The Berkeley Undergraduate Journal

Fall 2012

To Leah Carroll
BUJ advisor 2003–2012

Contents

Letter from the Editor's desk

This fall at the Berkeley Undergraduate Journal has been a semester of change. I couldn't be more proud of this journal and all of the people whose hard work made this semester's issue possible.

First and foremost, I want to thank our faculty sponsor, Kathleen Donegan, whose guidance and support allowed the BUJ to offer its first ever Decal.

Thank you to the BUJ editors—your enthusiasm for Oxford commas is infectious. Thank you to Minerva Ramirez, the BUJ's Layout Editor who created new, beautiful designs for our publication. Thank you to Natalie Oveyssi, Assistant Editor for your tenacious editing skills and awesome work ethic.

In addition to the hard work and talent of our faculty advisor and editorial staff, I am thankful to have the guidance of two excellent staff advisors at the UC Berkeley Office of Undergraduate Research, Leah Carroll and Mary Crabb. This semester we are honoring Leah's commitment to the BUJ. For the past eleven years, her guidance, advice and wisdom has been truly valuable to the BUJ editors—thank you.

And lastly, I would like to thank all of our fantastic authors. After twenty-five years of publishing exceptional undergraduate research papers in the social sciences and humanities at UC Berkeley, the BUJ is proud to present another diverse collection of excellent undergraduate work.

Sincerely,

Emma Lundberg
Editor in Chief
Berkeley Undergraduate Journal

Five-Year-Old Children Integrate Jointly Across Probabilistic and Social Domains When Inferring Preferences in Others

By Vidya Balakrishnan

Abstract

Human learners regularly face the challenging task of inferring unobservable psychological states in others. Sensitivity to relevant cues when inferring a psychological state –such as another's preference—is an invaluable skill: accurate inference of underlying states allows one to understand and predict another's behavior. Research has shown that 18-month-old children can use affective cues when asked to infer an agent's preference (Repacholi & Gopnik, 1997). Recent studies have also demonstrated that children from 16 months to 4 years can also use probabilistic cues to infer another's preference (Kushnir, Xu & Wellman, 2010; Ma & Xu, 2011). However, single cues are limited in the kinds of inference they allow and the inferential certainty they provide. While there is reasonable evidence that children can use a variety of single cues to infer preference, less attention has been paid to children's ability to integrate across multiple cues. The current study investigated whether children could rationally integrate both probabilistic and social cues to predict an agent's preference. 64 three- to five-year-old children were presented with probabilistic and social cues through a puppet agent who picked toys out of a jar. After watching the agent sample objects out of a jar and express either joy or disgust, the child was asked to offer the agent one toy he liked to play with. We found that children's toy choices were sensitive to both types of cues, suggesting that by five years of age children can integrate across multiple cues to support their social reasoning.

I. Introduction

In our social world, we regularly face the challenging task of inferring underlying mental states in others. One such important mental state is another person's preference. Preference inference and the ability to understand desires in others is a valuable social skill since it helps one to accurately make sense of and interpret otherwise ambiguous behaviors exhibited by another individual in normal social settings. As adults and probably quite early in life, we are aware of others' preferences, discuss these preferences, and use our preference knowledge in an assortment of social decisions. This is remarkable considering that we often cannot directly see another's preference we must infer it.

This naturally leads to the question of *how* we typically infer preference in others. One way we ascertain preferences may be from observing choice behaviors in others. Previous studies have investigated the role of verbal, statistical, and affective cues in inferring someone's preference. A number of these studies have focused on using children as a general model for understanding and exploring the development of preference inference. Our research follows this spirit.

Preference inference by children using *single* cues has already received some empirical attention. Repacholi and Gopnik[1] provided evidence that very young children have some understanding of the subjective nature of preferences. They tested 14- and 18-month-old infants to see if they understood that others may have preferences different from their own and if they could infer the preferences of others from simple affective cues. The authors used a food-requesting paradigm to measure whether 14 and 18-month-old infants could correctly identify the food (either goldfish crackers or broccoli) preferred by an experimenter based on her affective display, even when the experimenter's preferred food differed from that of the infant. The authors found that the 14-month-olds acted egocentrically by offering the food they themselves preferred. In contrast, 18-month-olds were able to use the experimenter's affective cues to offer her the food she preferred, even when it was different from their own preference. Thus by 18 months of age, infants can use simple affective displays to infer an agent's preference in a food preference task.

Prior research also suggests that children can use statistical evidence to infer another's preferences. Kushnir, Xu, and Wellman[2] argued that even young infants are sensitive to simple statistical relationships and expect randomly drawn samples to reflect their underlying populations and vice versa. Kushnir, Xu, and Wellman[3] subsequently proposed that children might be able to use this statistical sensitivity to infer agent preference, when the sampling evidence conflicted with the distribution of the population from which the sample had been drawn. In their first experiment, the authors tested whether three- to four-year-old children could predict an agent's preference when provided evidence of random sampling violations. An agent intentionally sampled five toys of the same type (e.g., frogs) from a box with two types of toys (e.g., frogs and fish). The proportion of toy frogs varied across the three conditions: 18%, 50%, or 100%. The agent maintained a positive affect across all three conditions. So if children only used the affective cue, they were predicted to infer identical preferences in the agent across

1 Betty M. Repacholi and Alison Gopnik, "Early Reasoning About Desires: Evidence From 14-and 18-Month-Olds," *Developmental Psychology* 33, no. 1 (1997): 12-14.

2 Tamar Kushnir, Fei Xu and Henry M. Wellman, "Young Children Use Statistical Sampling to Infer the Preferences of Other People," *Psychological Science* 21, no. 8 (2010): 1134.

3 Kushnir, Xu, and Wellman, "Statistical Sampling," 1135.

conditions and regardless of the toy proportion. Following the sampling, children were asked which toy the agent preferred. They found that children took evidence for violations of random sampling as implying a preference on the part of the agent, and that the extent to which an inference was drawn was proportional to the strength of the probability cue. In a second experiment, the authors tested whether 20-month-old infants could complete a similar task. The study used a live agent instead of a puppet agent (to present more salient affective cues), used a less memory-intensive procedure, and only consisted of two conditions of 18% and 82% of target objects. As with the first experiment, infants inferred a preference on the agent's behalf when the sample was drawn from the 18% condition (i.e., unlikely to have been chosen by chance), but not when it was drawn from the 82% condition. In a follow up to the work by Kushnir and colleagues, Ma & Xu[4] found that by two years of age, children were sensitive to non-random sampling as a cue for a different, subjective preference in another. Both studies demonstrate a possible role for statistical inference in early social reasoning of preference.

