

**Overcoming individual limitations through distributed computation:  
Rational information accumulation in multi-generational populations**

Mathew D. Hardy<sup>1,\*</sup>, P. M. Krafft<sup>2,\*</sup>, Bill Thompson<sup>1,3</sup>, Thomas L. Griffiths<sup>1,3</sup>

<sup>1</sup>Department of Psychology, Princeton University

<sup>2</sup>Creative Computing Institute, University of the Arts London

<sup>3</sup>Department of Computer Science, Princeton University

\*These authors contributed equally

## Abstract

Many of the computational problems people face are difficult to solve under the limited time and cognitive resources available to them. Overcoming these limitations through social interaction is one of the most distinctive features of human intelligence. In this paper, we show that information accumulation in multi-generational social networks can be produced by a form of distributed Bayesian inference that allows individuals to benefit from the experience of previous generations while expending little cognitive effort. In doing so, we provide a criterion for assessing the rationality of a population that extends traditional analyses of the rationality of individuals. We tested the predictions of this analysis in two highly controlled behavioral experiments where the social transmission structure closely matched the assumptions of our model. Participants made decisions on simple categorization tasks that relied on and contributed to accumulated knowledge. Success required these micro-societies to accumulate information distributed across people and time. Our findings illustrate how in certain settings, distributed computation at the group level can pool information and resources, allowing limited individuals to perform effectively on complex tasks.

**Overcoming individual limitations through distributed computation:  
Rational information accumulation in multi-generational populations**

Humans' success and survival depends on our ability to make intelligent choices and judgments. Like other animals, we must do so under hard constraints on the resources available to make our decisions (Griffiths, 2020). Not only is our cognition constrained by our restricted brain capacity, but we also must learn from limited and often incomplete data. This is true not only in our finite childhoods and lifespans, but also in the limited opportunities we have to learn many crucial facts—to survive, we need to learn about what life-threatening dangers exist in our immediate environments (from chemicals to electric outlets to poisonous insects and animals).

Given these constraints, understanding how people routinely form accurate beliefs on complex topics is a central focus of research in cognitive science, psychology, and economics (Simon, 1990). However, people rarely develop beliefs alone—rather, they learn from the knowledge, experiences, and opinions of other people. By doing so, individuals can obtain useful information while expending little physical or cognitive effort. A particularly important focus of previous research has been on the effects of social learning when repeated over successive groups of individuals, such as child-parent learning, formal education, and other domains where knowledge is transmitted from older members of a population to younger learners. In these multi-generational settings, knowledge can accumulate over time in a population, allowing individuals to extend their cognitive skills by learning from others (see e.g. Almaatouq, Alsobay, Yin, & Watts, 2021; Almaatouq et al., 2020; Belikov, Rzhetsky, & Evans, 2020; Caldwell, Atkinson, & Renner, 2016; Frey & Goldstone, 2018; Galesic, Olsson, Dalege, van der Does, & Stein, 2021; Goldstone, Wisdom, Roberts, & Frey, 2013; Haçla, Jadbabaie, Mossel, & Rahimian, 2021; Kempe & Mesoudi, 2014a; Mesoudi, 2016; Mesoudi & Thornton, 2018; Miton & Charbonneau, 2018; Riedl, Kim, Gupta, Malone, & Woolley, 2021; Rzhetsky, Foster, Foster, & Evans, 2015; Salhab, Ajorlou, & Jadbabaie, 2020; Wisdom, Song, & Goldstone, 2013; Wojtowicz & DeDeo, 2020, for recent overviews and related studies). Crucially, the constraints and structure of interpersonal transmission

often lead collective knowledge and learning to differ from individual outcomes (Kirby, Tamariz, Cornish, & Smith, 2015; Ravignani, Thompson, Grossi, Delgado, & Kirby, 2018; Silvey, Kirby, & Smith, 2019).

The accumulation of collective knowledge through sequential social learning is known as *cultural transmission* (Boyd & Richerson, 1985), and is thought to underpin cumulative cultural evolution (Mesoudi, 2011). Cultural transmission has been studied experimentally in a number of paradigms, including the evolution of simple technologies such as knots (Muthukrishna, Shulman, Vasilescu, & Henrich, 2014), virtual fishing nets (Derex, Beugin, Godelle, & Raymond, 2013), stone tools (Morgan et al., 2015), or arrowheads (Mesoudi & O'Brien, 2008); artificial languages (Kirby, Cornish, & Smith, 2008); jigsaw puzzles (Kempe & Mesoudi, 2014b); and social phenomena like stereotypes (Martin et al., 2014).

Theories of cultural evolution have primarily been underpinned by a Darwinian framework grounded in parallels and disanalogies with biological evolution (Boyd & Richerson, 1985; Mesoudi & Whiten, 2008; Nettle, 2020; Smolla et al., 2020). These frameworks have shown that the accumulation of knowledge and technology can be understood as a form of evolution, helping to situate cultural evolution within the biological and evolutionary sciences (Laland, Sterelny, Odling-Smee, Hoppitt, & Uller, 2011). While successful, evolutionary frameworks have been difficult to connect to psychological theory (Heyes, 2018), and are therefore difficult to connect with the concepts of computation that are central to the study of cognition and intelligence (Nettle, 2020). Similarly, while research on collective intelligence has extensively documented the advantages of groups over individuals, the computational structure of the social processes that lead to long-term, open-ended collective intelligence remain unclear (Krafft et al., 2016).

One way to measure the accumulation of knowledge in populations is by viewing collective behavior as distributed computation, a process that allows groups to “store and process the cumulative innovations and collaborations of generations of individuals”

(Smaldino & Richerson, 2013). From this perspective, knowledge accumulation becomes a problem of *distributed* Bayesian inference, extending probabilistic models of inference in individuals to the group setting (Chater, Tenenbaum, & Yuille, 2006; Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; Harper, 2009). This perspective establishes a formal connection between cultural evolution and statistical models of social learning used in cognitive science (Cushman & Gershman, 2019) and economics (Acemoglu, Dahleh, Lobel, & Ozdaglar, 2011).

In this paper, we build on this work and show how social learning can facilitate rational action that goes beyond the direct experience of individuals. That is, we show how social interactions allow limited individuals to improve their cognition without modifying their time and resource constraints. We do this in part by offering a formal criterion for *population rationality*: the probability that any individual in a network makes a particular decision is the same as the probability of that decision under a Bayesian posterior distribution conditioned on *all* the information observed by the population (Foster, 2018, cf.).

To develop our account, we first formulate a model of individual social learning based on a simple heuristic that requires only limited social observation. We then show that under certain conditions, groups of individuals following this heuristic will accumulate information through distributed Bayesian inference. This model offers insight into how social learning can extend the cognitive abilities of limited individuals—while conditioning on all accumulated information would be too complex for any individual, social interactions allow individuals to benefit from this information while expending little cognitive effort. That is, population rationality can be achieved by networks of highly limited individuals.

We tested our model in two large-scale experiments where participants made basic categorization decisions in simple multi-generational networks. In both experiments, the decisions of participants at one generation were propagated to those in the next generation, allowing us to study belief accumulation and transmission in a controlled

laboratory setting. We found that a substantial proportion of the data from these experiments are well-approximated by our model. Furthermore, by comparing participants’ choices with the Bayesian posterior distribution conditioned on the information observed by the entire population, we were able to quantify accumulation in our experimental networks relative to the Bayesian ideal. Our findings thus offer a clear demonstration of how limited individuals can use simple social learning heuristics to make intelligent inferences.

### Information accumulation through social sampling

Here, we describe a simple social learning heuristic we call *social sampling* that yields distributed Bayesian inference at the population level. That is, populations of individuals following social sampling will accumulate information in a way that is consistent with Bayesian inference, allowing individuals to offload computation to the group and make accurate inferences with little cognitive effort. We focus on multi-generational settings, where individuals are organized in discrete “batches” and learn from those in the previous generation. We also assume individuals observe others’ true beliefs, and do not need to discard or modify any observations.

### Model

Information accumulation by rational agents is specified by Bayes’ rule, which indicates how a probability distribution over hypotheses  $\theta$  (known as the *prior* distribution) should be updated (to the *posterior* distribution) in light of evidence  $D$ :

$$p(\theta|D) = \frac{p(D|\theta) \cdot p(\theta)}{p(D)}. \quad (1)$$

In this paper, we extend this characterization of optimal belief updating from individuals to groups (see Chater, Oaksford, Hahn, & Heit, 2010; Griffiths, Kemp, & Tenenbaum, 2008; Tenenbaum, Kemp, Griffiths, & Goodman, 2011, for reviews of Bayesian models of individual cognition). Viewing a population as a single agent is a perspective with roots in many traditions, including philosophy (Easwaran, 2019), economics (Gale & Kariv, 2003), economics (Hayek, 1945), organization science

(Argote, 2012), cognitive science (Krafft, Shmueli, Griffiths, Tenenbaum, & Pentland, 2020), anthropology (Hutchins, 1995), collective intelligence (Engel et al., 2015; Woolley, Chabris, Pentland, Hashmi, & Malone, 2010), ethology (Sasaki & Pratt, 2011), and computer science (Lynch, 1996; Shoham, Leyton-Brown, et al., 2009). Crucially, as we will show, rational action at the population level requires only limited computation by individuals—while each individual agent may follow a simple heuristic, the information accumulated by the population as a whole makes it possible for those individuals to act rationally.

We consider the simple case where individuals reason about a set of that can be either true or false, represented as a binary feature vector  $\mathbf{x}$ . Individuals beliefs are shaped by pieces of evidence  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_T)$  that arrive over time. At each time  $t$ , individuals observe  $J$  pieces of information about feature  $i$ , with

$P_i = P(y_{ijt} = 1 | x_i = 0) = \theta_{i0}$  if  $x_i = 0$  and  $P_i = P(y_{ijt} = 1 | x_i = 1) = \theta_{i1}$  if  $x_i = 1$ . For simplicity, we let  $S_{it} = \sum_{j=1}^J \mathbb{1}(y_{ijt} = 1)$  be the number of positive observations made about feature  $i$  at time  $t$ .

