Young children's use of statistical sampling evidence to infer the subjectivity of preferences

Lili Ma [a,*], Fei Xu [b]

[a] Department of Psychology, Ryerson University, Canada
[b] Department of Psychology, University of California, Berkeley, USA

ABSTRACT

A crucial task in social interaction involves understanding subjective mental states. Here we report two experiments with toddlers exploring whether they can use statistical evidence to infer the subjective nature of preferences. We found that 2-year-olds were likely to interpret another person's nonrandom sampling behavior as a cue for a preference different from their own. When there was no alternative in the population or if the sampling was random, 2-year-olds did not ascribe a preference and persisted in their initial beliefs that the person would share their own preference. We found similar but weaker patterns of responses in 16-month-olds. These results suggest that the ability to infer the subjectivity of preferences based on sampling information begins to emerge between 16 months and 2 years. Our findings provide some of the first evidence that from early in development, young children can use statistical evidence to make rational inferences about the social world.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Human social interaction depends on the understanding of our own and other people's minds. A competent participant in the social world must explain and predict other people's behavior by reference to subjective mental states, such as desires, emotions, intentions, preferences, and beliefs (e.g., Flavell, 1999; Perner, 1991; Wellman & Gelman, 1992). In particular, an important task in social interaction involves understanding the subjectivity of preferences. As a source of motivation that enables an agent's choice of one option over another, preferences are subjective and often person-specific – different people can have different attitudes toward the same entity. On many occasions, we must bear in mind this subjective nature of preferences in order to maintain amicable social interactions.

An intriguing question concerns how young children come to understand the subjectivity of preferences. One possibility is that children use verbal information that directly addresses the preferences of others, especially statements in the form of a comparison, such as "I like this one better than that one". However, verbal cues are not always available. In everyday social interaction, people do not always make explicit statements about what they like or dislike, especially when their preferences are in conflict with those of others (e.g., DePaulo & Bell, 1996; Saarni, 1984; Talwar & Lee, 2002). Another possibility is that children rely on emotional cues, given that the expression of a preference is often associated with relatively positive affects toward one object but neutral or negative affects toward another option. Research has shown that by about 18 months, infants can infer from emotional cues that another person may have a preference different from their own (Repacholi & Gopnik, 1997).

In addition to verbal or emotional cues, statistical patterns in the choices that other people make may also help children understand that preferences are subjective.

* Corresponding author. Address: Department of Psychology, Ryerson University, 350 Victoria Street, Toronto, Ontario, Canada M5B 2K3. Tel.: +1 416 979 5000x2694; fax: +1 416 979 5273.

E-mail address: lilima@psych.ryerson.ca (L. Ma).

Excluding accidental or involuntary behavior, human actions at large are statistically nonrandom in that they are goal-directed or intentional (e.g., White, 1988; Woodward, 2003). Imagine the following scenario: a child picks out five blocks from a box that contains mostly toy cars and just a few blocks. His intentional act to choose only the blocks is a violation of random sampling: given the base rates of the two types of objects, if the child sampled randomly, he would most likely get a few toy cars and perhaps a couple or none of the blocks. Such nonrandom sampling behavior may suggest that the child prefers the blocks to the toy cars, even though toy cars may seem inherently more interesting. Our goal in this paper is to examine whether young children are able to use such statistical evidence to infer the subjectivity of preferences.

Previous research has shown that young children can use statistical evidence in various areas of non-social reasoning. For example, preverbal infants use transitional probability in segmentation of speech streams or visual shapes (e.g., Kirkham, Slemmer, & Johnson, 2002; Saffran, Aslin, & Newport, 1996). When generalizing novel words, preschoolers are sensitive to the sampling processes that generated the labeled examples (Xu & Tenenbaum, 2007). Young children can also use patterns of probabilities to infer the causal structures of events (e.g., Gopnik & Schulz, 2004). In addition, preverbal infants can make inferences about a randomly drawn sample from information about the population, and vice versa (e.g., Xu & Garcia, 2008).

Little attention has been paid to young children's ability to use statistical evidence in social reasoning, despite the fact that in the social world children encounter various types of statistical information, such as behavioral violations of social norms, actions that are less probable, and consistency or contingency across and within individual actions. Only recently have researchers begun to study this area. Xu and Denison (2009) found that 11-month-olds considered intentions when interpreting sampling events: if a person expressed a preference beforehand, infants expected the outcome of her sampling behavior to match her preference, regardless of whether or not the sample was representative of the population. This finding suggests the possibility that a violation of random sampling can indicate a preference. Indeed, Kushnir, Xu, and Wellman (2010) found that preschoolers and 20-month-olds interpreted an agent's nonrandom sampling behavior as expressing a preference for one type of objects over another.