To reiterate, all of the aforementioned preference inference studies have researched children's use of single cues in the inferential process. It is true in some instances that a single cue, if strong and reliable, may serve as sufficient evidence upon which one can judge a preference in another. For instance, at younger ages we are less inhibited and may not hesitate to explicitly express our likes and dislikes by using a verbal cue to make our preference known to another (e.g. "Mommy, I *hate* broccoli! I won't eat it!"). Here, it is not painfully difficult for the mother to infer her child's distaste for broccoli since the single verbal cue is clear enough by itself and the child unambiguously and intentionally expresses what she does not want and presumably has no reason here to be deceptive. However, as earlier noted, in many instances our likes and dislikes may not be so transparent. In fact, circumstantially, social convention may deem it inappropriate to communicate our true preferences (e.g. some may consider it rude to tell your grandmother to her face that you don't like the itchy sweater she lovingly knits you for your birthday every year). What are some other reasons using single cues independently in the inferential process can be problematic? For one, some cues are not always available or appropriate to use in a situation, or a single cue's meaning can vary with context. Furthermore, some single cues may even be more reliable or heavily weighted than others.

Especially in instances where available evidences conflict, using a *single* cue in the inferential process can lead one to draw less accurate or even erroneous conclusions than by attending to and integrating across *all* relevant cues available. As adults, we recognize that in the real world multiple cues are often available and that mental states in others (such as another person's preference) are often difficult to directly observe. In turn, we believe that adults intuitively integrate across *all* available cues to allow for more accurate inferences. While it is known that young children can infer an agent's preference given independent affective or statistical cues, it is unknown whether children are able to integrate across multiple cues in this inferential process as adults almost certainly can.

The present study aimed to further previous research by investigating whether children possessed the ability to combine affective cues with probabilistic cues in drawing inferences

4 Lili Ma and Fei Xu, "Young Children's Use of Sampling Evidence to Infer the Subjectivity of Preferences," *Cognition* 120, no. 3 (2011): 403-411.

about an agent's preference, both when the cues complemented and conflicted. Affective cues were provided as in Repacholi & Gopnik[5], with the agent's affective display manipulated during sampling. Statistical cues were provided as in Kushnir, Xu & Wellman[6] and Ma & Xu[7] by manipulating the relative proportions of two toy types within the population from which the agent sampled. The current design combined these two types of cues to test each child under one of four probabilistic social cue conditions. Following the agent's sampling, the child was allowed to offer the agent one toy to play with as a measure of the child's inference of preference in the agent. We hypothesized that children's toy choices would reflect the integration of both social and probabilistic cues across all four conditions. First, we generally predicted that children's inference patterns would be clearest in cases when the probability and social cues aligned (e.g., LowHappy condition) and become progressively more unclear or "mixed" as the cues conflicted (e.g., HighDisgust condition). Second, we made specific predictions for each condition as to which toy(s) seemed most in line with the condition's social-probabilistic evidence and so would be the most likely selection for children if they were making use of that evidence. Condition-specific predictions are delineated in *Individual Toy Choice Within Conditions* under *Results*.

II. Method

A. Participants

Participants were 64 children (Mean age: 4.53 years, SD: 0.58 years, Range: 3.3-5.8 years). This sample included 30 girls and 34 boys. 16 children were assigned to each of four experimental conditions. Within each condition, there were approximately the same number of girls and boys, and no significant differences in participant mean age. Children were recruited from local preschools, a local museum, and lab testing facilities associated with the University of California, Berkeley. Seven additional children were excluded, due to insufficient English exposure (two), experimenter error (three), experiment interruption by another child (one), or inability to code child's toy choices due to shyness (one).

B. Materials

A puppet armadillo named Andy was used as the agent during the study. (See Figure 1.) Two of the experimenter's fingers could be used simultaneously to move the agent's arms and allow sampling of the toys. The sampled toys were classified broadly as animals and balls. The animal toy set consisted of three types of rubber finger puppets: an elephant, a giraffe, and a zebra. (See Figure 2.) The ball toy set consisted of three types of rubber balls: a baseball, a basketball, and a soccer ball. (See Figure 3.) During introduction and at test, the exemplars for a toy set (i.e., one baseball, one basketball, and one soccer ball) were displayed in a white, rectangular cardboard presentation box with squared-out slots for holding each toy in place. (See Figure 4.)

The population jars always consisted of two toy types. (See Figure 5.) The toy the agent sampled was denoted as the *target*. The toy the agent did not sample was denoted the *alternate*. The

5 Repacholi and Gopnik, "Early Reasoning About Desires,"14.
6 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1135-1136.
7 Ma and Xu, "Sampling and Preference Subjectivity,"405.

third toy—present in the test box but absent in the populations—was denoted as the *distracter*. The child only saw the distracters during the toy introduction period and at test. The populations were contained in clear cylindrical jars, and the ratio of the two types of toys in each jar was 7:31 (18% target or 82% target) as in Experiment 2 of Kushnir, Xu, & Wellman[8].

Though specific location and materials varied, the experimenter always made use of a table for presentation of test stimuli and chairs placed on opposite sides of the table—one for the experimenter and one for the child. Experimental sessions were filmed using a single standard digital video camera placed behind and to the side of the child or directly to the side of the table.

C. Procedure

The child and parent or guardian were shown to a quiet waiting room. Either a research assistant or the experimenter verbally provided the parent with a brief overview of the testing procedure and obtained informed consent. (See Figure 7.) The child was then seated at a table. The experimenter turned on the pre-placed video camera, took her seat at the table opposite the child, and started the test procedure.

The testing procedure can be outlined as introduction to the agent, introduction to the toys, sampling, and test. (See Figure 6.) The toy-sampling-test portion of the sequence was repeated twice, once for each toy set. Slight deviations from the study script were occasionally required to keep the child engaged in the task. (See Figure 8.)

Once seated, the experimenter brought out and introduced the child to the agent. The experimenter then put the agent aside and brought out the presentation box for the toy introduction. The presentation box contained three toys—one for each kind of toy in the first toy set (e.g., balls). The experimenter put the presentation box on the table between herself and the child and asked the child to label each of three toy kinds. If the child chose not to provide labels or did not know the name of one or more toys initially, the experimenter provided labels and subsequently confirmed that the child was able to distinguish the toys from one another by asking the child to repeat the toys' respective names.

