Sensitivity to shared information in social learning

Andrew Whalen

School of Biology, University of St Andrews

Thomas L. Griffiths

Department of Psychology, University of California, Berkeley

Daphna Buchsbaum

Department of Psychology, University of Toronto

Author Note

Preliminary versions of Experiments 1 and 3 were presented at the Annual Meeting of the Cognitive Science Society and published in the proceedings (Whalen, Buchsbaum, & Griffiths, 2013).

This work was supported by grant number IIS-0845410 from the National Science Foundation (NSF), John Templeton Foundation Grant #40128, a NSF Graduate Research Fellowship to DB, and a SURF/Rose Hills Fellowship to AW.

Correspondence should be addressed to School of Biology, University of St Andrews, Harold Mitchell Building, St Andrews, Fife, UK, KY16 9TH. Email: aczw@st-andrews.ac.uk

Abstract

Social learning has been shown to be an evolutionarily adaptive strategy, but can be implemented via many different cognitive mechanisms. The adaptive advantage of social learning depends crucially on the ability of each learner to obtain relevant and accurate information from informants. The source of informants' knowledge is a particularly important cue for evaluating advice from multiple informants; if the informants share the source of their information or have obtained their information from each other then their testimony is statistically dependent and may be less reliable than testimony from informants who do not share information. In this paper, we use a Bayesian model to determine how rational learners should incorporate the effects of shared information when learning from other people, conducting three experiments that examine whether human learners behave similarly. We find that people are sensitive to a number of different patterns of dependency, supporting the use of a sophisticated strategy for social learning that goes beyond copying the majority, and broadening the situations in which social learning is likely to be an adaptive strategy.

Sensitivity to shared information in social learning

Social learning is a key factor in humans' ability to adapt to a wide variety of environments and plays an important role in the cultural transmission of information (Boyd & Richerson, 1985, 2005). Formal models have shown that social learning is an evolutionarily adaptive strategy, able to out-compete individual learning (Laland, 2004; Henrich & Boyd, 1998; Rendell, Fogarty, & Laland, 2010; Rendell et al., 2011), however these evolutionary models do not tell us the mechanisms that individuals may be using. Understanding the mechanisms that underlie social learning is particularly important because how individuals learn from others may have a drastic impact on the spread of novel beliefs through populations and the development of human culture (Boyd & Richerson, 1985).

One of the key results from these evolutionary models is that learners must be selective in who and how they copy (Rogers, 1988; Boyd & Richerson, 1985; Rendell et al., 2011; Enquist, Eriksson, & Ghirlanda, 2007); in order for social information to be useful, it needs to provide accurate and relevant information. Selective social information use has been experimentally found in human adults (Efferson, Lalive, Richerson, McElreath, & Lubell, 2008; Morgan, Rendell, Ehn, Hoppitt, & Laland, 2011), children (for some reviews see Mills, 2013; Harris, 2012), and other animals (Pike & Laland, 2010; Hoppitt & Laland, 2013). In addition, when learning from unfamiliar individuals, it can pay to use the number of informants as an alternate cue, and follow the majority (Boyd & Richerson, 1985; Henrich & Boyd, 1998), a strategy that has received cross-species empirical support (Asch, 1956; Efferson et al., 2008; Pike & Laland, 2010). However, following the majority may not always be beneficial, and may even lead individuals to make the wrong decision (Asch, 1956; Corriveau & Harris, 2010; Haun & Tomasello, 2011), or allow less effective behaviors to spread through or be maintained in a population (Bikhchandani, Hirshleifer, & Welch, 1992; Henrich & Boyd, 1998).

In general, learning from multiple individuals increases the amount of evidence the group provides. However when a group of individuals make their decisions based on the same, or shared information, this decreases the amount of evidence the group provides. Being unaware of shared information, or unable to utilize it when making decisions based on social information may lead learners to place trust in larger groups, whether or not the group actually provides more information. For example, imagine hearing from two friends that they thought it was going to rain tomorrow. While it may seem like hearing the same thing from two different people provides additional support for the probability of rain, this is not the case if both friends are basing their testimony on the same news broadcast. In comparison if the two friends each looked at separate forecasts, then the second friend might be an additional, and confirmatory source of information. Therefore, whether or not individuals are sensitive to the shared information a group of informant's provides may impact the probability that learners adopt incorrect behaviors or beliefs, or alternately, might increase the chance a novel beneficial behavior might spread through a population even though it is initially in the minority.

In this paper we analyze whether or not learners are sensitive to shared information multiple informants use to make a decision. To analyze the impact of both shared social and asocial information more precisely, we first develop a Bayesian model of learning from others to analyze what inferences a rational learner should draw when statistical independence is violated in different ways. We ran three behavioral experiments and examined people's sensitivity to different forms of dependency, comparing human performance to our rational model's predictions. In Experiment 1, we analyzed whether individuals are sensitive to the dependency between informants who share social and asocial information. In Experiment 2, we replicated the findings of Experiment 1 and analyze a set of new experiments where individuals only hear directly from one of the informants. In both experiments we found that individuals are sensitive to shared asocial information, but may not be sensitive to shared social information. In Experiment 3, we modified the task to obtain larger predicted differences between conditions and find that individuals are also sensitive to shared social information. These experiments give insights into the mechanisms that underlie human social learning, and may help explain the processes that have shaped human culture.

