RESEARCH ARTICLE

Bayesian Decision Making in Human Collectives with Binary Choices

Víctor M. Eguíluz[1]*, **Naoki Masuda[2]**, **Juan Fernández-Gracia[1,3]**

1 Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), E07122 Palma de Mallorca, Spain, **2** Department of Engineering Mathematics, University of Bristol, Merchant Venturers Building, Woodland Road, Clifton, Bristol BS8 1UB, United Kingdom, **3** Instituto Mediterráneo de Estudios Avanzados IMEDEA (CSIC-UIB), E07190 Esporles, Spain

* victor@ifisc.uib-csic.es

Abstract

Here we focus on the description of the mechanisms behind the process of information aggregation and decision making, a basic step to understand emergent phenomena in society, such as trends, information spreading or the wisdom of crowds. In many situations, agents choose between discrete options. We analyze experimental data on binary opinion choices in humans. The data consists of two separate experiments in which humans answer questions with a binary response, where one is correct and the other is incorrect. The questions are answered without and with information on the answers of some previous participants. We find that a Bayesian approach captures the probability of choosing one of the answers. The influence of peers is uncorrelated with the difficulty of the question. The data is inconsistent with Weber's law, which states that the probability of choosing an option depends on the proportion of previous answers choosing that option and not on the total number of those answers. Last, the present Bayesian model fits reasonably well to the data as compared to some other previously proposed functions although the latter sometime perform slightly better than the Bayesian model. The asset of the present model is the simplicity and mechanistic explanation of the behavior.

Introduction

The process of information aggregation in social systems gives rise to emergent phenomena like the wisdom of crowds [1, 2]. In order to understand such phenomena a quantitative understanding of the mechanisms by which information is aggregated and used in opinion formation and decision making is needed. In the case of the wisdom of crowds, which refers to having a better estimation of the solution to a question when the opinions of multiple heterogeneous agents are aggregated, it has been shown that social interaction can lead to misleading estimations [3]. The issue of information aggregation is a hot topic which is expected to give insights into the solution of many societal problems. For example, 2014's World Economic Forum's meeting has the title "Leveraging collective intelligence for unprecedented challenges".

Models of opinion dynamics are based on assumptions on the decision making process on interacting individuals. Simple decision making rules employed in these models include proportional imitation (i.e., the rate of the opinion conversion is proportional to the number of peers possessing the different opinion), employed in the voter model, majority rules (i.e., the same rate is a superlinear function), thresholding rules (i.e., thresholding function), reinforcement rules (i.e., adaptive function depending on experiences of agents), and homophily rules (i.e., similar individuals more likely interact) [4–6]. The type of the employed decision making rule affects the possibility, final state, speed, and other dynamical phenomena of collective opinion formation. However, in physics and even social sciences literature, justification of these different types of models is at best based on a qualitative assessment of human behavior. Beyond opinion dynamics, social dilemmas, which in many cases are based on binary decision making, also offer an opportunity to bridge theory to experiments [7, 8].

For animals in groups, recent work in this direction has identified Bayesian inference as a mechanism behind their collective behavior [9–21]. In humans, experimental evidence of Bayesian inference has been provided in the realm of perceptual and cognitive domains [22, 23]. Effects of Bayesian types of inference on collective behavior have been investigated with the use of mathematical and individual-based models [16, 24–29]. Toward quantitative understanding of social decision making of humans, the seminal experiment by Milgram and colleagues [30] designed to assess the probability to stop by a group of bystanders has recently been reproduced [31] whose results are fitted by a heuristic function. There are also other recent studies attempting to fit Bayesian (see the references above), evolutionary dynamical [32], and other [33, 34] models to behavioral data. The wisdom of crowds when interaction among participants is allowed is also a target of recent experimental studies [3, 35, 36]. However, a unifying quantitative framework to infer models of social decision making on the basis of behavioral data of humans is still lacking and much preceded by accumulating modeling frameworks for social animals [11, 15, 17].

In the present study, we address the potential of the Bayesian approach to explain human decision making under social interaction. We focus on subjects answering questions with binary options, one of which is correct. This situation contrasts with that of the previous studies on the wisdom of crowds that allowed virtually real values of answers [3, 35, 36]. We examine binary choices because many options in nature are discrete, as exemplified by voting, purchasing, and deciding where to live. In many of such situations, extrapolation from continuous settings is not obvious. We use previously published data sets in which the participants first answer in the absence of social information and later with the information about the answers submitted by the r previous respondents; r gradually increases for the same question [33, 34]. The participants answer in a sequence, the situation akin to that for previous Bayesian models of the emergence of herd behavior [24, 37]. We show that simple Bayesian models reasonably explain the behavioral data.

Materials and Methods

Model

We denote the two options of a question by A and B Without loss of generality, we assume that A and B are the correct and wrong answers of the question q, respectively. We label the N agents $1, \ldots, N$ and denote the option that agent i ($i = 1, \ldots, N$) selects in question q by $x_i(q) \in \{A, B\}$. We denote by $P[x_i(q) = A]$ the strength of the belief (hereafter, simply the belief), with which agent i believes in A. A parallel definition is applied to $P[x_i(q) = B]$. Note that $P[x_i(q) = A]$, $P[x_i(q) = B] \geq 0$, and $P[x_i(q) = A] + P[x_i(q) = B] = 1$.

We update the agent i's belief as follows. We assume that the answer of the previous respondent j, i.e., $x_j(q)$, is generated according to the probability specified by the belief of agent j, i.e., $P[x_j(q) = A]$, which equals $1 - P[x_j(q) = B]$. Then, by using the Bayes' theorem, agent i is assumed to update the belief on the basis of the old belief and $x_j(q)$. The posterior belief of agent i is given by

$$P[x_i(q) = A]_{\text{post}}$$

$$= \frac{P\left[x_j(q) = A | x_i(q) = A\right] P[x_i(q) = A]_{\text{pre}}}{\sum_{X_i = A,B} P\left[x_j(q) = A | x_i(q) = X_i\right] P[x_i(q) = X_i]_{\text{pre}}}$$

$$= \frac{cP[x_i(q) = A]_{\text{pre}}}{cP[x_i(q) = A]_{\text{pre}} + (1-c)P[x_i(q) = B]_{\text{pre}}},$$

(1)

where $P[x_i(q) = A]_{\text{pre}}$ and $P[x_i(q) = B]_{\text{pre}}$ are prior beliefs summing up to unity. Parameter $c \equiv P[x_j(q) = A | x_i(q) = A]$ ($1/2 \leq c < 1$) represents the flexibility of agent i in response to agent j's answer. If c is close to unity, $P[x_j(q) = B | x_i(q) = A] = 1 - c$ is small such that $1 - P[x_i(q) = A]_{\text{post}}$, i.e., $P[x_i(q) = B]_{\text{post}}$ is large once agent i observes $x_j(q) = B$ for a given $P[x_i(q) = A]_{\text{pre}}$. If c is close to $1/2$, $P[x_i(q) = A]_{\text{post}}$ is insensitive to $x_j(q)$. By symmetry, we assumed that $P[x_j(q) = B | x_i(q) = B] = c$ such that $P[x_j(q) = A | x_i(q) = B] = 1 - P[x_j(q) = B | x_i(q) = B] = 1 - c$.