For the sampling portion, the experimenter then removed the presentation box from the table and brought out a clear plastic bin filled with the target and alternate toys for the agent to sample from. The experimenter then asked the child if Andy should play with the toys and then proceeded (as Andy) to sample the target toy five times from the population. Each toy sampling took approximately 3-4 seconds during which Andy made a show of digging randomly around in the population bin. Following each sampling, the experimenter (again, as Andy) exclaimed either "yay!" or "yuck!" (depending on the experimental condition) before briefly placing the toy on the table for the child to see. The experimenter then removed the sampled toy from the table—using the hand not occupied by Andy—and placed it out of the child's view before moving on to sample the next toy. The population was visible to the child for the entire duration of sampling. After sampling, the experimenter placed the five target toys back on the table and requested the child to help return the toys to the jar.

At test, the experimenter expressed that Andy was ready to play again. The experimenter then placed the presentation box on the stage again and asked the child, "Can you give Andy the toy that he likes?" The child's toy choice was recorded as the first toy touched.

8 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1137-1138.

As previously noted, this same sequence of events (except for the introduction to agent) was then repeated for the second toy set.

The target toy sampled from the jars (minority or majority) and Andy's affective display (positive or negative) were combined to result in four experimental conditions (LowHappy, LowDisgust, HighHappy, HighDisgust). (See Figure 5 for depictions of example conditions.) The proportion of toys in the jar functioned as a probabilistic cue by offering either ambiguous or strong evidence that the Andy intended to select a specific toy. Andy's affective display functioned as a social cue for Andy's feelings towards the sampled toy and varied between (but not within) children. For instance, a child in a LowHappy condition who saw Andy repeatedly sample the less frequent ball and express happiness each time *also* saw Andy repeatedly sample the less frequent animal and express happiness each time. Each child thus saw a single probabilistic-social cue combination twice—once for each toy set.

Several items were counterbalanced to control for potentially confounding effects: the toy set presented first, toy order within the presentation box, and the two toy types within the population jars.

III. Results

A. Descriptive Statistics

Children's toy choices for each of the four conditions were collected as both frequencies and proportions. (See Table 1 or Graphs 1-4.) Between conditions, we also examined the overall number of times children switched in the toy type they offered the agent across the two test trials as a possible indirect measure of children's "confidence" or consistency in their toy choices. (See Table 1: Overall Switches [OS].) We also examined switching to test how children's inferences for the agent's preferred toy changed when they used the statistical cue independently and presumably ignored the affective cue; this measurement was intended to replicate the findings of Kushnir et al[9]. This was accomplished by determining frequencies for when a child switched either to or from the *target* toy for one trial or another across all four conditions. (See Table 1: Target Switches [TS].) If children ignore the affective cue, the default assumption to be made is that the agent is only sampling the toy he prefers (i.e., the minority in Low conditions and the majority in High conditions). In turn, in the LowHappy and LowDisgust conditions, switches from the target (minority) toy type were measured. In the HighHappy and HighDisgust, switches from the target (majority) were measured. Analyses of both types of switch data are provided in *Additional Analyses*.

Preliminary analysis found no evidence for effects of age or gender. Age and gender were not included in any follow-up analyses, and male and female data were collapsed together.

B. Interactions

The traditional point to begin a statistical investigation is to test for interactions among the factors of interest. The factors of interest here are the agent's emotional expression that accompanied each sampling event and the probability of the sample produced. The current

9 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1136-1139.

sample size and distribution, however, were poorly suited for the statistical techniques required for assessing interactions in our data—a generalized linear model, specifically a multinomial logit regression. This analysis heavily depends on a large overall sample size and a reasonable allotment of this sample to the possible combinations of the factors and their outcomes (for instance, the number of children who selected the target object in the LowHappy condition should be greater than 0). For the purpose of the current thesis—given the complexity of the required analysis and the difficulty in applying it to the current data—interactions were instead inferred from the patterning of the following analyses, especially those analyses exploring simple effects. The proposed interaction and evidence for this interaction are reviewed in the *Discussion* section.

C. Main Effects

We tested for evidence of main effects for either affective or statistical cues.

To test for a main effect of affective cue, the LowHappy and HighHappy choice data were collapsed together and the LowDisgust and HighDisgust choice data were collapsed together, allowing us to test Happy conditions versus Disgust conditions. The choice patterns (majority, minority, and distracter choice frequencies) in the Happy and Disgust conditions were then compared using a Fisher's exact test which approached but did not reach significance ($p=0.10$).

To test for a main effect of statistical cue, the LowHappy and LowDisgust choices were collapsed together and the HighHappy and HighDisgust choice data were collapsed together, allowing us to test Low conditions versus High conditions. The choice patterns (again, between the majority, minority, and distracter choice frequencies) in the Low and High conditions were then compared using a Fisher's exact test, which was significant ($p=0.01$).

D. Simple Effects Between Conditions

We tested for differences in choice patterns between the four experimental conditions. This comparison also made use of a two-tailed Fisher's exact test. (See Table 2.) There was evidence for a significant simple effect for statistical cue between the LowHappy and HighHappy conditions. (Note that this at least partially replicated the findings of Kushnir, Xu, & Wellman[10]. See *Discussion* for more on this.) No significant simple effects of affective cue were found when comparing the LowHappy and LowDisgust conditions, or the HighHappy and HighDisgust conditions, respectively (though comparisons between these pairs also trended towards significance). We did not find evidence for a simple effect for statistical cue between the LowDisgust and HighDisgust conditions.

We also tested two specific pairwise comparisons of interest between conditions using two-tailed Binomial exact tests. The comparison for the number of times children chose the target (minority) in the LowHappy condition versus the LowDisgust condition was not significant ($p=0.18$). The comparison for the number of times children chose the target (majority) in the HighHappy condition versus HighDisgust condition was also not significant ($p=0.36$).

10 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1136-1139.

E. Simple Effects Within Conditions

We tested for differences in choice patterns from chance within each of the four experimental conditions. This comparison used a Pearson's Chi-square Goodness-of-fit test. (See Table 3.) Children's toy choices for the agent were significantly different than predicted by chance in the LowHappy and HighHappy conditions. Children's toy choices for the agent were not significantly different than predicted by chance in the LowDisgust and HighDisgust conditions.

F. Individual Toy Choice Within Conditions

To explore children's preferred toy choice in each condition, we tested individual toy choice against chance using two-tailed Binomial exact tests. (See Table 4.) Chance was assumed to be 0.33, the anticipated distribution if children were selecting randomly among the available toys to offer the agent. A Bayesian model[11] based on the rational model of preference learning in Lucas et al.,[12],[13] was used to generate predicted toy type frequencies for each condition in a manner sensitive to both social and probabilistic cues. The highest frequency toy for each condition as predicted by the model and in agreement with experimenter intuition, was taken as the "most likely" or "best" choice for each condition.

