How do you know that? Sensitivity to statistical dependency in social learning

Andrew Whalen (awhalen@gmail.com)
Daphna Buchsbaum (daphnab@berkeley.edu)
Thomas L. Griffiths (tom_griffiths@berkeley.edu)
Department of Psychology, University of California, Berkeley, Berkeley, CA 94720 USA

Abstract

Social learning has been shown to be an evolutionarily adaptive strategy, but can be implemented via many different cognitive mechanisms. Sensitivity to statistical dependency in the behavior of other people is a factor that discriminates between possible mechanisms: simple rule based strategies may be unaffected by dependency, while more sophisticated social learning strategies should take it into account. We use a Bayesian model to determine how rational agents should incorporate the effects of statistical dependency when learning from other people, conducting two experiments that examine whether human learners behave similarly. We find that people are sensitive to two different patterns of dependency, supporting the use of a sophisticated strategy for social learning.

Introduction

Social learning is a key factor in the human ability to adapt to a wide variety of environments and plays an important role in cultural transmission of information (Boyd & Richerson, 1988, 2005). Formal models have shown that social learning is an evolutionarily adaptive strategy, able to outcompete individual learning (Laland, 2004). However, there are many possible mechanisms by which social learning could be implemented, ranging from blind imitation to making sophisticated inferences about the beliefs that underlie that behavior. While evolutionary models tell us that social learning should be favored, they don't tell us which mechanism human learners might be using. This question is particularly important given results showing that both adults and children sometimes "overimitate", reproducing another's unnecessary actions (e.g., Lyons, Young, & Keil, 2007; Nielsen & Tomaselli, 2010; McGuigan, Makinson, & Whiten, 2011).

In this paper, we explore the mechanisms behind human social learning by examining how sensitive people are to statistical dependency in the behavior of other people. For example, imagine hearing from two friends that they visited a particular restaurant. Taken at face value, this seems like strong evidence that the restaurant might be a good place to eat. But if you discover that one friend went there after finding out that the other had been, the two pieces of information are no longer statistically independent and the evidence they provide about the quality of the restaurant is reduced. And if one friend had taken the other there, it is reduced even further.

Examining whether human social learning is sensitive to statistical dependency provides an opportunity to discriminate between social learning strategies. Simpler rule-based approaches such as "imitate the majority" should be insensitive to the subtleties of how other people's behavior is generated, focusing just on the behavior itself. In contrast, if social learning is based on rational inferences from the available data, the way in which those data are generated should matter a great deal (for an example see Buchsbaum, Gopnik, Griffiths, & Shafto, 2011). Determining the consequences of dependencies in other's behavior involves reasoning about their mental states and the factors that contribute to their decisions, requiring a sophisticated approach to social learning.

To assess whether people are appropriately sensitive to statistical dependency in the behavior of others, we developed a Bayesian model for a simple social learning task. The model indicates what inferences a rational agent should draw when statistical independence is violated in different ways. We ran two behavioral experiments using this social learning task, finding that people are sensitive to two forms of dependency. These results support the idea that human social learning is based on reasoning about the mental states of other people, rather than simpler strategies such as imitating the majority.

Learning from others

Before presenting our model and experiments, we will summarize some of the key theoretical and experimental results on cultural evolution and social learning. These results break down into three areas of research. At the largest scale, models of cultural evolution have examined how the learning strategies adopted by individuals impact the spread of different behaviors between generations. Within generations, models of what are called "information cascades" have been used to analyze the rapid spread of novel innovations among populations. Finally, a number of studies have explored how individual people learn from informant testimony.

Cultural evolution

Theoretical studies of cultural evolution have shown that social learning has adaptive advantages (Boyd & Richerson, 1988, 2005; Laland, 2004). However, many of these studies analyze systems where individuals are faced with the choice of either learning from the environment or learning socially (Rogers, 1988). In reality, learners are likely to combine both environmental and social information when making a decision. Perreault, Moya, and Boyd (2012) modeled agents who choose a behavior based on a Bayesian learning algorithm which integrates social and environmental information. In this model, agents assumed that the social cues provided by other agents were independent from one another. This assumption was justified by the fact that all behavioral transmission happened between generations where the probability that the informants learned from each other was low. However, many behaviors are transmitted within generations, where informants are likely to share information.