We assume that the popularity of a belief among members of the population acts as a prior distribution over beliefs for new individuals drawing inferences about the environment. Specifically, an individual  $a$  first chooses a member of the population  $a'$  to learn from uniformly at random.<sup>1</sup> The learner  $a$  then accepts or rejects their companion's decision  $d_{a',i,t-1}$  with probability proportional to

$(\theta_{i,d_{a',i,t-1}})^{S_{it}} (1 - \theta_{i,d_{a',i,t-1}})^{J-S_{it}}$ , the likelihood of the evidence that  $a$  observes at time  $t$  based on the beliefs of  $a'$ .

In an infinite population of individuals following this strategy, the probability that an individual makes a categorization decision about feature  $i$  is equal to the posterior probability of that decision conditioned on the evidence observed by the *entire*

---

<sup>1</sup> This assumption may seem implausible in large populations, but it is satisfied if each person chooses from a small number of others and those others are an unbiased sample from the full population.

population:

$$p_{i,t+1} = P(x_i = 1 \mid \mathbf{y}_{i,\cdot,\leq t}). \quad (2)$$

This result establishes that social sampling is a valid algorithm for distributed Bayesian inference in infinitely large populations, satisfying our criterion for population rationality (see SEM for proof). The finite population case is more complex, but social sampling can be seen to be formally equivalent to a class of sequential Monte Carlo algorithms known as particle filters (Crisan & Doucet, 2002; Murphy, 2012). Social sampling can thus support Bayesian inference in expectation in finite populations, allowing groups of bounded individuals to accumulate knowledge over time.

### Measuring accumulation relative to optimal inference

The model we have outlined illustrates how a simple social decision-making heuristic can lead to optimal information aggregation over generations. This suggests that in certain contexts the Bayesian posterior distribution can be used to construct a valid upper limit on information accumulation, because posterior distributions are information-theoretically and decision-theoretically optimal belief representations (Bernardo & Smith, 2000; Jaynes, 2003; Ortega, 2011). To quantify ideal information accumulation, we use the sufficient statistics of the Bayesian posterior distribution over the environment features given the evidence observed by the entire population. We assume a uniform prior, and so the posterior probability that a feature categorization is correct given the evidence received by the group up to generation  $t$  is

$$P(x_{it} = 1 \mid \mathbf{y}_{i,\cdot,\leq t}) = \frac{(\theta_{i1})^{S_{i,\leq t}} (1 - \theta_{i1})^{t \cdot J - S_{i,\leq t}}}{(\theta_{i1})^{S_{i,\leq t}} (1 - \theta_{i1})^{t \cdot J - S_{i,\leq t}} + (\theta_{i0})^{S_{i,\leq t}} (1 - \theta_{i0})^{t \cdot J - S_{i,\leq t}}}, \quad (3)$$

where  $S_{i,\leq t} = \sum_{k=1}^t \sum_{j=1}^J \mathbb{1}(y_{ijk} = 1)$  is the total positive evidence observed for feature  $i$  up to time  $t$ .

The sufficient statistics of this posterior distribution are  $\frac{S_{i,\leq t}}{t \cdot J}$ , or the proportion of pieces of evidence favoring a feature categorization. Because these statistics can be used to exactly compute the posterior distribution, we can measure accumulation in observed networks by comparing this statistic with the popularity of a belief, or proportion of



people who believe it. We express this formally by letting  $p_{it} = \frac{1}{N} \sum_a \mathbb{1}(d_{a,i,t-1} = 1)$  be the proportion of individuals who believe  $x_i = 1$  at time  $t - 1$ , with  $d_{ait}$  indicating whether individual  $a$  at generation  $t$  chooses category  $i$ , and  $N$  representing the number of individuals in each generation. If the popularity is close to  $\frac{S_{i,\leq t}}{t \cdot J}$  at time  $t$ , we can conclude that information is being effectively accumulated in the population.

## Summary

This analysis demonstrates that individuals in multi-generational networks can improve their inferences by following a simple social learning heuristic. To perform social sampling, an individual first randomly selects a person in the previous generation to learn from. They then evaluate their companion’s belief against observed evidence—the more aligned the belief with the evidence, the higher the chance the individual will accept it as their own, otherwise they will continue searching and sample another person. While similar social learning models have been explored in multi-armed bandit problems (Celis, Krafft, & Vishnoi, 2017; Krafft, 2017; Krafft et al., 2020), they have not been tested in controlled laboratory settings.

Social sampling is simple to perform and only requires limited computation by individuals. More precisely, at any time the popularity of a hypothesis will approximate the true Bayesian posterior probability conditioned on all the evidence observed by the population, and so the population itself performs distributed Bayesian inference. Because this distributed inference matches the inference problems individuals face, our model allows us to be precise about *how* social learning extends people’s cognitive abilities.

In our model, individuals make unbiased judgments and cannot choose which information to transmit. Instead, every person’s true beliefs can be observed by those in the next generation. While this setup—and assumptions of uniform sampling of others’ true beliefs—is highly simplified, it reflects a common approach to modeling individual cognition as a two-stage decision-making process (Howard & Sheth, 1969; Krumme, Cebrian, Pickard, & Pentland, 2012; Payne, 1976; Pratt, Sumpter, Mallon, & Franks,

2005; Seeley & Buhrman, 1999; Vul, Goodman, Griffiths, & Tenenbaum, 2014) that has also been applied to modeling iterated learning in populations (Bonawitz, Denison, Gopnik, & Griffiths, 2014; Kalish, Griffiths, & Lewandowsky, 2007; Mozer, Pashler, & Homaei, 2008; Sanborn & Griffiths, 2008; Vul & Pashler, 2008), and allows us to make progress in understanding knowledge accumulation in a way that can be studied empirically and extended to more naturalistic settings.

While our model shows how group-level Bayesian inference is possible, can we identify and quantify accumulation in real populations? To investigate, we ran two behavioral experiments where participants made categorization decisions in multi-generational networks. In both experiments, we assessed population rationality by comparing observed popularity with the Bayesian ideal, and evaluated our social sampling model by comparing it with several alternative models.

### **Experiment 1: Gem classification**

In Experiment 1, participants completed a categorization task that was framed as gemstone classification. This experiment was designed to closely mirror an idealized setting in which optimal information aggregation is possible through our proposed Bayesian social sampling mechanism.

In this experiment, participants played the role of technicians classifying gemstones in a certain shift (i.e., generation).<sup>2</sup> Each gemstone could have up to eight potential classifications, with each classification being a randomized nonsense word (e.g., “pesho”, “ivil”, “thyun”). On each trial, participants observed social information in the form of classification judgments drawn from the previous shift of technicians, as well as non-social information in the form of a set of four new laboratory results. The laboratory results presented stochastic positive or negative evidence for each potential gemstone classification. We recruited three independent networks of participants, with

---

<sup>2</sup> See <https://github.com/pkrafft/>

`Overcoming-individual-limitations-through-distributed-computation` for experiment code and data.

each network including 10 shifts of 20 technicians (i.e., 200 participants per network). The interface for the experiment is illustrated in Figure 1.

## Methods

**Participants.** We recruited 600 participants from Amazon Mechanical Turk. Recruitment was restricted to participants in the United States with an Amazon Mechanical Turk approval rating of 95 or above. We paid participants \$1.75 as compensation, plus a performance-based bonus payment of up to \$0.50. The task typically took participants under ten minutes to complete. Prior to starting the experiment, participants completed an attention and comprehension check which included questions about details of the study, including the probability of different kinds of evidence appearing in individual tests. Participants who failed to answer these questions correctly in three attempts were excluded from taking part in the study.

**Stimuli.** On each round of the experiment, participants observed a cartoon gemstone positioned at the top of the screen on a white background. Each gem had eight possible binary classifications, and participants performed lab tests to make these classifications. Each test resulted in either positive or negative evidence for each classification. Figure 1 shows an example participant view.

Classifications were given labels from an artificial vocabulary. These labels were different on each trial a participant completed (i.e. none of the labels recurred across trials for a participant), and were presented in randomized order. After each test, participants indicated which of the eight classifications they thought were true for this gem, and could modify these classifications throughout the trial.

Participants in generations 2 through 10 observed the final classification decisions of a randomly sampled participant from the previous generation of the same network. Participants could choose to resample from the previous generation at any time during the trial. When participants chose to resample, the new sample was selected at random with replacement from the twenty participants in the previous generation of the same network. Social information (an earlier participant's classification decisions) was