In the LowHappy condition, we predicted children would offer the agent the target significantly more often than the distracter, and that the distracter would be offered slightly more often than the alternate. The results were consistent with this prediction as only the target toy (minority) was selected above chance, while both other toy types were selected less often than chance.

In the LowDisgust condition, we predicted children would offer the agent the alternate most often, followed by the distracter, and that the target would be offered least often. Consistent with the Chi-square analysis above, but contrary to our predictions, we found no statistical evidence for one any one toy type being selected above chance.

In the HighHappy condition, we predicted children would offer the agent the target most often, followed by the distracter, and that the alternate would be offered least often—in a similar but weaker pattern to the LowHappy condition. The results were consistent with this prediction as only the target toy (majority) was selected above chance.

In the HighDisgust condition we predicted children would offer the agent the distracter and alternate roughly equally often, and that both would be offered significantly more often than the target toy. Consistent with the earlier Chi-square analysis, but again contrary to our predictions, we found no statistical evidence for one object being selected above chance.

G. Additional Analyses

Overall, switch comparisons between all four conditions were not significant, according to two-tailed Binomial exact tests, which yielded $p > 0.20$ across all comparisons. It is worth

11 Chris Lucas, "Modeling notes for Onion project," 1-2.

12 Chris Lucas, Thomas L. Griffiths, Fei Xu, and Christine Fawcett. "A Rational Model of Preference Learning and Choice Prediction by Children," *Advances in Neural Information Processing Systems* 21 (2009): 1-8.

13 Chris Lucas, Thomas L. Griffiths, Fei Xu, Christine Fawcett, Alison Gopnik, Tamar Kushnir, and Lori Markson. "The Child as Econometrician: A Rational Model of Preference Understanding in Children," *University of California at Berkeley* n.d.

noting, however, that there was a weak trend towards children in the high probability conditions. In the high probability conditions, there were 16 switches total across the happy and disgust conditions combined, with 8 in each of happy and disgust. However, in the low probability conditions there were only 12 switches total across happy and disgust combined, with 6 switches in each of happy and disgust. (See Table 1.) Target switch comparisons between all four conditions were not significant either. Two-tailed Binomial exact tests yielded $p > 0.20$ across all comparisons.

IV. Discussion

We asked the question of whether children could integrate across both probabilistic and social cues when inferring preferences in an agent. We hypothesized that children's toy choices would reflect the integration of both social and probabilistic cues across all four conditions. It was anticipated that children's inference patterns would be clearest in cases when the probability and social cues aligned (e.g., LowHappy condition) and would become progressively unclear as the cues conflicted (e.g., HighDisgust condition).

Firstly, children's toy choice data reflect evidence that children are sensitive to both cue types independently. This is seen most clearly in looking at main effects, individually, for agent affect and probability distribution of the sampled population. We tested for a main effect of statistical cue, and this was significant ($p=0.01$), replicating Kushnir et al.[14] Children interpret strength of statistical cues as an independent indicator of preference in another. As previously stated, while no main effect for agent affect was found, the Fisher's exact test that was used did in fact approach significance. A main effect for agent affect would provide evidence that children can also independently use affect to interpret preference in an agent, thereby replicating Repacholi & Gopnik.[15] A suggestion to bring the affect main effect to statistical significance is provided under *Future Work*.

Secondly, after confirming children were in fact sensitive to both cues independently, we proceeded to test for interactions by looking at patterns of simple effects within conditions, to consider whether children could integrate across both cue types simultaneously in the inference task. We observed differences in choice patterning between the High and Low conditions when the accompanying emotional expression was Happy, but did not observe these differences when the accompanying social cue was Disgust. Specific patterning by condition is interpreted subsequently.

In the two Happy conditions, we saw consistent choice patterning. Children tended to select the same toy that was most frequently sampled by the agent. This makes sense for the LowHappy condition and replicates the Kushnir et al.[16] findings. Children demonstrated that they consider both cues by realizing that if a toy is infrequently present in the population yet drawn repeatedly by the agent and followed by a joyful expression, it cannot be a chance occurrence and the agent must be picking out the toy he likes. At first glance, however, this is odd for the HighHappy condition: we would expect weaker choice patterning in this case than we observed. In the HighHappy condition, children consistently chose the majority

14 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1136-1139.
15 Repacholi and Gopnik, "Early Reasoning About Desires,"16-18.
16 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1136-1139.

toy even though the probabilistic evidence was weak and did not easily hint at whether the agent should prefer the majority or minority toy. One explanation is that, coupled with the Kushnir et al. results[17], the majority toy gets a "bump" above and beyond what the probabilistic evidence would suggest alone when coupled with a strong, positive social cue. In the presence of positive social cues, the children seem to give the social cue particularly increased weight. Even so, there are weak patterns to support the idea that the probabilistic evidence is still being considered, such as the majority is selected less often than the distracter in the LowHappy condition or the minority is selected more often than the distracter in the HighHappy condition. We propose a possible way of testing whether children are in fact using the probabilistic evidence at all in the current task under *Future Work*.

As stated above, in the Disgust conditions, we found no statistical evidence for patterning. In the LowDisgust condition, it was predicted children would interpret the agent's removals of the target toy (minority) as evidence that he was picking out the toy he did not like, though child choice data fail to reflect these predictions. However, when we asked children to explain, post-test, why they gave the agent a particular toy, we found anecdotal evidence with individual children providing explanation of specifically choosing *away* from the target toy, as per our prediction. For example, at test one child in this condition pointed to the target (minority) toy type in the presentation box and volunteered: "Certainly not that one!" It was also predicted in the LowDisgust condition that children would have slightly more difficulty than in the LowHappy condition in judging which toy type (between the two remaining toy types of alternate and distracter) the agent *actually wanted*. The HighDisgust condition was hypothesized to be the "messiest" of the four experimental conditions. A conceivable explanation for children's arbitrary offerings in this condition (and in LowDisgust) is that the agent's sampling appeared intentional to children; children then presumably had difficulty reconciling the agent's intentional selection of a toy coupled with his negative affect. As a result of this confusion, children seemingly offered the agent any of the three toy types equally and arbitrarily across both test trials. It may well be that children's choice patterns only reflected confusion in the Disgust conditions because the agent's sampling appeared intentional, but conflicted with his affect.