The current study examines whether 16-month-olds and 2-year-olds can use sampling evidence to infer the subjectivity of preferences, using a behavioral-choice paradigm modified after Repacholi and Gopnik (1997). In Experiment 1, we first measured children's own preference, as well as their initial beliefs about an adult's preference, between two types of objects, one highly interesting and the other one boring. Then children observed as the adult sample six boring objects from a population. In the experimental condition (Boring 13%), the population consisted of mostly interesting and just a few boring objects, so the sampling was nonrandom and indicated a preference for the boring objects. In the control condition (Boring 100%), the population consisted of only the boring objects

with no other option, so the sampling behavior did not license the inference for a preference. In both conditions, after the sampling event children were asked again to indicate the adult's preference. We were interested in whether children's inference about the adult's preference after the sampling event would differ from their initial beliefs.

Given our intuition that generally people would prefer an interesting object to a boring one in the absence of other cues, we predicted that initially children at both ages would expect the adult to share their preference for the interesting objects. After the sampling event, if children were rational learners, their response patterns would differ in accord with the sampling evidence. In the Boring 13% condition, we predicted that 2-year-olds would interpret the evidence of nonrandom sampling as a cue for a preference different from their own, in which case they would be more likely to offer the boring objects to the adult after the sampling event, based on the findings of Repacholi and Gopnik (1997) and Kushnir et al. (2010). In the Boring 100% condition, because the sampling behavior did not license a preference, we predicted that 2-year-olds would disregard the sampling information and offer the same interesting objects to the adult after the sampling event. Since this study is the first attempt to use a behavioral-choice paradigm to examine statistical inference in children younger than 20 months, we had no specific predictions with the 16-month-olds.

2. Experiment 1

2.1. Method

2.1.1. Participants

Participants were 32 sixteen-month-olds (mean age 16; 14 [months; days]; range 15; 16 to 17; 22; 17 girls) and 32 two-year-olds (mean age 26; 11; range 24; 2 to 29; 19; 16 girls). Children were recruited from the greater Vancouver area in Canada, and parents were contacted by telephone. Eighteen additional children were excluded, due to unwillingness to respond (13), difficulty following the instructions (2), parental interference (2), or experimenter error (1).

At each age, 16 children (approximately half each gender) participated in the Boring 13% condition and 16 in the Boring 100% condition (16-month-olds: mean age 16; 14 in both conditions; 2-year-olds: mean ages 26; 12 and 26; 10, respectively).

2.1.2. Materials

Four transparent jars were used to contain the populations, two in each condition. In the Boring 13% condition, each jar contained 13% of boring objects and 87% of interesting objects: 9 white wooden cubes and 60 orange slinkies with pumpkin-face print, or 9 plain wooden balls and 60 purple–green spin tops with rattling sounds. In the Boring 100% condition, each jar contained only the boring objects (69 cubes or 69 wooden balls). All the objects were of similar sizes (approximately 2.5 cm × 2.5 cm). A transparent container was used to display the sample drawn from the jar (i.e., six cubes or six wooden balls).

The test items were four sets of objects, presented in pairs. One pair consisted of four cubes and four slinkies (the "cube vs. slinky" pair). The other pair consisted of four wooden balls and four spin tops (the "ball vs. spin top" pair). Each set of objects was contained in a small paper bowl.

2.1.3. Design and procedure

The experiment employed a 2 (age) × 2 (condition) between-subjects design. Each child was randomly assigned to one of the two conditions. There were two test trials in each condition, one with the "cube vs. slinky" pair and the other with the "ball vs. spin top" pair. The order of the two trials was counterbalanced across the participants in each condition.

In the waiting room, after children felt at ease, the experimenter initiated a game, in which she offered children two small toys to play with and then asked them to share, by placing one hand with palm facing up and asking, "I like to have a toy to play with. Can I have one?" This was to ensure that children understood the hand gesture as a request for a desired object (Repacholi & Gopnik, 1997; Tomasello & Haberl, 2003). Most children responded immediately by placing a toy on the experimenter's hand; a few 16-month-olds did so after one or more prompts. The experimental procedure then followed.

Each child was tested individually in a quiet room. Children were seated in a high chair opposite the experimenter at a table. Their parents sat next to them and were instructed not to interfere with the procedure. A video camera recorded the entire session. Each of the two test trials involved four phases: child preference, baseline, sampling, and test.