Iterative application of Equation (1) leads to

$$P[x_i(q) = A]$$

$$= \frac{c^{n_A - n_B} P_0[x_i(q) = A]}{c^{n_A - n_B} P_0[x_i(q) = A] + (1-c)^{n_A - n_B} \{1 - P_0[x_i(q) = A]\}}$$

(2)

and $P[x_i(q) = B] = 1 - P[x_i(q) = A]$, where n_A and n_B are the accumulated numbers of A and B responses of the previous respondents observed by agent i, respectively. The initial belief of agent i in option A is denoted by $P_0[x_i(q) = A]$. It should be noted that the order in which the previous responses are observed does not affect i's behavior. The belief of each agent i is uniquely determined by $n_A - n_B$ and the initial belief. We can rewrite Equation (2) as

$$P[x_i(q) = A] = \frac{1}{1 + ps^{n_A - n_B}} ,$$

(3)

where $p = \{1 - P_0[x_i(q) = A]\}/P_0[x_i(q) = A]$ and $s = (1-c)/c$. Previous studies used Equation (3) to account for consensus decision making by fish [38, 39].

Data set

In the present study, we use the two data sets collected in Refs. [33, 34]. The first data set, which we denote by D_1, consists of two sets of face-to-face experiments [33]. Data set D_1 consists of the results obtained from two populations of subjects each of which contains $N = 31$ subjects (KUE-A and KUE-B in Ref. [33]). Each subject went through 100 questions. Each question allowed binary options, one being correct and the other being incorrect. Generally speaking, the subjects were asked to answer each question more than once under different information conditions. We refer to a sequence of answering sessions under a given question q ($1 \leq q \leq 100$) and information condition parameterized by r as a round. Subjects went through several rounds for each question in general.

The number of rounds that a subject experienced for each question depends on the subject. The N subjects in a population were randomly assigned labels $1, 2, \ldots, N$. In the first round, all subjects answered the question without referring to others' responses. This is the memoryless

condition ($r = 0$). If everybody answered within the allocated time, there were N data points for each population and question.

The second round was implemented as follows. First, subject 1 left this question without participating in the second and following rounds. Second, subject 2 observed the answer of subject 1 in the first round and possibly updated the private answer. Similarly, subject i observed subject ($i − 1$)'s answer in the first round and possibly updated the answer, where i runs from $i = 3$ to $i = N$ in an ascending order. In the best case whereby everybody answered, $N − 1$ data points were collected in the second round. The collected data correspond to information condition $r = 1$.

The third round, corresponding to $r = 2$, was implemented as follows. First, subject 2 left without participating in the third and further rounds. Second, subject 3 observed the number of answers (n_A, n_B) submitted most recently by the previous $r = 2$ respondents and answered the question again. It should be noted that (n_A, n_B) = (2, 0), (1, 1), or (0, 2). To calculate (n_A, n_B), the answer of subject 1 in the first round and that of subject 2 in the second round were used. This is because subject 1 already left the question before the second round. In other words, the answer of subject 1 is assumed to be quenched to that made in the first round in the subsequent (i.e., second and later) rounds. Third, subject i answered after observing (n_A, n_B) calculated on the basis of the most recent choice of subjects $i − 1$ and $i − 2$, where i runs from 4 to N. There are at most $N − 2$ answers obtained from the third round.

After the third round was completed, further rounds were carried out with $r = 3, 5, 7, 9$, and ∞ in this order, where $r = \infty$ implies that the subjects can refer to the most recent answers of all the preceding respondents. Subject 3 had left before the fourth round, corresponding to $r = 3$, started. Subjects 4 and 5 had left before the fifth round, corresponding to $r = 5$, started. There are eight rounds in total. The labels of the subjects were fixed throughout the 100 questions.

The second data set, which we denote by D_2, consists of two sets of web-based experiments. They are denoted by HUE-A and HUE-B in Ref. [33] and the O and C treatments, corresponding to $r = 0$ and $r > 0$, respectively, in Exp-II in Ref. [34]. Data set D_2 consists of the results obtained from two subject populations each of which contains $N = 52$ subjects. Each subpopulation of subjects went through 120 questions. In D_2, each subject experienced up to 6 rounds, i.e., $r = 0, 1, 5, 11, 21$, and ∞ for each question. The labels of the subjects were randomly shuffled in the beginning of each question.

Results

Let us first consider the aggregate results for each experiment. As described previously, a subject answers a question after observing the number of the correct answer, n_A, and that of the incorrect answer, n_B, from the last $r = n_A + n_B$ respondents. By the aggregate results we mean that we aggregate the number of correct answers across questions for the same condition (n_A, n_B). We denote by $R(n_A, n_B)$ the number of answers obtained under condition (n_A, n_B), summed over respondents i and questions q. Out of these answers, the number of answer A, denoted by $N_A(n_A, n_B)$, is given by

$$N_A(n_A, n_B) = \sum_{i,q} x_i(q, n_A, n_B). \tag{4}$$

The fraction of A answers under condition (n_A, n_B) is given by $N_A(n_A, n_B)/R(n_A, n_B)$. This fraction for various (n_A, n_B) pairs is plotted in Fig. 1(a) and (b) for D_1 and D_2, respectively. We fit $P[x_i = A]$ given by Equation (3) to the experimental data, where we suppress q in the argument of x_i because we have aggregated the data over the questions. We estimate the values of p and s
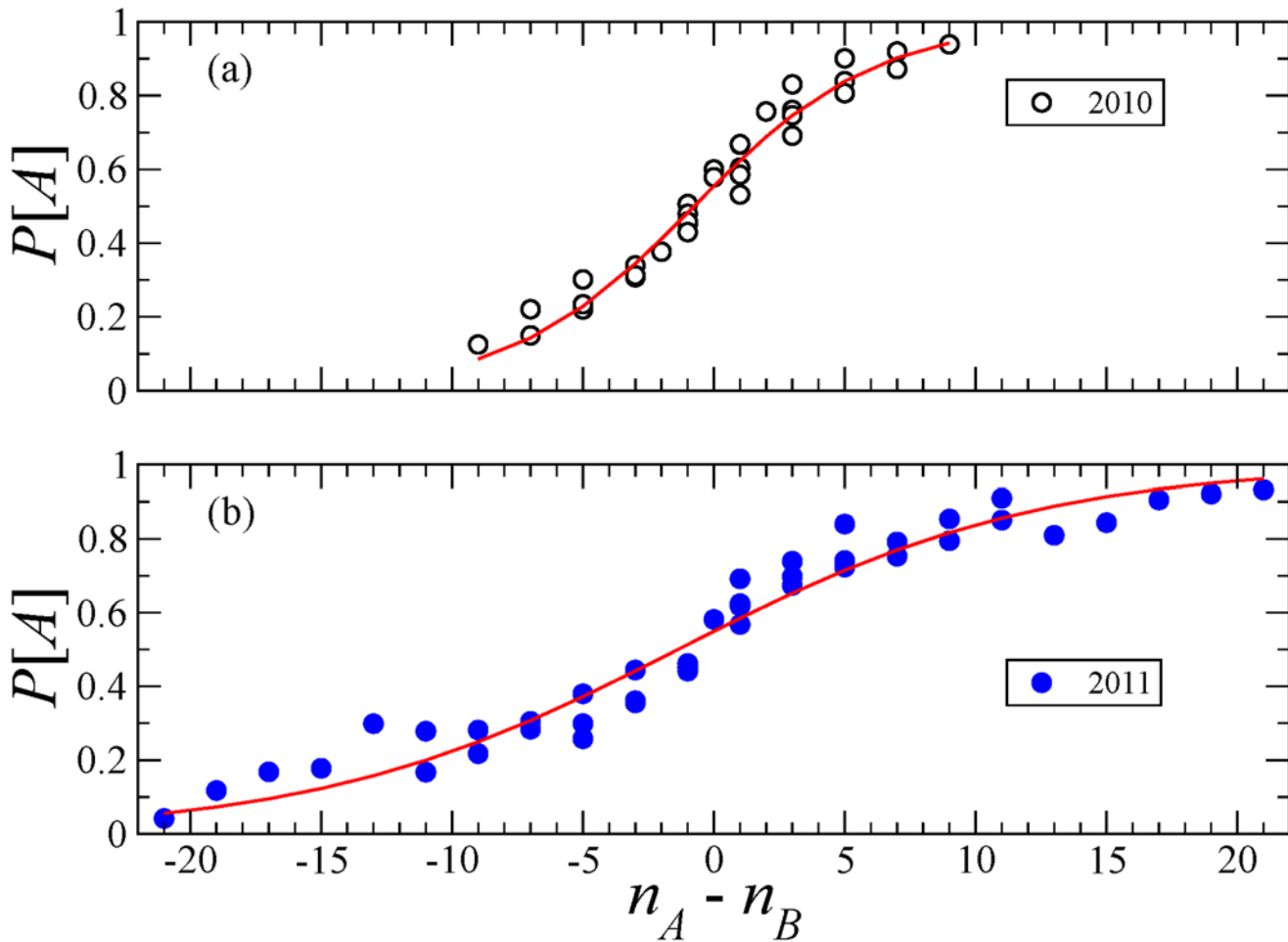
Fig 1. Bayesian inference and experimental data. We plot the probability to report a correct answer A as a function of $n_A - n_B$ for various (n_A, n_B) pairs: (a) Data set D_1, (b) Data set D_2. The circles correspond to the data. The solid curves indicate the best fits of Equation (3): $(p, s) = (0.80, 0.75)$ in (a) and $(0.82, 0.87)$ in (b).