Recent piloting testing with college undergraduates shows that the largest difference between children and adults is that adults seem to have consistent patterning in their choices on the agent's behalf. Adults are further able to reconcile the intention-affect conflict. For instance, in the LowDisgust condition, adults appear to not only select *away* from the target, but also seem to use the probabilistic evidence as a signal that the agent had a particular sampling strategy to get rid of the rare toy that he disliked. Adults' more consistent choice patterning is possibly also because adults take the pragmatics of an experimental session more seriously. Therefore, they, more than children, may feel the need to commit to using one cue more heavily than another when the two cues conflict; we found that in most cases adults committed to using the affective cue over the statistical cue. Nonetheless, adults, unlike children, appear to maintain their original interpretations they come up with to explain the agent's sampling behaviors. Additional testing is needed to confirm if this intention-affect conflict is in fact responsible for limiting child performance in the Disgust conditions. A proposed means of suggesting to children that the agent's sampling is unintentional in order to find clear patterning in Disgust conditions in follow-up studies is listed under *Future Work*.

17 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1136-1139.

Thirdly, given that children appeared to integrate across both cues, we next sought to understand whether children appeared to make *accurate* inferences according to our hypothesis. This was achieved by comparing the individual toy type choice frequencies in each condition against chance. (See *Individual Toy Choice Within Conditions* for predictions.) In both Happy conditions, comparisons between the target toy types and chance was significant, providing evidence that is in line with our predictions for children's "most likely" toy choice. In the HighHappy condition, children's toy choices reflected accurately for the most part, their belief that the agent would prefer the target (majority) more often than the distracter. However, children appeared to be unsure of the agent's attitude towards the alternate—unlike with the target or distracter. In the Disgust conditions, we found no evidence for differences in toy choices. Again, we attribute this to the possibility of children possessing an inability to reconcile the intention-affect conflict. Perhaps children, while appearing to use both cue types, are more hesitant to commit to a single toy type to consistently offer to the agent since the cues conflict. That is, children appear to be at a loss for explaining why the agent would repeatedly choose a statistically infrequent toy unless he did not in fact have a preference for it. Within the HighDisgust condition, all three pairwise comparisons between toy types offered were not significant, though predictions were that the distracter and majority should each be offered significantly more often than the target. Again, we attribute this to children's possible confusion regarding the intentionality of the agent's sampling process and suggest a possible remedy to this below.

V. Future Work

In the current study, children's choice patterning in response to negative cues was ambiguous; the data shows that children's inferences appear to break down in the Disgust conditions. The most pressing questions at present, then, are whether: (1) children are using the probabilistic evidence at all, and (2) why children display more confusing choice behaviors in the Disgust conditions.

There is a strong trend towards a main effect of affect in the current data. However, this trend did not reach statistical significance. Accordingly, we propose that a main effect for agent affective display can be found by using a slightly larger total sample size to see if the effect held and achieved significance. However, we hesitate to suggest this as a good solution: main effects are not always observed in the presence of an interaction and this may be the case here. The shift from clear choice patterns to relatively flat choice patterns may make it difficult to detect a main effect of affect. Instead, we note that this is strong evidence for an interaction and thus evidence that affect is indeed playing a role in children's inferences.

More interesting than increasing the sample size would be t test conditions with neutral affect from the agent (resulting in two experimental conditions of LowNeutral and HighNeutral) to gather more evidence as to the possible impact of affect relative to probabilistic cues. In attempting to replicate the findings of Kushnir et al.[18] by using a neutral affect from the agent, we would see if children's choice patterns more strongly reflect the use of probabilistic evidence alone. Put more generally, the interaction may occur like so: In the absence of strong social cues, children appear to make use of probabilistic evidence. In the presence of strong social cues, children's inferences may weight that social information more strongly than probabilistic cues, making the effect of the latter subtle.

18 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1136-1139.

Still, even if we clarify the relationship between the cue types in children's inferences, children's difficulty in making inferences in the Disgust conditions remains a mystery. As previously discussed, one possibility is that children have difficulty reconciling the intentional selection of a toy coupled with negative agent affect. If this is the case, removing the intentional nature of the selection should allow children to make accurate inferences from the social evidence alone. To indicate to children that the agent's sampling process in the follow-up study is unintentional (since the apparent intentional nature of sampling could explain the null pairwise comparisons in both Disgust conditions), future testing can incorporate the use of a scoop in the sampling portion of the experiment. This may help children better interpret what may currently appear to them as the agent's peculiar behavior of intentionally yet repeatedly choosing a toy that he does not appear to like, but that can also be drawn by chance due to being the more abundant toy type. In other words, if a scoop is added to the sampling portion in future testing to suggest to children that the agent's sampling policy is unintentional, children will presumably gravitate towards offering the agent either the alternate or distracter significantly more often than chance (when compared to the target) than they do presently. This would be in line with the adult pilot data discussed above. We presume removing the intention-affect conflict should push children's choices to be more adult-like. We propose testing with children six years or older to determine precisely when in development the ability to reconcile this conflict may emerge.

To summarize the last two suggestions, adding in the neutral conditions will help us address the former question of whether children appear to use the probabilistic evidence and allow us to confirm the interaction. Adding in the "unintentional sampling" conditions will help us address the latter question regarding children's confusing choice patterns and allow us to observe children's inferences when the possible intention-disgust conflict is removed. At the very least, we may have evidence that children, when making inferences about an agent, have a difficult time reconciling the fact that agent can intend to do something, yet dislike it.

VI. Conclusion

We begin to infer preferences in others as early as 18 months of age and continue to do so for the rest of our lives. While past studies show that young children can independently use simple affective cues[19] from an agent and statistical cues[20,21] in inferring others' preference, little research has been done on children's abilities to combine both cue types in the inferential process. This is surprising given the common knowledge that inferences drawn by integrating across any and all available cues will naturally lead to more accurate inferences being drawn than by attending to single sources of information. We tested whether children could rationally integrate both probabilistic and social cues to predict an agent's preference and found that children's toy choices were sensitive to both types of cues; children's inferences were strongest when the cues complemented, as in the LowHappy condition, and were weakest when the cues conflicted, as in either of the Disgust conditions.