2.1.3.1. Child preference. At the beginning of each trial, children were asked to indicate their own preference between two types of objects. The experimenter first introduced the contents of two small bowls to children: four interesting objects vs. four boring ones. Then she placed the two bowls on the table (about 90 cm from each other; the positions of the two bowls were counterbalanced), and the child was midway between them. Each bowl was at a distance out of children's reach. While looking at children, the experimenter asked them to choose between the two bowls, "Which one do you like to play with? Just choose one!" After children made a choice by either pointing or reaching, the experimenter pushed the chosen bowl to them and let them play with the objects inside. Children were also given a chance to examine the objects in the other bowl. All objects were then put back in the bowls.

2.1.3.2. Baseline. This phase was to assess children's prior beliefs about the experimenter's preference. The experimenter switched the positions of the two bowls and placed them on the table within children's reach. She then asked children about her own preference, by placing one hand, palm facing up, exactly midway between the two bowls and asking, "I like to have a toy to play with. Can I have the one I like?" The object that children placed in her hand was coded as children's choice. Occasionally children put

both an interesting object and a boring one in her hand, in which case the experimenter put the objects back and asked children to choose again and only choose one object (e.g., "Which one do you think I like better? Just choose ONE."). The experimenter then thanked children for sharing and removed everything from the table.

2.1.3.3. Sampling. During this phase, the experimenter sampled six boring objects from the population as children observed. She first brought out a jar and directed children's attention to the objects inside (e.g., "Look! I have a big jar. There are two kinds of things in it [Boring 13% condition]/ there is only one kind of things in it [Boring 100% condition]. I am going to get some!"). Then she picked six boring objects from the jar, one at a time, and placed them in the display container. In both conditions, she spent the same amount of time and effort sampling the objects and her facial expressions were neutral. At the end of the sampling event, she directed children's attention to both the population and the sample, "Look! This many (holding the jar), and I got six of this one (holding the display container)." Then the jar and the sample were removed from the table and remained invisible to children during the test phase.

2.1.3.4. Test. This phase was to examine whether the sampling information would affect children's judgment of the experimenter's preference. Immediately after the sampling event, the experimenter asked children about her own preference a second time, in the same manner as in the baseline phase. The positions of the two bowls were switched as compared to the baseline phase. The object that children placed in the experimenter's hand was coded as their choice.

2.1.4. Coding and reliability

The experimenter coded children's object choices on each trial. A research assistant who was blind to the hypothesis independently coded the complete sample from videotapes. The two coders agreed 97% of the time. Disagreement was resolved with discussion.

2.2. Results

Preliminary analyses revealed no significant effects of gender, trial order, or trial type ("cube vs. slinky" or "ball vs. spin top") on children's performance in both experiments. These factors were thus not included in the main analyses. All reported p values are 2-tailed.

2.2.1. Children's own preference

During the child-preference phase, when asked to indicate their own preference, children chose the interesting objects over the boring ones most of the time (94% of the trials; 60 out of 64 trials in each condition). That is, they chose the slinkies over the cubes on one trial and the spin tops over the wooden balls on the other trial. In each condition, the 16-month-olds chose the interesting objects on 29 out of 32 trials, as did the 2-year-olds on 31 out of 32 trials. When children chose the cubes or the balls (8 out of 128 trials), we considered these objects as the "interesting" ones from the children's perspectives and changed the populations accordingly, so the sampling event consisted

of drawing six slinkies or six spin tops (i.e., the "boring" objects for these children). In the rest of Experiment 1, the term "interesting objects" referred to the type of objects that children chose for themselves during the child-preference phase.

2.2.2. Children's initial beliefs about another person's preference

During the baseline phase, when the experimenter asked children to hand her the type of objects that she would prefer, most of the time children gave her the same interesting objects that they had chosen for themselves (see Table 1). There was no significant age or condition difference (t ranges from .38 to 1.82, $p > .05$, independent-samples t-tests). Occasionally children offered the boring objects to the experimenter at baseline. One possibility is that children actually thought that the experimenter would prefer the boring objects. Alternatively, children might be unwilling to share the interesting objects, as some children claimed that the interesting objects were "all mine" and pushed all the boring ones toward the experimenter.

2.2.3. Children's beliefs about another person's preference after the sampling event

Table 1 also shows the mean proportions of trials on which children gave the interesting objects to the experimenter during the test phase. With test phase (baseline vs. test) as the within-subjects factor, condition and age as the between-subjects factors, a mixed-design ANOVA revealed a significant main effect of test phase, F $(1, 60) = 14.24$, $p < .001$, $\eta^2 = .19$. Children were less likely to offer the interesting objects to the experimenter after the sampling event than at baseline ($M_{Test} = .67$, $SD = .36$, and $M_{Baseline} = .84$, $SD = .23$, respectively).