View of Participants in **Generations 1 and 2**

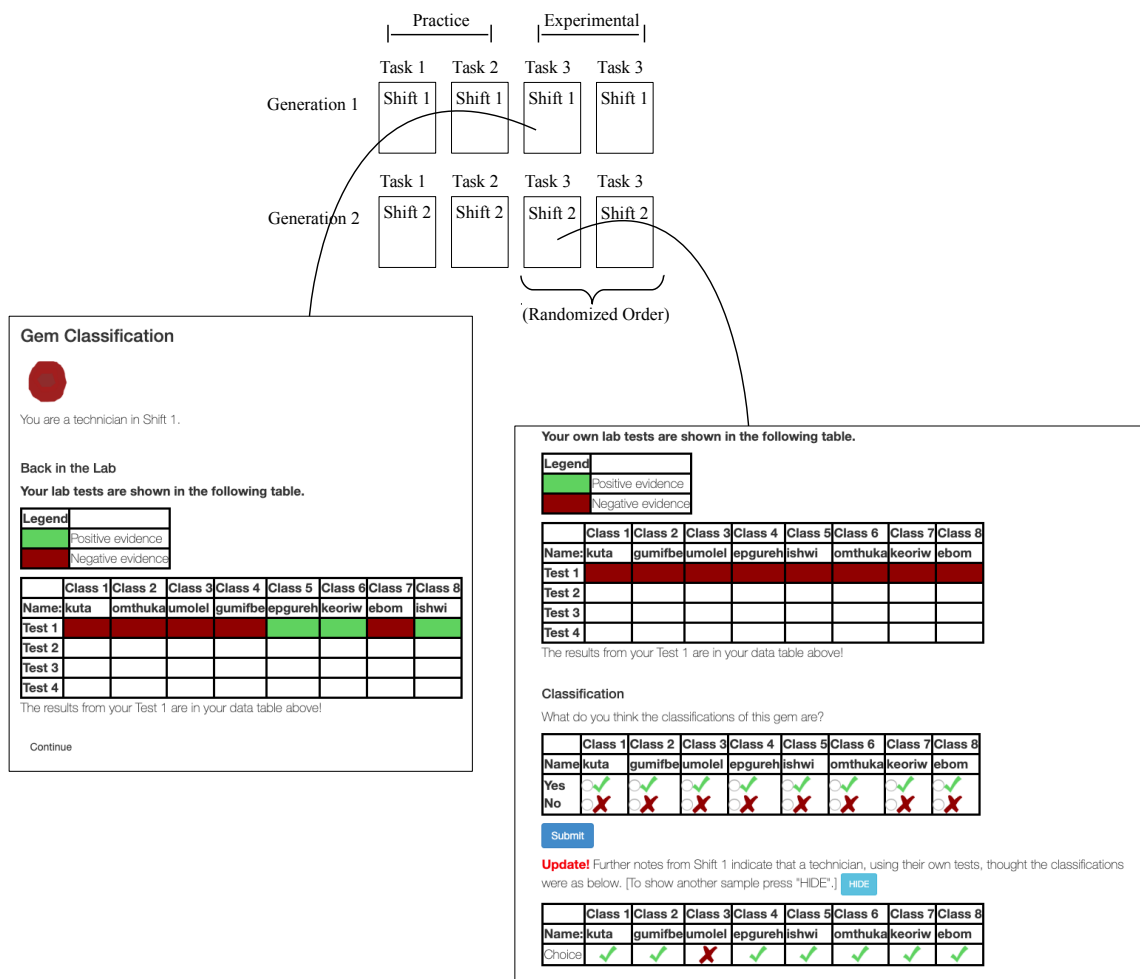


Figure 1. Participant interface for Experiment 1. Each categorization task consisted of a different set of gem classifications that were given random nonsense words. Each participant belonged to a “generation” of participants, called a “shift” in this experiment. In each generation after the first, participants could view the gem classifications made by participants in the last generation. Each participant made their own classification decisions for a gem four times after observing four pieces of non-social evidence.

presented in a feedback table below the participant’s own feedback table. After completing the fourth lab test, the results of the test were displayed and participants moved on to the next trial.

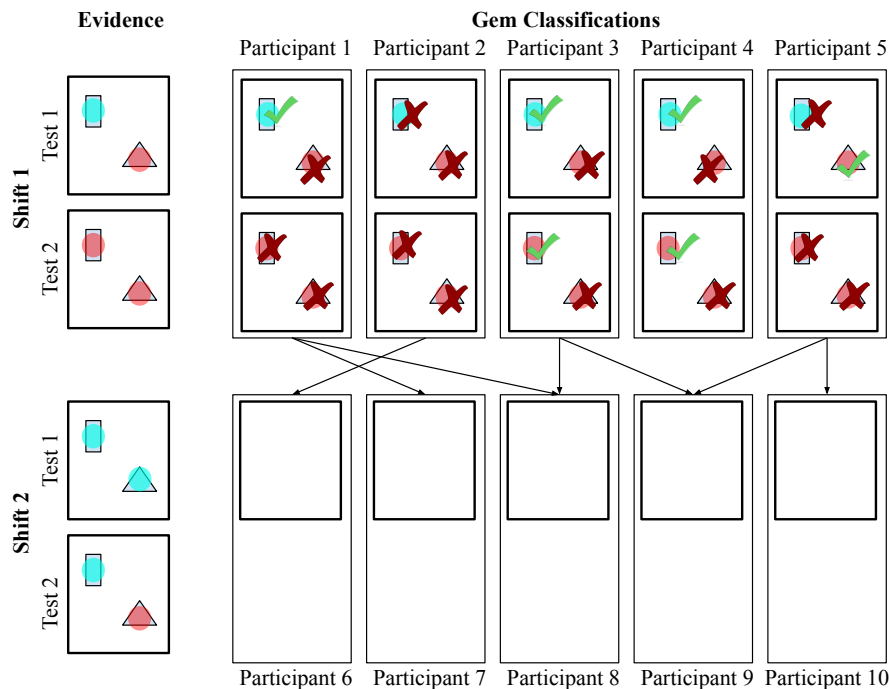
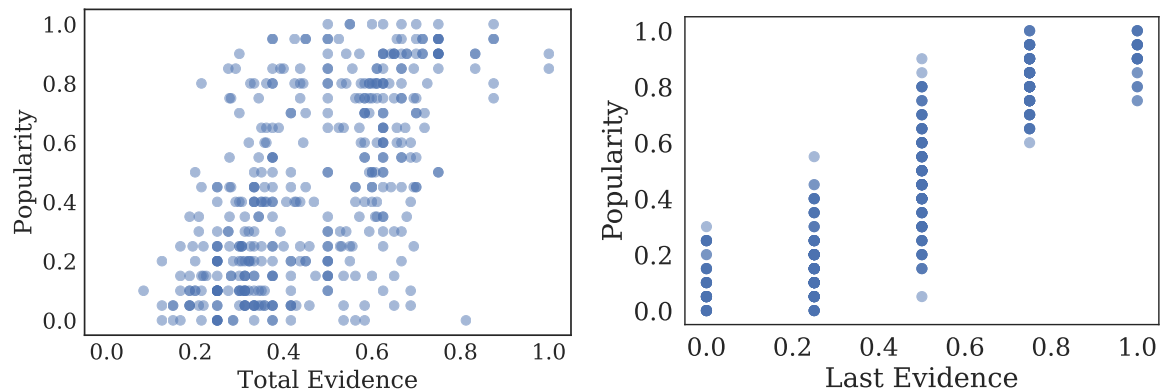


Figure 2. Illustration of the evidence correlation structure used in Experiment 1. If a gem classification test was positive for one participant, it was positive for all participants in that generation on that classification task. Positive evidence about a classification is illustrated here using turquoise circles, and negative evidence with red circles. The classification judgments made by participants are shown as checkmarks and x-marks. Participants could change their classifications after each new test.

**Procedure.** Participant completed four trials (i.e., gem classifications). The first two trials were practice rounds and were presented in the same order for all participants. The last two tasks were experimental trials and were ordered using simple randomization. We limited our analysis to data from the two experimental trials.

Mirroring the social sampling model, participants were organized into discrete shifts, or “generations”. We recruited three independent networks of participants, with each network consisting of 10 generations with 20 participants per generation (see Figure 2). Participants in generations 2 through 10 observed the classifications of a randomly sampled member of the previous generation, and could choose to draw additional samples as many times as they wished.



*Figure 3.* Experiment 1 results. The graph on the left shows the relationship between the total evidence available for each gem classification and the proportion of participants who selected that classification (i.e., the popularity of the classification). The graph on the right plots the relationship between the most recently observed evidence and the popularity. Each point represents a gem classification for a single shift, trial, and replication.

At the beginning of the experiment, the true classifications for each task were chosen uniformly at random. To reduce variance between networks, the lab test results for a given classification were the same for all participants in a single generation (see Figure 2). The probability of positive evidence in a lab test result was 0.6 for true classifications and 0.3 for false classifications (participants were informed of both probabilities).

## Results and Discussion

We limited our analysis to participants in generations 2-10, and preregistered our statistical analyses and sample size before the experiment.<sup>3</sup> These analyses, however, were adjusted and expanded over the course of revisions. This included adding model comparisons, and using all the test data in the regression described below (rather than only one decision in generations 6-10) for consistency with our analysis for Experiment 2. Limiting our regression to the preregistered subset does not change our findings.

<sup>3</sup> Preregistration available at

[https://osf.io/7qx8e/?view\\_only=53781c6aa4c742e2a9843fcef533dfd7](https://osf.io/7qx8e/?view_only=53781c6aa4c742e2a9843fcef533dfd7).

Table 1

*Regression results for Experiments 1 and 2.*

Variable	Regression model	
	Experiment 1	Experiment 2
Intercept	-0.10** (-3.15)	0.29*** (12.69)
Total evidence	0.24*** (4.54)	0.05* (2.44)
Last evidence	0.91*** (29.26)	0.25*** (18.74)
Dependent Variable	Popularity	Popularity
Degrees of freedom	424	2270
Observations	432	2304

t statistics in parentheses

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ 

Note that fixed effects for each network (Experiments 1 and 2) and each condition (Experiment 2) were included in the regressions but are not shown.

As predicted by our social sampling model, we observed a strong correlation between the popularity of each gem classification and both the total evidence ( $r(430) = 0.6, p < 0.0001$ ) and the most recent evidence ( $r(430) = 0.88, p < 0.0001$ , see Figure 3). Furthermore, total evidence was significantly predictive of popularity in a regression that included the total evidence, most recent evidence, and fixed effects for each network ( $t(424) = 4.54, p < 0.0001$ , see Table 1 for full regression results). The correlation between popularity and total evidence shows that the proportion of social sampling was high enough to facilitate significant information aggregation over time.