doi:10.1371/journal.pone.0121332.g001

by an exhaustive sampling in the parameter space. For each sampled (p, s) pair, we calculate the error by the total square distance between Equation (3) and the empirical values summed over the available (n_A, n_B) pairs. The parameter values yielding the smallest error are adopted. The results of the best fitting are shown by the solid curves in Fig. 1(a) and (b) for D_1 and D_2, respectively. For D_1, the best fit is obtained for $p = 0.81$, $s = 0.75$ which lead to a root mean squared error RMSE ≈ 0.042. For D_2, we obtain $p = 0.82$, $s = 0.87$ leading to RMSE ≈ 0.059. Fig. 1 indicates that Equation (3) fits both data sets reasonably well. The value of the RMSE as a function of both parameters is shown in Fig. 2.

An alternative hypothesis of collective decision making is that $P[x_i = A]$ obeys Weber's law such that it is a function that only depends on $(n_A - n_B)/(n_A + n_B)$, or equivalently, $n_A/(n_A + n_B)$ [15, 17]. To test this hypothesis, we aggregate the data over q and i using the same aggregation as that used in Fig. 1, but separately for r to examine the effect of r on the decision making, and plot $P[x_i = A]$ as a function of $n_A/(n_A + n_B)$. The results are shown in Fig. 3(a) and (b) for D_1 and D_2, respectively. Each color corresponds to a value of $r = n_A + n_B$. If Weber's law holds true, all curves collapse on a single curve. Fig. 3 indicates that it is not the case. To be more
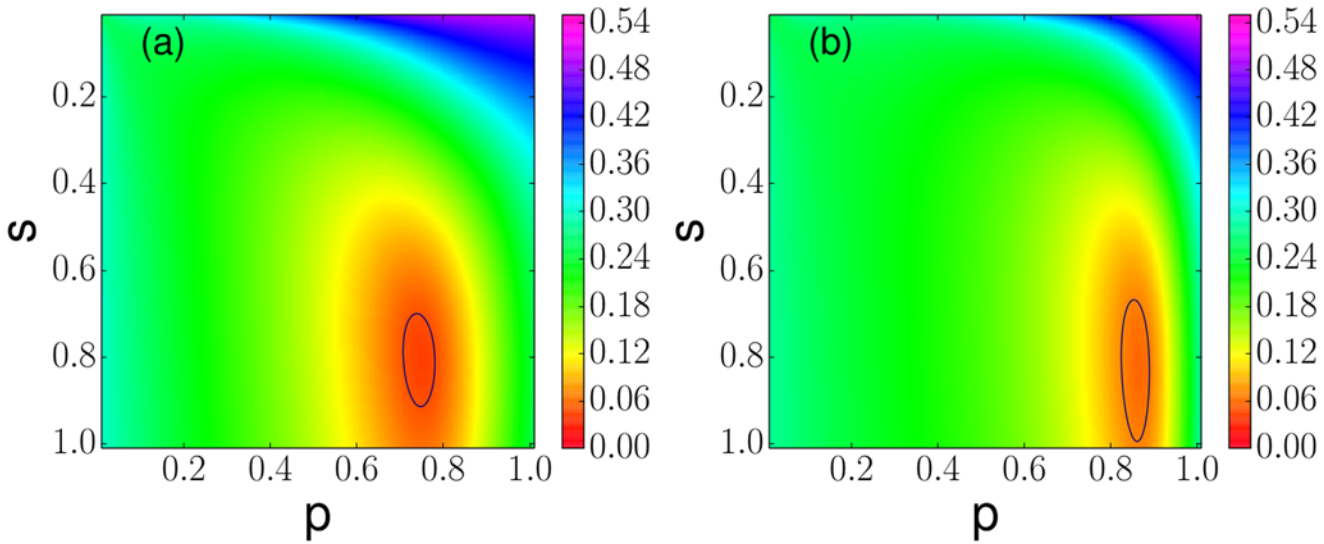
Fig 2. Parameter estimation. Root mean squared error associated to the fitting of the model given by Equation (3) to data sets (a) D_1 and (b) D_2. The contour line shows a level of 0.05 and 0.07 in (a) and (b), respectively.

quantitative, in Fig. 3(c), we plot the slope of the curves obtained by applying the least square method to the data shown in Fig. 3(a) and (b). The figure indicates that the slope increases with $r(= n_A + n_B)$ and seems to saturate. That would mean that Weber's law is correct for sufficiently large r values. Nevertheless, for the r values accessed by the experiment, Weber's law does not hold. With data for larger r values one could assess if Weber's law holds and from which r value on.

