base-rate comprehension check given to children to ensure adults were paying attention to the stories. To facilitate comparison with child participants, adults provided a forced-choice response to the test question and indicated how confident they were in their response using a three-point scale (very confident, slightly confident, or not very confident). After they finished both stories, participants completed two memory checks (i.e., "What activities did the kids complete during free time?" and "What exhibits did the kids see at the museum?"). To ensure that humans were making responses, participants also provided a response to two open-ended statements ("In a few words, describe the weather today" and "Please indicate if you used a calculator or pencil and paper to solve the problems"). The order of the stories presented (i.e., activity story first, museum story first), the majority group, the individuating information used, the placement of the pictures in the base-rate array, and the group introduced first were counterbalanced across participants.

Results

Three hundred ninety-two trials were included in the final analyses, as each participant completed two trials (see Figure 3 for overall means, and Table 3 for descriptive statistics). Preliminary analyses

Table 3
Descriptive Statistics and Significance Tests for Experiment 3, by Condition and Story Type

	Individuating information				Inference scores	
Story type	*M*	*SD*	Wald χ^2	*p*	*M*	*SD*
Activity story						
Skewed base-rate, more individuating	.83	.38	16.66	<.001	4.74	1.45
Skewed base-rate, less individuating	.82	.39	16.97	<.001	4.76	1.46
Equal base-rate, more individuating	.92	.27	21.95	<.001	5.22	1.23
Equal base-rate, less individuating	.90	.31	21.24	<.001	5.04	1.29
Museum story						
Skewed base-rate, more individuating	.94	.25	20.25	<.001	5.38	1.05
Skewed base-rate, less individuating	.82	.39	16.97	<.001	4.92	1.45
Equal base-rate, more individuating	1	0			5.76	.48
Equal base-rate, less individuating	1	0			5.37	.73

indicated there was no effect of problem order on responses, though there was an unexpected effect of story type, Wald's $\chi^2(df = 1) = .8.54$, $p = .003$. Adults used the individuating information significantly more when given a museum story ($M = .94$, $SD = .24$) compared with the activity story ($M = .87$, $SD = .34$).

Thus, we ran separate logistic regressions including base-rate information (skewed base-rate, equal base-rate), individuating information (more individuating, less individuating) and their interaction for each story type. These factors were not significant for the activity story: base-rate information, Wald's $\chi^2(df = 1) = 1.84$, $p = .17$; individuating information, Wald's $\chi^2(df = 1) = .02$, $p = .90$; interaction, Wald's $\chi^2(df = 1) = .05$, $p = .82$. However, there was a main effect of base-rate information for the museum story, Wald's $\chi^2(df = 1) = 4.45$, $p = .04$, in that scores were lower (less individuating information use, more base-rate use) when base-rates were skewed than when they were equal. The main effect of individuating information, Wald's $\chi^2(df = 1) = 3.15$, $p = .08$, and the interaction, Wald's $\chi^2(df = 1) = 0$, $p = 1$, were not significant. Follow-up intercept-only GEEs indicated adults' performance was significantly above chance in each condition for each story type (see Table 3).

We next examined adults' inference scores. 392 trials were included in the final analyses, as each participant completed two trials (see Figure 4 for overall means, and Table 3 for descriptive statistics). Preliminary analyses indicated there was no effect of problem order on responses, though the unexpected effect of story type was also present with inference scores, Wald's $\chi^2(df = 1) = 16.13$, $p < .001$. Thus, we ran a separate Generalized Linear model including base-rate information (skewed base-rate, equal base-rate), individuating information (more individuating, less individuating) and their interaction for each story type. For the activity story, there was a main effect of base-rate information, Wald's $\chi^2(df = 1) = 3.80$, $p = .05$. The main effect of individuating information, Wald's $\chi^2(df = 1) = .18$, $p = .67$, and the interaction, Wald's $\chi^2(df = 1) = .25$, $p = .62$, were not significant. For the museum story, there was a main effect of base-rate information, Wald's $\chi^2(df = 1) = 8.35$, $p = .004$, and individuating information, Wald's $\chi^2(df = 1) = 8.99$, $p = .003$. The interaction was not significant, Wald's $\chi^2(df = 1) = .06$, $p = .80$.