In contrast to adult pilot data, these findings suggest that though children's inferences are sensitive to simple social cues, children may have difficulty reconciling cases where strong

19 Repacholi and Gopnik, "Early Reasoning About Desires,"14-18.
20 Kushnir, Xu, and Wellman, "Statistical Sampling and Preference," 1136-1139.
21 Ma and Xu, "Sampling and Preference Subjectivity,"404-410.

preference evidence and apparent intention conflict, such as when an agent repeatedly selects something he dislikes. We believe that this intuition, along with some other limitations of our design, require further testing before we can draw certain conclusions about children's behavior in our experiments. Nonetheless, we believe our study provides some of the first evidence suggesting that by five years of age, children can integrate across multiple cues to support their inferences in a social reasoning task.

Bibliography

Gweon, Hyowon, Joshua B. Tenenbaum, and Laura E. Schulz. "Infants Consider Both the Sample and the Sampling Process in Inductive Generalization." *Proceedings of the National Academy of Sciences* 107, no. 20 (2010): 9066-9071.

Kushnir, Tamar, Fei Xu, and Henry M. Wellman. "Young Children Use Statistical Sampling to Infer the Preferences of Other People." *Psychological Science* 21, no. 8 (2010): 1134-1140.

Lucas, Chris, Thomas L. Griffiths, Fei Xu, and Christine Fawcett. "A Rational Model of Preference Learning and Choice Prediction by Children." *Advances in Neural Information Processing Systems* 21 (2009): 1-8.

Lucas, Chris. "Modeling Notes for Onion Project." (2011): 1-2.

Lucas, Chris, Thomas L. Griffiths, Fei Xu, Christine Fawcett, Alison Gopnik, Tamar Kushnir, and Lori Markson. "The Child as Econometrician: A Rational Model of Preference Understanding in Children." University of California at Berkeley n.d.

Ma, Lili, and Fei Xu. "Young Children's Use of Sampling Evidence to Infer the Subjectivity of Preferences." *Cognition* 120, no. 3 (2011): 403-411.

Repacholi, Betty M., and Alison Gopnik. "Early Reasoning About Desires: Evidence From 14- and 18-Month-Olds." *Developmental Psychology* 33 no. 1 (1997): 12–21.

Vaish, Amrisha, and Amanda Woodward. "Infants Use Attention But Not Emotions to Predict Others' Actions." *Infant Behavior and Development* 33 no. 1 (2010): 79-87.

Xu, Fei, and Stephanie Denison. "Statistical Inference and Sensitivity to Sampling in 11-Month-Old Infants." *Cognition* 112 no. 1 (2009): 97–104.

Xu, Fei, and Vashti Garcia. "Intuitive Statistics by 8-Month-Old Infants." *Proceedings of the National Academy of Sciences* 105, no. 13 (2008): 5012-5015.

Appendix A: Materials and Procedure Overview

Figure 1: Andy (Puppet Agent)

Figure 2: Animal Toy

Figure 3: Ball Toy Set

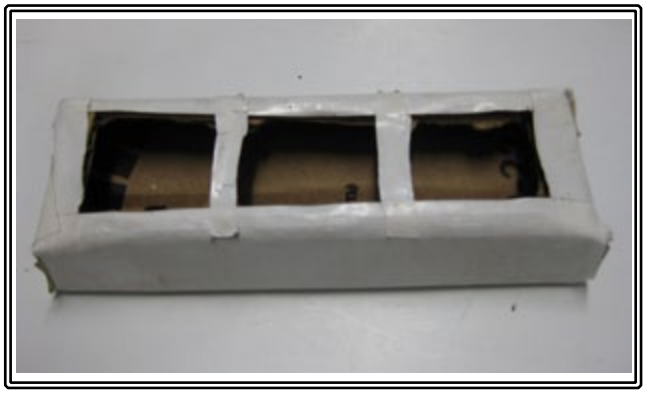
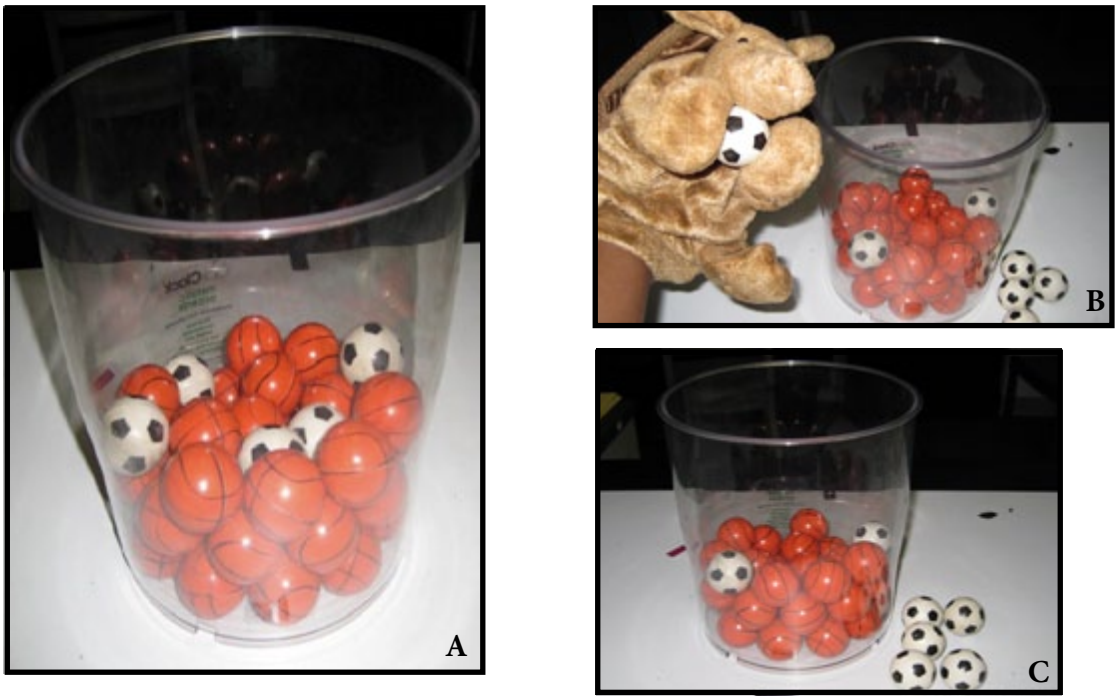
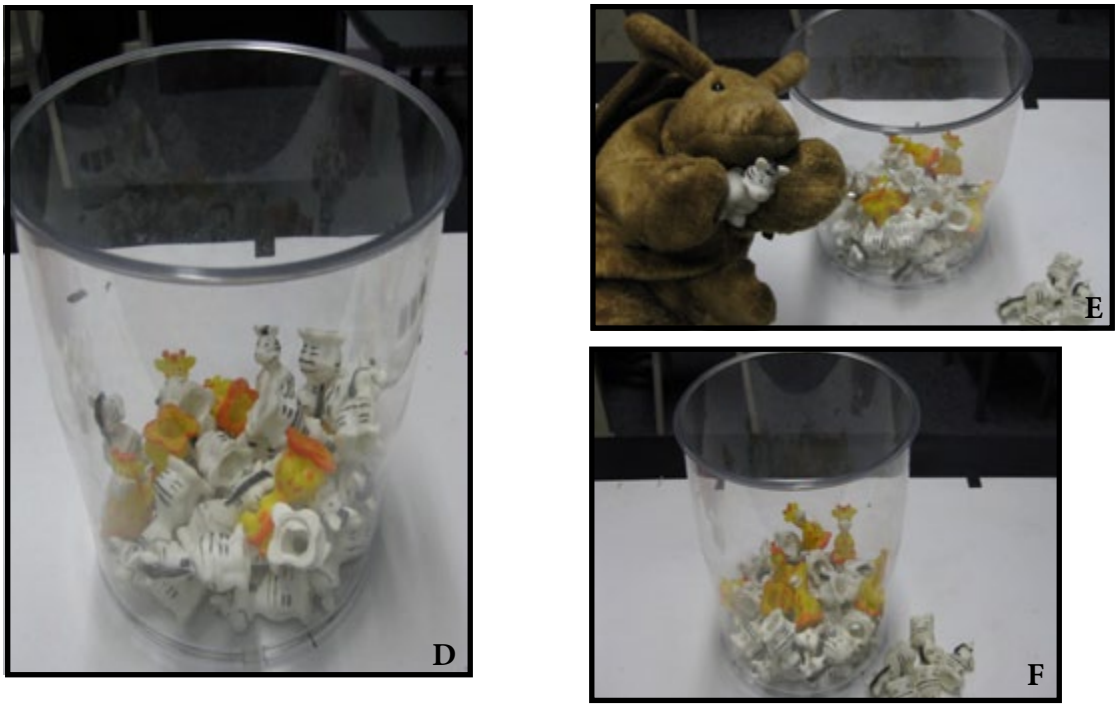


