Learning to reason about desires: An infant training study

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Learning to reason about desires: An infant training study

Tiffany Doan\textsuperscript{a}(1), Stephanie Denison\textsuperscript{a}(1), Christopher G. Lucas\textsuperscript{a,c}(1,c.lucas@ed.ac.uk), & Alison Gopnik\textsuperscript{a} (gopnik@berkeley.edu)

\textsuperscript{a}University of Waterloo, Department of Psychology
\textsuperscript{b}University of Edinburgh, School of Informatics
\textsuperscript{c}University of California, Berkeley, Department of Psychology

Abstract

A key aspect of theory of mind is the ability to reason about other people's desires. As adults, we know that desires and preferences are subjective and specific to the individual. However, research in cognitive development suggests that a significant conceptual shift occurs in desire-based reasoning during infancy and early childhood, as from person to person (Gopnik & Wellman 1994).

Introduction

As social creatures, we are constantly trying to figure out what other people are thinking. The ability to infer others' mental states, such as their desires and beliefs, serves a number of important functions. It allows us to please or refuse to share on the grounds of fairness (Repacholi & Gopnik, 1997). The development of desire understanding of preferences by testing them in a modified version of Repacholi & Gopnik's (1996) Goldfish/Broccoli task. All infants were tested in the critical unmatched trial where infant's preferences conflict with the experimenter's desires. Only infants who failed to produce other signals to suggest that their preferences are inconsistent with one another or with the infants themselves.

The present study examined the shift in reasoning about desires during training. The present experiments examine a shift that occurs in infants' desire-based reasoning, specifically in their reasoning about preferences. The paradigm is based on a study that asked whether infants understand that preferences can serve as an underlying cause of people's behaviors (Repacholi & Gopnik, 1997). Fourteen- and eighteen-month-olds were tested in two different conditions: a “matched” trial (where the experimenter gave the preferred goldfish crackers), and a “unmatched” trial (where the experimenter gave broccoli), depending on experimental conditions.

Keywords: Theory of mind; Desire-based reasoning; Infant learning; Social cognition; Preferences.

Experiment 1: Methods

Participants

Infants in both experiments were recruited by phone and email from the California East Bay Area and Southwestern Ontario. In Experiment 1, 55 infants were tested. We used the strict criterion that only infants who did not share the correct item on an initial pre-test (described below) were included in the final analysis. Thirty infants completing training did not already know that preferences are diverse. Twenty infants per condition were tested in the full training procedure. The mean age of the DDT: mean age = 15.7 months; Range = 14.1 months to 17.5 months. The sets of toys were 4 trucks and 4 dogs, 4 planes and 4 monkeys. The age range of the infants varied from 14 months to 16 months (Lucas et al., 2014).

Procedure, Design and Predictions

While this is an important distinction we will discuss it further, because both processes result in identical behavior in our task.

Understanding of preferences by testing them in a modified version of Repacholi & Gopnik's (1996) Goldfish/Broccoli task. All infants were tested in the critical unmatched trial type, wherein the experimenter’s preferences conflicted with the infant’s preferences. Only infants who failed to produce other signals to suggest that their preferences are inconsistent with one another or with the infants themselves.

The other half completed a “Non-Diverse Desires” Training condition (henceforth, N-DDT), where they observed multiple training trials with two experimenters demonstrating different preferences from one another. The other half completed a “Diverse Desires” Training condition (henceforth, DDT) where they observed multiple training trials with two experimenters demonstrating the same preferences. Following training, infants were tested again on two unmatched test trials, one directly after training and the other approximately 24 hours later.

The present experiments examine a shift in reasoning about desires during training. The idea that infants might shift from a simple to a more complex model was formalized as part of a broader look into whether children learn preferences in a way that is rational or optimal under certain assumptions (Lucas et al., 2014).

The present experiments examine a shift in reasoning about desires during training. The idea that infants might shift from a simple to a more complex model was formalized as part of a broader look into whether children learn preferences in a way that is rational or optimal under certain assumptions (Lucas et al., 2014).

The present experiments examine a shift in reasoning about desires during training. The idea that infants might shift from a simple to a more complex model was formalized as part of a broader look into whether children learn preferences in a way that is rational or optimal under certain assumptions (Lucas et al., 2014).
ensure that they could share with the experimenters. The warm up consisted of each experimenter passing a toy (e.g., a doll or toy keys) to the infant and asking her to pass it back by placing it in the experimenters’ hands.