Discussion

We presented adults with our problems to examine the impact of information strength on their use of base-rates and individuating information. Adults showed a strong preference for the individuating information, generally regardless of its length (though they appeared sensitive to length, only when examining inference scores, in just one problem type). Although they were not *heavily* swayed by the changing base-rates, they did show *sensitivity* to this information, as they relied on the individuating more often when it was presented with an equal (and thus, uninformative) base-rate. This supports the notion that, while adults clearly rely heavily on representativeness, their reasoning is not entirely impervious to base-rate information, and that use of this statistical information may be somewhat intuitive (De Neys, 2012).

Interestingly, the inference scores diverged from the binary scores in some respects for adults, which was not the case for children. It appears that the inference scores, as opposed to the binary measure alone, captured that at least some adults were sensitive to the conflict between the individuating and base-rate information in the skewed conditions and there was some sensitivity to different amounts of individuating information in at least one story. It is possible that children were either less sensitive to these features or that they were less competent at indicating their confidence. In any case, this is encouraging for continuing to include inference scores in future work.

General Discussion

In three experiments, four-year-olds, six-year-olds, and adults were given base-rate and individuating information in child-friendly versions of the lawyer-engineer problem. Between the ages of 4 and 6, children's ability to reconcile base-rate and individuating information underwent developmental change. Although four-year-olds provided responses that were not statistically different from the 8:2 base-rates in Experiment 1 (which might appear quite rational), their responses in subsequent experiments were difficult to interpret and suggested that a tenuous understanding of how to use individuating information to predict behavior plays a role. Six-year-old children opted to rely on the individuating information when they were given a strong description that conflicted with the base-rates. However, when the individuating information was not as strong, six-year-olds used the base-rates in their judgments. From these findings, it seems that six-year-olds have developed a general preference for individuating information in their judgments, though they rationally attend to the strength of the information when deciding whether to rely on it over the base-rates. Surprisingly, six-year-olds weighed the information *more efficiently* than adults. That is, across base-rate manipulations, adults showed a strong preference for the individuating information regardless of its strength (though they did show some sensitivity to the base-rate information, as they relied on the individuating information significantly more when the base-rate was equal).

At age 4, children relied on the skewed base-rate information in their judgements in Experiment 1, providing responses that did not differ statistically from the base-rate value and that did differ from chance. These findings are in line with previous work showing that four-year-olds provide responses that are more aligned with observed statistical data, suggesting that they use a more data-driven approach than older children and adults (Gualtieri & Denison, 2018; Lucas et al., 2014; Seiver et al., 2013). However, four-year-olds were not sensitive to manipulations of individuating information strength, and their performance was inconclusive when base-rates were uninformative in Experiment 2a. Thus, at first glance, four-year-olds' base-rate use in Experiment 1 *looks* like a normative response, but it might be attributable to a nonnormative process that does not adequately account for the individuating information. This early proficiency in base-rate use, as opposed to individuating information, is feasible because base-rate use appears to be phylogenetically ancient (Tecwyn et al., 2017), infants are sensitive to it at 6 months (Denison et al., 2013), and toddlers use base-rates to make social inferences in other paradigms (Kushnir et al., 2010; Ma & Xu, 2011). Future work that more incrementally manipulates the base-rate proportions and examines other kinds of social information that four-year-olds use more readily could reveal whether this early base-rate use results from an intuitive, heuristic process (i.e., base-rates as an intuitive response, Pennycook et al., 2014) or a more analytic one.