Figure 4: Toy Presentation Box

Figure 5: Example Conditions



Example LowHappy Condition: Soccerball (Target), Basketball (Alternate), Baseball (Distracter)
5b: "Yay!"



Example HighDisgust Condition: Zebra (Target), Giraffe (Alternate), Elephant (Distracter)
5e: "Yuck!"

Figure 6: Procedure Outline

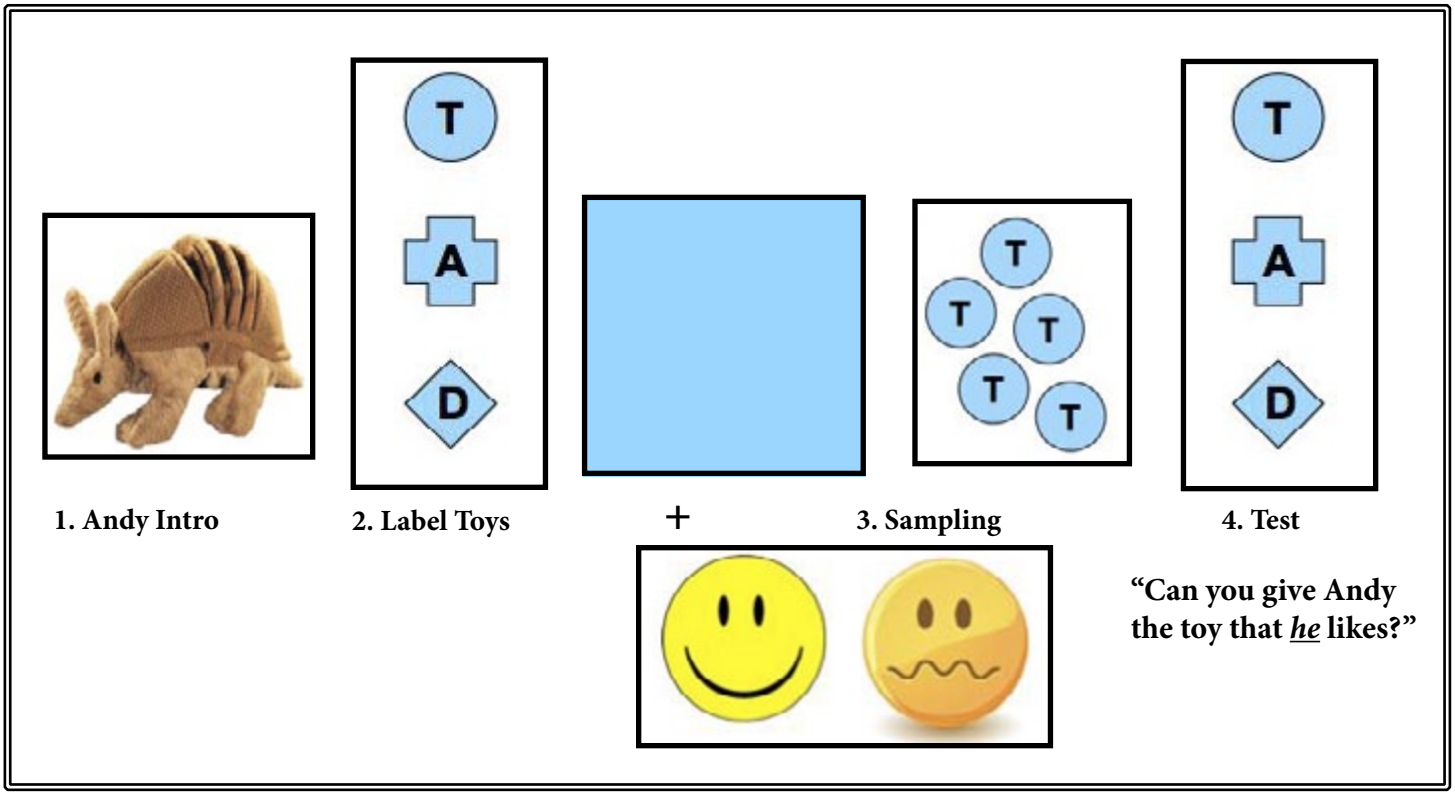


Figure 8: Study Script

Introduction to Agent

E: Are you ready to meet Andy?
(Andy pops up from behind the stage)
E: Look! There's Andy! He's waving hi!
(Experimenter puts Andy aside and out of view)