We have a reasonable fit of the data to Equation (3) even without aggregation over the questions. To show this, for a given question, we calculate the fraction of the correct answers $N_A^q(n_A, n_B)/R^q(n_A, n_B)$, where $R^q(n_A, n_B)$ is the number of answers to question q obtained under condition (n_A, n_B), and $N_A^q(n_A, n_B) \equiv \sum_i x_i(q, n_A, n_B)$ is the corresponding number of answer A. The relationship between $P[x_i(q) = A]$ and $z = (n_A - n_B) \ln s_q + \ln p_q$ for different questions is plotted in Fig. 4(a). If Equation (3) holds true, the results for different questions should collapse on a single curve $P[x_i(q) = A] = [1 + \exp(z)]^{-1}$ shown by the solid line. The results for the different questions do roughly collapse on this curve. The estimated values of p_q and s_q for individual questions are shown in Fig. 4(b). As before, we obtained parameter values p_q and s_q by sampling the parameter space and finding the values giving the smallest error. Fig. 4(b) shows that the estimated parameter values depend on the question to a large extent. For some questions, $p > 1$, implying that the initial belief in the correct answer is worse than the random coin flip, i.e., $P_0[x_i(q) = A] < 0.5$. For a majority of questions, however, the initial belief is better than the random coin flip, and for some questions, it is quite accurate (for example, $p = 0.1$ corresponds to $P_0[x_i(q) = A] = 0.91$). Another remark is that p and s are apparently uncorrelated. This implies that the flexibility of the opinion change does not depend on the difficulty of the question.

In the literature one can find different models that propose different functional forms for $P[A|n_A, n_B]$. Following [17], we fitted several of them [17, 39–42] to the current data. The quality of fitting is shown in Figs. 5 and 6 for data sets D_1 and D_2, respectively, in different colors for different values of n_B. For D_1, the best results are produced with the model in Ref. [42] with a
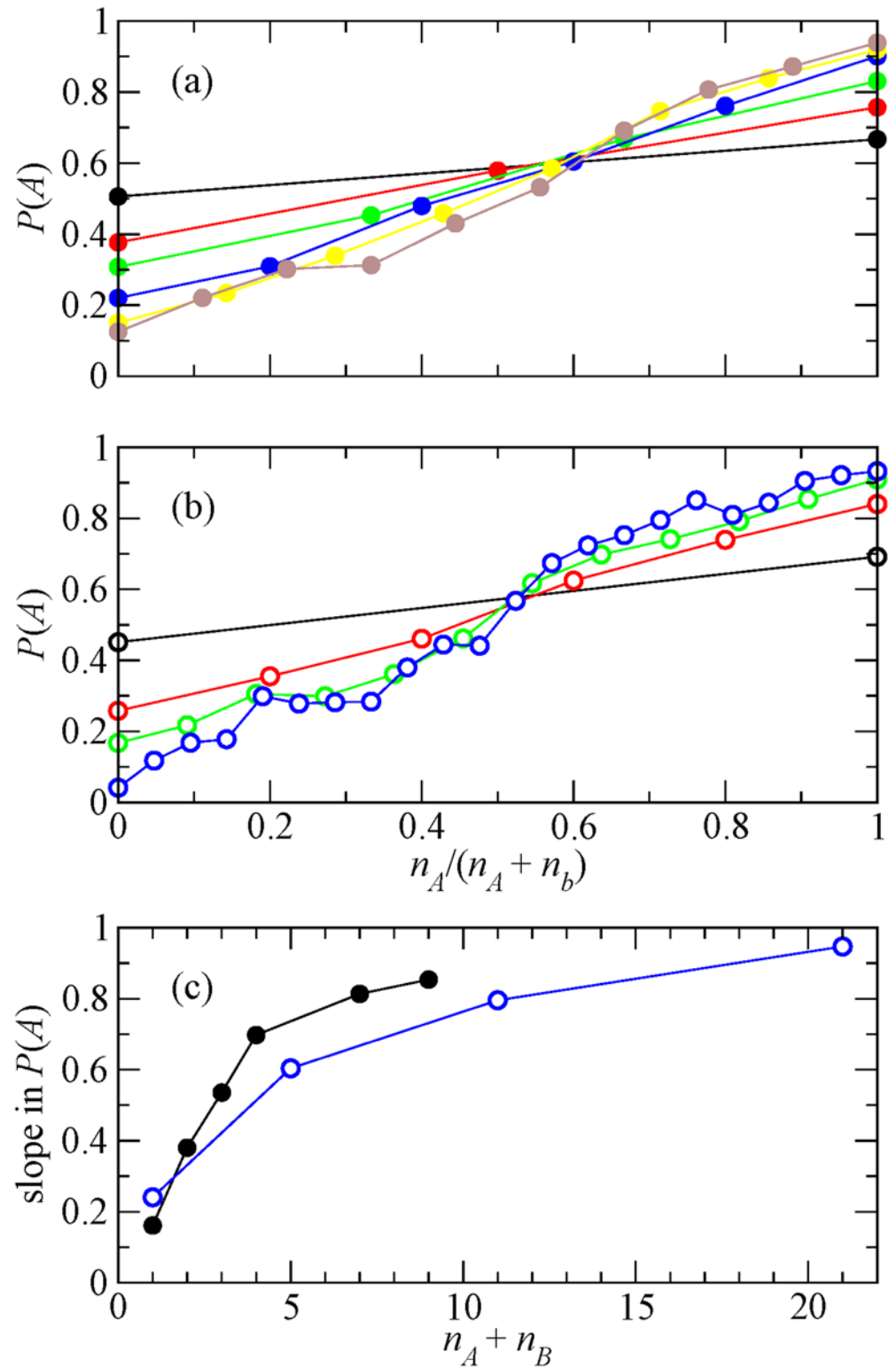
Fig 3. Dependence on the fraction of correct answers. (a) Probability to answer correctly as a function of the fraction of correct answers of the previous respondents for data set D_1. Black $r = 1$, red $r = 2$, green $r = 3$, blue $r = 4$, yellow $r = 7$, and brown $r = 9$. (b) Same results for data set D_2. Black $r = 1$, red $r = 5$, green $r = 11$, and blue $r = 21$. (c) Slope as a function of r obtained by the least square method applied to the plots in panels (a) and (b). The closed and open circles correspond to D_1 and D_2, respectively.

