Supporting Information

Miller et al. 10.1073/pnas.1217513110

SI Text

Derivation of Theoretical Distributions

To determine whether groups in our experiment were able to integrate the information held by their different members, we compared the distributions of fish in each of the three reward arms to the distributions expected if the fish were ignoring social information and choosing a reward arm based solely on their personal information (Fig. 2*C*). Here we derive the theoretical distributions.

If each individual chose an arm based only on its personal information, we would expect the proportion of the group that entered any given arm to be modeled by a binomial distribution. To parameterize this distribution we require only the probability with which individuals choose that arm based on their personal information. However, our test trials were conducted using mixed groups, consisting of equal numbers of individuals from training groups A and B. These individuals, due to their different training histories, would be expected to choose between the three arms with different probabilities. Thus, the probability of finding a given number of the mixed group in any arm must take into account how many of those individuals were from each training group. To find this probability, we sum the binomial distribution over all possible combinations of members of each training group. Thus, the probability that n individuals chose any arm Y is given by

$$P(n_Y) = \begin{cases} \text{If } n \le \frac{N}{2},\ \sum_{i=0}^{n} \binom{N/2}{i} A_Y^i (1-A_Y)^{\left(\frac{N}{2}-i\right)} \times \binom{N/2}{n-i} B_Y^{n-i} (1-B_Y)^{\left(\frac{N}{2}-n+i\right)} \\ \text{If } n > \frac{N}{2},\ \sum_{i=n-N/2}^{N/2} \binom{N/2}{i} A_Y^i (1-A_Y)^{\left(\frac{N}{2}-i\right)} \times \binom{N/2}{n-i} B_Y^{n-i} (1-B_Y)^{\left(\frac{N}{2}-n+i\right)}, \end{cases} \quad \text{[S1]}$$

where N is the total size of the group (always 16 in our experiment) and A_Y and B_Y are the probabilities that individuals from training groups A and B, respectively, choose arm Y based on their personal information. The first term of the sum gives the probability that i individuals from training group A chose arm Y; the second term gives the probability that the remaining $n - i$ individuals from training group B chose the same arm.

We estimate the values of A_Y and B_Y from the choices of the lead individual in each test trial (rank 1) because these individuals have no social information and so are most likely to choose an arm based solely on their personal information. Because lead fish chose one of their preferred arms on about 80% of test trials (*Results*), we set the probability of choosing a preferred arm at 0.4; the probability of choosing a nonpreferred arm is therefore 0.2. These values give the distributions shown in Fig. 2*C* (solid lines). Note that for the consensus arm, which both training groups prefer, Eq. **S1** reduces to a simple binomial, because A_Y and B_Y are identical for this arm.

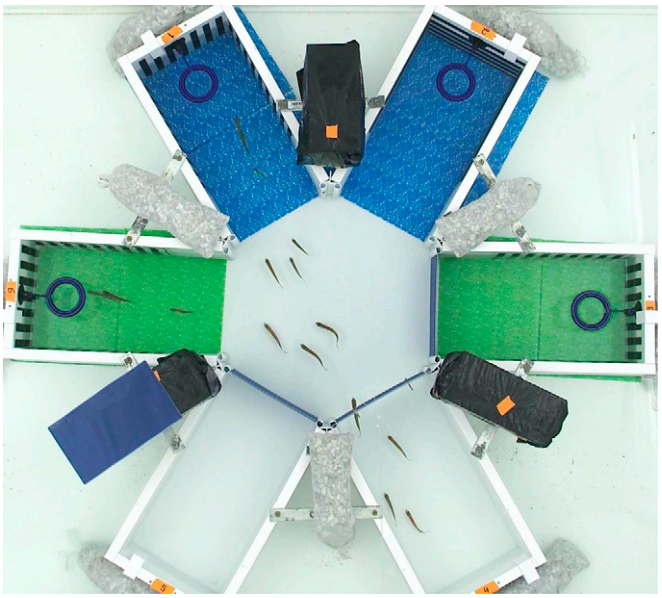


Fig. S1. Sample video frame from a test trial. The fish are shown leaving the start arm at the bottom right (under the opened door) and arriving at the reward arms at the top (stripe arm at left; consensus arm, center; and color arm, right). The barred reward arm at far right was used on alternate trials to counterbalance turning directions. Blue circles are floating feeding rings in which food was placed during training trials. Black rectangles are bricks placed on top of the maze to weight it down in the water.