Regardless of the inconclusive nature of four-year-olds' performance, in a period of two years, children develop a preference for individuating information, as evidenced by six-year-olds' reliance on the strong piece of individuating information that conflicted with a base-rate. It is possible that six-year-olds have a more general preference for individuating information due to having developed a stronger prior belief regarding the importance of traits and personality descriptions in predicting behavior (Boseovski et al., 2013; Boseovski & Lee, 2006; Kalish & Shiverick, 2004), because they have also shown preferences for personality information over statistical data in prior work (Gualtieri & Denison, 2018; Seiver et al., 2013). Notably, six-year-olds' preference for individuating information was affected by the strength of the information. Whereas they relied on the strong individuating information when it conflicted with the base-rate, they used the base-rates in their inferences when it conflicted with a weaker piece of individuating information that contained representative information. This flexible use of base-rate and individuating information suggests that six-year-olds attempt to weigh information in their decisions. Thus, whereas previous findings with six-year-olds suggested that perhaps they were already showing very strong stereotyped responses that were impervious to other information (such as base-rates; see Gualtieri & Denison, 2018), it appears that they may instead be using information quite flexibly and rationally. This is in contrast to adults, who showed a more pronounced preference for individuating information. Whereas six-year-old children were sensitive to the strength of the individuating information when it conflicted with the base-rate, adults showed very little sensitivity to strength. Thus, there is a strong, perhaps reflexive, preference for individuating information by adulthood.

Limitations and Future Directions

One potential limitation of this work is in the design of the experimental tasks, because there are necessary trade-offs in design decisions when bridging two research traditions. In the adult judgment and decision-making work, base-rates are always presented first, because they serve as *background* statistical information, which is ecologically valid in the sense that this is typically how these scenarios are encountered in the real world (we know, for example, the typical base-rate of teachers vs. librarians when we come to decide whether someone is likely to fall into one category or the other). We opted to retain this feature in creating "child-friendly" versions of these tasks. However, in developmental work, it is more typical to counterbalance information presentation (presenting each source first for half of the children). Thus, our findings are open to the interpretation that four-year-olds may have responded based on a primacy effect, using the first piece of information given, which always corresponds to base-rates. Although we urge future researchers to consider counterbalancing this information, we think this is a less likely interpretation of our findings than the one we present. First, this cannot explain the rational responding of six-year-olds, who used each type of information flexibly. Second, in a highly similar task that examines the classic "cab problem", which pits testimony (a very well-practiced social learning mechanism in early childhood, Harris et al., 2018) against base-rate information, four-year-olds reliably respond with the testimony information, which was also always presented last in that paradigm (Gualtieri et al., 2020).

A methodological goal of the current experiments involved examining the validity of inference scores as a dependent measure. These scores were created from participants' responses to the forced-choice questions and their confidence in their responses. The findings from the inference scores confirmed the general trends captured by the binary dependent measure, which aided in establishing the validity of the forced-choice scores. However, it is unclear whether, and to what extent, this measure illuminated any nuances in children's ability to integrate information. It is possible that four- and six-year-old children have difficulty reporting the conflict they detect in our paradigm via confidence judgments, as prior work suggests that explicit conflict detection on base-rate neglect problems does not emerge until 11 years of age (De Neys & Feremans, 2013). To thoroughly examine children's ability to integrate base-rate and individuating information, future work should continue to investigate ways to obtain graded responses from children.

Moreover, much future work is needed to elucidate the potential mechanism driving these age-related changes in performance. Many aspects of cognition develop between the ages of 4 and 6, which could contribute to differences in their approach to solving these problems. Children's executive functions undergo drastic development during the preschool years (see Diamond, 2013, for a review). It is possible that development in children's executive functions aids in their ability to integrate information and/or hold multiple pieces of competing information in mind, accounting for age differences observed in children's sensitivity to the strength of the information. That is, six-year-olds' ability to weigh the strength of the information in their decisions may be due in part to their better-developed working memory and cognitive flexibility, because these abilities may allow them to hold multiple pieces of information in mind while they consider different decisions. However, these general advancements cannot explain why we see *more* heuristic use in adults compared with six-year-olds. Indeed, adults' better working memory and attention abilities aid their ability to integrate information. By adulthood, individuating information may be so practiced that adults forego weighing the information and stick to ignoring or underusing the base-rate information.