doi:10.1371/journal.pone.0121332.g003
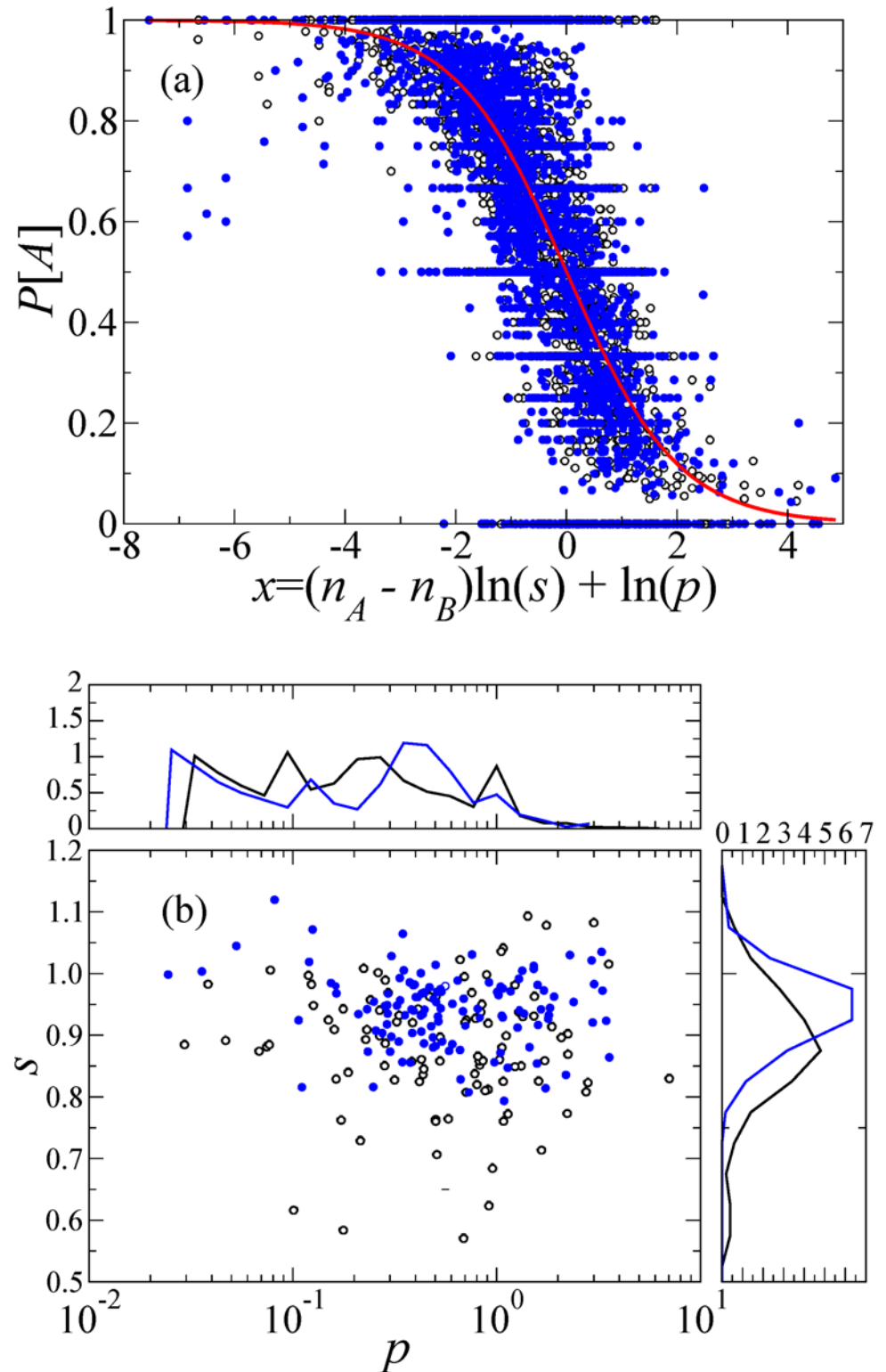
Fig 4. Dependence on the question. (a) Probability of correct answers as function of rescaled accumulated answers of previous respondents $(n_A - n_B) \ln(s) + \ln(p)$. Each symbol represents a question. Black, open symbols correspond to D_1, and blue filled circles to D_2. We estimated the s and p values for each question by applying the least square method to the data for the corresponding question. (b) Estimated p and s values for different questions. The top and side panels show the distributions of p and s, respectively.

doi:10.1371/journal.pone.0121332.g004
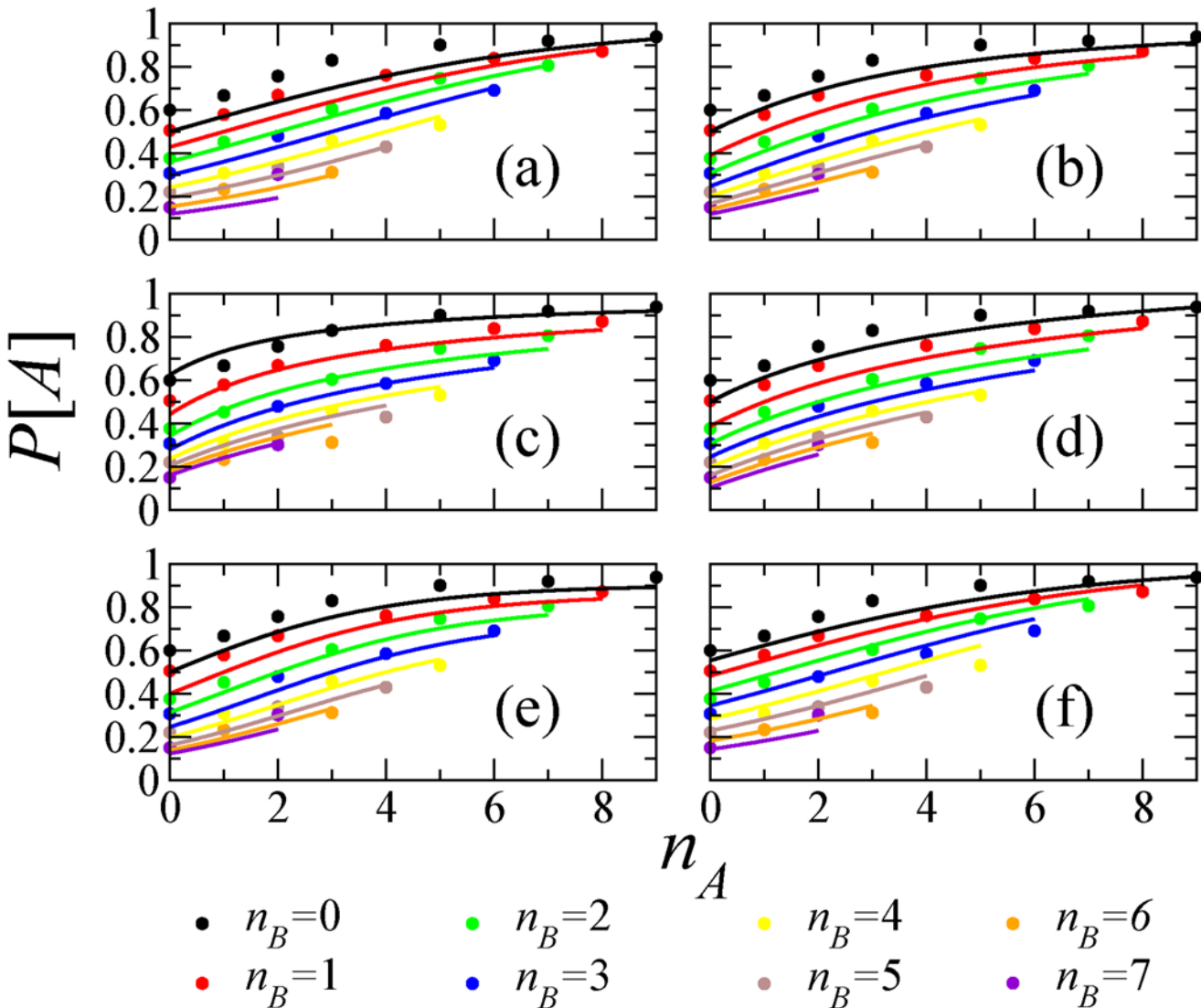
Fig 5. Testing different models for data set D_1. (a) $P[A] = \delta^{n_A}/(\delta^{n_A} + \delta^{n_B})$, $\delta = 1.33$, RMSE = 0.06, (b) $P[A] = (\delta + n_A)^{\varepsilon}/[(\delta + n_A)^{\varepsilon} + (\delta + n_B)^{\varepsilon}]$, $\delta = 3.90$, $\varepsilon = 1.95$, RMSE = 0.05, (c) $P[A] = (\delta + \varepsilon n_A)/[1 + \varepsilon(n_A + n_B)]$, $\delta = 0.63$, $\varepsilon = 0.41$, RMSE = 0.03, (d) $P[A] = 1/2 + \delta(n_A - n_B)/(n_A + n_B + \varepsilon)$, $\delta = 0.68$, $\varepsilon = 5.06$, RMSE = 0.06, (e) $P[A] = [1 + (1 + \delta\varepsilon^{-n_A})/(1 + \delta\varepsilon^{-n_B})]^{-1}$, $\delta = 8.32$, $\varepsilon = 1.60$, RMSE = 0.05, (f) $P[A] = (1 + ps^{n_A - n_B})^{-1}$, $p = 0.81$, $s = 0.75$, RMSE = 0.04. The different colors correspond to $n_B = 0$ (black), 1 (red), 2 (green), 3 (blue), 4 (yellow), 5 (brown), 6 (grey), 7 (violet).