Bayesian Learning from Statistically Dependent Social Information

In some cases it can be difficult to understand how shared information may impact a learner's assessment of an informant's testimony. To address this issue, we construct a Bayesian model that allows for the explicit incorporation of different patterns of shared information, modeled as a form of statistical dependency between learners, and allows for the integration of both social and asocial sources of information. This model makes direct predictions that we can test experimentally, while having no free parameters. The goal of this model is to examine how a rational learner might incorporate information about shared social information to see if people incorporate social information in similar ways.

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

$$p(h|d, t_1, \ldots, t_n) \propto p(t_1, \ldots, t_n|d, h)p(d|h)p(h) \tag{1}$$

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. 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 when generating their testimony.

For clarity, we have split the evidence that individuals receive into "asocial" and "social" categories, corresponding if the evidence came from the world (asocial) or from an informant (social). However, in our formulation social information is treated in the same way as asocial information. Both asocial and social information are integrated based on the evidence they provide (similar to Perreault, Moya, & Boyd, 2012), as opposed to presupposing a separate social learning mechanism as often assumed in other models (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Rendell et al., 2010). Equation 1 could be rewritten in a more traditional format where the individual pieces of evidence the learner receives, $d, t_1, ..., t_n$ are combined into a single "evidence" vector, e.

Independent testimony

If the informants' testimonies are independent of one another given h (i.e. based on separate sources of information), then the probability of a series of testimonies is equal to the product of the probability of the individual testimonies:

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

If the testimony produced by the informants is based on their own private data, d_i we can marginalize over possible sets of private data to obtain $p(t_i|h)$,

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

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 $h_i = \arg\max_h p(d_i|h)p(h)$ (in the case where the posterior probabilities are equal we assume that the informant gives testimony consistent with the data they received). This is typically assumed in models of information cascades (a type of run-away cultural process where a maladaptive behavior can spread through a population even if all of the learners are rational, see e.g., Bikhchandani et al., 1992; Easley & Kleinberg, 2010). Alternatively, empirical (Vulkan, 2000) and theoretical (Luce, 1977; 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 the predictions of both 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 of the others. The lack of independence between informants can be modeled as a form of statistical dependency: in each of the cases analyzed informants are more likely to produce similar testimony than if the informants were independent of each other. 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. The first case is an example of shared social information, where each of the informants also receives social testimony from the preceding informants to make their decision. The latter case is an example of shared asocial information where the informants are making their decisions based on mutually observed evidence from the world.

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, which is then given to the next informant in the chain. One of the difficulties in using sequentially shared social information is that as more and more informants produce information, it can make the testimony of future informants less informative. Later informants will be heavily influenced by the increasing amount of social testimony they receive, and comparatively less influenced by the private information they observe. If learners are sensitive to the way in which informants make their decisions, they should then take the social information each previous informant has provided into account when evaluating the testimony of the last informant.

We can see this occur as a consequence of Bayes' rule, where the likelihood of the testimony is,

$$p(t_1,\ldots,t_n|h) = p(t_1|h)\prod_{i=2}^{n} p(t_i|t_1,\ldots,t_{i-1},h). \quad (4)$$

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.

Shared private data. If the informants all provided testimony based on a single shared piece of data, instead of each observing their own private data, then the probability of this testimony can be found by marginalizing over this shared private data. Denoting the shared data d', we have

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

where the probabilities $p(t_i|d',h)$ are calculated as before. Here we assume that $p(t_i|t_1,...,t_{i-1},d',h) = p(t_i|d',h)$. Since the only source of information that each informant has is d', once that is observed, previous testimony provides no new information, and so even though informants may hear each other give testimony, this will not influence the testimony they give.

Reasoning about balls and urns

The consequences of different forms of dependency for social learning can be hard to imagine in abstract, so we will work through a concrete example 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), which 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. The experimenter then asks the participant to decide which urn was used to fill the bag.

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 likely received red balls and three red balls outweigh the participant's single blue ball.

More technically, in this case, $p(t|h)$ is given by the probability that the informant gives testimony t depending on which urn was used to fill the bag. If the informants maximizes, the probability that the informant supports the correct urn is $p = \frac{5}{6}$, where p is the proportion of balls in the urn. If the informants probability match, this probability is instead $p = \frac{5}{6}^2 + \frac{1}{6}^2$, where the first term accounts for the likelihood they received a majority colored ball from the urn, and say the color of the ball, and the second term accounts for the likelihood they received a minority colored ball and say the opposite color. The probability that the learner receives a ball consistent with the majority ball color is $q = \frac{5}{6}$. If the learner observes three informants who say the same thing, and a ball that conflicts with that testimony, Equation 3 simplifies down to:

$$p(h|t) = \frac{p^3(1-q)}{(p^3(1-q)+(1-p)^3 q}, \tag{7}$$

where h is the hypothesis that the majority is correct.