Information cascades

Unlike the cultural evolution literature, the literature on information cascades, a within-generation model of social decision-making developed by economists, takes into account the statistical dependency between agents (Bikhchandani, Hirshleifer, & Welch, 1992). The basic scenario has a sequence of agents each making a decision by combining the information provided by the decisions made by previous agents with that provided by a small amount of private data. An information cascade occurs when agents adopt the majority belief, regardless of their own private information. The cascade persists as more agents enter the population and adopt the majority belief. Bikhchandani et al. (1992) analyzed how rational agents who took into account dependencies in previous responses would act in this situation, and showed that information cascades are surprisingly common. This result provides a potential explanation for the adoption and spread of fads and fashions, as well as boom-bust cycles in the economy. Information cascades have been tested in the laboratory using a simple scenario that provided the inspiration for the experiments we present later (Anderson & Holt, 1997), but this previous work did not examine the consequences of manipulating people's beliefs about statistical dependency.

Individual decision making

Social learning, and imitation in particular, have been studied extensively by psychologists. This work has generally demonstrated that adults and even young children are sensitive to many aspects of the knowledge and mental states of their social informants (for some well known examples see Meltzoff, 1995; Gergely, Bekkering, & Kiraly, 2002; Carpenter, Call, & Tomasello, 2005). Related work on how children learn from testimony has similarly found that children take many factors into account, including the prior accuracy (e.g., Koenig & Harris, 2005; Corriveau, Meints, & Harris, 2009), expertise (e.g. Jaswal, 2006; Sobel & Corriveau, 2010) and certainty (Jaswal & Malone, 2007; Tenney, Small, Kondrad, Jaswal, & Spellman, 2011) of informants. However, other work has found that in some situations, people appear to simply copy the beliefs of others. Adults often disregard their own judgments when socially pressured (for a review see Cialdini & Goldstein, 2004), and both adults and children may sometimes conform to a majority opinion that conflicts with their own direct perceptions (Asch, 1956; Corriveau & Harris, 2010). Looking at the effect of statistical dependency can help us determine if this conformity is the result of a simple rule-based strategy, or a more sophisticated inference process.

Rational social learning

Many inferences that people make rely upon a combination of their own experience and the behavior of other people. In order to determine how a sophisticated agent should combine these forms of information, we developed a Bayesian model that can incorporate different patterns of dependency. This model makes direct predictions that we can test experimentally, having no free parameters.

We assume that agents receive some directly observed data about the state of the world, d, and testimony from n informants $t_1, \ldots, t_n$. To make a decision, learners evaluate a potential hypothesis, h, using Bayes' rule,

$$p(h|d, t_1, \ldots, t_n) \propto p(t_1, \ldots, t_n|d, h) p(d|h) p(h) \qquad (1)$$

where $p(h|d, t_1, \ldots, t_n)$ is the posterior probability of h, the degree of belief assigned to h after receiving the data and testimony, and $p(h)$ is the prior probability of h, the degree of belief assigned to h before receiving any evidence.

In order to estimate the probability of the testimony, $p(t_1, \ldots, t_n|d, h)$, the learner should consider the sources of information that each informant had access to. If the data the learner observes, d, is unknown to the informants, then $p(t_1, \ldots, t_n|d, h) = p(t_1, \ldots, t_n|h)$. We will assume that this is the case, since it simplifies calculations and is consistent with the task we use in our experiments. The form of $p(t_1, \ldots, t_n|h)$ depends on how the informants generate their testimony. We first consider the case of independent testimony, and then discuss two different patterns of dependency.

