Preschoolers Use Sampling Information to Infer the Preferences of Others

Tamar Kushnir (tkushnir@umich.edu)
Department of Psychology, University of Michigan
520 Church St. Ann Arbor, MI 48109

Fei Xu (fei@psych.ubc.ca)
Department of Psychology, University of British Columbia
2406-2136 West Mall, Vancouver, B.C., Canada V6T 1Z4

Henry M. Wellman (hmw@umich.edu)
Department of Psychology, University of Michigan
520 Church St. Ann Arbor, MI 48109

Abstract
We investigate whether an apparent violation of random sampling drives children to infer that an agent who is selecting the sample is expressing a preference. Preschool children were shown a box containing two types of toys in some proportion: 100% to 0, 50% to 50%, or 18% to 82%. They were introduced to an agent (a puppet) who looked inside the box and chose 5 toys, all of them of one type. Children were then asked infer the puppet’s preference based on his choices. We found that children used the statistical information to infer the agent’s preference: the more the sample deviated from the population of objects in the box, the more likely the children were to infer that the agent preferred his chosen type of toy. We provide the first evidence that children’s social understanding interacts with statistical information in preference attribution.

Keywords: Preference attribution, statistical inference, sampling.

Introduction
In recent years, a number of researchers have investigated how infants and young children are able to use patterns of statistical evidence to make meaningful inferences about the world. For instance, studies of children’s causal learning have shown that young children are able to use patterns of probabilities to infer the causal structure of events (Gopnik et al. 2004; Kushnir & Gopnik, 2005). Studies of children’s word learning have shown that infants and young children generalize new words differently depending on the sampling variability in past usage (Xu & Tenenbaum, 2007). Several studies have shown that children can revise existing beliefs based on new, often limited, evidence (Kushnir & Gopnik, 2007; Schulz & Gopnik, 2004; Schulz, Bonawitz & Griffiths, 2007), thus their learning can be described as a form of rational, statistical inference (Xu & Tenenbaum, 2007).

Another area in which children make fast and accurate inferences is in their understanding of other people (Wellman, 1990). In particular, by the time children are in preschool they are able to reason in quite sophisticated ways about the causes of intentional actions - about goals, desires, beliefs and other mental states (Flavell, Green, & Flavell, 1995; Gergely & Csibra, 2003; Meltzoff, 1995; Perner, 1991). To date, however, little is known about the relevance that statistical information might play in shaping children’s inferences about the underlying psychological causes of actions.

The present study looks at how one particular type of statistical cue – namely that sampling proportions are usually representative of their underlying population proportions – might play a role in children’s learning about mental states. There is recent evidence which shows that preschoolers expect samples to be representative of the larger population when they are randomly drawn. For example, if 3- and 4- children are shown a sample of 4 red balls and 1 white ball, they expect the population to contain mostly red balls and some white ones and not visa-versa. Similarly, if children are shown a population of mostly red balls and some white ones, they expect the sample to contain many red balls and some white ones (Denison et al. 2006).

What if a sample of objects drawn from a population is non-representative? For example, consider a scenario in which children are shown a box containing a population of mostly red and some white balls. They then see an agent who deliberately draws only white balls from the box – a non-representative sample. Since the sampling was done by intentional action, children should reason that the agent’s behavior was selective rather than random, and thus may have been caused by a preference for white balls.

In the following study, we investigate preschoolers’ inferences based on a violation of random sampling which occurs as a consequence of someone’s actions. We used a modified version of the experimental paradigm used by Denison et al. (2006), in which a small set of toys (the sample) is chosen from a box containing a larger set of toys (the population). The critical modification in the current experiment is that the agent sampled intentionally (with eyes open rather than eyes closed).

