

The integration of social influence and reward: Computational approaches and neural evidence

Damon Tomlin¹ • Andrea Nedic² • Deborah A. Prentice³ • Philip Holmes^{4,5,6} • Jonathan D. Cohen^{3,4}

Published online: 24 May 2017 © Psychonomic Society, Inc. 2017

Abstract Decades of research have established that decisionmaking is dramatically impacted by both the rewards an individual receives and the behavior of others. How do these distinct influences exert their influence on an individual's actions, and can the resulting behavior be effectively captured in a computational model? To address this question, we employed a novel spatial foraging game in which groups of three participants sought to find the most rewarding location in an unfamiliar two-dimensional space. As the game transitioned from one block to the next, the availability of information regarding other group members was varied systematically, revealing the relative impacts of feedback from the environment and information from other group members on individual decision-making. Both reward-based and socially-based sources of information exerted a significant influence on behavior, and a computational model

Electronic supplementary material The online version of this article (doi:10.3758/s13415-017-0512-1) contains supplementary material, which is available to authorized users.

Damon Tomlin dtomlin@uccs.edu

- ¹ Department of Psychology, University of Colorado, 1420 Austin Bluffs Parkway, Colorado Springs, CO 80918, USA
- ² Princeton Consultants, New York, NY, USA
- ³ Department of Psychology, Princeton University, Princeton, NJ, USA
- ⁴ Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA
- ⁵ Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA
- ⁶ Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ, USA

incorporating these effects was able to recapitulate several key trends in the behavioral data. In addition, our findings suggest how these sources were processed and combined during decision-making. Analysis of reaction time, location of gaze, and functional magnetic resonance imaging (fMRI) data indicated that these distinct sources of information were integrated simultaneously for each decision, rather than exerting their influence in a separate, all-or-none fashion across separate subsets of trials. These findings add to our understanding of how the separate influences of reward from the environment and information derived from other social agents are combined to produce decisions.

Keywords Decision-making \cdot Computational model \cdot Social cognition \cdot Neuroimaging

Introduction

In order to survive, human beings must navigate a complex world – a world filled with rich sensory stimuli and an enormous collection of possible behavioral responses. From this assortment of potential behaviors, we must select a small number (possibly a single action) that we deem most appropriate based upon our circumstances. Further complicating this issue is the fact that we must base our decisions upon multiple stimuli that exist in a variety of domains. How do human beings process these separate sources of information in order to arrive at a decision?

One domain that has a demonstrable impact on behavior is that of rewards – material incentives that organisms expend time and effort to acquire. Thanks to psychological and neuroscientific research, we are beginning to understand rewardbased behavior in both animals (e.g., Herrnstein, 1961; Schultz, Dayan, & Montague, 1997; Sugrue, Corrado, & Newsome, 2005) and the human brain (e.g., Berns et al., 2001; Breiter et al., 2001; Delgado et al., 2003; Knutson et al., 2008; McClure, Berns, & Montague, 2003; O'Doherty et al., 2006). Previous work has also examined reward-based decisionmaking from a computational perspective in order to determine how the brain tracks rewards and adjusts the decision process in response to environmental constraints (Bogacz et al., 2006; Montague, Dayan, & Sejnowski, 1996). Importantly, recent accounts of decision-making have emphasized the need for a "common currency," which would transform outcomes so that their values can be compared and combined (Montague & Berns, 2002; Sugrue, Corrado, & Newsome, 2005).

In a similar vein, disparate sources of information that influence actions (rather than directly predict outcomes of such decisions) must also be compared and combined, even when the domains of those sources do not easily lend themselves to comparison and may actually provide conflicting indications of which action should be chosen. One such domain is that of social influence (Ruff & Fehr, 2014). Previous work has shown that, under the proper circumstances, groups of interacting decision-makers exhibit performance better than that exhibited by isolated individuals (Mason & Watts 2012; Surowiecki 2005). However, the impact of others' actions (e.g., Asch, 1951; Milgram, 1963) and inaction (e.g., Darley & Latane, 1968; Latane & Darley, 1968) on an individual's behavior is not always positive, and evidence for such social influence is extensive and robust. One mechanism at play here is the strong tendency toward social imitation, which allows individuals to converge on a given behavior. This behavior can be productive when it helps a child learn appropriate social behaviors (Meltzoff & Moore, 1977), but mutual imitation also has a destructive counterpart in social herding, in which individuals converge on a behavior based on mutual imitation, without an accompanying evaluation of their experience. At the aggregate level, these influences contribute to the development of convergent social behavior within an interacting group. Research has suggested that this phenomenon may have an impact as far reaching as the global economy (Bikhchandani & Sharma, 2000), and recent work has emphasized that this behavior is ripe for economic, psychological, and neuroscientific investigation (Baddeley, 2010; Raafat, Chater, & Frith, 2009). Furthermore, it has been proposed that computational models of social behavior, when informed by measures of neural activity, could yield valuable insights regarding these important social processes in both the population at large (Stanley & Adolphs, 2013; Toelch & Dolan, 2015) and in populations exhibiting social dysfunction (Montague et al., 2012).