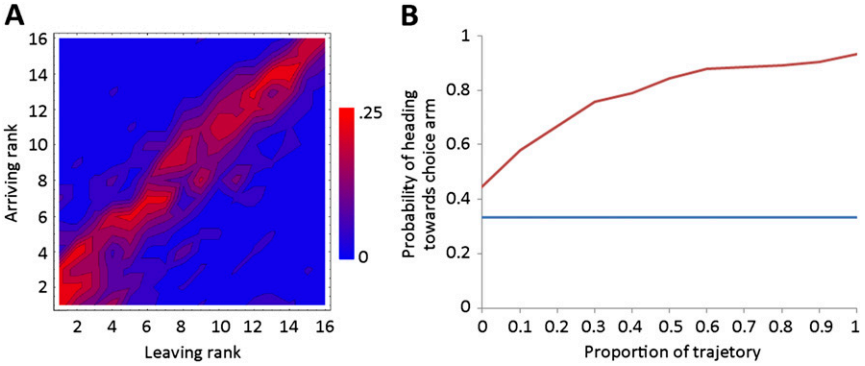


Fig. S2. (*A*) Density plot of choice rank at the start and end of a trajectory. Data are the probability of being at a particular rank when entering a reward arm, given a certain rank when leaving the start arm. Most of the data cluster along the diagonal, demonstrating that individuals usually maintain their rank between leaving the start arm and entering a reward arm. (*B*) Probability that an individual is heading toward the arm that it eventually chooses, as a function of the proportion of its trajectory it has already traversed (red line). Trajectories are the entire path of the fish from the moment it leaves the start arm to the moment it first enters one of the reward arms. The blue line indicates chance levels.

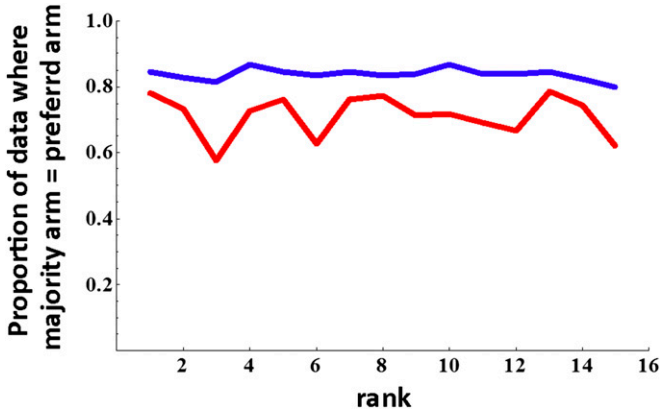


Fig. S3. Proportion of data where social and personal information correlate. The figure shows the proportion of the data at each choice rank in which the majority arm was also an individual's preferred arm in the mixed tests (red) and the unmixed control tests (blue). Due to the partial correlation of the preferences of individuals, even in mixed tests, there were relatively few cases in which the majority arm was a nonpreferred arm.

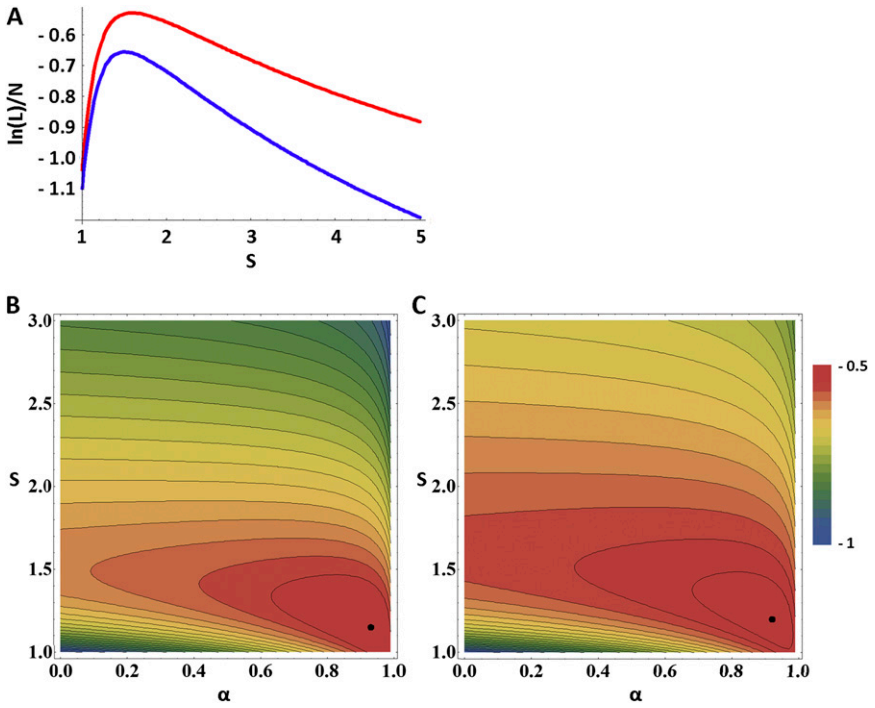


Fig. S4. Goodness-of-fit for the pure Bayesian and isolation-averse Bayesian models. (*A*) Pure Bayesian model fit to the mixed test (blue) and unmixed control test (red) data, as a function of *s*, and isolation-averse Bayesian model fit to the mixed test (*B*) and unmixed control test (*C*) data, as a function of its two parameters, *s* and α. The black circle marks the optimal parameter values for each dataset; $s = 1.15$, $\alpha = 0.93$ for the mixed data (*B*), $s = 1.2$, $\alpha = 0.92$ for the unmixed data (*C*). All the plots show the maximum likelihood estimate (MLE) for the models; MLE = ln(*L*)/*N*, where *N* is the number of data points and *L* is the likelihood score (the likelihood is divided by *N* to permit comparison of the two differently sized datasets).