Also critical to the current investigation is the fact that the population proportions were varied across groups of subjects. Thus, children were randomly assigned to one of
three conditions. In the 100% condition, boxes contained only one type of object. In the 50% condition, boxes contained two types of objects – one which was selected by the agent and one which was not – in a 1:1 ratio. In the 18% condition the minority of objects in the box (only 18%) were of the type selected by the agent. Moreover, in each condition the agent exhibited behaviors which on their own might signify a preference: he picked multiple toys (five) of the same type, he spent the same amount of time selecting objects in all conditions, and he played with the set of objects for the same length of time in all conditions. Nonetheless, we hypothesized that children’s inferences about the preferences of an agent can be driven by statistical evidence alone (i.e. the probability that the sample is non-random), controlling for the quantitative (intentional, emotional) properties of the actions themselves. Thus, we expected that children will infer that the agent is expressing a preference most in the 18% condition and least in the 100% condition and the 50% condition should fall somewhere in between.

Method

Participants

Seventy-three preschoolers participated in the experiment (33 3-year-olds and 40 4-year-olds; overall mean age = 4 years, 1 month). Children were randomly assigned to one of three conditions with the stipulation that the average ages match across conditions: 25 children were assigned the 100% Target Condition, 24 were assigned to the 50% Target Condition, and 24 were assigned to the 18% Target Condition.

Design and Materials

There were two sets of toys used in the study. In each set, one type of toy was designated the “target” which was chosen by the agent. A second type of toy was present in the box in the 50% and 18% conditions. A third type of toy was presented at beginning and end of the task along with the box in the 100% condition, they saw 2 trials each with a box full of 100% of one type of object. The set of toys presented first, foam shapes or balls, was counterbalanced across subjects. The “sample” was always 5 toys of one type, counterbalanced along with the box contents. An additional set of 5 toys of each type were kept in 3 separate bowls to be used for the preference question at the end of the procedure (see below). Order of placement of the bowls (left, right, center) was also counterbalanced.

Procedure

Each child was tested individually in a quiet corner of a daycare classroom. The child sat opposite the experimenter at a low table.

Step 1: Introduction of agent. The experimenter began by introducing a puppet (a Squirrel) to the child, “This is my friend Squirrel. Can you say hi? He really wants to play with some toys today. Squirrel really likes some toys, and other toys he doesn’t like at all. Should we let him play?”

Step 2: Familiarization with objects. The squirrel was then put behind the table, (“Bye Squirrel! See you soon!”) so that the experimenter could familiarize the child with the three types of toys for the trial. She did this by taking out the three bowls, placing them on the table in a row, and asking the child to describe what they saw. For example, the experimenter took out three bowls of foam shapes and asked the child, “Can you tell me what these are?” She prompted the child to describe the types of objects by name if the child had not done so spontaneously (e.g., blue flowers, red circles, baseballs, basketballs). The familiarization was included to insure that, later on, when tracking what was familiar to and novel for the puppet, the novelty/familiarity of the toys for the child herself wouldn’t interfere.

Step 3: Introduction of population. After the familiarization, the experimenter removed the bowls from the table and took out the first box. She opened up the top so the child could look inside and touch the objects if she/he wanted to. The experimenter said, “I have some toys in this box. Can you tell me what you see inside this box?” She prompted the child to describe the objects by name without reference to the ratios of the objects.

Step 4: Sample selection. Then experimenter said to the child, “Maybe Squirrel would like to play with these toys. Let’s bring him back and see.” The experimenter brought Squirrel back and asked, “Squirrel, would you like some toys to play with? You can take some from this box.” Squirrel took 5 objects of one type (the “target”) placed them on the table in a pile. After the 5th item, the experimenter closed the box and Squirrel “played” with the toys then left again. The amount of time spent selecting the toys and them playing with them was kept equal across the conditions. After Squirrel left, the experimenter and child...
Step 5: Preference question. The experimenter then brought out the 3 bowls and placed them in front of the child. Squirrel then came back and the experimenter said to the child, “Squirrel is back! He wants some toys again. Can you give him the ones that he likes?” The child’s first choice was recorded.

After putting away the toys, the experimenter began the second trial. Steps 2-5 were repeated with the other set of objects (the balls, for example, if the foam shapes were used in trial 1).