We then compared the performance of our Bayesian social sampling model to three asocial models and two social models—a probability matching social sampling

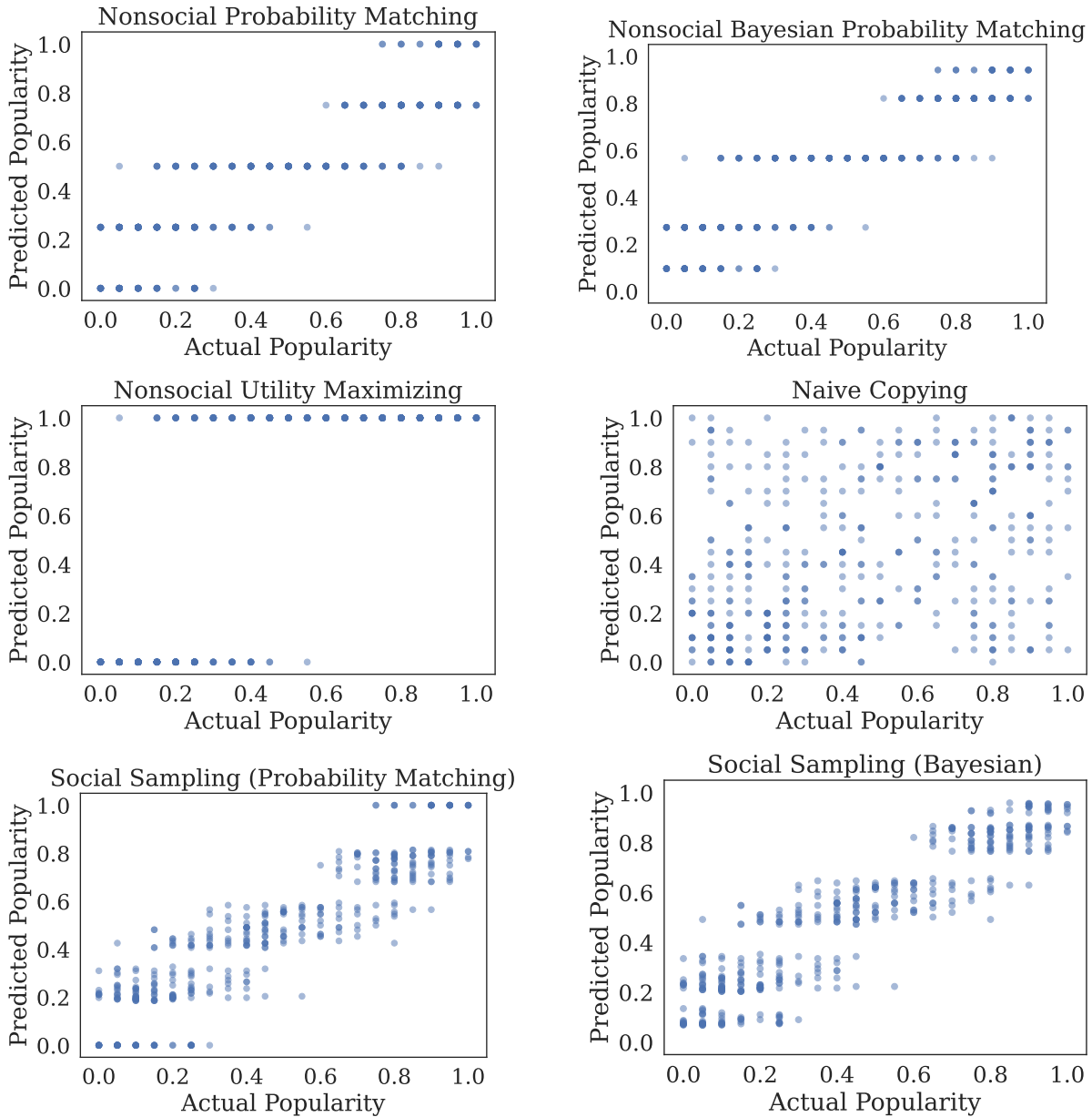


Figure 4. Model predictions for the three non-social models and three social models we compared in Experiment 1. As in Figure 3, 432 points are shown in each plot.

model and naive copying model (models are described in detail in the SEM). In the probability matching social sampling model, individuals accept candidate hypotheses with probability equal to the fraction of most recent positive evidence, and in the naive copying model, individuals simply copy others’ categorizations. To account for non-social learners, both social sampling models include a term capturing the proportion of social decision making. We estimated this value as the proportion of trials where participants chose to resample at least once, and found participants did so on



Table 2

*Mean squared error (MSE) of predicted and observed popularity for each model in Experiment 1.*

Model	MSE
Non-social Prob. Matching	0.0213
Non-social Bayesian Prob. Matching	0.0220
Non-social Utility Maximizing	0.1220
Naive Copying	0.1260
Social Sampling (Prob. Matching)	<b>0.0181</b>
Social Sampling (Bayesian)	0.0182

19% of trials. This is a lower bound on the true proportion of social sampling, since participants may have used the initial social information without resampling. However, it is close to both a separate estimate of the asocial learning proportion based on a qualitative coding of participants' strategy descriptions, and the estimated regression coefficient on total evidence available, which should correspond to the level of social sampling (see Table 1).

No free parameters are used in any of these models, and so we evaluate each model using their performance on all the experimental data. We found that the probability matching social sampling model achieved the lowest mean squared error in predicting the popularity of each gem classification ( $MSE : 0.0.0181$ ), followed closely by the Bayesian social sampling model ( $MSE : 0.0.0182$ , see Figure 4 and Table 2 for results). Because both social sampling models predict accumulation, these results support our other analyses from Experiment 1, suggesting that networks of participants accumulated information about each categorization across time.

### Experiment 2: Spaceship construction

We constructed Experiment 1 to have a high correspondence with the assumptions of our social sampling model. Experiment 2 was designed to test our model in a less

idealized context. Most notably, in Experiment 2l all information was social—participants received feedback on the previous generation’s decisions, rather than their own. Furthermore, we did not allow participants to resample from the previous generation, and feedback was censored based on the observed participant’s choices. That is, participants did not always observe evidence on every possible categorization.

In Experiment 2, participants designed spaceships by choosing which components to include from an inventory of alternatives (see Figure 5).<sup>4</sup> Each component was either a good part that rarely failed, or a bad part that failed often. On each trial, participants observed spaceships designed in the previous generation, along with the success or failure of each of the included components on one or more flights. We organized participants into five conditions, varying the number of participants in each generation (Large vs. Small), the amount of evidence presented to each participant (High vs. Low), and the strength of that evidence (High vs. Low) (see Table 3).

## Methods

**Participants.** We recruited 1,000 participants from Amazon Mechanical Turk. Recruitment was restricted to participants living in the United States. The task typically took participants around 5-10 minutes to complete, and participants earned \$1.80 as compensation.

**Stimuli.** Experiment 2 consisted of a spaceship design categorization task inspired by prior experiments on social learning (Brand, Brown, & Cross, 2018). Example stimuli from this categorization task are shown in Figure 5. In this task, participants designed spaceships by selecting which parts to include from a set of available options. A different set of eight parts was available on each trial. Each of the eight items in a trial’s set of spaceship parts could be included in that trial’s spaceship

---

<sup>4</sup> See <https://github.com/pkrafft/>

Overcoming-individual-limitations-through-distributed-computation for experiment code and data.

Table 3

*The parameter settings associated with each condition of Experiment 2.*

<b>Experiment</b>	<b>Reps</b>	$N$	$J$	$\theta$
Condition 1	2	20	4	0.6
Condition 2	1	20	1	0.6
Condition 3	1	20	4	0.8
Condition 4	2	5	4	0.6
Condition 5	2	5	1	0.6

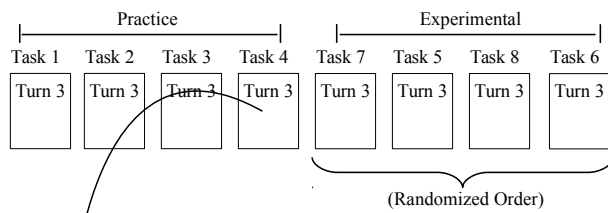
Note: The repetitions (reps) is the number of repetitions of the parameter settings we ran.  $N$  is the number of participants per generation.  $J$  is the number of flights, i.e. the amount of evidence, shown per spaceship design trial.  $\theta$  gives the probability of success of good parts, and the probability of failure of bad parts—that is, the strength of evidence.

design. Participants could select or remove a part in a trial’s design by clicking on an icon of the part.

Participants were organized into 10 discrete generations, with each generation composed of a different set of participants. In the first generation, each participant had no information about part failures and had to simply guess which parts might be good or bad. After selecting which parts to include, a participant was given feedback from a number of “flights” of their constructed spaceship. On each flight, spaceships parts could either succeed or fail. Each flight was shown in sequence, and we did not allow participants to change the spaceship design between flights.

In all generations except the first, participants viewed the spaceship design and flight outcomes from a spaceship selected uniformly at random from the designs made in the previous generation. Participants viewing another participant’s design could select which of the prior flights to examine, as different parts may have failed on different flights. A prior design was displayed in the same way that a participant’s own