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. In this case, the model takes into account that the third individual may not have received a red ball, but instead after hearing the previous two people support the red urn, she may disregard her own private evidence and vote in favor of the majority. This possibility makes the model predict individuals will also go with the majority, but less often than in the independent condition.

We can see this occur in the model most easily in the probability matching condition. The behavior of the first informant is captured by Equation 7. In the scenario we present above, the learner observes the second informant agree with the first informant. If the second informant receives a ball that conflicts with the testimony the first informant says, they will support the first informant with probability:

$$p(h|t) = \frac{p(1-q)}{(p(1-q)+(1-p)q},\tag{8}$$

or if they receive a ball that agrees with the testimony of the first informant, they will agree with the first informant with probability

$$p(h|t) = \frac{pq}{(pq+(1-p)(1-q)},\tag{9}$$

where h again is the hypothesis that the majority is correct.

Since the learner does not know which color ball the informant received, they must marginalize over possible ball colors, given a state of the world, and so

$$p(t_2|t_1,h) = (1-q)\frac{pq}{(pq+(1-p)(1-q)} + q\frac{p(1-q)}{(p(1-q)+(1-p)q}.\tag{10}$$

A similar set of equations can be derived for a third (or further) informant.

Shared private data. Consider what happens if all three informants are sitting at the same table and also all observe the same ball. The three informants probably received a single red ball, while the participant received a blue ball, providing equal evidence for either urn being used to fill the bag. In this case, the model is evenly split between the two urns, providing the least support for the majority compared to the other two conditions.

Because the informant see the same ball (which is the only information they receive to make their decision), they can safely ignore each other's testimony. In the maximizing condition, the probability that the informants all say the same as the color is 1, and the probability their testimony is correct is p. In the probability matching condition, the probability they all say the same thing is p^n, and the probability they are correct is

$$p(t|h) = p(p^n) + (1-p)(1-p)^n\tag{11}$$

where the first term stands for the probability that they get a ball that is of a majority ball color, and all say the color of the ball, and the second term is if they get a minority ball and all say the opposite ball color.

To compare our model predictions with human behavior, we next present a series of experiments to see how people incorporate their own understanding of the evidence each informant used to generate 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 cases of independent testimony, sequential testimony, and shared private data.

Methods

A total of 123 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.

Procedure. The experiment was a web-administered survey involving only text and pictures. First a woman named Jane (the *experimenter*) introduced an opaque red urn and blue urn. The introduction was given as text beneath a cartoon image. 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, 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 friends. The friends then told the participant which urn they thought 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. The actual colors were randomized, so half the participants received testimony favoring the blue urn and then saw a red ball.

In the independent testimony condition the participant was shown three wooden doors, and text below the image told the participant that one informant was waiting in each room. In the room, each informant sat behind a desk and said, e.g. “I looked at my ball and I think the bag was filled from the red Urn. I have not talked to Mary or Ann nor did I see their balls.” In the sequential testimony condition all the informants sat together behind a single 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 their decision, and that they agreed with the previous informants’ testimony (“I looked at the ball that Jane gave me, and I thought about what Sue said. I agree with Sue. I think the bag was filled from the red urn”). 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 (“I looked at the ball that Jane gave me and I think that the bag was filled from the red urn. I saw the same ball as Mary.”). In all conditions, 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 from the red urn or the blue 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

Ratings were placed on a consistent scale, corresponding to agreement with the majority. The mean ratings for all conditions are shown in Figure 1(c). The ordering of the means of each condition is consistent with the model predictions. Both the maximizing and the probability matching model provided a very good model fit to the data, Pearson’s $r = .83$ and $r = .91$ respectively.

We found a significant effect of condition on participants’ rating of the majority, one-factor ANOVA $F(2,120) = 7.749$, $MSE = 49.56$, $p < 0.001$. We explored the differences between conditions using pre-planned t-tests. Participants supported the majority significantly more in the independent condition, two-sample t-test, $t(80) = 3.88$, $d = 0.8$, $p < 0.01$, and the sequential testimony condition compared to the shared private data condition, two-sample t-test, $t(84) = 2.66$, $d = 0.55$, $p < 0.01$. The difference between the sequential testimony and independent testimony conditions was not significant, two-sample t-test, $t(76) = 0.96$, $d = 0.22$, $p = 0.34$.

Discussion

The difference between the shared private data condition and the independent and sequential conditions suggests that participants were sensitive to the informant's source of knowledge, using shared information to discount the majority's testimony based on the fact that all three informants shared asocial information – a single, shared ball. We also found that the ordering of the means of each condition was consistent with the model predictions.