Independent testimony

If the informants' testimonies are independent of one another given h, then the probability of a series of testimonies is equal to the product of the probability of the individual testimonies:

$$p(t_1, \ldots, t_n|h) = \prod_{i=1}^{n} p(t_i|h). \qquad (2)$$

If the testimony produced by the informants is based on their own experiences, this needs to be taken into account in calculating the probability that they would produce their testimony. More formally, if we assume that informant i received private data d_i, we obtain $p(t_i|h)$ by marginalizing over d_i,

$$p(t_i|h) = \sum_{d_i} p(d_i|h) p(t_i|d_i), \qquad (3)$$

where $p(t_i|d_i)$ is the probability that the informant produces testimony t_i after observing d_i. One possibility is that informants deterministically give testimony that supports the hypothesis with the highest posterior probability, with $p(t_i = h_i|d_i) = 1$ for $h_i = \arg\max_h p(d_i|h) p(h)$. This is typically assumed in models of information cascades (e.g., Bikhchandani et al., 1992). Alternatively, empirical (Vulkan, 2000) and theoretical (Luce, 2005; Shepard, 1958) results in psychology suggest that in many cases people "probability match", so that informants would give testimony in support of a hypothesis proportional to the informant's posterior probability of the hypothesis, with $p(t_i = h_i|d_i) \propto p(d_i|h_i) p(h_i)$. We evaluate both the maximizing and probability matching models.

Dependent testimony

If multiple informants give testimony based on shared information, then the probability of any single testimony is not

independent from the others. We consider two cases: where informants give their testimony sequentially, with each informant hearing the preceding testimony, and where informants base their testimonies on shared private data.

Sequential testimony   Much of the theoretical work on information cascades assumes that informants give their testimony sequentially. Each informant uses their own private information, and the testimony of previous informants to make a decision as to which option to support. Formally, the testimony of informant i is based on the previous testimony of the previous informants, $t_1, \ldots, t_{i-1}$, and their own private data, d_i. The probability of $t_1, \ldots, t_n$ is then

$$p(t_1, \ldots, t_n | h) = p(t_1 | h) \prod_{i=2}^{n} p(t_i | t_1, \ldots, t_n, h). \quad (4)$$

The value $p(t_i | t_1, \ldots, t_{i-1}, h)$ can be found recursively by finding the values for $p(t_1 | h)$ up to $p(t_{i-1} | t_1, \ldots, t_{i-2}, h)$:

$$p(t_i | t_1, \ldots, t_{i-1}, d_i, h) \propto (\prod_{j=1}^{i-1} p(t_j | t_1, \ldots, t_{j-1}, h)) p(d_i | h) p(h). \quad (5)$$

As in the case of independent informants, we can find $p(t_i | t_1, \ldots, t_{i-1}, h)$ by marginalizing over the private information, d_i, and assuming informants apply Bayes' rule and then either maximize or probability match.

Shared private data   If the informants all provided testimony based on a single piece of data (e.g., they all went to the restaurant together), then the probability of this testimony can be found by marginalizing over this shared private data. Denoting the shared data d', we have

$$p(t_1, \ldots, t_n | h) = \sum_{d'} p(d' | h) \prod_i p(t_i | d', h) \quad (6)$$

where the probabilities $p(t_i | d', h)$ are calculated by applying Bayes rule and assuming either maximizing or probability matching to the posterior, as above.

Reasoning about balls and urns

The consequences of different forms of dependency for rational social learning can be hard to imagine in abstract, so we will work through a concrete example in detail. One of the simplest examples that illustrates these consequences is the "ball and urn" scenario used in the information cascade experiment conducted by Anderson and Holt (1997). This scenario is also the basis for our own experiments.

Imagine there are two colored urns. One of the urns is colored red, the other urn is colored blue. An experimenter explains that in the red urn $\frac{5}{6}$ of the balls are red, and the rest of the balls are blue. In the blue urn the proportions are reversed. In secret, the experimenter pours one of the urns into a bag. She then shows a ball to each of three informants, and one to the participant. The informants say which urn they think was used to fill the bag, based on the information available to them. The experimenter then asks the participant to decide
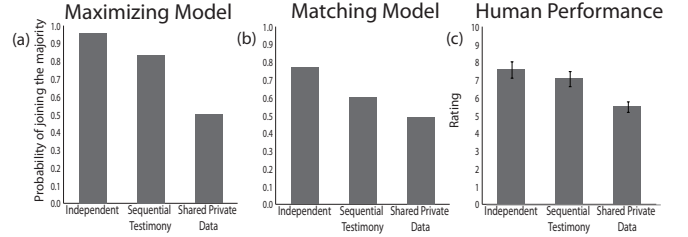