As stated earlier, we predicted that children would choose the target object the least in the 100% condition and the most in the 18% condition, and the 50% condition should fall somewhere in between. A linear contrast (with coefficient -1, 0, 1) revealed a statistically significant trend, \( t(70) = 3.485, p = .001 \). One-tailed t-tests were performed comparing the three conditions. There was a marginally significant difference between the 100% and the 50% conditions, \( t(47) = 1.506, p = .07 \); there was a significant difference between the 50% and the 18% conditions, \( t(46) = 1.956, p = .03 \); and there was a significant difference between the 100% and 18% conditions, \( t(47) = 3.59, p = .0005 \).

Table 1 shows the distribution of responses in each condition. The consistency across trials – the number of children who gave the Squirrel the sampled (target) toy on both trials – shows the same pattern as the ANOVA above. In the 18% condition, 17 children (71%) chose the target object on both trials. In the 50% condition, 11 children (46%) chose the target object on both trials, and in the 100% condition 7 children (28%) chose the target object on both trials.

Moreover, table 1 indicates that, on trials where children did not give Squirrel the target toy, their responses were roughly equally distributed between the alternate toy in the box and the novel toy which was not in the box. That is, they did not seem to be inferring that Squirrel was avoiding the other toy in the box, but rather that he was expressing a preference for the target. Further studies would be needed, however, to confirm this interpretation of the results.

As stated earlier, we predicted that children would choose the target object the least in the 100% condition and the most in the 18% condition, and the 50% condition should fall somewhere in between. A linear contrast (with coefficient -1, 0, 1) revealed a statistically significant trend, \( t(70) = 3.485, p = .001 \). One-tailed t-tests were performed comparing the three conditions. There was a marginally significant difference between the 100% and the 50% conditions, \( t(47) = 1.506, p = .07 \); there was a significant difference between the 50% and the 18% conditions, \( t(46) = 1.956, p = .03 \); and there was a significant difference between the 100% and 18% conditions, \( t(47) = 3.59, p = .0005 \).

Table 1 shows the distribution of responses in each condition. The consistency across trials – the number of children who gave the Squirrel the sampled (target) toy on both trials – shows the same pattern as the ANOVA above. In the 18% condition, 17 children (71%) chose the target object on both trials. In the 50% condition, 11 children (46%) chose the target object on both trials, and in the 100% condition 7 children (28%) chose the target object on both trials.

Moreover, table 1 indicates that, on trials where children did not give Squirrel the target toy, their responses were roughly equally distributed between the alternate toy in the box and the novel toy which was not in the box. That is, they did not seem to be inferring that Squirrel was avoiding the other toy in the box, but rather that he was expressing a preference for the target. Further studies would be needed, however, to confirm this interpretation of the results.

Table 1: Distribution of responses across the two trials in the 100%, 50% and 18% conditions.

<table>
<thead>
<tr>
<th></th>
<th>100%</th>
<th>50%</th>
<th>18%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target twice</td>
<td>7</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>Target + alternate toy*</td>
<td>NA</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Target + novel toy</td>
<td>10</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Alternate + novel toy*</td>
<td>NA</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Alternate toy twice*</td>
<td>NA</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Novel toy twice</td>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

* Alternate refers to the other toy in the box; 50% and 18% conditions only. When children in the 100% condition chose non-targets, their choices were roughly equally distributed among the two novel toys.

Conclusion

Our results showed that preschoolers were able to use sampling information to infer the preferences of an intentional agent. When children were presented with an apparent violation of random sampling, they used the statistical evidence to infer that the agent was expressing a preference for one type of object. The more the sample deviated from a random sample of the population, the more likely children were to infer a preference. To our knowledge, these results provide the first evidence that, armed with their knowledge of intentional actions, preschool children are capable of making inferences about...
social/psychological phenomena based on statistical information.

Acknowledgments
Thanks to Emily Hass for her help in data collection. TK was supported by the McDonnell Collaborative Initiative on Causal Learning and a NICHD Post-Doctoral Fellowship.

References