More importantly, there was a significant interaction between test phase and condition, F $(1, 60) = 23.06$, $p < .001$, $\eta^2 = .28$. In the Boring 13% condition, children chose to give the interesting objects to the experimenter significantly less often after the sampling event than at baseline, t $(31) = -5.58$, $p < .001$ (paired-samples t-test). In the Boring 100% condition, the effect of test phase was not significant and children chose to give the interesting objects to the experimenter both before and after the sampling event, t $(31) = .83$, n.s. The other main effects or interactions were not significant.

Our main goal was to see whether children could learn from violations of random sampling that preferences are subjective, based on the premise that children initially

expect others to like the same interesting objects as themselves. Across the two conditions, 20 out of 64 children offered the boring object to the experimenter at baseline on one trial, possibly because they thought she had a different preference or because they were unwilling to share the interesting objects, which would confound the main results about children's *learning* from the sampling evidence. Three additional children offered both types of objects to the experimenter at baseline on one trial. To be conservative, we conducted further analyses to examine children's performance after the sampling event, with only the children who offered the interesting objects to the experimenter at baseline on *both* trials ($N = 41$).

We submitted children's interesting-object choice (proportion) during the test phase to an ANOVA, with condition and age as the between-subjects factors. A significant main effect of condition emerged, F $(1, 37) = 17.50$, $p < .001$, $\eta^2 = .32$. After the sampling event, children in the Boring 13% condition were less likely to offer the interesting objects to the experimenter than children in the Boring 100% condition ($M_{Boring\ 13\%} = .50$, $SD = .35$; $M_{Boring\ 100\%} = .91$, $SD = .20$). The age effect and the interaction were not significant.

For exploratory purposes, we conducted post hoc tests with these children to examine the condition effect by age. The 2-year-olds in the Boring 13% condition chose the interesting objects significantly less often than their peers in the Boring 100% condition, t $(18) = -3.83$, $p = .001$. With the 16-month-olds, this condition effect was marginally significant, t $(19) = -2.04$, $p = .056$ (see Fig. 1).

2.3. Discussion

Experiment 1 indicated that by about 26 months, children were sensitive to sampling evidence when deciding whether another person had the same preference as their own. Before the sampling event, when the person asked for the objects that she would prefer, children gave her the same interesting objects that they had chosen for themselves, suggesting their initial beliefs that another person would share their own preference. After watching the person choose some boring objects from a population of mostly interesting and just a few boring objects (i.e., nonrandom sampling), 2-year-olds were less likely to offer her the interesting objects in response to her request. Instead they were more likely to give her the boring objects, suggesting that children were sensitive to the evidence of nonrandom sampling and took it as a cue for a preference

Table 1
Mean proportion of trials on which children gave the interesting objects to the experimenter, by age, condition, and test phase (SD in parentheses; Experiment 1).[a]

Age	Boring 13%			Boring 100%		
	N	Baseline	Test	N	Baseline	Test
16-month-olds	16	.94 (.17)	.56 (.31)	16	.81 (.25)	.91 (.20)
2-year-olds	16	.88 (.22)	.47 (.39)	16	.75 (.26)	.75 (.37)
Total	32	.91 (.20)	.52 (.35)	32	.78 (.25)	.83 (.30)

[a] Three 16-month-olds chose both types of objects at baseline on one trial. One 2-year-old did not make a clear choice during the test phase on one trial. Thus mean proportions of trials were reported.
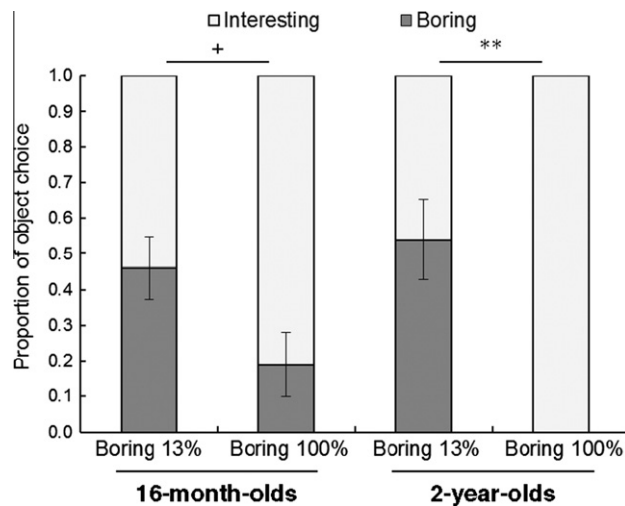
Fig. 1. Children's object choice during the test phase, by age and condition (Experiment 1). ^+p = .056, $^{**}p$ = .001.

different from their own. In contrast, after watching the person choose a few boring objects from a population of only boring objects with no alternative, 2-year-olds offered her the same interesting objects as they did before the sampling event.