doi:10.1371/journal.pone.0121332.g005

RMSE ≈ 0.035, followed closely by the model presented in this paper (RMSE ≈ 0.042). For D_2, the best fitting (RMSE ≈ 0.046) is produced with the model in Ref. [42], and followed closely by the models in Refs. [40, 41] (RMSE ≈ 0.053) and Ref. [17] (RMSE ≈ 0.054), with none of them being the one in this paper. See Table 1 for more information.

It should be noted that the first model in Table 1 is equivalent to a special case of our model (i.e., $p = 1$). Therefore, the fitting cannot be better than the present model. Note also that we fitted the model in Ref. [17] with $k = 0$. The result of fitting with k as a free parameter gives rise to very small values of k ($k = 0.04$ for D_1 and $k = 0.065$ for D_2), in the order of 10^{-2}. The parameter ϵ is insensitive to the small value of k being different to 0 ($\epsilon = 1.60$ for D_1 and $\epsilon = 1.30$ for D_2), while parameter δ is a much more sensitive ($\delta = 7.04$ for D_1 and $\delta = 8.72$ for D_2) (compare to
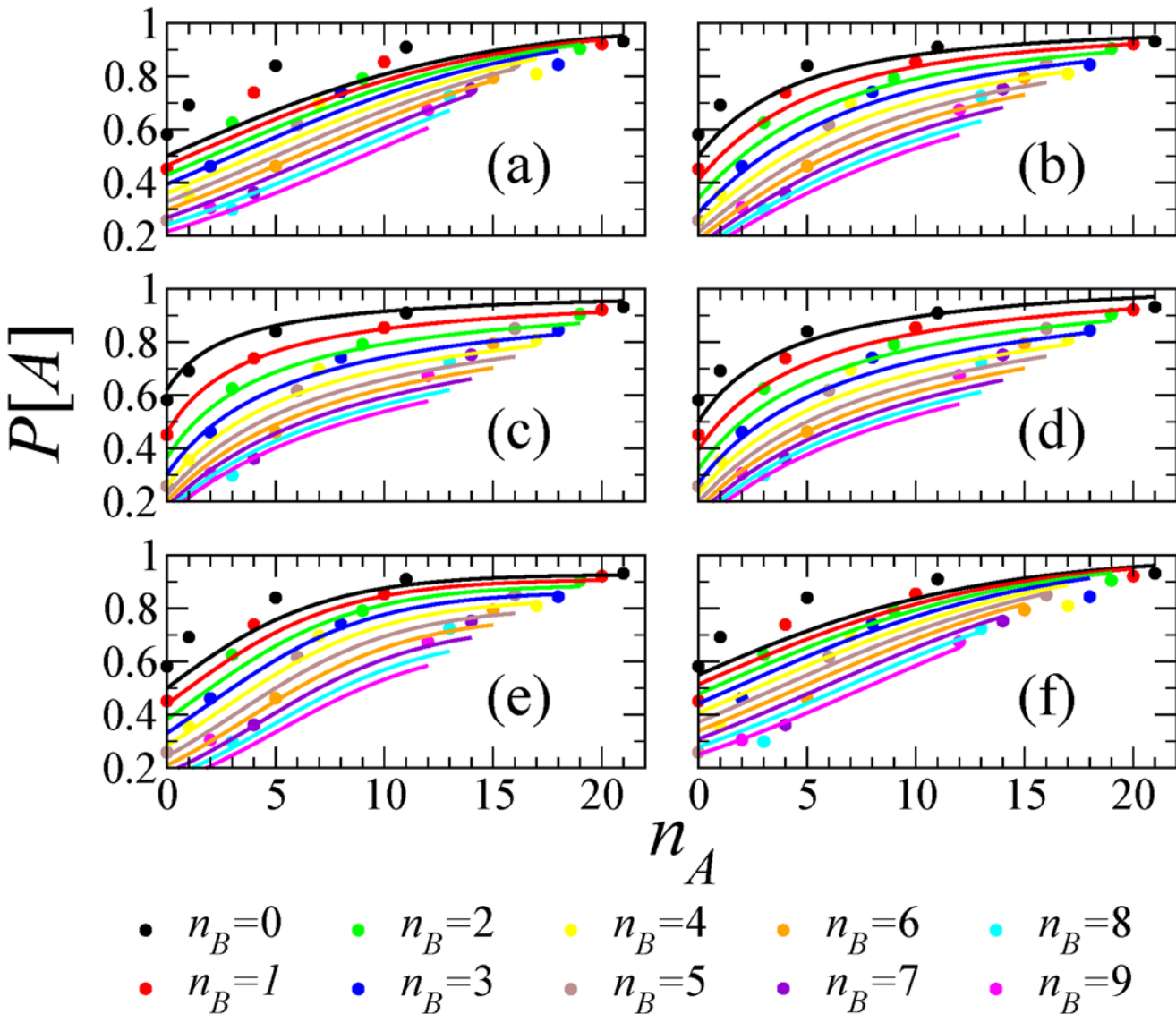
Fig 6. Testing different models for data set D_2. (a) $P[A] = \delta^{n_A} / (\delta^{n_A} + \delta^{n_B})$, $\delta = 1.15$, RMSE = 0.07, (b) $P[A] = (\delta + n_A)^\varepsilon / [(\delta + n_A)^\varepsilon + (\delta + n_B)^\varepsilon]$, $\delta = 3.66$, $\varepsilon = 1.52$, RMSE = 0.05, (c) $P[A] = (\delta + \varepsilon n_A)/[1 + \varepsilon(n_A + n_B)]$, $\delta = 0.62$, $\varepsilon = 0.35$, RMSE = 0.05, (d) $P[A] = 1/2 + \delta(n_A - n_B)/(n_A + n_B + \varepsilon)$, $\delta = 0.57$, $\varepsilon = 4.39$, RMSE = 0.06, (e) $P[A] = [1 + (1 + \delta\varepsilon^{-n_A})/(1 + \delta\varepsilon^{-n_B})]^{-1}$, $\delta = 12.01$, $\varepsilon = 1.30$, RMSE = 0.05, (f) $P[A] = (1 + ps^{n_A - n_B})^{-1}$, $p = 0.82$, $s = 0.87$, RMSE = 0.06. The different colors correspond to $n_B = 0$ (black), 1 (red), 2 (green), 3 (blue), 4 (yellow), 5 (brown), 6 (grey), 7 (violet), 8 (cyan), 9 (pink).

doi:10.1371/journal.pone.0121332.g006

results in Table 1), as the minimum in the optimization is flatter in the direction of the δ parameter, as happens also for parameter p in the present model (see Fig. 2). The quality of the fittings is of the same order as when using $k = 0$ (RMSE = 0.054 for D_1 and RMSE = 0.054 for D_2). This also happens for the zebrafish data in Ref. [17].