Figure 1: Probability of agreeing with the majority opinion in the simple balls and urns task used in Experiment 1, for both the Bayesian model and human participants.

which urn was used to fill the bag, based on the testimony of the informants and the ball seen by the participant.

If all three informants agreed with each other and thought the bag was filled from the red urn, but the participant got a blue ball, what should the participant say? We will analyze three conditions, corresponding to the three cases presented in the previous section. The predictions for the three conditions are shown in Figure 1(a) for the maximizing model and in Figure 1(b) for the probability matching model, using the true probabilities of red and blue balls for $p(d|h)$ and assuming both hypotheses are equally likely for $p(h)$.

Independent testimony

Imagine that the three informants are all in separate rooms and each receive a different ball sampled from the bag, making their testimony completely independent. In this case, the model predicts that the participant should agree with the social testimony, picking the red urn. The model infers that all three informants all probably received red balls and three red balls outweigh the participant's single blue ball.

Sequential testimony

In this case, all three informants might be sitting at the same table and each receive a different ball, but have the opportunity to hear the answer given by the previous informants before providing their testimony. This is the situation that was analyzed in Anderson and Holt's (1997) experiment. If the informants give their testimony sequentially, the model again predicts that the participant should agree with the social testimony. However, the model places less weight on the hypothesis that the red urn was used to fill the bag. The model takes into account the fact that the three informants shared information. If the first two people received red balls and the third person received a blue ball, they may still all agree that the red urn was used to fill the bag even if the third person goes against the private evidence she received – a mere blue ball against two likely red balls – and votes in favor of the majority. This possibility makes the model less sure of its decision compared to the independent condition.

Shared private data

Now, consider what happens if all three informants are sitting at the same table and all observe exactly the same ball, rather than each seeing a separate ball drawn from the bag. If all three informants saw the same ball, the model is evenly

split between the two urns. On the one hand, the three informants probably received a single red ball, but the participant received a blue ball. With one red ball and one blue ball on the table, the balls provide equal evidence for either urn being used to fill the bag.

Summary

Even in a simple scenario with two hypotheses and three informants, a rational social learner should act differently in response to different patterns of statistical dependency. To compare our model with human behavior, we ran an experiment to see how people incorporate their own understanding of the information each informant used to give their testimony.

Experiment 1: Consistent informants

Experiment 1 used the scenario presented in the previous section, with three informants providing consistent testimony that went against the private data received by the participant. There were three conditions corresponding to the independent testimony, sequential testimony, and shared private data.

Methods

Participants A total of 120 participants were recruited through Amazon Mechanical Turk (http://www.mturk.com). Participants were compensated $0.25 for their time. They were randomly assigned to one of three experimental groups: the independent condition ($n = 37$) or the shared testimony ($n = 41$), or shared-data ($n = 45$). No participants were dropped from the analysis.

Stimuli The experiment was a web-administered survey involving text and pictures. A cartoon of a brown haired woman was the experimenter. Three cartoon women were the informants. The informants differed in terms of hair color, hair style, skin color, and shirt color. Each urn was a picture of a red or blue opaque urn. The balls were colored red and blue.

Procedure First a woman named Jane (the *experimenter*) introduced the urns. She explained that five-sixths of the balls in the red urn were red, and one-sixth were blue. The opposite was true for the blue urn. She introduced her three friends (the *informants*), and explained that she was going to pour one of the urns into a bag and give a ball from the bag to each of her three friends. The friends would then tell the participant which urn they think the bag was filled from. In all three conditions the three informants agreed that the bag was filled from the red urn. The participant then saw a blue ball.[1]

In the independent testimony condition the participant was shown three doors, and was told that one informant was waiting in each room. Inside, each informant sat behind a desk.