As with the 2-year-olds, 16-month-olds showed a similar but weaker pattern of results. Thus, a tentative conclusion can be drawn that between the ages of 16 and 26 months, children begin to show sensitivity to sampling information and use the evidence of nonrandom sampling to infer another person's preference that is different from their own.

A possible alternative explanation of the results is that in the Boring 13% condition, when children gave the boring object to the experimenter after the sampling event, they might not consider the ratio of the objects in the population. Rather, they might have offered the experimenter the boring object because she had repeatedly chosen it over an alternative. In the Boring 100% condition, there was no other choice in the population, so the results could not rule out this possibility. To address this, in Experiment 2 we contrasted a Boring 12% condition with a Boring 88% condition. The Boring 12% condition was to replicate the Boring 13% condition in Experiment 1. In the Boring 88% condition, each population consisted of mostly boring and just a few interesting objects, and the experimenter randomly sampled six boring ones. There was an alternative in the population, but the sampling was random and did not license a preference. If children responded differently across the two conditions as in Experiment 1, the alternative explanation mentioned above would be ruled out.

The second goal of Experiment 2 was to examine whether children would persist in their own preference after the sampling event. When children inferred from the evidence of nonrandom sampling that an adult might have a preference different from their own, would they be swayed by the adult's preference or would they persist in their own initial preference? If it were the latter, we would have more evidence that children clearly understand the subjectivity of preferences. To examine this

question, at the end of Experiment 2 children were asked a second time about their own preference between the two types of objects.

3. Experiment 2

3.1. Method

Participants were 32 sixteen-month-olds (mean age 16; 16; range 15; 10 to 17; 17; 16 girls) and 32 two-year-olds (mean age 26; 17; range 24; 2 to 30; 8; 16 girls). None participated in Experiment 1. At each age, 16 children participated in the Boring 12% condition and 16 in the Boring 88% condition (16-month-olds: mean age 16; 16 in both conditions; 2-year-olds: mean ages 26; 19 and 26; 15, respectively). Five additional 16-month-olds were excluded because of unwillingness to respond.

The same materials from Experiment 1 were used, except that the ratio of the two types of objects in each population was different. In the Boring 12% condition, each jar contained 7 boring objects and 53 interesting ones. In the Boring 88% condition, each jar contained 53 boring objects and 7 interesting ones.

The design and the procedure were the same as in Experiment 1, with the addition of a fifth phase after the test phase, in which the experimenter asked children a second time about their own preference ("child-preference2"), in the same manner as in the initial child-preference phase.

The experimenter coded children's choices. Later a research assistant who was blind to the hypothesis independently coded the complete sample from videotapes. The two coders agreed 98% of the time. Disagreement was resolved with discussion.

3.2. Results and discussion

Experiment 2 replicated the main findings of Experiment 1. During the initial child-preference phase, when asked about their own preference, most of the time

Table 2

Mean proportion of trials on which children gave the interesting objects to the experimenter, by age, condition, and test phase (SD in parentheses; Experiment 2).[a]

Age	Boring 12%			Boring 88%		
	N	Baseline	Test	N	Baseline	Test
16-month-olds	16	.91 (.20)	.66 (.30)	16	.94 (.17)	.84 (.24)
2-year-olds	16	.88 (.22)	.44 (.40)	16	.91 (.20)	.81 (.25)
Total	32	.89 (.21)	.55 (.37)	32	.92 (.18)	.83 (.24)

[a] One 2-year-old chose both types of objects at baseline on one trial. Six children (3 each age) chose both types of objects during the test phase on one trial. Thus mean proportions of trials were reported.

children chose the interesting objects over the boring ones (84% of the trials; Boring 12% condition: 53 out of 64 trials; Boring 88% condition: 55 out of 64 trials). In the Boring 12% condition, the 16-month-olds chose the interesting objects on 26 out of 32 trials, as did the 2-year-olds on 27 out of 32 trials. In the Boring 88% condition, the 16-month-olds chose the interesting objects on 28 out of 32 trials, and the 2-year-olds on 27 out of 32 trials. As in Experiment 1, when children chose the cubes or the balls (20 out of 128 trials), we considered these objects as the interesting ones from the children's perspectives and changed the populations accordingly. In the rest of this paper, the term "interesting objects" referred to the type of objects that children chose for themselves during the initial child-preference phase.

During the baseline phase, when asked about the experimenter's preference, most of the time children chose the same interesting objects that they themselves had opted (see Table 2). There was no significant age or condition effect (*t* ranges from .42 to .47, *p* > .05). Occasionally children offered the boring objects to the experimenter at baseline. Perhaps children did so either because they thought that the experimenter liked the boring objects better or because they were unwilling to share the interesting objects.