It is also possible that six-year-olds' preference for individuating information may be due in part to a stronger sensitivity to the pragmatics of the paradigm. Adults' use of individuating information is affected by the framing of the problem, suggesting that their sensitivity to pragmatic features influences their weighing of the information. For instance, if adults are told the individuating information was randomly drawn by a computer, they are more inclined to use the base-rates in their inferences than if it was chosen by a psychologist (Schwarz et al., 1991). Children's sensitivity to conversational pragmatics develops between the ages of 4 and 6 (Matthews, 2014). In the current study, six-year-olds' varied responses across manipulations were interpreted as a development in sensitivity to information strength. However, these data could also be viewed as evidence for a pragmatic account. For instance, a six-year-old child who is more attuned to conversational pragmatics may wonder why the experimenter provided such a detailed description of the individual if she did not want them to consider the individuating information in their decision. Because the lower quality individuating information was also much shorter, it might be seen as less heavy-handed, which could reduce its use if children are attuned to this pragmatic feature.

A final future direction involves exploring the interaction between children's stereotype familiarity and information quality in their weighing of individuating information. Recent developmental

investigations of the representativeness heuristic have gone to great lengths to ensure that children are familiar with the group information given to them in the problem (i.e., presenting them with gender information they are familiar with; see De Neys & Vanderputte, 2011). In the current experiments, we conceptually replicated previous findings with less culturally embedded information, which allowed us to explore whether stereotype familiarity was leading six-year-olds to use individuating information. However, it is possible that children would continue to use less descriptive individuating information in their decisions if it described a social category that they were quite familiar with. Future work manipulating the familiarity and relative entrenchment of group information in tandem with information quality would be pivotal in determining the features that lead young children to rely on individuating information. Further, an approach that integrated these manipulations with examinations of individual differences in cognitive style, cognitive flexibility, and working memory would illuminate the factors that might predict flexible information use as opposed to overreliance on different cues.

Implications and Conclusions

Within a period of two years, children employ different strategies when reconciling base-rate and individuating information. By the age of 6, children show a preference for individuating information, although they weigh the strength of the information and rely on base-rates when the individuating information is less descriptive. These findings provide insight on nuances in the emergence and strengthening of children's developing ability to integrate information that were previously unexplored, suggesting a period in development where children are able to flexibly weigh information before they are strongly biased toward individuating information.

References

Boseovski, J. J., Chiu, K., & Marcovitch, S. (2013). Integration of behavioral frequency and intention information in young children's trait attributions. *Social Development*, *22*(1), 38–57. https://doi.org/10.1111/sode.12008

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, *42*(3), 500–513. https://doi.org/10.1037/0012-1649.42.3.500

Davidson, D. (1995). The representativeness heuristic and the conjunction fallacy effect in children's decision making. *Merrill-Palmer Quarterly*, *41*(3), 328–346.

De Neys, W. (2007). Nested sets and base-rate neglect: Two types of reasoning? *Behavioral and Brain Sciences*, *30*(3), 260–261. https://doi.org/10.1017/S0140525X07001719

De Neys, W., & Feremans, V. (2013). Development of heuristic bias detection in elementary school. *Developmental Psychology*, *49*(2), 258–269. https://doi.org/10.1037/a0028320

De Neys, W., & Vanderputte, K. (2011). When less is not always more: Stereotype knowledge and reasoning development. *Developmental Psychology*, *47*(2), 432–441. https://doi.org/10.1037/a0021313

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, *7*(1), 28–38. https://doi.org/10.1177/1745691611429354

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology*, *49*(2), 243–249. https://doi.org/10.1037/a0028278