At first glance, the null difference between the independent testimony and sequential testimony conditions suggests that people may not be sensitive to dependency due to shared social testimony, but only dependency due to shared asocial data. However the magnitude of the difference between these two conditions predicted by the model is relatively small. This suggests instead that the scenario presented in previous experiments on information cascades may not be sufficient to distinguish between how people use independent testimony over sequential testimony, a limitation we return to in Experiment 3.

First, we build on the results of Experiment 1 by examining statistical dependency in cases where individuals only hear testimony from a subset of informants, but where the informants may still hear testimony from each other. This is a case of shared social information with particular real-world relevance, since in many situations we learn from others without knowing exactly what other people told them.

Informants with unheard testimony

In Experiment 2, we examine three new conditions, where the learner learns from informants who whisper testimony to each other. We assume that the learner is placed in the ball and urn scenario used in Experiment 1, but instead of hearing from all three informants, the learner hears testimony from only the last informant. The two other informants whisper their testimony to the next informant in the chain. We assume that the participant receives a ball that is the opposite color from the urn supported by the informant's testimony.

The three cases differ in the quality of information that each informant receives and what they whisper. In the *color whispering* case, each informant receives their own ball and whispers the color of that ball (and the color of all the balls they heard about). In the *sequential whispering* case, each informant receives their own ball and whispers which jar they thought was used to fill the bag, similar to the sequential testimony condition in Experiment 1. In the *shared private data whispering* case, each informant receives access to a single, shared ball, and whispers which jar they thought was used to fill the bag, similar to the shared private data condition in Experiment 1.

These conditions provide an additional route to understand how participants take the information an informant has into account when the informant generates testimony. In the case the sequential testimony condition in Experiment 1, participants must go beyond assuming that each informant provides an equal amount of information and must instead evaluate the quality of the information each informant individually provides. The color whispering and sequential whispering conditions provide two additional examples where participants must take into account not only what data each informant saw, but how each informant combined the data they observed with the testimony they received. This is particularly interesting in the sequential whispering condition and the color whispering conditions where even though the participant receives the same testimony from the final informant, the model predicts that the participant should treat that testimony differently based on what type of information was passed at each step of the chain.

To incorporate whispered testimony into the rational model, we account for the possibility of unheard testimony by marginalizing over possible options for testimony. Suppose an informant produces testimony t_n based on unknown testimony $t_1, ..., t_{n-1}$. We can calculate the likelihood of the heard testimony by marginalizing over the unheard testimony,

$$p(t_n|h) = \sum_{t_1,\ldots,t_{n-1}} p(t_n|t_1,\ldots,t_{n-1},h)p(t_{n-1}|t_1,\ldots,t_{n-2},h)...p(t_1|h). \quad (12)$$

Just as before the values for $p(t_{n-1}|t_1,\ldots,t_{n-2},h)$ can be calculated recursively.

Model Predictions

Model predictions for how likely participants are to go with either the majority testimony (in the non whispering conditions) or single piece of testimony (in the whispering conditions) in the three cases used in Experiment 1, *independent, sequential* and *shared data* and these three new cases are presented in Figure 2(a) and (b). In these predictions we assume that the proportion of red balls in the red urn is $\frac{9}{10}$ and is reversed in the blue urn. The increased proportion of balls in each urn also has the property of increasing the predicted difference between the sequential testimony and independent conditions.

We find that in the case of color whispering the information that the last informant receives is equivalent to receiving three balls (unknown to the learner), leading the model to predict that the learner will go with the informant's choice over their own ball. Compared to the previous non-whispered conditions, the model predicts that color whispering provides less information than the independent condition, but about the same information as the sequential testimony condition.

In the case of sequential whispering, the model predicts that the additional social information the final informant receives will not substantially alter their testimony. The key reason for this is that in this condition informants only pass along which urn they thought was used to fill the bag, and do not mention the color of their ball (as in color whispering). If an informant's private information disagrees with the whispered social information, the informant has no opportunity to express their uncertainty. They will either whisper that they thought the bag was filled from the urn that was the same color as the ball they received, or the color that social information suggested. Both options lead to roughly equivalent accuracy (about a single ball's worth of information). This leads the model to predict that information will not accumulate across the chain, and that the final informant provides about the same amount of information as seeing a single ball drawn from the urn.

In the case of shared ball whispering, the social information that the final informant receives provides her no new information; the final informant has access to all the information previous informants had to make their decision, leading the model to predict that the learner will be split between their own private data and the testimony.

Experiment 2: Whispered Testimony

Experiment 2 examined how individuals responded to the three new cases of whispered testimony. To insure that individuals understood the proportion of balls in each urn, we changed the stimuli from being opaque urns to translucent jars and increased the proportion of balls in each urn to 90% red and 10% blue (or the reverse). Because of this change we also replicated the three conditions in Experiment 1.

Methods

A total of 450 participants were recruited through Amazon Mechanical Turk. Participants were compensated \$0.25 for their time. They were randomly assigned to one of six experimental groups, independent ($n = 67$), sequential testimony ($n = 65$), shared private data ($n = 64$), color whispering ($n = 69$), sequential whispering ($n = 69$), and shared private data whispering ($n = 67$).