View of Participant in **Generation 3**



**1. View action of Generation 2 participant and their results, and choose own action**

**2. View results of own action for this game**

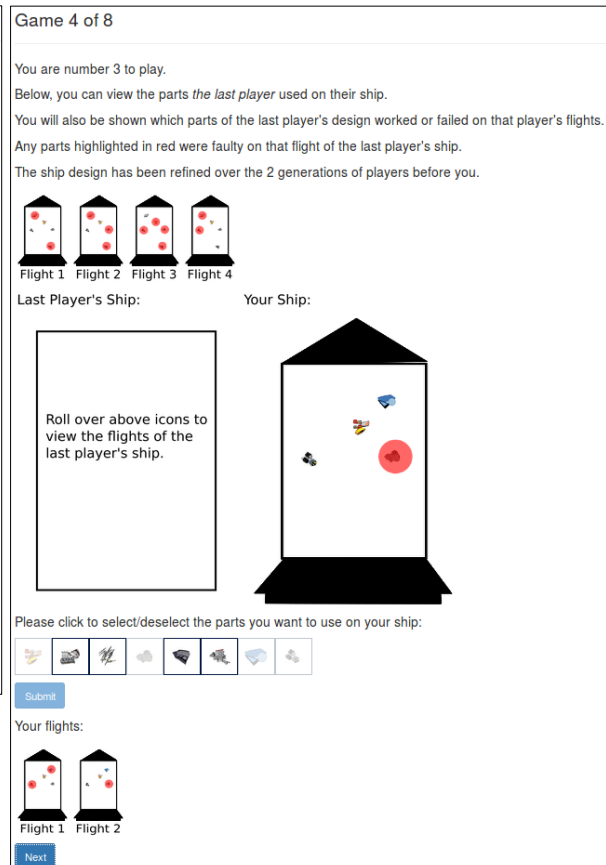
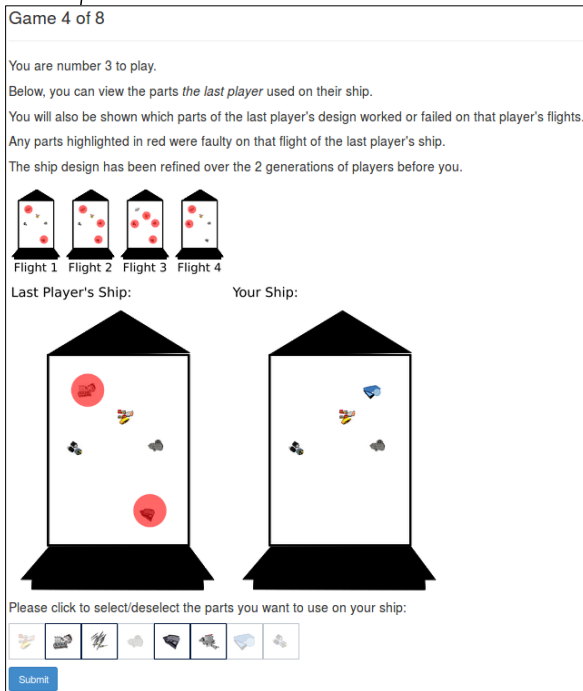
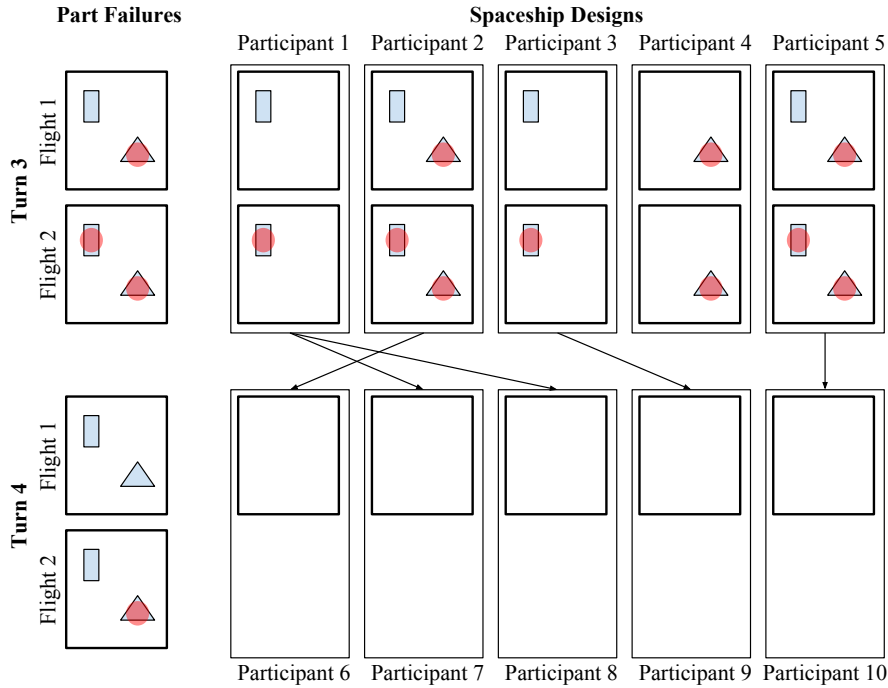


Figure 5. Example of the participant interface for the Experiment 2 spaceship design task. Each trial was associated with a different set of spaceship parts (using different visual icons). Participants were organized into discrete generations, or “turns”. In each generation after the first, a participant started by viewing an example spaceship design from a participant of the last generation. The participant then chose their own parts for that task and observed the part failures on each flight of that design.

spaceship design was displayed, except that parts could not be modified on a previous design and part failures were shown at the start of the trial. Unlike Experiment 1, participants could not resample a choice from the previous generation.



*Figure 6.* Illustration of the evidence constraints in Experiment 2. As in Experiment 1, evidence was correlated across participants: If a part failed on one flight, it failed for all participants in that generation on that flight (part failures are displayed in the figure as red circles). In contrast to Experiment 1, participants could not change their spaceship designs in between flights, and evidence about a particular spaceship part was only shown if the part was chosen in the sample design shown to the participant.

**Procedure.** Participants completed eight categorization tasks. The first four trials participants completed were practice trials and the last four were test trials. Practice trials were displayed in the same order for all participants, and test trials were ordered using simple randomization. We did not inform participants of this practice-test distinction, and we limited our analysis to data from the four test trials.

In order to mirror the structure of the social sampling model, we recruited participants in discrete generations. Generation  $t$  was recruited after all participants in generation  $t - 1$  completed the experiment. In the first generation, good and bad parts were chosen at random with probability 0.5. As in Experiment 1, spaceship failures were perfectly correlated across designs for a single flight in a given generation to reduce variance between networks. That is, if two different participants both used the same

part in a certain generation, then the part would either succeed or fail for both participants on a given flight (see Figure 6).

**Results and Discussion.** As in Experiment 1, all statistical tests were two-tailed with an alpha level of 0.05 and we excluded the initial generation of participants from our analysis. We ran experiment conditions separately, and so no between-subject randomization into conditions was used. Instead, participants could only participate in the experiment once and thus could not complete multiple experimental conditions.

Table 4

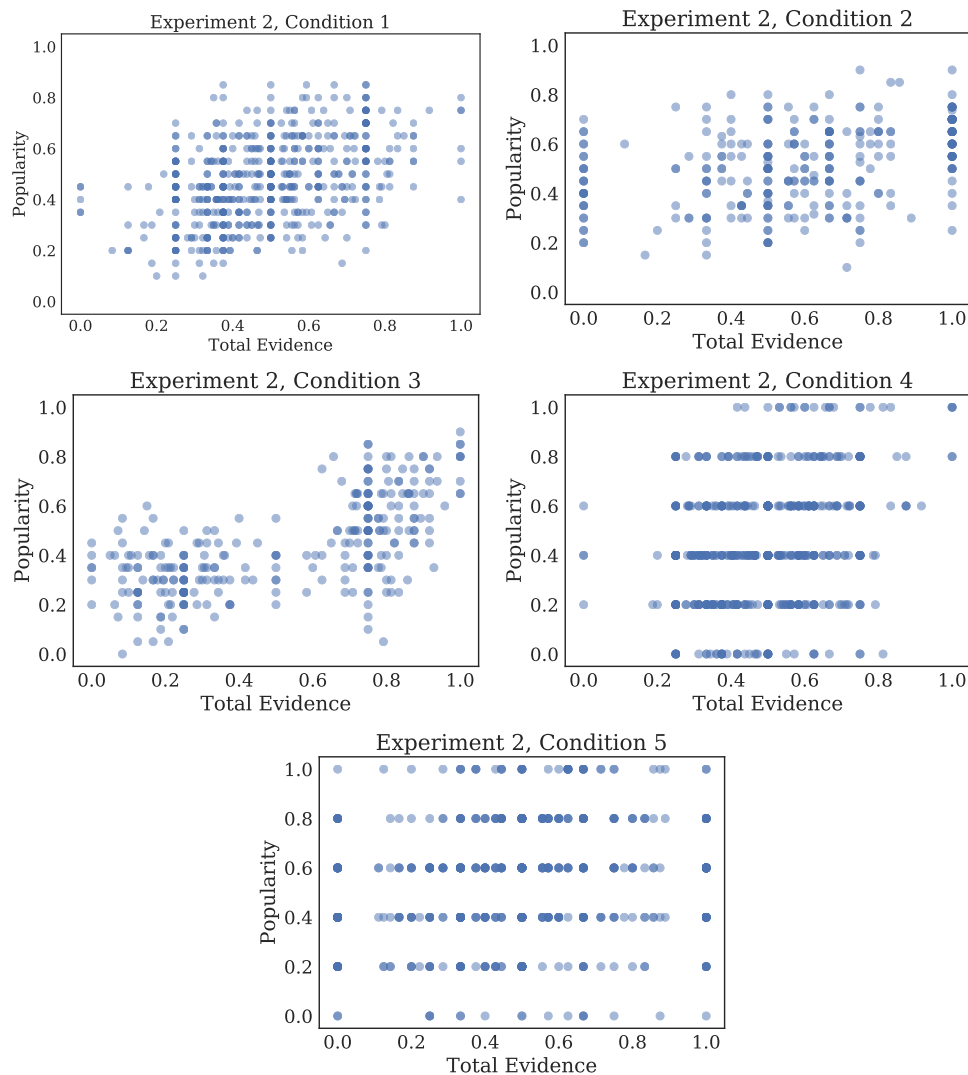
*Overview of the five conditions of Experiment 2 and the correlation between popularity and total evidence within each condition.*

<b>Condition</b>	<b>Reps</b>	<b>Population Size</b>	<b>Evidence Amount</b>	<b>Evidence Strength</b>	<b>Correlation Between Popularity and Total Evidence</b>
Condition 1	2	Large	High	Low	0.43***
Condition 2	1	Large	Low	Low	0.39***
Condition 3	1	Large	High	High	0.65***
Condition 4	2	Small	High	Low	0.29***
Condition 5	2	Small	Low	Low	0.16***

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

Replicating our findings from Experiment 1, we found that total evidence was significantly predictive of popularity, controlling for the most recent evidence observed in a regression that included the total evidence, most recent evidence, and fixed effects for each network and condition ( $t(2270) = 2.44, p = 0.015$ , full results in Table 1).

While we found significant positive correlations between popularity and total evidence in each of our conditions, the strength of this correlation varied between 0.16 in Condition 5 and 0.65 in Condition 3 (see Table 4 and Figure 7). Indeed, popularity tended to be less correlated with total evidence in networks with small generation sizes,



*Figure 7.* Plots showing the relationship between the total evidence available for each particular spaceship part and the proportion of participants who selected that part in Experiment 2. Each point represents one part (of eight) in one social generation (of nine) on one trial (of four) in a single repetition.

low evidence strength, and low evidence amount.