Discussion

We showed that the simple Bayesian model provides a quantitative agreement with behavioral data of humans sequentially answering questions with binary options. At least two other studies used the same model as ours to be fit to data in different contexts. In Ref. [39], sequential

Table 1. Fitting results for different models.

Model $P[A]$	Fitted parameters and RMSE				[Refs.]
	D_1	RMSE	D_2	RMSE	
$\frac{\delta^{n_A}}{(\delta^{n_A}+\delta^{n_B})}$	$\delta = 1.33$	0.061	$\delta = 1.15$	0.070	[39]
$\frac{(\delta+n_A)^\epsilon}{(\delta+n_A)^\epsilon+(\delta+n_B)^\epsilon}$	$\delta = 3.90, \epsilon = 1.95$	0.054	$\delta = 3.66, \epsilon = 1.52$	0.053	[40, 41]
$\frac{\delta+\epsilon n_A}{1+\epsilon(n_A+n_B)}$	$\delta = 0.63, \epsilon = 0.41$	0.035	$\delta = 0.62, \epsilon = 0.35$	0.046	[42]
$\frac{1}{2}+\delta\frac{n_A-n_B}{n_A+n_B+\epsilon}$	$\delta = 0.68, \epsilon = 5.06$	0.056	$\delta = 0.57, \epsilon = 4.39$	0.057	[15]
$\left(1+\frac{1+\delta\epsilon^{-n_A}}{1+\delta\epsilon^{-n_B}}\right)^{-1}$	$\delta = 8.32, \epsilon = 1.60$	0.054	$\delta = 12.01, \epsilon = 1.30$	0.054	[17]
$(1+ps^{n_A-n_B})^{-1}$	$p = 0.81, s = 0.75$	0.042	$p = 0.82, s = 0.87$	0.059	[here]

Results of fitting different models for $P[A]$ to data sets D_1 and D_2.

choices by fish between two identical refugia are modeled. Depending on whether the two refugia are identical or nonidentical (i.e., only one arm was with a replica predator), the unbiased prior ($p = 1$ in our notation) or a biased one ($p \neq 1$) is used, respectively. In both unbiased and biased prior cases, the authors concluded $s \approx 0.4$ (and the results are robust for $0.25 \leq s \leq 0.5$), translating into $c = 1/(s + 1) \approx 0.7$ in our notation. In another experiment with a different fish species, where fish individuals chose one of the two arms of a maze to avoid replica predators, Ward and colleagues [38] estimated $s \approx 1/e^{0.478} \approx 0.62$, translating into $c \approx 0.62$. In contrast, our results indicate $s \approx 0.7 - 0.8$ and hence $c \approx 0.56 - 0.59$. This difference may result from different species; humans may have lower responsivity to social stimuli (i.e., c value closer to 0.5) than fish (see Ref. [32] for related experiments). The type of the task may also contribute to this difference. In the current study, the data sets used are quizzes asking general knowledge of the participants. By contrast, in the fish experiments, each fish chose between two pathways that were identical except for the possible presence of a replicator predator.

Quantitatively, some models fit better to our data than the present model does, in particular for data set D_2 (Table 1). However, it should be noted that some of these previous models were proposed as fits, without particular mechanistic derivation [40–42]. Another model, i.e., the fourth model in Table 1 [17], which results from the Taylor expansion of the model proposed in Ref. [15], has mechanistic underpinning. However, the model is derived from ant's random walk on a specific arena [15]. In particular, the exit point that corresponds to the decision of one of the two alternatives is literally the spatial exit point of the animal. That may be why this model [15, 17] does not fit well to the present data. Compared to Arganda's model [17] (fifth model in Table 1), the present model fits better to data set D_1 and worse to D_2.

A way to differentiate between models is to have data on the behavior for large number of information sources (large r). In that limit the different models provide different functional forms for $P[x = A]$. Therefore, the models from Table 1 give rise to different limits $r \to \infty$. The first and the last one (model used in this paper) give rise to a step function. The second model converges to $x^\epsilon/(x^\epsilon + (1 - x)^\epsilon)$, where x is the fraction of A responses, which coincides with Weber's law for $\epsilon = 1$. However, the values of ϵ estimated for our data are much larger than unity. The third function for large r approximates the fraction of A responses. The fourth function gives $1/2 + \delta(2x - 1)$, which is a good approximation of the previous model given that the fitting parameter $\delta \simeq 1/2$ for our data sets. The fifth model gives a constant value $P[x = A] = 1/2$ in the limit $r \to \infty$. More experimental data for large r would enable the further validation of models.

There are some limitations of the present study. First, we ignored the individuality of the respondents. In fact, for each question, there should be those who know the correct answer and those who do not. Such personal knowledge can be incorporated to models for sequential answering [24, 33]. Clarifying this issue warrants future work. Second, we tried to incorporate the information about the previous responses into our model. However, the design of the experiment makes it difficult to cope with this issue. The answers offered to subject i in each round are not a random sample from the pool of responses in the previous round, but are the responses of the previous respondents $i − 1$, $i − 2$, ..., $i − r$ as initially labeled, which represents a biased sampling. Together with the influence of the history of self-responses on the new decision, these features affect the decision making process of the subjects and thus the evolution of the fraction of correct answers. Indeed in many situations individuals are not making decision from a tabula rasa but they are shaping decisions continuously from social interactions and external signals. Future developments of the theory are expected to incorporate these ingredients to deal with more realistic situations. Besides, large scale experiments taking advantage of the new technologies available would be welcome to confront with decision making theories.

Acknowledgments

Author Contributions

Conceived and designed the experiments: VME NM. Analyzed the data: VME NM JFG. Contributed reagents/materials/analysis tools: VME NM JFG. Wrote the paper: VME NM JFG.

References

1. Galton F.Vox populi. Nature 1907; 75: 450–451. doi: 10.1038/075450a0

2. Surowiecki J.The Wisdom of Crowds. New York NY: Anchor Books; 2005.

3. Lorenz J, Rauhut H, Schweitzer F, Helbing D. How social influence can undermine the wisdom of crowd effect. Proc Natl Acad Sci USA 2011; 108: 9020–9025. doi: 10.1073/pnas.1008636108 PMID: 21576485

4. Axelrod R.The Complexity of Cooperation. Princeton NJ: Princeton University Press; 1997.

5. Castellano C, Fortunato S, Loreto V. Statistical physics of social dynamics. Rev Mod Phys 2009; 81: 591–646 doi: 10.1103/RevModPhys.81.591

6. Galam S.Sociophysics: A Physicist's Modeling of Psycho-political Phenomena. New York NY: Springer; 2012.

7. Helbing D, Wenjian Y (2010) The future of social experimenting. Proc Natl Acad Sci USA 107: 5265–5266. doi: 10.1073/pnas.1000140107 PMID: 20231431

8. Szolnoki A, Wang Z, Perc M. Wisdom of groups promotes cooperation on evolutionary social dilemmas. Sci Rep 2012; 2: 576 doi: 10.1038/srep00576 PMID: 22893854.