In the sequential testimony condition the informants sat behind a long table. The informants gave their testimony in order down the table and acknowledged that they had used their own ball and the testimony of previous informants to make

---

[1]The actual colors were randomized, so half the participants received testimony favoring the blue urn and then saw a red ball.

their decision, but did not see anyone else's ball. Each informant agreed with the previous informants' testimony.

The shared private data condition was the same as the sequential testimony condition, except that a single ball was shared between the informants, and each informant said that they saw the same ball as the other informants. The experimenter then showed the participant a single blue ball, contrary to the three informants' testimony.

Finally, the experimenter asked participants to rate how likely it was that the bag was filled each urn. Participants responded to the survey on an 11-point scale, with 0 corresponding to "definitely the blue urn", 10 to "definitely the red urn", and 5 to "equally likely the blue urn or red urn".

Results and Discussion

Ratings were placed on a consistent scale, corresponding to agreement with the majority, by recoding a rating x to $10 - x$ if testimony favored the blue urn. The mean rescaled ratings for all conditions are shown in Figure 1(c). Overall, participants sided with the informants' testimony over their own private information most in the independent condition, second in the sequential testimony condition, and least in the shared private data condition. The ordering of the means are consistent with the model predictions. The matching model provided a good model fit to the data (Pearson's $r = .90$).

We analyzed the effect of condition on participant responses using an ANOVA. The effect of condition was significant ($F(2, 99.1) = 7.749$, $MSE = 49.56$, $p < 0.001$). We explored the differences between conditions using planned t-tests. The difference between the independent and shared private data conditions was significant (two sample t-test, $t(80) = 3.88$, $p < .001$) as well as the difference between the sequential testimony and shared private data conditions (two sample t-test, $t(84) = 2.66$, $p < .01$). The difference between the sequential testimony and independent testimony conditions was not (two sample t-test, $t(76) = 0.96, p = .34$).

The difference between the shared private data condition and the independent condition suggests that participants were able to use their knowledge of what information informants received to evaluate the informants' testimony. Because the three informants received the same ball and gave the same testimony, participants were able to weigh their judgments against their own conflicting ball.

However, both the maximizing and the probability matching models predict that in the shared private data condition the probability of the bag having mostly red balls is approximately 50%: less than the participant's average value of 60%. Even though this difference was not significant (one sample t-test, $t(44) = 1.31$, $p > .05$), participants may place more weight on informant testimony than the model predicts.

At first glance, the null result between the sequential testimony and independent testimony conditions suggests that people respond similarly in the cases of independent testimony and sequential testimony. However, the magnitude of the difference between these two conditions predicted by the model is relatively small. This suggests instead that the sce-
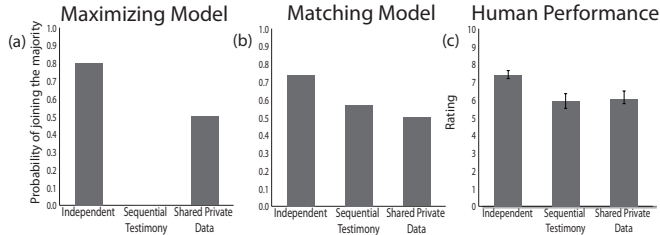
Figure 2: Probability of agreeing with the majority opinion in the more complex task used in Experiment 2, for both the Bayesian model and human participants.

nario used in previous experiments on information cascades may not be sufficient to distinguish between how people use independent testimony over sequential testimony, a limitation that we address in our second experiment.

Experiment 2: Dissenting informant

The ball-and-urn scenario presented above does not result in situations where there is a large expected difference between the independent testimony and sequential testimony conditions. In order to assess whether people are sensitive to the difference between these two patterns of dependency, we changed the scenario by having the third informant dissent from the previous two informants. To give a reason why the informant would dissent, a single diagnostic ball (either white or black) was added to each of the two urns. Since each diagnostic ball was present in only one of the two urns, any informant who received the diagnostic ball would know exactly which urn was used to fill the bag. We also made two other changes. First, the participant did not receive their own ball, having to make a judgment based purely on the testimony of the informants. Second, in the shared private data condition only the first two informants received the same ball. This was done so that the dissenter received her own ball, providing an explanation for why she might dissent.