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gualtieri, S., Buchsbaum, D., & Denison, S. (2020). Exploring information use in children's decision-making: Base-rate neglect and trust in testimony. *Journal of Experimental Psychology: General*, *149*(8), 1527–1536. https://doi.org/10.1037/xge0000726

Gualtieri, S., & Denison, S. (2018). The development of the representativeness heuristic in young children. *Journal of Experimental Child Psychology*, *174*, 60–76. https://doi.org/10.1016/j.jecp.2018.05.006

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251–273. https://doi.org/10.1146/annurev-psych-122216-011710

Jacobs, J. E., & Potenza, M. (1991). The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development*, *62*(1), 166–178. https://doi.org/10.2307/1130712

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus & Giroux.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237–251. https://doi.org/10.1037/h0034747

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.004

Kalish, C. W., & Shiverick, S. M. (2004). Children's reasoning about norms and traits as motives for behavior. *Cognitive Development*, *19*(3), 401–416. https://doi.org/10.1016/j.cogdev.2004.05.004

Koehler, J. J. (1996). The base rate fallacy reconsidered: Normative, descriptive and methodological challenges. *Behavioral and Brain Sciences*, *19*, 1–53. https://doi.org/10.1017/S0140525X00041157

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*(8), 1134–1140. https://doi.org/10.1177/0956797610376652

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284–299. https://doi.org/10.1016/j.cognition.2013.12.010

Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition*, *120*(3), 403–411. https://doi.org/10.1016/j.cognition.2011.02.003

Matthews, D. (2014). *Pragmatic development in first language acquisition*. John Benjamins Publishing Company. https://doi.org/10.1075/tilar.10

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2012). Are we good at detecting conflict during reasoning? *Cognition*, *124*(1), 101–106. https://doi.org/10.1016/j.cognition.2012.04.004

Pennycook, G., & Thompson, V. A. (2012). Reasoning with base rates is routine, relatively effortless, and context dependent. *Psychonomic Bulletin & Review*, *19*(3), 528–534. https://doi.org/10.3758/s13423-012-0249-3

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2014). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 544–554. https://doi.org/10.1037/a0034887

Schwarz, N., Strack, F., Hilton, D., & Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of "Irrelevant" information. *Social Cognition*, *9*(1), 67–84. https://doi.org/10.1521/soco.1991.9.1.67

Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development*, *84*(2), 443–454. https://doi.org/10.1111/j.1467-8624.2012.01865.x

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011). The complexity of developmental predictions from dual process models. *Developmental Review*, *31*(2-3), 103–118. https://doi.org/10.1016/j.dr.2011.07.003

Statistics Canada. (2017). *Kitchener - Cambridge - Waterloo [Census metropolitan area], Ontario and Ontario [Province] (table)*. Census Profile. 2016 Census. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E

Tecwyn, E. C., Denison, S., Messer, E. J., & Buchsbaum, D. (2017). Intuitive probabilistic inference in capuchin monkeys. *Animal Cognition*, *20*(2), 243–256. https://doi.org/10.1007/s10071-016-1043-9

Thevenow-Harrison, J. T., & Kalish, C. W. (2016). Learning from non-representative instances: Children's sample and population predictions. *Journal of Experimental Child Psychology*, *152*, 335–342. https://doi.org/10.1016/j.jecp.2016.07.013

Toplak, M. E. (2018). The development of rational thinking: Insights from the heuristics and biases literature and dual process models. In L. J. Ball & V. A. Thompson (Eds.), *The Routledge international handbook series. The Routledge international handbook of thinking and reasoning* (pp. 542–558). Routledge/Taylor & Francis Group.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Walker, C. M., Bridgers, S., & Gopnik, A. (2016). The early emergence and puzzling decline of relational reasoning: Effects of knowledge and search on inferring abstract concepts. *Cognition*, *156*, 30–40. https://doi.org/10.1016/j.cognition.2016.07.008