We then compared children's judgment of the experimenter's preference before and after the sampling event (i.e., baseline vs. test). With test phase as the within-subjects factor, condition and age as the between-subjects factors, a mixed-design ANOVA revealed a significant main effect of test phase, $F(1, 60) = 20.07$, $p < .001$, $\eta^2 = .25$. Children were less likely to offer the interesting objects to the experimenter after the sampling event than at baseline ($M_{Test} = .69$, $SD = .34$ and $M_{Baseline} = .91$, $SD = .20$, respectively). The main effect of condition was also significant, $F(1, 60) = 13.70$, $p < .001$, $\eta^2 = .19$. Children in the Boring 12% condition were less likely to offer the interesting objects to the experimenter than children in the Boring 88% condition, but this effect was significant only during the test phase, $t(62) = -3.62$, $p = .001$ (independent-samples *t*-test). In addition, there was a significant interaction between test phase and condition, $F(1, 60) = 6.55$, $p = .01$, $\eta^2 = .10$. In the Boring 12% condition, children were less likely to offer the interesting objects to the experimenter after the sampling event than at baseline, $t(31) = -4.34$, $p < .001$ (paired-samples *t*-test). In the Boring 88% condition, the effect of test phase was not significant, $t(31) = -1.65$, n.s. The main effect of age and the other interactions were not significant.

As in Experiment 1, we conducted further analyses to examine children's performance after the sampling event, with only the children who offered the interesting objects to the experimenter at baseline on *both* trials ($N = 52$). We submitted children's interesting-object choice (proportion) to an ANOVA, with condition and age as the between-subjects factors. A significant main effect of condition emerged, $F(1, 48) = 11.63$, $p = .001$, $\eta^2 = .20$. Children in the Boring 12% condition were less likely to give the interesting objects to the experimenter than children in the Boring 88% condition ($M_{Boring\ 12\%} = .52$, $SD = .39$, and $M_{Boring\ 88\%} = .81$, $SD = .25$, respectively). The main effect of age was also significant, $F(1, 48) = 4.38$, $p = .04$, $\eta^2 = .08$. The 16-month-olds were more likely to offer the interesting objects to the experimenter than were the 2-year-olds. The interaction between condition and age was not significant.

For exploratory purposes, post hoc tests were conducted to examine the condition effect by age. After the sampling event, when asked about the experimenter's preference, the 2-year-olds in the Boring 12% condition chose the interesting objects significantly less often than their peers in the Boring 88% condition, $t(23) = -2.79$, $p = .01$. With the 16-month-olds, this condition effect was marginally significant, $t(25) = -1.91$, $p = .067$ (see Fig. 2).

During the "child-preference2" phase, when asked again about their own preference, most of the time children chose the same interesting objects that they had favored during the initial child-preference phase (87% of the trials; Boring 12% condition: 53 out of 60 trials; Boring 88% condition: 49 out of 57 trials).[1] Of particular interest is whether children would change their own preference after they offered the boring objects to the experimenter during the test phase. In the Boring 12% condition, after children chose to give the boring objects to the experimenter, they themselves still preferred the interesting objects most of the time (16-month-olds: 82% of the trials; 2-year-olds: 94% of the trials); so did children who incorrectly offered the boring objects to the experimenter in the Boring 88% condition (60% and 83% of the trials, respectively).

To summarize, as in Experiment 1, 2-year-olds were sensitive to the evidence of nonrandom sampling and took it as a cue for a preference different from their own. They were not swayed by this inconsistency, however, and most

---

[1] In the Boring 12% condition 4 children chose the same interesting object as they did during the initial child-preference phase on one trial but both types of objects on the other trial; so did 7 children in the Boring 88% condition.
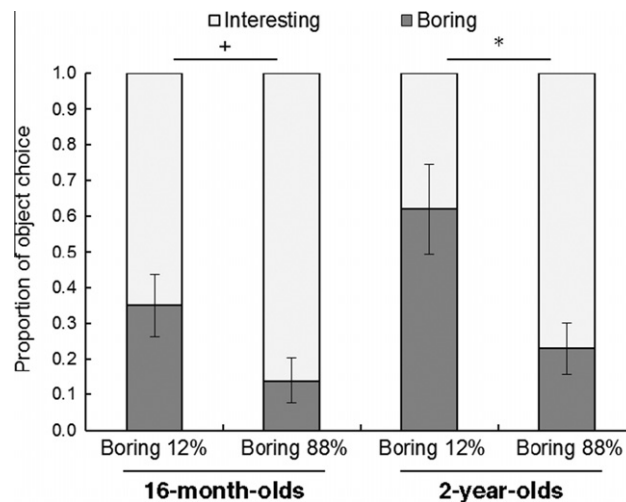
Fig. 2. Children's object choice during the test phase, by age and condition (Experiment 2). ^+p = .067, $^*p$ = .02.

of the time they still preferred the same objects that they had initially favored. When the sampling appeared random, children did not infer a preference and offered the same interesting objects to the person as they did before the sampling event. Thus, children did pay attention to the ratio of the objects in the population, and they based their choices on whether or not the sampling behavior was an obvious violation of random sampling. These results help to rule out the possible alternative explanation for the results of Experiment 1. The comparisons of the two age groups again indicate that the ability to use sampling evidence as the basis for understanding subjective preferences begins to emerge between 16 and 26 months.