The model predictions are given in Figure 2(a), for maximizing, and Figure 2(b), for probability matching. The addition of a low-probability diagnostic ball does not substantially change the model predictions in the independent condition. However, it makes an important change to the predictions in the sequential testimony condition, most dramatically in the maximizing model. The model predicts that the last informant will dissent only if she received a diagnostic ball. Since she does dissent, she most likely received a diagnostic ball and so the participant should side with the dissenter over the majority (a similar but somewhat more subtle effect occurs for the probability matching model). Finally, in the shared private data condition, the dissenter probably received a different colored ball than the two informants in the majority. This provides equal evidence for either urn.

Methods

A total of 124 participants were recruited through Amazon Mechanical Turk. Participants were compensated $0.25 for their time. They were randomly assigned to one of three experimental groups: the independent condition ($n = 41$) or the

shared testimony ($n = 41$), or shared private data ($n = 42$). No participants were dropped from the analysis.

Stimuli The stimuli were identical to those in Experiment 1, except for the urns shown. Instead of using opaque colored urns, the urns were replaced with a picture of a clear jar filled with a mix of red and blue balls. A single diagnostic ball (either white or black) was placed in each urn. Each urn was labeled either "Jar A" or "Jar B".

Procedure The procedure was the same as Experiment 1, except for the following changes. References to the "red urn" and the "blue urn" were replaced by references to "Jar A" and "Jar B". In all three conditions the last informant dissented from the previous informants, and supported the belief that the bag was filled from the other urn. In the shared private data condition, the first two informants received the same ball. The last informant received a different ball. At the end of the experiment the participant did not see their own ball and made their judgments based solely on the informants' testimonies. Responses were made on the same 11-point scale as in Experiment 1, changing the names of the urns appropriately.

Results and Discussion

Ratings were rescaled as in Experiment 1. The mean rescaled ratings are shown in Figure 2(c). Participants sided with the majority testimony most in the independent testimony condition, second in the sequential testimony condition, and least in the shared private data condition. The means and order of the results are consistent with the probability matching model predictions, but not the maximizing model predictions. The probability matching model provides a good fit for the experimental data (Pearson's $r = .94$).

We analyzed the effect of condition on participant responses using a one-way ANOVA. The effect of condition was significant ($F(2, 54.3) = 5.561$, $MSE = 27.13$, $p < 0.005$). We explored the differences between the conditions using planned t-tests. The difference between the independent testimony and sequential testimony conditions was significant (two sample t-test, $t(80) = 3.12$, $p < .005$) as well as the difference between the independent testimony and shared private data conditions (two sample t-test, $t(81) = 3.16$, $p < .001$). The difference between the sequential testimony and shared private data conditions was not significant (two sample t-test, $t(81) = 0.22, p > .82$).

The difference between the independent testimony and sequential testimony conditions suggests the learning mechanism that participants use is sensitive to social information that is shared between informants. The difference between the shared private data condition and the independent testimony condition supports our conclusion from Experiment 1 that people are sensitive to non-social shared information.

Qualitatively, participants' performance resembles the probability matching model more than the maximizing model used in earlier work on information cascades. However, in both the sequential testimony and the shared private data con-

ditions participants sided with the majority slightly more than the probability matching model predicts, suggesting that even though individuals are able to utilize shared information that informants use to make their judgments, they may place more trust in the informants' testimony. This difference was significant in the Shared Private Data condition (one sample t-test, $t(41) = 2.867$, $p < .01$), but not significant in the sequential testimony condition (one sample t-test, $t(40) = .54$, $p > .05$).