Procedure 1 is based on Repacholi & Gopnik (1997). Experimenter 1 slid a plate of food consisting of a few pieces of vegetables and snacks (e.g., raw broccoli and Goldfish crackers) towards the infant and encouraged the infant to try some. The experimenter gave the infant a 45 second time frame to taste the foods and the experimenter determined which of the two foods the infant preferred. We used the same coding as in Repacholi & Gopnik (1997) to determine food preferences on all trials (pre- and post-tests). Inter-coder agreement for preferences was 91%. When the infant’s preference was determined, the experimenter took out a container consisting of the same foods the infant had tried. The experimenter then demonstrated that she liked the food that the infant did not show a preference for and was disgusted by the food that the infant preferred. The experimenter showed her preferences by saying, e.g., “Eww! Crackers! I hated the crackers! Eww!”, and “Mmm! Broccoli! I tasted the broccoli! Mmm!” The experimenter showed a liking and disliking towards each food three times and she did this using facial expressions based on the descriptions of Ekman & Friesen (1975). Next, the experimenter placed broccoli on one side of a tray and Goldfish crackers on the other, placed her hand with her palm up towards the infant, said, “can you give me some?” and slid the tray towards the infant. The infant was given 45s to pass food to the experimenter. If the infant gave the experimenter the food that the experimenter showed a preference towards, then the infant passed the pre-test. If the infant gave the experimenter the food that she disliked, or did not provide the experimenter with any food, then the infant failed the pre-test.

Training Trials. Infants who failed the pre-test were introduced to either the DDT condition or the N-DDT condition. Infants in the DDT condition saw two experimenters liking and disliking different toys and infants in the N-DDT condition saw two experimenters liking and disliking the same toys.

Training proceeded as follows: Training trial 1 occurred right after the pre-test. During training trial 1, Experimenter 1 put a jar of toys (e.g., dogs and trucks) onto the table and subsequently pulled out three toys of one type (e.g., dogs) and placed them on the table. Then, the experimenter pulled out three toys of the other type (e.g., trucks) and expressed dislike towards them. The dialogue and facial expressions were used during the pre-test. The experimenter expressed her preferences by saying, “Yay! A dog! I got a dog! Yay!” and “Eww! A truck! I picked up a truck! Eww!”. Once Experimenter 1 expressed her emotions for each type of toy three times, Experimenter 2 took over. Experimenter 2 showed liking and disliking towards the same toys as Experimenter 1 if the infant was in the N-DDT condition (e.g., liked dogs and disliked trucks) and she showed liking and disliking towards the opposite toys as Experimenter 1 if the infant was in the DDT condition (e.g., liked trucks and disliked dogs).

Training trial 2 involved Experimenter 2 and the infant. It was similar to the pre-test, except that it involved a different set of food (e.g., the toy and fruit) for each item that each food item was part of the pre-test. Experimenter 2 gave the infant a plate of food and determined which food the infant preferred within 45s. In the DDT condition, the experimenter then demonstrated that she preferred the food that the infant disliked and disliked the food that the infant preferred. In the N-DDT condition, the experimenter demonstrated that she liked and disliked the same foods as the infant. The infant was not asked to share any food with the experimenter, as this was a training trial and not a test.

Training trial 3 was identical to training trial 1, but with a different set of toys (e.g., monkeys and planes). Experimenter 1 expressed liking to one type of toy and dislike towards the other type of toy. Experimenter 2 had a turn expressing her emotions towards each of the toys. She expressed happiness and dislike towards the same toys as Experimenter 1 if the infant was in the N-DDT condition and expressed happiness and dislike towards the opposite toys as Experimenter 1 if the infant was in the DDT condition. After Experimenter 2 finished her demonstrations, infants completed training task 1. Experimenter 2 put one of each type of toy on both sides of a tray (e.g., a monkey on right, a plane on left), placed her palms face up towards the infant, pushed the tray towards the infant and asked the infant to share one with her. The infants were given 45s to share a toy with the experimenter. Once the infant shared a toy with Experimenter 2, Experimenter 1 had a chance to ask the infant to share with her the toy that she liked.

Training trial 4 was a repetition of training trial 3 and included a training task that was identical to the one completed after training trial 3. The purpose of the training tasks, where infants were asked to share one of two toys with each experimenter, was simply to ensure that the infants did not get bored and continued to share throughout the study. We did not expect that infants would remember the toy and in fact we found that infants did not reliably remember the experimenters’ preferences in either condition of Experiment 1 or in Experiment 2 (all p’s > .25 for ANOVA’s examining infants’ passing behavior on the experimenters’ side; training test 1 performance was identical across training conditions (DDT: 7/20 correct; N-DDT: 7/20 correct). Post-training test 2 was administered by training the post-training test based for post-training test 2 (DDT: 15/20 correct; N-DDT: 7/20 correct; X²(1, N = 34) = 6.46, p = .01). Only the performance on post-training test 2 for infants in the DDT condition was significantly above chance (p = .04, binomial).