9. McNamara JM, Houston AI. The application of statistical decision theory to animal behaviour. J Theor Biol 1980; 85: 673 – 690. doi: 10.1016/0022-5193(80)90265-9 PMID: 7442286

10. D Kahneman AT, Tversky A. Subjective probability: a judgment of representativeness. In: Kahneman D, Slovic P, Tversky A, editors. Judgment under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press; 1982. p. 32–47.

11. McNamara JM, Green RF, Olsson O. Bayes' theorem and its applications in animal behaviour. Oikos 2006; 112: 243–251. doi: 10.1111/j.0030-1299.2006.14228.x

12. Martins ACR, de B Pereira C, Vicente R. An opinion dynamics model for the diffusion of innovations. Physica A 2009; 388: 3225–3232. doi: 10.1016/j.physa.2009.04.007

13. McKay R, Efferson C. The subtleties of error management. Evol Hum Behav 2010; 31: 309 – 319. doi: 10.1016/j.evolhumbehav.2010.04.005

14. Trimmer PC, Houston AI, Marshall JAR, Mendl MT, Paul ES, McNamara JM. Decision-making under uncertainty: biases and bayesians. Anim Cogn 2011; 14: 465–476. doi: 10.1007/s10071-011-0387-4 PMID: 21360119

15. Perna A, Granovskiy B, Garnier S, Nicolis SC, Labédan M, Theraulaz G, et al. Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). PLoS Comput Biol 2012; 8: e1002592 doi: 10.1371/journal.pcbi.1002592 PMID: 22829756.

16. Andreoni J, Mylovanov T Diverging opinions. Am Econ J Microecon 2012; 4: 209–232 doi: 10.1257/mic.4.1.209

17. Arganda S, Pérez-Escudero A, de Polavieja GG. A common rule for decision making in animal collectives across species. Proc Natl Acad Sci USA 2012; 109: 20508–20513 doi: 10.1073/pnas.1210664109 PMID: 23197836

18. Johnson DD, Blumstein DT, Fowler JH, Haselton MG. The evolution of error: error management, cognitive constraints, and adaptive decision-making biases. Trends Ecol Evol 2013; 28: 474 – 481. doi: 10.1016/j.tree.2013.05.014 PMID: 23787087

19. Marshall HH, Carter AJ, Ashford A, Rowcliffe JM, Cowlishaw G. How do foragers decide when to leave a patch? a test of alternative models under natural and experimental conditions. J Anim Ecol 2013; 82: 894–902. doi: 10.1111/1365-2656.12089 PMID: 23650999

20. Marshall JAR, Trimmer PC, Houston AI, McNamara JM. On evolutionary explanations of cognitive biases. Trends Ecol Evol 2013; 28: 469 – 473. doi: 10.1016/j.tree.2013.05.013 PMID: 23790393

21. Pérez-Escudero A, Miller N, Hartnett AT, Garnier S, Couzin ID, de Polavieja GG. Estimation models describe well collective decisions among three options. Proc Natl Acad Sci USA 2013; 110: E3466–E3467. doi: 10.1073/pnas.1309867110 PMID: 23898214

22. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 2004; 27: 712–719. doi: 10.1016/j.tins.2004.10.007 PMID: 15541511

23. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND. How to grow a mind: statistics, structure, and abstraction. Science 2011; 331: 1279–1285. doi: 10.1126/science.1192788 PMID: 21393536

24. Banerjee AV.A simple model of herd behavior. Q J Econ 1992; 107: 797–817. doi: 10.2307/2118364

25. Orléan A.Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion. J Econ Behav Organ 1995; 28: 257–274. doi: 10.1016/0167-2681(95)00035-6

26. Martins ACR.Continuous opinions and discrete actions in opinion dynamics problems. Int J Mod Phys C 2008; 19: 617–624. doi: 10.1142/S0129183108012339

27. Binmore K.Rational Decisions. Princeton NJ: Princeton University Press; 2008.

28. Acemoglu D, Ozdaglar A. Opinion dynamics and learning in social networks. Dyn Games Appl 2011; 1: 3–49. doi: 10.1007/s13235-010-0004-1

29. Nishi R, Masuda N. Collective opinion formation model under bayesian updating and confirmation bias. Phys Rev E 2013; 87: 062123 doi: 10.1103/PhysRevE.87.062123

30. Milgram S, Bickman L, Berkowitz L. J Pers Soc Psychol 1969; 13: 79. doi: 10.1037/h0028070

31. Gallup AC, Hale JJ, Sumpter DJT, Garnier S, Kacelnik A, Krebs JR, et al. Proc Natl Acad Sci USA 2012; 109:7245. doi: 10.1073/pnas.1116141109 PMID: 22529369

32. Traulsen A, Semmann D, Sommerfeld RD, Krambeck HJ, Milinski M. Human strategy updating in evolutionary games. Proc Natl Acad Sci USA 2010; 107: 2962–2966. doi: 10.1073/pnas.0912515107 PMID: 20142470

33. Mori S, Hisakado M, Takahashi T. Phase transition to a two-peak phase in an information-cascade voting experiment. Phys Rev E 2012; 86: 026109. doi: 10.1103/PhysRevE.86.026109

34. Mori S, Hisakado M, Takahashi T. Collective adoption of max–min strategy in an information cascade voting experiment. J Phys Soc Japan 2013; 82: 084004. doi: 10.7566/JPSJ.82.084004

35. King AJ, Cheng L, Starke SD, Myatt JP. Is the true' wisdom of the crowd' to copy successful individuals? Biol Lett 2012; 8: 197–200. doi: 10.1098/rsbl.2011.0795 PMID: 21920956

36. Moussaïd M, Kämmer JE, Analytis PP, Neth H. Social influence and the collective dynamics of opinion formation. PLoS ONE 2013; 8: e78433 doi: 10.1371/journal.pone.0078433

37. Bikhchandani S, Hirshleifer D, Welch I. A theory of fads, fashion, custom, and cultural change as informational cascades. J Polit Econ 1992; 100: 992–1026. doi: 10.1086/261849

38. Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J. Fast and accurate decisions through collective vigilance in fish shoals. Proc Natl Acad Sci USA 2011; 108: 2312–2315. doi: 10.1073/pnas.1007102108 PMID: 21262802

39. Pérez-Escudero A, de Polavieja GG. Collective animal behavior from bayesian estimation and probability matching. PLoS Comput Biol 2011; 7: e1002282. doi: 10.1371/journal.pcbi.1002282 PMID: 22125487

40. Goss S, Aron S, Deneubourg JL, Pasteels JM. Self-organized shortcuts in the Argentine ant. Naturwissenschaften 1989; 76: 579–581. doi: 10.1007/BF00462870

41. Deneubourg JL, Aron S, Goss S, Pasteels JM. The self-organizing exploratory pattern of the Argentine ant. J Insect Behav 1990; 3: 159–168. doi: 10.1007/BF01417909

42. Meunier H, Leca JB, Deneubourg JL, Petit O. Group movement decisions in capuchin monkeys: the utility of an experimental study and a mathematical model to explore the relationship between individual and collective behaviours. Behaviour 2006; 143: 1511–1527 doi: 10.1163/156853906779366982