General Discussion

The goal of this research is to determine whether human social learning is based on a sophisticated strategy that appropriately takes into account dependencies in the behavior of other people. To answer this question, we developed a Bayesian model that indicates how patterns of dependency should affect social learning. The model makes clear predictions about two kinds of dependency – sequential testimony and shared private data – which we tested through two experiments. Experiment 1 showed that people are sensitive to shared private data, using a task that has been employed in previous experiments on information cascades. Experiment 2 showed that people are sensitive to sequential testimony, using a task that is more sensitive to this kind of dependency. However, in both experiments people's judgments were influenced by dependency less than they should have been.

Our results have implications for models of cultural evolution and information cascades. For models of cultural evolution, they offer insights into the mechanisms that underlie social learning, and suggest that patterns of dependency should be taken into account in contexts where agents might encounter dependent social cues. While models of information cascades typically assume sequential testimony, our results show that people are sufficiently sensitive to patterns of dependency that information cascades will be even more probable if it is assumed that informants provide independent testimony. In addition, the matching model provided a closer qualitative and quantitative fit to human performance than the maximizing model. This empirical evidence conflicts with the assumption that informants maximize their posterior used in previous work on information cascades (e.g. Bikhchandani et al., 1992) and helps explain some of the patterns of "errors" observed in the experiments by Anderson and Holt (1997).

Taken together, our findings suggest that human social learning mechanisms are fairly sophisticated. People do not just use simple rule-based imitation strategies. Instead they are able to integrate their own private information with informants' testimony, and take into account how each informant decided upon their testimony. This implies that human cultural evolution is not simply a result of individuals making a trade-off between acquiring their information socially or through trial-and-error learning, but is instead the result of complex decisions that draw on beliefs about informants' sources of information. When learning from testimony, learners are asking themselves the question: "and just how do you know that?"

References

Anderson, L., & Holt, C. (1997). Information cascades in the laboratory. *The American economic review*, 847–862.

Asch, S. E. (1956). Studies of independence and conformity: I. a minority of one against a unanimous majority1. *Psychological Monographs: General and Applied*, *70*(9), 1-70.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 992–1026.

Boyd, R., & Richerson, P. (1988). *Culture and the evolutionary process*. University of Chicago Press.

Boyd, R., & Richerson, P. (2005). *The origin and evolution of cultures*. Oxford University Press, USA.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, *120*(3), 331-340.

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13-F20.

Cialdini, R., & Goldstein, N. (2004). Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, *55*, 591–621.

Corriveau, K. H., & Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. *Developmental Psychology*, *46*(2), 437-445.

Corriveau, K. H., Meints, K., & Harris, P. L. (2009). Early tracking of informant accuracy and inaccuracy. *British Journal of Developmental Psychology*, *27*(2), 331–342.

Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*, 755.

Jaswal, V. K. (2006). Preschoolers favor the creator's label when reasoning about an artifact's function. *Cognition*, *99*, B83-B92.

Jaswal, V. K., & Malone, L. S. (2007). Turning believers into skeptics: 3-year-olds' sensitivity to cues to speaker credibility. *Journal of Cognition and Development*, *8*(3), 263-283.

Koenig, M., & Harris, P. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*(6), 1261–1277.

Laland, K. (2004). Social learning strategies. *Learning & behavior*, *32*(1), 4–14.

Luce, R. (2005). *Individual choice behavior: A theoretical analysis*. Dover publications.

Lyons, D., Young, A., & Keil, F. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, *104*(50), 19751–19756.

McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology*, *102*(1), 1–18.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838-850.

Nielsen, M., & Tomaselli, K. (2010). Overimitation in kalahari bushman children and the origins of human cultural cognition. *Psychological Science*, *21*(5), 729–736.

Perreault, C., Moya, C., & Boyd, R. (2012). A bayesian approach to the evolution of social learning. *Evolution and Human Behavior*.

Rogers, A. (1988). Does biology constrain culture? *American Anthropologist*, *90*(4), 819–831.

Shepard, R. (1958). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *Journal of Experimental Psychology*, *55*(6), 509.

Sobel, D. M., & Corriveau, K. H. (2010). Children monitor individuals' expertise for word learning. *Child Development*.

Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology*, *47*(4), 1065-1077.

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*, 101-118.