For the first 10 infants in both training conditions, the food on post-training test 2 was identical to the food on training trial 2 (which the infant used with Experimenter 2 on Day 1 but did not share). We switched this to a new food type to ensure that any improvement in infants’ performance on Day 2 in DDT could not be explained by already being familiar with these foods.

**Experiment 1: Discussion**

Our results suggest that the type of information provided during training was crucial to infants’ learning of diverse desires. When infants were provided with a large number of instances indicating that two different people can like different things, they appeared to share the likes that they disliked but the experimenter preferred. However, infants’ performance did not improve when they saw preferences that were not diverse: infants in the N-DDT condition did not share the correct food with the experimenter on any post-training tests. This suggests that training with appropriate evidence can result in significant changes to children’s explicit Theory of Mind. But infants in the DDT condition only demonstrate advances in understanding on Day 2 of the experiment, during the second post-training test? We see at least two possible explanations. One possibility is that post-training test 1 served as a final training trial, giving infants the minimum number of examples required to change their model of how preferences work (i.e., to learn that they apply to the individual). A second possibility is that a night of sleep resulted in improved learning of this general knowledge about other’s minds, allowing infants to pass the test on Day 2 but not on Day 1. We will address these possibilities more fully in the General Discussion.

Before we can speculate as to why children appeared to learn something new about preferences in the DDT condition, we must first investigate an alternative interpretation of the Experiment 1 data. It is possible that the infants in the DDT condition did not learn that preferences are diverse, but instead learned something less conceptually powerful like, “In this game I’m playing, people always get opposite things when I should give the other person the thing that I didn’t take.” If this is the case, then the participants did not learn that preferences are specific to the individual; they may play a game of opponents where you run a second experiment to tease apart these explanations.
test 2, even though the experimenter demonstrates that she likes the food that the infant also prefers. We maintained the exact same procedure as in the DDT condition of Experiment 1, including using an “unmatched” trial type for post-training test 1, as the effect was observed only in post-training test 2 and so every aspect of the experimental session must remain the same until that point.

Experiment 2: Methods

Participants

Participants were 29 infants and, as in Experiment 1, only children who failed to give the correct food on the initial pre-test continued to training with 20 infants tested in the full training procedure (mean age = 15.5 months; Range = 14.4 months to 17.0 months). An additional 10 infants were tested but not included in data analyses due to failing to complete the study because of fussiness (1), parental interference (1) or refusing to share anything with the experimenters on all test trials (8).

Materials

Food. The food was the same as in Experiment 1 except that the wheel-shaped crackers were replaced with Animal Crackers. This was done because we could no longer find the wheel-shaped crackers.

Toys. The sets of toys were 4 hippos and 4 trucks, and 4 cats and 4 planes. Again, all of the toys within an individual type were slightly different in shape and/or color.

Procedure and Design

The experimental procedure, counterbalancing and randomization were identical to Experiment 1 DDT.

Predictions

We predicted that infants would perform at chance on post-training test 1, as they did in Experiment 1. If infants give the experimenter the correct food on post-test 2 (the food that both the experimenter and the infant like), then this will suggest that infants in Experiment 1 did not simply learn to play a game of opposites but instead learned that preferences are diverse.

Experiment 2: Results

Again we replicated the findings from Reepchol & Gopnik. Again, 9/30 infants passed the pre-test (p = .06, binomial, marginally significantly fewer than chance), 18 infants shared the incorrect food and 2 infants shared nothing.

Six out of 20 infants were correct on post-training test 1 and 13 out of 20 were correct on post-training test 2, both significantly different from chance (p = .02 and p = .03, respectively).

The critical comparison is between infants’ performance on post-training test 2 in the Experiment 1 DDT condition and in Experiment 2. This comparison addresses whether infants in Experiment 1 simply learned to play a game of opposites and would have shared the opposite food type to their own preference regardless of what the experimenter demonstrated on post-test 2. For this analysis, we coded infants’ performance in terms of whether they gave the experimenter the opposite food to what the infant preferred (which is correct in Exp 1 DDT but incorrect in Exp 2). We gave infants a score of 1 for sharing the opposite food and a score of zero for sharing the same (non-opposite) food. This comparison is 4 score of 7/20 was significantly different from chance in Experiment 2 and 15/20 on post-training test 2 in the DDT condition of Experiment 1. Using a Fisher’s Exact test, we found that infant’s performance on these trials was significantly different from one another, X²(1; N = 40) = 6.46, p = .01, suggesting that infants in Experiment 1 were more likely to share the opposite food than infants in Experiment 2, where they would have been incorrect in doing so.

Experiment 2: Discussion

Overall, most infants gave the experimenter the food that they preferred (and that the infant also preferred) on post-training test 2 (this was not significantly different from chance using a binomial test). Though we would have expected infants to share the correct food at a higher than chance levels in this “matched” trial, we suspect that the non-significant result is due to a lack of statistical power caused by having relatively few participants for binomial statistics. In general, the percentage of infants offering the correct, “matched” food on this trial is very similar to the percentage of younger infants who did so in Reepchol & Gopnik (1997) (65% vs. 72%, respectively).