Toy Introduction

E: Okay, so now I'm going to show you some of Andy's toys! Okay?
(When the child agrees, move on; if trouble getting agreement, just move on)
E: Ooo... look at these! Can you tell me what these are?
(If child just names general type – i.e., animals – ask child to name each object; confirm that child distinguishes the objects)

Toy Sampling

E: Good job! Guess what? Andy has a whole bucket of toys!
(Put presentation box aside and put toy bucket on table)
E: Look at these!
(Turn bucket around so that child gets to see all sides)
E: I have an idea. Maybe Andy will play with the toys! Should we do that?
(when the child agrees, move on; if trouble getting agreement, just move on)
(Andy comes out again; samples the Target toy five times and makes an appropriate sound - "yay" or "yuck" each time; each sampling should take about 3-4 seconds)
(When sampling is finished put Andy aside)
E: Okay! That was fun! How about we clean up?
(Experimenter and child put toys back into bucket – have child help to get them used to touching the toys; put bucket aside)

Test

E: Hey, Andy's ready to play again! Can you do me a favor? Can you give Andy the toy that he likes?
(Take out the presentation box and push the presentation box towards the child; encourage him/her to make a choice; Andy takes the toy and "runs off" without making any sounds)

Second half of experiment

E: Good job! Would you like to look at some more toys?
(repeat the above with the second set of toys)

After second half of experiment

E: Good job! How did you know that Andy wanted the [insert toy selection here]?

Appendix B: Tables and Graphs

Note: LH=LowHappy, LD=LowDisgust, HH=HighHappy, HD=HighDisgust

TABLE I
Frequency of Toy Offered
(by condition)

condition	majority	minority	distractor	OS	TS
LH	5 (0.16)	18 (0.56)	9 (0.28)	6	6
LD	9 (0.28)	10 (0.31)	13 (0.41)	6	5
HH	18 (0.56)	9 (0.28)	5 (0.16)	8	4
HD	12 (0.38)	8 (0.25)	12 (0.38)	8	4

TABLE II
Simple Effects
(between conditions)

condition	Fisher's Exact
LH - LD	p = 0.15
HH - HD	p = 0.13
LH - HH	**p < 0.01**
LD - HD	p = 0.74

TABLE III
Simple Effects
(within conditions) ; df = 2

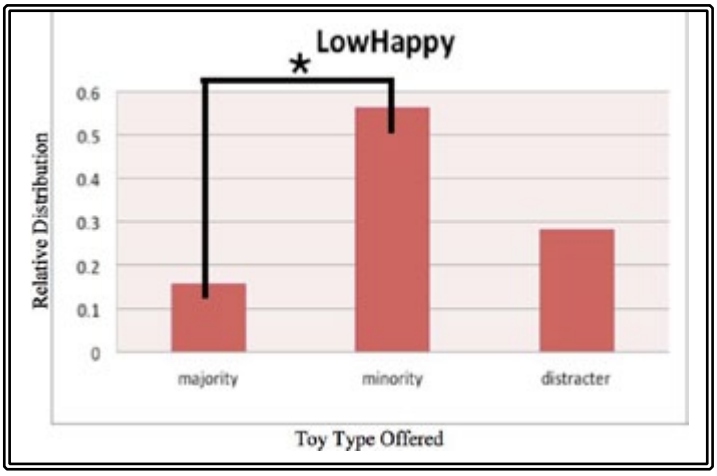
condition	ChiSQ GOF	p-value
LH	8.31	**0.016**
LD	0.81	0.66
HH	8.31	**0.016**
HD	1	0.61

TABLE IV
Individual Toy Choice Compared to Chance
(within conditions)

condition	majority	minority	distractor
LH	**p = 0.03**	p > 0.20	p > 0.20
LD	p > 0.20	p > 0.20	p > 0.20
HH	p > 0.20	**p = 0.03**	p > 0.20
HD	p > 0.20	p > 0.20	p > 0.20

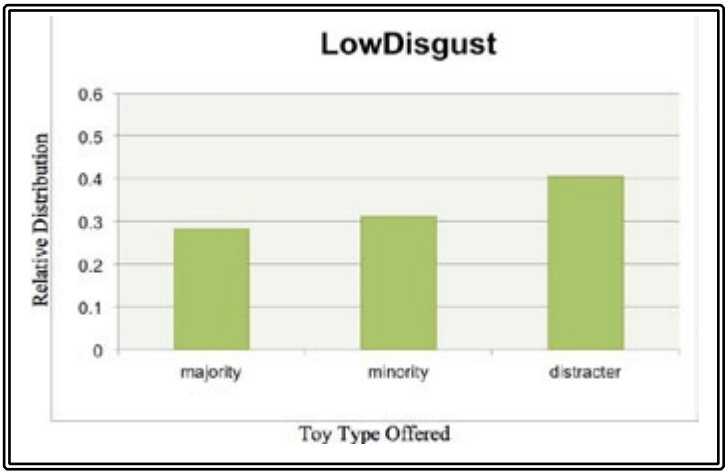
GRAPH I

Relative Distribution of Toy Type Offered for LowHappy Condition



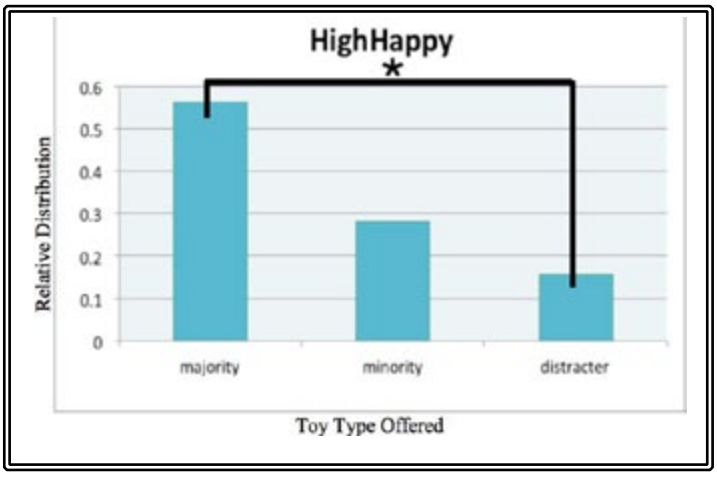
GRAPH II

Relative Distribution of Toy Type Offered for LowDisgust Condition



GRAPH III

Relative Distribution of Toy Type Offered for HighHappy Condition



GRAPH IV

Relative Distribution of Toy Type Offered for HighDisgust Condition