All decisions in Experiment 2 were social, and so we limited model comparison to the three social models used in Experiment 1 and set the proportion of asocial learners to zero. As in Experiment 1, each model was assessed by comparing the popularity predicted by the model at each generation with the observed popularity. We found that in four of the five experimental conditions, the Bayesian social sampling model achieved the lowest mean squared error in predicting observed popularity. However, in the

Table 5

*Mean squared error (MSE) of predicted and observed popularity for the three social models we compared for each of the conditions (C1-C5) in Experiment 2.*

Model	Mean Squared Error				
	C1	C2	C3	C4	C5
Naive Copying	0.0433	0.0424	<b>0.0371</b>	0.1078	0.1030
Social Sampling (Matching)	0.0542	0.1703	0.0825	0.1069	0.2196
Social Sampling (Bayesian)	<b>0.0366</b>	<b>0.0333</b>	0.1188	<b>0.0969</b>	<b>0.0960</b>

remaining condition, where the evidence strength was high, the naive copying model outperformed both the Bayesian and probability matching social sampling models (Bayesian social sampling MSE: 0.1188; probability matching social sampling MSE: 0.0825; naive copying MSE: 0.0371; see Table 5).

## General Discussion

Understanding how people make intelligent decisions under limited time and cognitive resources is a central focus of research in cognitive science, psychology, and economics. Social interactions offer a way for people to overcome these limitations by distributing computation across a group. By observing and learning from others, individuals do not need to perform complex computations or to condition on large amounts of data to make rational inferences. Instead, they can use simple heuristics that leverage accumulated social information.

To show how distributed inference can emerge in populations, we derived a social sampling model of individual decision-making in multi-generational networks. We then showed that social sampling at the individual level can lead to distributed Bayesian inference at the population level. We tested the predictions of our social sampling model in two highly controlled behavioral experiments where participants made simple categorization decisions. While the transmission structure we used in both experiments was idealized and highly simplified compared to real-world social networks, this allowed



us to directly quantify the degree to which information accumulated across time relative to the Bayesian ideal.

Although we observe information accumulation in both experiments, the performance of the Bayesian social sampling model varied in different conditions. Most notably, in Experiment 1 the probability social sampling model achieved a slightly lower MSE than the Bayesian social sampling model (see Table 2). Furthermore, in Experiment 2 the naive copying model had a lower MSE than both social sampling models in large networks with high levels of strong evidence (see Tables 4 and 5). These results suggest that individuals may adapt their social learning strategies to different domains, and may be more likely to use social sampling in noisy, low-information environments (Toyokawa, Whalen, & Laland, 2018).

It is important to emphasize that for information aggregation to occur, the precise details of how social sampling takes place are less important than that individuals' decisions incorporate social information and new evidence using a probabilistic rule. Indeed, our social sampling models in Experiment 1 included a mixture of social and non-social decision-makers, replicating related findings on under-exploitation of social information in behavioral experiments (Mercier & Morin, 2019). Our analyses are thus not intended to show that social sampling is a definitive description of human behavior, but that a social decision-making model derived from a normative Bayesian standard can help make sense of the extent of information aggregation we observe in specific experimental contexts.

While we find that information aggregation can be robust to individual differences, accumulation can fail if people's decision-making differs systematically from a probabilistic belief adoption-rejection strategy. Indeed, related research on the wisdom of crowds has shown that in certain contexts, social interactions and observations can actually decrease group performance by reducing the diversity and independence of people's beliefs (Jenness, 1932; Lorenz, Rauhut, Schweitzer, & Helbing, 2011; Myers & Bishop, 1971). These dynamics—that is, whether social interactions

improve or worsen group outcomes—appear to vary depending on people’s social learning strategies (Toyokawa et al., 2018). For example, accumulation can be disrupted if people are utility-maximizing rather than probabilistic in their responses to evidence (Anderson & Holt, 1997; Bikhchandani, Hirshleifer, & Welch, 1992). While probability matching in individual decision-making is observed in certain domains (Shanks, Tunney, & McCarthy, 2002; Vul et al., 2014; Vulkan, 2000), our experiments show that this strategy can be extended to social settings: when a substantial proportion of individuals incorporate social information into their probability matching behavior, information accumulates across people and time. However, people may be less likely to follow a social sampling strategy in domains with low levels of probability matching, such as tasks with large financial incentives or consistent feedback (Shanks et al., 2002).

Our findings complement other approaches to understanding rationality in limited individuals. For example, previous work on adaptive heuristics (Gigerenzer & Goldstein, 1996) has shown that people use simple decision-making rules to exploit the structure of the environment and make decisions on complex tasks. Our work suggests that this framework can be extended to social decision-making. Furthermore, by providing a rational motivation for people’s social heuristics, our work offers a way to connect adaptive heuristics with resource-rational analysis (Griffiths, Lieder, & Goodman, 2015) in simple multi-generational populations. Integrating our model and findings with other frameworks for studying bounded rationality, such as models that utilize quantum probability theory (Pothos et al., 2021), should be addressed in future work.

We tested our model on a simple binary-choice decision-making task where the transmission structure was explicitly designed to match our model assumptions. These simplifications limit our ability to draw general conclusions about information accumulation in natural populations. For example, while enforcing uniform sampling of the previous generation gave us tight control over the transmission dynamics, people’s sampling and exploration strategies may depend on the task (Oaksford & Chater, 1994) and are likely considerably more complex in naturalistic domains and networks (Latora, Nicosia, & Russo, 2017). Similarly, people may not always transmit their true

beliefs—instead, they may give noisy, incomplete, or misleading accounts of their opinions and decisions to others (Xu & Griffiths, 2010).

Despite these limitations, the general framework we have outlined and the principle of identifying the computational structure of population dynamics could therefore be extended to more complex and naturalistic domains in future work. For example, individuals could be given greater control over which (if any) beliefs to transmit or sample from others. Models of learning in structured representational domains, such as language of thought models (Goodman, Tenenbaum, & Gerstenberg, 2015) or posterior sampling in general Markov decision processes (Agrawal & Jia, 2017; Osband, Van Roy, & Russo, 2013), could also be extended to the population setting in a way that is analogous to our extension of simple categorization decisions. Testing our model and quantifying knowledge accumulation in more naturalistic domains is an exciting challenge for future work.

### **Acknowledgments**

This work was funded in part by NSF grant 1456709, DARPA Cooperative Agreement D17AC00004, and the support of the NOMIS Foundation.

## References

- Acemoglu, D., Dahleh, M. A., Lobel, I., & Ozdaglar, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, *78*(4), 1201–1236.
- Agrawal, S., & Jia, R. (2017). Posterior sampling for reinforcement learning: worst-case regret bounds. In *Advances in neural information processing systems* (pp. 1184–1194).
- Almaatouq, A., Alsobay, M., Yin, M., & Watts, D. J. (2021). Task complexity moderates group synergy. *Proceedings of the National Academy of Sciences*, *118*(36).
- Almaatouq, A., Noriega-Campero, A., Alotaibi, A., Krafft, P., Moussaid, M., & Pentland, A. (2020). Adaptive social networks promote the wisdom of crowds. *Proceedings of the National Academy of Sciences*, *117*(21), 11379–11386.
- Anderson, L. R., & Holt, C. A. (1997). Information cascades in the laboratory. *The American Economic Review*, 847–862.
- Argote, L. (2012). *Organizational learning: Creating, retaining and transferring knowledge*. Springer Science & Business Media.
- Belikov, A. V., Rzhetsky, A., & Evans, J. (2020). Detecting signal from science: The structure of research communities and prior knowledge improves prediction of genetic regulatory experiments. *arXiv preprint arXiv:2008.09985*.
- Bernardo, J. M., & Smith, A. F. (2000). *Bayesian theory*. Wiley.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, *100*(5), 992–1026.
- Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating Bayesian inference. *Cognitive Psychology*, *74*, 35–65.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.
- Brand, C. O., Brown, G. R., & Cross, C. P. (2018). Sex differences in the use of social

- information emerge under conditions of risk. *PeerJ*, 6, e4190.
- Caldwell, C. A., Atkinson, M., & Renner, E. (2016). Experimental approaches to studying cumulative cultural evolution. *Current Directions in Psychological Science*, 25(3), 191–195.
- Celis, L. E., Krafft, P., & Vishnoi, N. (2017). A distributed learning dynamics in social groups. In *Proceedings of the acm symposium on principles of distributed computing* (pp. 441–450).
- Chater, N., Oaksford, M., Hahn, U., & Heit, E. (2010). Bayesian models of cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), 811–823.
- Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: Conceptual foundations. *Trends in cognitive sciences*, 10(7), 287–291.
- Crisan, D., & Doucet, A. (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on signal processing*, 50(3), 736–746.
- Cushman, F., & Gershman, S. (2019). Editors' introduction: Computational approaches to social cognition. *Topics in cognitive science*, 11(2), 281–298.
- Derex, M., Beugin, M.-P., Godelle, B., & Raymond, M. (2013). Experimental evidence for the influence of group size on cultural complexity. *Nature*, 503(7476), 389–391.
- Easwaran, K. (2019). The Concept of Rationality for a City. *Topoi*, 1–13.
- Engel, D., Woolley, A. W., Aggarwal, I., Chabris, C. F., Takahashi, M., Nemoto, K., . . . Malone, T. W. (2015). Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures. In *Proceedings of the 33rd annual acm conference on human factors in computing systems (chi)* (pp. 3769–3778).
- Foster, J. G. (2018). Culture and computation: Steps to a probably approximately correct theory of culture. *Poetics*, 68, 144–154.
- Frey, S., & Goldstone, R. L. (2018). Cognitive mechanisms for human flocking dynamics. *Journal of Computational Social Science*, 1(2), 349–375.
- Gale, D., & Kariv, S. (2003). Bayesian learning in social networks. *Games and*

- Economic Behavior*, 45(2), 329–346.
- Galesic, M., Olsson, H., Dalege, J., van der Does, T., & Stein, D. L. (2021). Integrating social and cognitive aspects of belief dynamics: towards a unifying framework. *Journal of the Royal Society Interface*, 18(176), 20200857.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650.
- Goldstone, R. L., Wisdom, T. N., Roberts, M. E., & Frey, S. (2013). Learning along with others. In *Psychology of learning and motivation* (Vol. 58, pp. 1–45). Elsevier.
- Goodman, N. D., Tenenbaum, J. B., & Gerstenberg, T. (2015). Concepts in a probabilistic language of thought. In Morgolis & Lawrence (Eds.), *The conceptual mind: New directions in the study of concepts*. MIT Press.
- Griffiths, T. L. (2020). Understanding human intelligence through human limitations. *Trends in Cognitive Sciences*.
- Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in cognitive sciences*, 14(8), 357–364.
- Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In *The Cambridge Handbook of Computational Psychology*. Cambridge University Press.
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in cognitive science*, 7(2), 217–229.
- Harper, M. (2009). The replicator equation as an inference dynamic. *arXiv preprint arXiv:0911.1763*.
- Hayek, F. (1945). The use of knowledge in society. *The American Economic Review*, 35(4), 519–530.
- Hazła, J., Jadbabaie, A., Mossel, E., & Rahimian, M. A. (2021). Bayesian decision making in groups is hard. *Operations Research*, 69(2), 632–654.