4. General discussion

Many studies have shown that young children can use statistical evidence to understand various aspects of the physical world (e.g., Gopnik & Schulz, 2004; Saffran et al., 1996; Xu & Garcia, 2008). Less is known about young children's ability to use statistical evidence in social reasoning. The present study provides some of the first evidence that by age 2 children are sensitive to statistical patterns in the choices that other people make and use such evidence to reason about subjective mental states.

Across two experiments, the 2-year-olds were likely to attribute a preference to another person only when her choice behavior was an obvious violation of random sampling. When the sampling behavior did not license a preference, such as when there was only one type of objects in the population or when the sampling behavior was random, children did not ascribe a preference and persisted in their initial beliefs that the adult would share their own preference for objects that are intrinsically interesting. These findings indicate an early achievement in making rational inferences about the social world based on statistical evidence.

One might argue that the 2-year-olds might have used a shortcut based on saliency when deciding which type of objects the adult would prefer. In three conditions, after the sampling event a correct response to the test question was to choose the minority objects in the population (i.e., choosing the boring objects in the Boring 13% and Boring 12% conditions but the interesting ones in the Boring 88% condition). In all three cases, the target objects "stood out" in the population given their minority status or relative scarcity. Children might have relied on saliency alone and chose the minority objects accordingly.

Three concerns arise regarding this alternative explanation. First, after the sampling event both the population and the drawn sample were removed from the table and remained invisible to the children during the test phase. The perceptual cues of saliency were thus not available in the immediate reality when children responded to the test question. Second, the saliency of the minority objects is an inherent characteristic of nonrandom sampling, in that nonrandom sampling entails selecting the objects that are relatively rare in the population. Thus, in the Boring 13% and Boring 12% conditions "choosing the boring objects by saliency" could be in line with children's sensitivity to the violation of random sampling. Third, if saliency of the minority objects were the only cue that children had picked up on, after the nonrandom sampling event they would have chosen the minority objects (i.e., the boring ones) for themselves when they were asked a second time about their own preference, which was not the case as the children in the Boring 12% condition persisted in their initial preference for the interesting objects.

To directly tease apart the "going by saliency" alternative explanation, in a follow-up study we could simply show children a population of mostly interesting and just a few boring objects and ask them to choose which type of objects another person (and they themselves) would prefer, in the absence of any sampling behavior. If children chose to give the interesting rather than the boring objects

to the person, the "going by saliency" possibility would be ruled out. This seems very likely, given that most children have the prior beliefs that another person would prefer an interesting object to a boring one unless that person's choice behavior suggests otherwise.[2]

Our results also provide evidence that 2-year-olds can revise their prior beliefs based on the sampling evidence. Most children initially thought that another person would share their preference for the interesting objects. When the sampling information did not license a preference, they persisted in their initial beliefs. When the sampling behavior was nonrandom and expressed a preference, they were more likely to revise their initial beliefs and infer that the person might have a preference different from their own. To our knowledge, this is the first demonstration that by age 2 children are able to revise their prior beliefs based on statistical evidence and infer the subjective nature of mental states.

Our findings thus far point to the possibility of a rational learning mechanism by which 2-year-olds make meaningful inferences about subjective mental states: when judging the preference of another person, children: (a) bring prior beliefs to the situation, (b) are sensitive to the patterns of statistical evidence in that person's choice behavior, and (c) decide whether to revise their prior beliefs based on the likelihood of the sampling behavior expressing a preference different from their own. This rational learning mechanism differs from existing associative models that have been used to explain how children learn the meanings of words. Associative learning models assume that children pick up on the statistical regularities among early lexical categories – they keep track of word-referent pairings, adjust the strengths of these associations based on repeated exposures, and form expectations about how to generalize novel words (e.g., Colunga & Smith, 2005; Regier, 2003). Such associative models might explain how children form prior beliefs about the preference of others (e.g., through repeated exposures to people favoring objects that are intrinsically interesting) and bring generalized expectations to the current situation. However, associative learning would not suffice to fully explain our findings with the 2-year-olds: when inferring the preference of another person from her choice behavior, children did not simply keep track of the frequency with which she chose one type of objects, as associative models would suggest. Instead, children made rational inferences, taking into account both the presence of an alternative and whether the object chosen was rare in the population.