At the end of the experiments participants took an attention check (question: "How many green balls were in the experiment", answer: 0). Participants who failed the memory check (n=49) were dropped from the analyses and were not included in the counts above.

Procedure. The stimuli were identical to those in Experiment 1, except that instead of opaque urns, we used clear jars filled with a mix of red and blue balls. There were 18 red balls and 2 blue balls in “Jar A”. The proportions were reversed for “Jar B”.

The procedure was the same as in 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”. The whispering conditions were presented in the same manner as the sequential testimony condition in Experiment 1: informants spoke in order and said that they either whispered the color of their ball, and in the case of the second informant, the previous informant’s ball, i.e. “I whispered to Ann the color of my ball and the color of Sue’s ball.” (color whispering) or which jar they thought the bag was filled from, i.e.“I whispered to Ann which jar I thought the bag was filled from.” (sequential whispering or shared ball whispering). The text, “*whispers*” also appeared above their heads. Only the final informant in the chain gave testimony i.e. "I looked at my ball, and I thought about what Mary told me. I think the bag was filled from Jar A.". Responses were made on the same 11-point scale as in Experiment 1, changing the names of the urns appropriately.

Results

Ratings were re-scaled as in Experiment 1. The mean re-scaled ratings are shown in Figure 2(c). We analyzed the effect of condition on participant responses using an ANOVA. The effect of condition was significant, one-factor ANOVA $F(5,395) = 20.13$, $MSE = 187.74$, $p < .01$. We explored the differences between conditions using planned two-sample t-tests.

In the three replicated conditions of Experiment 1, *independent, sequential* and *shared private data*, we found a similar pattern of results as in Experiment 1. Individuals sided significantly less with the informants' testimony in the shared private data condition than either the independent, two-sample t-test $t(129) = 4.35$, $d = 0.71$, $p < 0.01$ or sequential testimony conditions, two-sample t-test $t(127) = 5.47$, $d = 0.87$, $p < 0.01$. We found no significant difference between the independent and the sequential testimony conditions, two-sample t-test $t(130) = 0.97$, $d = 0.17$, $p = 0.33$.

In the three whispered conditions, we found that, as predicted, participants sided with the final spoken piece of testimony significantly more often in the color whispering condition than the shared private data whispering condition, two-sample t-test $t(134) = 6.88$, $d = 1.02$, $p < 0.01$, and in the sequential whispering than the shared private data whispering condition, two-sample t-test $t(134) = 5.23$, $d = 0.82$, $p < 0.01$, but found no significant difference between the color whispering and sequential whispering condition, two-sample t-test $t(136) = 1.47$, $d = 0.25$, $p = 0.14$.

Comparing between whispered and non whispered conditions, we found that participants sided with the testimony more in the color whispering two-sample t-test $t(131) = 3.12$, $d = 0.53$, $p < 0.01$ and marginally more in the sequential whispering condition, two-sample t-test

$t(131) = 1.72$, $d = 0.3$, $p = 0.09$ than the shared private data (non-whispering) condition. Individuals sided with the informant's testimony least in the shared ball whispering condition, and significantly less than they did in the shared private data (non-whispering) condition, two-sample t-test $t(129) = 3.06$, $d = 0.52$, $p < 0.01$. Finally, contrary to our model predictions, participants sided with the testimony more often in the sequential testimony condition than the color whispering condition, two-sample t-test $t(132) = 2.64$, $d = 0.45$, $p < 0.01$.

Overall we find that participants are sensitive to the total amount of information the informants received (one ball or three balls), and the number of informants the participants heard from (one or three). We confirmed this finding with a 2 way ANOVA, finding that both the number of balls $F(1,398) = 73.13$, $MSE = 682.5$, $p < .01$, and the number of informants $F(1,398) = 24.16$, $MSE = 225.5$, $p < .01$ were significant predictors.

When comparing the model predictions to observed data, we find some agreement with the maximizing model, Pearson's $r = .63$, but better agreement with the matching model, Pearson's $r = .81$.

Discussion

The results of this experiment suggest that people are appropriately sensitive to the number of informants they hear information from (e.g., comparing shared private data versus shared private data whispering), and the amount of information the group of informants has collectively received (e.g., color or sequential whispering versus shared private data whispering). We find overall agreement with the model predictions, and find that the matching model provides closer support to the data than the maximizing model. We find one small deviation from the model predictions – the model predicts that the mean rating of the color whispering and sequential testimony conditions should be similar, whereas here we find a significant difference between them. We return to this point in the general discussion.

As in Experiment 1, we find no significant differences between participants' inferences in the sequential testimony and independent conditions. However, the difference predicted between these conditions remained small. To investigate the question of whether people are sensitive to statistical dependency due to shared social information more closely, we analyze a case where this difference is predicted to be much larger.