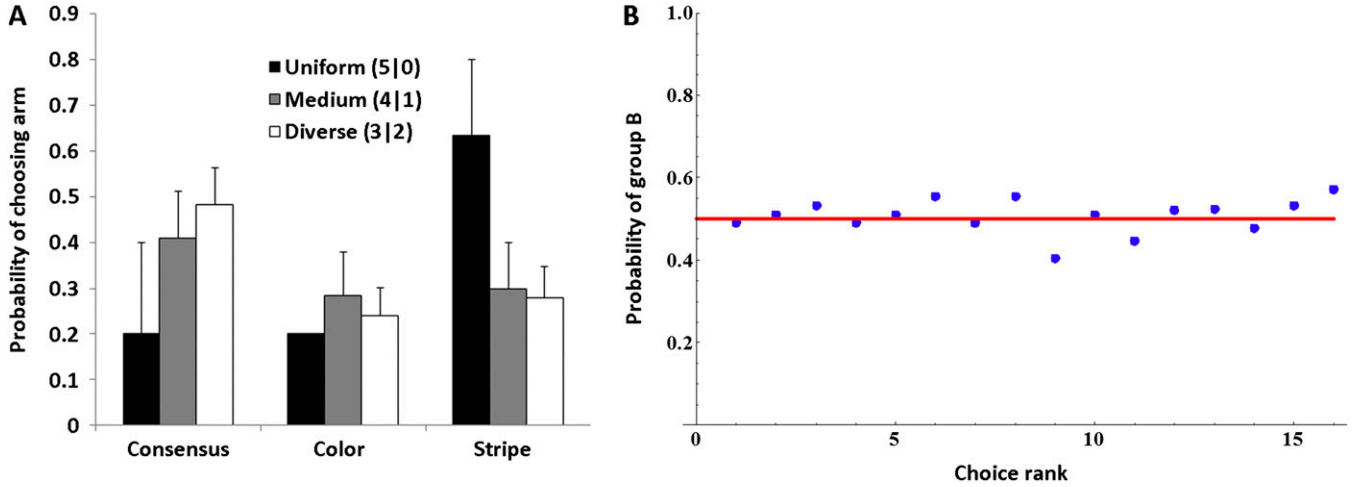
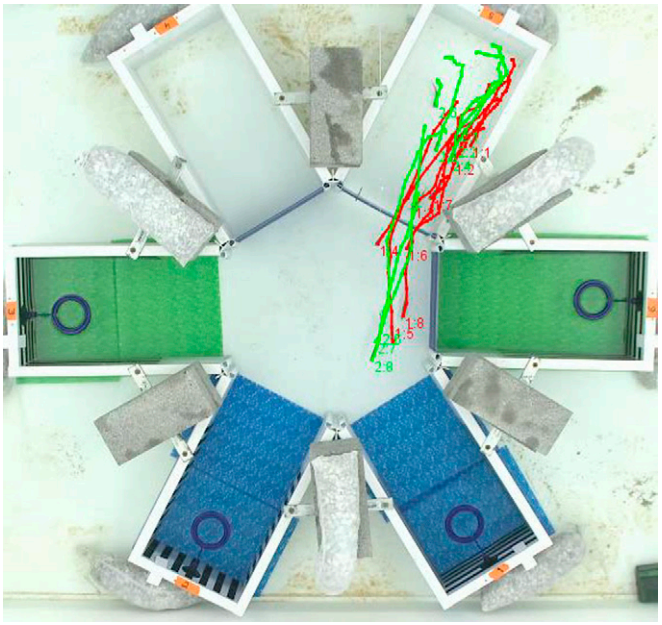


Fig. S5. The importance of informational diversity and mixing. (*A*) Distributions of arm choices in mixed tests as a function of the diversity of information in the lead five ranks. Trials were separated by whether the first five fish to make a choice were all from the same training group (Uniform), four were from the same training group (Medium), or whether there were three from one training group and two from the other (Diverse). More diverse groups exhibit a stronger bias toward the consensus arm. However, uniform and medium groups are rare, and the comparisons were therefore not significant [paired-sample *t* tests: uniform vs. medium, $t(15) = 0.71$, $P = 0.49$; uniform vs. diverse, $t(30) = 0.88$, $P = 0.39$; medium vs. diverse, $t(43) = 0.52$, $P = 0.61$]. (*B*) Probability of finding an individual from group B at any choice rank (blue circles) in mixed tests. The red line indicates $P = 0.5$. These data demonstrate that our groups were well-mixed, with neither training group (A or B) more likely to be found at any choice rank, and explain why uniform and medium groups so rarely occurred.



Movie S1. Sample test trial video. The video shows the first few seconds of one test trial with the tracks of all the fish overlaid on the image. Green tracks indicate fish from Group A (trained, in this example, to blue floors), red tracks fish from Group B (trained to horizontal wall stripes). The consensus arm is at bottom right, the color arm at bottom left, and the stripe arm at far left.

Movie S1