- Heyes, C. (2018). Enquire within: Cultural evolution and cognitive science. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1743), 20170051.
- Howard, J. A., & Sheth, J. N. (1969). *The theory of buyer behavior*. MIT Press.
- Hutchins, E. (1995). *Cognition in the Wild* (No. 1995). MIT press.
- Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.
- Jenness, A. (1932). The role of discussion in changing opinion regarding a matter of fact. *The Journal of Abnormal and Social Psychology*, 27(3), 279.
- Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review*, 14(2), 288–294.
- Kempe, M., & Mesoudi, A. (2014a). Experimental and theoretical models of human cultural evolution. *WIREs Cogn Sci*, 5, 317–326.
- Kempe, M., & Mesoudi, A. (2014b). An experimental demonstration of the effect of group size on cultural accumulation. *Evolution and Human Behavior*, 35(4), 285–290.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(31), 10681–6.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.
- Krafft, P. (2017). *A rational choice framework for collective behavior*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.
- Krafft, P., Shmueli, E., Griffiths, T. L., Tenenbaum, J. B., & Pentland, A. (2020). Bayesian collective learning emerges from heuristic social learning. *Cognition*.
- Krafft, P., Zheng, J., Pan, W., Della Penna, N., Altshuler, Y., Shmueli, E., . . .



- Pentland, A. (2016). Human collective intelligence as distributed bayesian inference. *arXiv preprint arXiv:1608.01987*.
- Krumme, C., Cebrian, M., Pickard, G., & Pentland, A. (2012). Quantifying social influence in an online cultural market. *PLoS ONE*, 7(5), e33785.
- Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011, 12). Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? *Science*, 334(6062), 1512–6. Retrieved from <http://www.sciencemag.org/content/334/6062/1512.abstract> doi: 10.1126/science.1210879
- Latora, V., Nicosia, V., & Russo, G. (2017). *Complex networks: principles, methods and applications*. Cambridge University Press.
- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the national academy of sciences*, 108(22), 9020–9025.
- Lynch, N. A. (1996). *Distributed algorithms*. Morgan Kaufmann Publishers Inc.
- Martin, D., Hutchison, J., Slessor, G., Urquhart, J., Cunningham, S. J., & Smith, K. (2014). The spontaneous formation of stereotypes via cumulative cultural evolution. *Psychological Science*, 25(9), 1777–1786.
- Mercier, H., & Morin, O. (2019). Majority rules: how good are we at aggregating convergent opinions? *Evolutionary Human Sciences*, 1.
- Mesoudi, A. (2011). *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. Chicago, IL: University of Chicago Press.
- Mesoudi, A. (2016). Cultural evolution: integrating psychology, evolution and culture. *Current Opinion in Psychology*, 7, 17–22.
- Mesoudi, A., & O'Brien, M. J. (2008). The cultural transmission of great basin projectile-point technology I: An experimental simulation. *American Antiquity*, 73(01), 3–28.
- Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution?

- Proceedings of the Royal Society Biological Sciences*, 285(1880), 20180712.
- Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509), 3489–3501.
- Miton, H., & Charbonneau, M. (2018). Cumulative culture in the laboratory: Methodological and theoretical challenges. *Proceedings. Biological Sciences*, 285(1879).
- Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., ... Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, 6(1), 6029.
- Mozer, M. C., Pashler, H., & Homaei, H. (2008). Optimal predictions in everyday cognition: The wisdom of individuals or crowds? *Cognitive Science*, 32(7), 1133–1147.
- Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.
- Muthukrishna, M., Shulman, B. W., Vasilescu, V., & Henrich, J. (2014). Sociality influences cultural complexity. *Proceedings. Biological sciences*, 281(1774), 20132511.
- Myers, D. G., & Bishop, G. D. (1971). Enhancement of dominant attitudes in group discussion. *Journal of personality and social psychology*, 20(3), 386.
- Nettle, D. (2020, 4). Selection, adaptation, inheritance and design in human culture: the view from the Price equation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1797), 20190358. Retrieved from <https://royalsocietypublishing.org/doi/10.1098/rstb.2019.0358> doi: 10.1098/rstb.2019.0358
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101(4), 608.
- Ortega, P. A. (2011). *A unified framework for resource-bounded autonomous agents interacting with unknown environments*. Unpublished doctoral dissertation,

University of Cambridge.

- Osband, I., Van Roy, B., & Russo, D. (2013). (more) efficient reinforcement learning via posterior sampling. *Advances in Neural Information Processing Systems*.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, *16*(2), 366–387.
- Pothos, E. M., Lewandowsky, S., Basieva, I., Barque-Duran, A., Tapper, K., & Khrennikov, A. (2021). Information overload for (bounded) rational agents. *Proceedings of the Royal Society B*, *288*(1944), 20202957.
- Pratt, S. C., Sumpter, D. J., Mallon, E. B., & Franks, N. R. (2005). An agent-based model of collective nest choice by the ant *Temnothorax albipennis*. *Animal Behaviour*, *70*(5), 1023–1036.
- Ravignani, A., Thompson, B., Grossi, T., Delgado, T., & Kirby, S. (2018). Evolving building blocks of rhythm: how human cognition creates music via cultural transmission. *Annals of the New York Academy of Sciences*, *1423*(1), 176–187.
- Riedl, C., Kim, Y. J., Gupta, P., Malone, T. W., & Woolley, A. W. (2021). Quantifying collective intelligence in human groups. *Proceedings of the National Academy of Sciences*, *118*(21).
- Rzhetsky, A., Foster, J. G., Foster, I. T., & Evans, J. A. (2015). Choosing experiments to accelerate collective discovery. *Proceedings of the National Academy of Sciences*, *112*(47), 14569–14574.
- Salhab, R., Ajorlou, A., & Jadbabaie, A. (2020). Social learning with sparse belief samples. In *2020 59th IEEE conference on decision and control (cdc)* (pp. 1792–1797).
- Sanborn, A., & Griffiths, T. L. (2008). Markov chain Monte Carlo with people. In *Advances in neural information processing systems* (pp. 1265–1272).
- Sasaki, T., & Pratt, S. C. (2011). Emergence of group rationality from irrational individuals. *Behavioral Ecology*, *22*(2), 276–281.
- Seeley, T. D., & Buhrman, S. C. (1999). Group decision making in swarms of honey

- bees. *Behavioral Ecology and Sociobiology*, *45*(1), 19–31.
- Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, *15*(3), 233–250.
- Shoham, Y., Leyton-Brown, K., et al. (2009). Multiagent systems. *Cambridge Books*.
- Silvey, C., Kirby, S., & Smith, K. (2019). Communication increases category structure and alignment only when combined with cultural transmission. *Journal of Memory and Language*, *109*, 104051.
- Simon, H. A. (1990). Bounded rationality. In *Utility and probability* (pp. 15–18). Springer.
- Smaldino, P. E., & Richerson, P. J. (2013). Human cumulative cultural evolution as a form of distributed computation. In *Handbook of human computation* (pp. 979–992). Springer New York.
- Smolla, M., Jansson, F., Lehmann, L., Houkes, W., Weissing, F., Hammerstein, P., . . . Enquist, M. (2020). Underappreciated similarities and differences between cultural and biological evolution, and the future of cultural evolution thinking.
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285.
- Toyokawa, W., Whalen, A., & Laland, K. N. (2018). Social learning strategies regulate the wisdom and madness of interactive crowds. *bioRxiv*, 326637. Retrieved from <https://www.biorxiv.org/content/early/2018/06/19/326637> doi: 10.1101/326637
- Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science*, *38*(4), 599–637.
- Vul, E., & Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, *19*(7), 645–647.
- Vulkan, N. (2000). An economist’s perspective on probability matching. *Journal of economic surveys*, *14*(1), 101–118.
- Wisdom, T. N., Song, X., & Goldstone, R. L. (2013). Social learning strategies in

networked groups. *Cognitive science*, 37(8), 1383–1425.

Wojtowicz, Z., & DeDeo, S. (2020). From probability to consilience: How explanatory values implement bayesian reasoning. *Trends in Cognitive Sciences*.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004), 686–688.

Xu, J., & Griffiths, T. L. (2010). A rational analysis of the effects of memory biases on serial reproduction. *Cognitive psychology*, 60(2), 107–126.

**Supplementary Electronic Materials for:  
Overcoming individual limitations through distributed computation**

## Model

Here, we describe our Bayesian social sampling in detail and prove Equation 2.

### Environment Model

We assume that the state of the environment which is to be inferred by the population of learners can be represented as a binary feature vector,  $\mathbf{x}$ . Features are related to observed evidence via simple logical relations. Observed pieces of evidence are represented in  $\mathbf{Y}$ . Evidence about the feature vector arrives over time, so  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_T)$ , where  $T$  is the time horizon. We assume that at each time  $t$ ,  $J$  new pieces of evidence about feature  $i$  are available,  $y_{ijt} \sim P_i$ , where  $P_i$  is a distribution specific to feature  $i$ . We assume that  $y_{ijt}$  are independent, so  $P_i$  is simply a component of a conditional probability table,  $P_i = P(y_{ijt} = 1 | x_i = 0) = \theta_{i0}$  if  $x_i = 0$  and  $P_i = P(y_{ijt} = 1 | x_i = 1) = \theta_{i1}$  if  $x_i = 1$ . Under these assumptions, evidence is either perceived as positive or negative, and pieces of evidence are uncorrelated across time.