Experiment 3: Dissenting Informant

In order to assess whether people are sensitive to shared social information in the case of sequential testimony, we modified the scenario presented in Experiment 1 to increase the predicted difference between the independent and sequential testimony conditions. We changed the informant testimony 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. These changes were necessary to create a situation where the participant would be expected to go against the majority in the sequential testimony condition, but not in the independent condition.

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, to provide a reason why the final informant might dissent in the shared private data condition, only the first two informants received the same ball and the dissenter received her own ball.

Model Predictions

The model predictions are given in Figure 3(a), for maximizing, and Figure 3(b), for probability matching. The addition of a low-probability diagnostic ball does not substantially change the model predictions in the independent or shared private data conditions. However, it makes an important change to the sequential testimony condition, where the probability that a learner will go with the majority is substantially less in both the maximizing and matching models. The difference is largest in the maximizing model: under this model, 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 learner should side with her against the majority. While less dramatic, the matching model also predicts an increased probability of going against the majority relative to the previous experiments.

Methods

A total of 124 participants were recruited through Amazon Mechanical Turk. Participants were compensated $0.25 for their time. They were randomly assigned to one of three experimental groups: the independent ($n = 41$), sequential testimony ($n = 41$), or shared private data conditions ($n = 42$). No participants were dropped from the analysis.

Procedure. The stimuli were identical to those in Experiment 2, except that a single diagnostic ball (either white or black) was added to each urn.

The procedure was the same as the non-whispered conditions of Experiment 2, except for the following changes. 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, only the first two informants received the same ball, the last informant received a different ball. 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

Ratings were rescaled as in Experiment 1 and 2. The mean rescaled ratings are shown in Figure 3(c). The maximizing and probability matching models both provide a good fit for the experimental data, Pearson's $r = .83$, and $r = .94$ respectively.

We analyzed the effect of condition on participant responses using a one-factor ANOVA. The effect of condition was significant, $F(2,121) = 5.56$, $MSE = 27.13$, $p < 0.01$. We explored the differences between the conditions using planned t-tests. Participants sided with the majority significantly more often in the independent testimony than the sequential testimony condition, two-sample t-test $t(80) = 3.12$, $d = 0.65$, $p < 0.01$, and than the shared private data condition, two-sample t-test $t(81) = 3.16$, $d = 0.66$, $p < 0.01$. The difference between the sequential testimony and shared private data conditions was not significant, two-sample t-test $t(81) = 0.22$, $d = 0.05$, $p = 0.83$.

Discussion

The difference between the independent testimony and sequential testimony conditions suggests that the learning mechanism participants use is sensitive to statistical dependencies between informants that are a result of shared social testimony, confirming the non significant trend seen in Experiment 1 with a stronger manipulation. Likely a key reason we only see this difference arise in Experiment 3, and not in Experiments 1 or 2, is that the magnitude of the expected difference between independent and sequential testimony conditions is much larger in Experiment 3, leading to greater statistical power even with the same number of participants. The difference between the shared private data condition and the independent testimony condition supports our conclusion from Experiments 1 and 2 that people are sensitive to shared non-social information.

As in Experiments 1 and 2 the model provides a very good fit to the experimental data across conditions, and the probability matching model provides a better fit than the maximizing model – particularly in the sequential testimony condition.

General Discussion

In this paper we examined how individuals evaluated social testimony from multiple informants who shared information. Experiment 1 showed that people are sensitive to shared private data amongst informants, using a task that has been employed in previous experiments on information cascades. Experiment 2 expanded on the results of Experiment 1 and found that participants were sensitive to the total amount of asocial information the informants received, as well as the number of informants participants heard from. Experiment 3 showed that people are also sensitive to sequential testimony, where informants have learned from each other and share social information, using a task that is more sensitive to this kind of dependency. In all three cases, a rational learner model provided a good fit to participant's responses.

In these experiments, we find that individuals are sensitive to the number of people they learn from, and the amount of information the group as a whole provides. The first finding supports previous accounts of majority copying behavior in humans, while the second finding suggests a more nuanced version of how this copying might operate. People in these experiments were not blindly copying the majority, as assumed in some models of cultural evolution (e.g Henrich & Boyd, 1998), but instead were copying the majority based on the amount of information the majority provides, evidence for a sophisticated social learning mechanism.

Furthermore, we find that throughout these experiments, participants' behaviors were well captured by a rational model of social learning. This model was able to capture shared information as a form of statistical dependency, and well accounted for how individuals handled different sources of information, including cases where they heard from a single or multiple informants who shared information. We examined two such models, one where individuals assume that informants are maximizing their beliefs to produce testimony, and a second where they assume individuals probability match. We find that the probability matching model provides better qualitative and quantitative fit to humans' behavior across all three experiments. This finding is significant given that many previous models of rational learning, particularly information cascades, have assumed that both participants and informants maximize (Bikhchandani et al., 1992; Anderson & Holt, 1997), which may change the situations in which information cascades are likely to occur.