The purpose of Experiment 2 was to eliminate the possible explanation that participants in the Experiment 1 DDT condition only learned to give the experimenter the opposite food of what they themselves wanted. Comparison of Experiments 1 and 2 suggest that this was not the case, as infants shared the food that they preferred in Experiment 2 and not reflexively give the experimenter the opposite food following training.

General Discussion

Together, these findings show that infants younger than 18 months can learn about the subjectivity of preferences when provided with appropriate training. Infants who were exposed to any training, they provided an adult with the food that they personally liked and not the food that the experimenter liked. This was probably because they incorrectly believed that preferences are universal. However, when provided with diverse preferences during training, infants were able to reason correctly about another person’s preferences, providing the experimenter with the food that she liked. In contrast, the infants who only saw congruent expressions of liking and disliking options did not learn to reason correctly about another person’s preferences, and continued to give the experimenter the food that they themselves preferred, regardless of the experimenter’s preference.

Experiment 2 helped to clarify these findings, providing evidence that infants did not simply learn to always give the experimenter the food that they liked (and the infant liked), we can be confident that infants in Experiment 1 learned that preferences are diverse. Taken together, this coding resulted in a score of 7/20 that was learned through training about the diversity of desires, moving from a less to a more sophisticated understanding of other’s preferences. On a broader level, these findings suggest that young children can learn from experience to make an important advance in explicit reasoning about Theory of Mind.

One concern regarding these data is the relatively low statistical power that results from our experimental design and the small sample size for each experiment. Although the results in the Experiment 1 DDT condition were significant, it will be prudent to replicate these findings. This replication experiment is currently underway in a larger sample.

An interesting finding in these experiments is that the participants performed identically during the pre-test and post-training test 1, but performed significantly above chance on post-training test 2 in Experiment 1. Both tests occurred after training and we had not predicted this pattern of results, so we now return to the question of why we only saw improvement on post-training test 2.

One possible explanation for this improved performance on post-training test 2 is that post-training test 1 might act as another piece of evidence to train the infants to better understand diverse preferences. That is, post-training test 1 gives infants yet another trial in which the experimenter demonstrates that she likes the opposite food to the infant. It is possible that this extra trial is what allows the infants to learn that preferences are subjective. This possibility can be examined by manipulating the number of training trials, to include an additional trial before post-training test 1 on Day 1. Related to this, we can also examine what type of evidence is most informative – evidence that involves first person experience such as training trial 2 and post-training test 1, or training trials that involve observing other actors display diverse preferences. By manipulating the number and type of training trials across various conditions in future experiments, we can answer these questions.

Another possible explanation for the improved performance only on Day 2 is the role of memory consolidation in sleep. Post-training test 2 occurs the following day, whereas post-training test 1 occurs on the same day as the training trials. Therefore, a potentially critical difference between the two tests is sleep. Research has shown that sleep is important for the consolidation of memories, and improvements in children’s and infants’ learning are seen with longer and more intense sleep (Wilhelm, Prehn-Kristensen & Born, 2012). For example, Hupbach, Gomez, Brand, and Nadel (2009) found that when 5-month-old infants napped after they were exposed to an artificial language, they were more likely to remember the general grammatical pattern of that language 24 hours later, compared to infants who did not nap. It is possible that the infants in our experiment performed better on post-training test 2 because they had slept. To address the sleep hypothesis, one could conduct an experiment similar to these here, except with the entire procedure occurring on the same day. After infants complete post-training test 1, half of the infants would take a nap and half would experience a similar delay without taking a nap. Follow that infants in all conditions would complete post-training test 2. If the infants who napped performed better than those who did not, then this would suggest that sleep consolidation is a crucial aspect of their improved performance.

Conclusion

Research on children’s desire-based reasoning has persisted for decades. Here we examined a prediction from a particular model of how children attribute preferences to others, namely that appropriate training regarding the diversity of desires could result in infants undergoing a significant shift in conceptual development (Lucas et al., 2014). We found that following exposure to different people demonstrating divergent desires, infants were able to move from a model of universal preferences to a model that allows for the individualization of preferences. The success of this training procedure more broadly suggests that early advances in Theory of Mind could be due to experience.

Acknowledgments

We thank Justine Hoch and Sophie Bridgers for an immense amount of help with data collection and infant recruitment. We also thank Elizabeth Attisano, Emily McNish, Meghan McGrath, Julia Henius, Christen Calabas, and Gina Mandreacha for assistance with data collection. Thanks also to the parents and infants for their participation.

References