### Social Sampling

To develop a normative model of behavior in this environment, we specify an individual-level, sample-based social heuristic that leads to population-rational, distributed Bayesian inference in aggregate. Crucially, group computation can emerge even when the social learning heuristic is simple and easy to compute—population rationality does not require unbounded, perfectly rational individuals. Indeed, individuals rely on social information *because of* the limited time and cognitive resources they have to make their decisions.

Our model is specified at the level of the individual in that we seek an explanation in terms of the behavioral mechanisms of individual people, and is sample-based as it relies on the assumption that people do not fully represent uncertainty in their mental representations. Instead, we assume that people often think of the world as simply being one way or the other at a particular point in time. Formally, we let  $N$  be the number of agents,  $d_{ait}$  be a discrete decision of agent  $a$  about the value of feature  $i$  at

time  $t$ , and  $p_{it} = \frac{1}{N} \sum_a \mathbb{1}(d_{a,i,t-1} = 1)$  be the proportion of agents who believe  $x_i = 1$  at time  $t - 1$ .  $\theta_{i1}$ ,  $\theta_{i0}$ , and  $y_{ijt}$  are defined as in the environment model above. For notational convenience, we also let  $S_{it} = \sum_{j=1}^J \mathbb{1}(y_{ijt} = 1)$ .

In our social sampling model, at each time step  $t$  each agent  $a$  uses a heuristic decision-making rule: The agent starts out undecided, and continues to sample possible decisions until decided. To sample a decision, agent  $a$  samples an agent  $a'$  who was decided on the last time step uniformly at random, and considers the choice of agent  $a'$ ,  $d_{a',i,t-1}$ . Agent  $a$  then decides  $d_{a',i,t-1}$  with probability proportional to  $(\theta_{i,d_{a',i,t-1}})^{S_{it}} (1 - \theta_{i,d_{a',i,t-1}})^{J-S_{it}}$ , which is the likelihood of the evidence that agent  $a$  has at time  $t$  under the state of the world indicated by agent  $a'$ . The number of samples agent  $a$  takes is a geometric random variable with parameter  $p_{it} \cdot (\theta_{i1})^{n_{S_{it}}} (1 - \theta_{i1})^{J-S_{it}} + (1 - p_{it}) \cdot (\theta_{i0})^{S_{it}} (1 - \theta_{i0})^{J-S_{it}}$ , and is finite with probability one.

### Evidence Accumulation

Because each iteration of the decision loop is independent, the probability of  $a$  deciding  $d_{ait} = 1$  is

$$\frac{(\theta_{i1})^{S_{it}} (1 - \theta_{i1})^{J-S_{it}} \cdot p_{it}}{(\theta_{i1})^{S_{it}} (1 - \theta_{i1})^{J-S_{it}} \cdot p_{it} + (\theta_{i0})^{S_{it}} (1 - \theta_{i0})^{J-S_{it}} \cdot (1 - p_{it})}. \quad (4)$$

In an infinite population of agents, for a particular feature  $i$ , we then have (by induction)

$$\begin{aligned} p_{i,t+1} &= \frac{(\theta_{i1})^{S_{it}} (1 - \theta_{i1})^{J-S_{it}} \cdot p_{it}}{(\theta_{i1})^{S_{it}} (1 - \theta_{i1})^{J-S_{it}} \cdot p_{it} + (\theta_{i0})^{S_{it}} (1 - \theta_{i0})^{J-S_{it}} \cdot (1 - p_{it})} \\ &= \frac{P(\mathbf{y}_{i,\cdot,t} | x_i = 1) \cdot p_{it}}{P(\mathbf{y}_{i,\cdot,t} | x_i = 1) \cdot p_{it} + P(\mathbf{y}_{i,\cdot,t} | x_i = 0) \cdot (1 - p_{it})} \\ &= \frac{P(\mathbf{y}_{i,\cdot,t} | x_i = 1) P(x_i = 1 | \mathbf{y}_{i,\cdot,<t})}{P(\mathbf{y}_{i,\cdot,t} | x_i = 1) P(x_i = 1 | \mathbf{y}_{i,\cdot,<t}) + P(\mathbf{y}_{i,\cdot,t} | x_i = 0) P(x_i = 0 | \mathbf{y}_{i,\cdot,<t})} \\ &= \frac{P(\mathbf{y}_{i,\cdot,\leq t} | x_i = 1) P(x_i = 1)}{P(\mathbf{y}_{i,\cdot,\leq t})} \\ &= P(x_i = 1 | \mathbf{y}_{i,\cdot,\leq t}) \end{aligned}$$

being the posterior distribution on  $x$  given all preceding evidence.



## Model comparison

We compared the performance of the Bayesian social sampling model with three alternative non-social models and two alternative social models. We assessed the fit of each model by comparing the aggregate popularity observed at each generation with the popularity predicted by the model.

### Non-social Probability Matching Models

Our first two non-social models differentiate between pure probability matching and Bayesian probability matching.

**Non-social Probability Matching.** The non-social probability matching model describes decision-makers who make a feature categorization with probability equal to the evidence they have directly observed for that categorization being correct. In this case, the predicted popularity of a categorization is the fraction of positive evidence for that categorization in the current generation,  $\frac{S_{it}}{J}$ . If no evidence is observed, the predicted popularity of the categorization is zero.

**Non-social Bayesian Probability Matching.** The non-social Bayesian probability matching model describes decision-makers who make feature categorizations with probability equal to the posterior probability of the categorization being correct, given the evidence they have directly observed. The predicted popularity of a categorization is thus equal to the posterior probability for that categorization being correct given the evidence in the current generation,  $P(x_{it} = 1 | \mathbf{y}_{i,t})$ .

### Non-social Utility Maximizing

Our third non-social comparison model describes individuals who follow a utility maximizing strategy based on only the evidence they have directly observed. A non-social utility maximizing agent makes a feature categorization if the posterior probability for that categorization is at least 0.5. In this case

$$P(d_{ait} = 1) = \begin{cases} 1 & \text{if } P(x_{it} = 1 | \mathbf{y}_{i,t}) > 0.5 \\ 0 & \text{if } P(x_{it} = 1 | \mathbf{y}_{i,t}) < 0.5. \end{cases} \quad (5)$$

That is, the predicted popularity of a categorization is 1 if the posterior for that categorization is 0.5 or higher, and 0 otherwise (ties are broken randomly).

### Naive Copying

Our fourth comparison model describes a purely social decision-maker who copies the categorization of a previously observed agent. In this model, the first generation makes their choices uniformly at random, and the predicted popularity of a categorization decision  $i$  at generation  $t$  is equal to  $p_{it}$ , the popularity of decision  $i$  at the previous generation.

### Social Sampling Models

In addition to our hypothesized Bayesian social sampling model, we consider an alternative social sampling model that is based on pure probability matching. While both this model and the Bayesian social sampling model lead to information accumulation over time, the Bayesian social sampling model allows us to view transmission as distributed inference, offering a closer correspondence to the Bayesian approach to modeling individual cognition.

To fit both of these models, we first observed that a large proportion of participants in Experiment 1 ignored social information. To derive predictions from our social sampling models on Experiment 1, we therefore assumed that there is a mixture of non-social sampling using the corresponding probability matching model and social sampling. We let  $p_{ns}$  be the proportion of non-social sampling.

**Social Sampling (Probability Matching).** In the probability matching social sampling model, individuals iteratively sample decisions at random from the previous generation and accept or reject the components of those decisions based on the latest observed evidence. In this case, the acceptance probability is the same as the decision

probability in the non-social probability matching model. The predicted popularity according to this combination of components is the previous popularity of the categorization times the fraction of most recent positive evidence, normalized by the previous unpopularity of the categorization decision times the most recent negative evidence. Introducing non-social agents into this model is equivalent to setting the prior popularity to the proportion of social agents times the previous popularity of the categorization plus a uniform prior times the proportion of non-social agents. The predicted popularity at each generation is thus

$$\frac{\frac{S_{it}}{J} \cdot (p_{it} \cdot (1 - p_{ns}) + 0.5 \cdot p_{ns})}{\frac{S_{it}}{J} \cdot (p_{it} \cdot (1 - p_{ns}) + 0.5 \cdot p_{ns}) + (1 - \frac{S_{it}}{J}) \cdot ((1 - p_{it}) \cdot (1 - p_{ns}) + 0.5 \cdot p_{ns})}. \quad (6)$$

In the edge case where the popularity of a decision is equal to one but its evidence is equal to zero, the estimated popularity is set to zero.

**Social Sampling (Bayesian).** The Bayesian social sampling model is our primary hypothesized model that was previously described in detail. In this model, individuals iteratively sample decisions at random from the previous generation and accept or reject the components of those decisions based on the Bayesian likelihood of the categorization being correct given the latest evidence. In this model, the acceptance probability is proportional to the decision probability in the non-social Bayesian probability matching model. The predicted popularity according to this combination of components is the previous popularity of a categorization times the likelihood of the categorization being correct given the most recent evidence, normalized by the previous unpopularity of the categorization times the likelihood of the categorization being incorrect given the most recent evidence. Integrating non-social decision-makers into this model is equivalent to modeling the prior popularity as the proportion of social agents times the previous popularity of a categorization, plus the proportion of non-social agents times a uniform prior. These components yield the predicted popularity at each generation to be

$$\frac{\gamma_{it1} \cdot (p_{it} \cdot (1 - p_{ns}) + 0.5 \cdot p_{ns})}{\gamma_{it1} \cdot (p_{it} \cdot (1 - p_{ns}) + 0.5 \cdot p_{ns}) + \gamma_{it0} \cdot ((1 - p_{it}) \cdot (1 - p_{ns}) + 0.5 \cdot p_{ns})}, \quad (7)$$

where  $\gamma_{itk} = \theta_{ik}^{S_{it}} (1 - \theta_{ik})^{J - S_{it}}$ .