These experiments then shed new light on traditional notions of conformity biases (Asch, 1956; Boyd & Richerson, 1985). In many cases, as illustrated by our model, it is rational to follow the majority. We find that across experiments, participants will follow the majority if they represent a greater source of knowledge (particularly in the sequential testimony or independent conditions). These findings suggest that some of the cases of conformity biased copying (as seen in, e.g. Asch, 1956; Haun & Tomasello, 2011), may actually be the product of a rational learning process where the learner believes that the informants have access to more information than the informants actually have, and so the learner disregards their own private evidence to follow the majority. However, individuals are following a more complex strategy than just copying the majority. We find that when the information that the majority provides is equal to the amount of information provided by the participants' own social information, participants go equally between the two options, or default to their own asocial information if their private data presents a stronger pool of knowledge. This provides an additional explanation for why individuals follow the majority that is likely not mutually exclusive with other well known factors (like desire to be part of a group) that influence a person to conform to a majority.

We do find some slight deviations from the rational model, where in Experiment 2 the model predicts that the color whispering and sequential testimony conditions should provide the learner with similar amounts of information, but we find that participants place more weight on the testimony in the sequential testimony condition than in the color whispering condition. This deviation may suggest that individuals have a bias to conform to a majority above and beyond what is rational: in the color whispering condition the participant only hears from the final informant, whereas in the sequential testimony condition the participant hears from all three informants. However, an alternative explanation may be that individuals place substantially more weight on the sequential testimony condition than predicted by the rational model due to a lower sensitivity to sequential testimony, rather than a bias towards following the majority. This lower sensitivity might reflect a bias to assume individuals are more independent than they actually are, which could be rational if, for example, informants brought outside knowledge to the task.

The findings of these experiments also give insights into the processes that have generated human culture. We find that participant's are sensitive to shared information between informants, a subtle cue to use for evaluating testimony, and do not just rely on the number of informants in a group, an assumption widely used in evolutionary models of social learning (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Wakano & Aoki, 2007). Our findings also impact our understanding of how cultural transmission may operate. By being sensitive to shared information individuals are able to more accurately assess when a majority provides accurate information decreasing the likelihood that they will follow an incorrect majority an issue highlighted by computational models e.g., Henrich and Boyd (1998), and perhaps decreasing the rate at which information cascades occur (Bikhchandani et al., 1992). Both of these effects would lead to social information being adaptive in a wider range of situations.

Taken together, our findings suggest that humans use a complex social learning mechanism that is sensitive to a wide number of cues, including subtle distinctions in how informants make decisions. We also find that individuals are able to easily integrate their own private information with informants' testimony, and that this integration is consistent with a rational model of individual learning. These results suggest that individuals are not just copying others, but relying on their understanding of where informants got their information from to make decisions.

References

Anderson, L., & Holt, C. (1997). Information cascades in the laboratory. *The American economic review*, *87*(5), 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. (1985). *Culture and the evolutionary process.* University of Chicago Press.

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

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

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, *29*(1), 56–64.

Enquist, M., Eriksson, K., & Ghirlanda, S. (2007). Critical social learning: a solution to rogers's paradox of nonadaptive culture. *American Anthropologist*, *109*(4), 727–734.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Cambridge, MA: Belknap Press/Harvard University Press.

Haun, D., & Tomasello, M. (2011). Conformity to peer pressure in preschool children. *Child Development*, *82*(6), 1759–1767.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and human behavior*, *19*(4), 215–241.

Hoppitt, W., & Laland, K. (2013). *Social learning: An introduction to mechanisms, methods, and* *models*. Princeton University Press.

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

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, *15*(3), 215–233.

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental psychology*, *49*(3), 404.

Morgan, T., Rendell, L., Ehn, M., Hoppitt, W., & Laland, K. (2011). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1729), 653–662.

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

Pike, T. W., & Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biology letters*, *6*(4), 466–468.

Rendell, L., Boyd, R., Enquist, M., Feldman, M. W., Fogarty, L., & Laland, K. N. (2011). How copying affects the amount, evenness and persistence of cultural knowledge: insights from the social learning strategies tournament. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1118–1128.

Rendell, L., Fogarty, L., & Laland, K. N. (2010). Roger's padadox recast and resolved: Population structure and the evolution of social learning strategies. *Evolution*, *64*(2), 534–548.

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.

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

Wakano, J. Y., & Aoki, K. (2007). Do social learning and conformist bias coevolve? Henrich and Boyd revisited. *Theoretical Population Biology*, *72*(4), 504–512.