We found similar but weaker results with the 16-month-olds. Compared to their peers who were presented with sampling information that did not license a preference, 16-month-olds who received the evidence of nonrandom sampling were more likely to infer that another person might have a preference different from their own, but they were less likely to do so than the 2-year-olds. These findings indicate that the ability to infer the subjectivity of preferences based on sampling evidence begins to emerge between the ages of 16 months and 2 years. It is intriguing to consider why the 16-month-olds were less competent than the 2-year-olds.

One possibility is that the 16-monht-olds have a more egocentric conception of mental states. They might judge another person's preference based only on objective reality (e.g., the slinkies are inherently interesting) and have not yet developed the understanding that different people can have different attitudes toward the same entity (Repacholi & Gopnik, 1997). The second possibility is that the statistical evidence given might not be sufficient for some toddlers to make interferences about mental states. Finally, 16-month-olds might simply have greater difficulty overriding their initial beliefs that their preference was shared and/or overriding a more general bias assuming that people should prefer an interesting object to a boring one. Future studies are needed to tease apart these possibilities.

In the current study, the experimenter's nonrandom sampling behavior was explicitly situated in a context about preferences. However, human actions that are statistically nonrandom may not always indicate a preference. Under some circumstances, nonrandom sampling behavior could be performed simply to achieve a neutral goal, without any positive or negative valence. For example, a person picks out only the science fiction novels from a pile of books that consists of mostly textbooks and just a few novels, because of his intention to organize the books into two categories, without expressing a preference. Under other circumstances, nonrandom sampling behavior could even indicate a dislike or disgust. For instance, a person picks out the only few slices of onions from her salad because she dislikes onions. How young children make meaningful inferences based on statistical evidence in these social contexts also deserves future examination.

To conclude, this study is the first attempt to explore whether young children can use statistical patterns in the choices that other people make to infer the subjective nature of mental states. The findings suggest that by age 2 children apprehend the subjectivity of preferences based on sampling evidence alone, in the absence of social-pragmatic cues. Future research on young children's use of statistical evidence in other social contexts will provide important insight into how the young learners become rational participants in the social world.

Acknowledgements

---

[2] Another path for future research concerns young children's preference attribution based solely on the relative frequency of objects. Previous studies suggest that when two types of objects are equally ambiguous on a target attribute (e.g., value), adults use relative frequency to judge object value and place greater value on objects that are relatively scarce (e.g., Brock, 1968; Hirshleifer, Glazer, & Hirshleifer, 2006). It is interesting to examine whether young children will also use this heuristic in preference attribution. For example, if children are shown two types of objects that are equally interesting (or boring) but differ in frequency, would they prefer and attribute greater value to the minority objects that are scarce and ascribe the same preference to others? An ongoing study is exploring this question.

References

Brock, T. C. (1968). Implications of commodity theory for value change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psycho-logical foundations of attitudes* (pp. 243–275). New York: Academic Press.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: The role of associative learning. *Psychological Review, 112*, 347–382.

DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: Lies are told to those who care. *Journal of Personality and Social Psychology, 71*, 703–716.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50*, 21–45.

Gopnik, A., & Schulz, L. E. (2004). Mechanisms of theory-formation in young children. *Trends in Cognitive Sciences, 8*, 371–377.

Hirshleifer, J., Glazer, A., & Hirshleifer, D. (2006). *Price theory and applications: Decisions, markets, and information*. New York: Cambridge University Press.

Kirkham, N., Slemmer, J., & Johnson, S. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*, B35–B42.

Kushnir, T., Xu, F., & Wellman, H. (2010). Young children use statistical sampling to infer the preferences of others. *Psychological Science, 21*, 1134–1140.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Regier, T. (2003). Emergent constraints on word learning: A computational review. *Trends in Cognitive Science, 7*, 263–268.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12–21.

Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behavior. *Child Development, 55*, 1504–1513.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Talwar, V., & Lee, K. (2002). Emergence of white lie telling in children between 3 and 7 years of age. *Merrill-Palmer Quarterly, 48*, 160–181.

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology, 39*, 906–912.

Wellman, H. W., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Reviews of Psychology, 43*, 337–375.

White, P. A. (1988). Causal processing: Origins and development. *Psychological Bulletin, 104*, 36–52.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11month-old infants. *Cognition, 112*, 107–114.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 5012–5015.

Xu, F., & Tenenbaum, J. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*, 288–297.