Whalen, A., Buchsbaum, D., & Griffiths, T. L. (2013). How do you know that? sensitivity to statistical dependency in social learning. In *Proceedings of the 35th annual conference of the cognitive science society.*

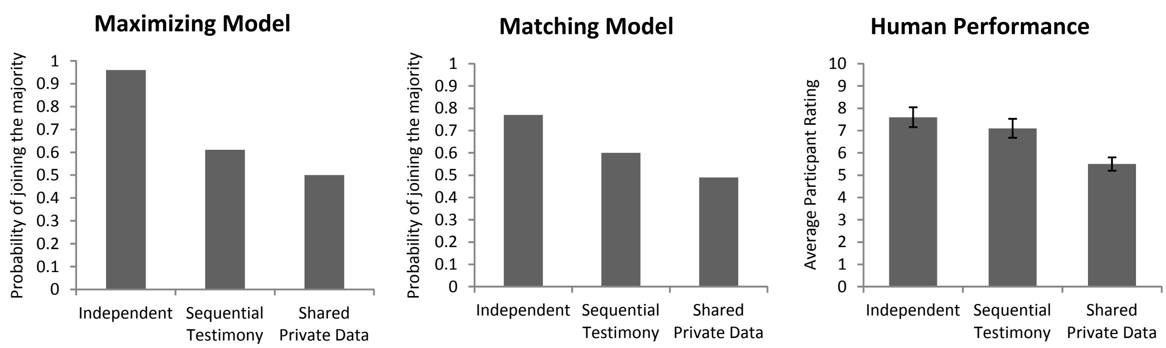

Figure 1. Results from Experiment 1 for the (a) maximizing model, (b) matching model, and c) human performance. Error bars represent ±1 SE.

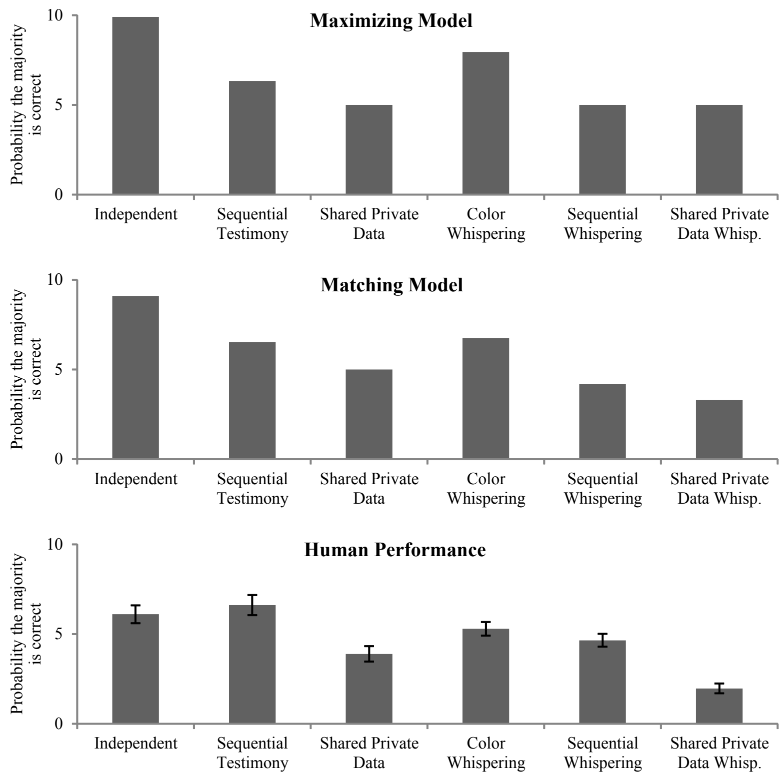

Figure 2. Results from Experiment 2 for the (a) maximizing model, (b) matching model, and c) human performance. Error bars represent ±1 SE.

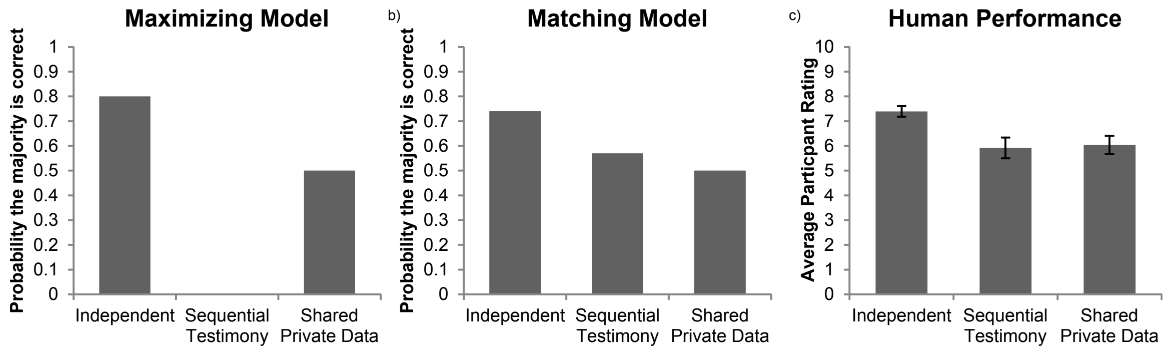

Figure 3. Results from Experiment 3 for the (a) maximizing model, (b) matching model, and c) human performance. Error bars represent ±1 SE. In the case of sequential testimony with the maximizing model, the model always predicts that individuals will go with the minority and so the bar is set at 0.