Several groups have elucidated potential neural substrates for social influence. In these studies, participants made decisions while witnessing the decisions or recommendations of other "group members." Participants made judgments and expressed preferences across a range of stimulus types: three-dimensional shapes (Berns et al., 2005), human faces (Klucharev et al., 2009), and abstract symbols (Mason, Dver, & Norton, 2009), to name a few. However, these previous studies have typically called for decisions for which no objective measure of performance was possible, or for which such objective feedback was withheld, thereby precluding the examination of how social information was combined with information derived from the participant's own experience. In addition, because such experiments typically involve one genuine participant (thereby relegating that participant to the role of "observer"), relatively little information exists regarding the dynamics of groups as their behavior evolves over the course of multiple decisions (Schilbach et al., 2013). The work presented here seeks to address this issue by using interacting groups of participants.

How do reward-based and socially-based influences combine within the human brain to produce behavior? A building convergence of evidence suggests that three cortical regions are particularly important for the evaluation of stimuli and actions: the orbitofrontal cortex (OFC), anterior cingulate cortex (ACC), and dorsolateral prefrontal cortex (dlPFC). Each of these regions appears to be at least partially specialized, with the OFC and ACC reflecting the subjective value of stimuli and actions, respectively (Rangel & Hare, 2010; Rudebeck et al., 2008), and the dIPFC modulating the values of stimuli (Camus et al., 2009; Hare, Camerer, & Rangel, 2009; Rangel & Hare, 2010). The ultimate goal of this framework is to understand how the brain evaluates environmental stimuli and potential actions in terms of value - a single numerical quantity that would render any two stimuli or actions readily comparable. But what happens when actions are assigned equal (or approximately equal) values? In addition to its role in action evaluation, the ACC appears to be critical for the detection of response conflict - the competition between mutually exclusive representations or behaviors when they are triggered simultaneously by environmental stimuli (Botvinick et al., 2001; Carter et al., 2000). Once such a conflict is detected, an additional function of the dlPFC, and one particularly critical for the work presented here, is to exert topdown cognitive control in order to select one of the competing actions (Carter & van Veen, 2007) - an intervention that entails additional processing time (Botvinick et al., 2001; Carter et al., 2000). The dIPFC's role as arbiter between conflicting behaviors also extends into the social domain, such that it serves a critical role in maintaining norms for social exchange (Knoch et al., 2006) and ethical decision-making (Greene et al., 2001). Thus, these regions seem to be responsible not only for evaluating potential actions, but for selecting an action when the evaluation process does not yield an obvious choice.

A 2008 study by Behrens et al. examined the case in which reward-based information and information from another social agent were simultaneously presented, with the social information in the form of a simple, binary recommendation that could be immediately evaluated in terms of its intent and its validity - that is, it was easily transformed into the same domain as the reward itself (Behrens et al., 2008). This study revealed that, under these circumstances, the two streams of information were combined into a single value in the ventromedial prefrontal cortex (vmPFC), rather than processed as separate, potentially conflicting, influences. It has also been suggested that this region not only integrates information regarding socially and reward-based sources in order to predict outcomes, but that the social information so integrated need not be derived from the context of the current decision (Hackel et al., 2015). In this experiment, the vmPFC responded not just to the momentarily available social information, but to more general properties of the agent – a long-term trait that yielded important information regarding expected outcomes. The vmPFC may also process socially-based and reward-based information in a different sense: by estimating rewardrelated effects such as prediction errors experienced by others, it may allow an individual to better anticipate the behavior of other social agents (Suzuki et al., 2012). Recent work by Boorman et al. has contributed to this body of work by examining how the relative influence of information from other social agents can shift over time (Boorman et al., 2013). A model-based analysis revealed that the medial prefrontal cortex and ACC were involved in updating an individual's beliefs regarding the reliability of information coming from other agents, especially when the agents were human. That is, the influence of information received from the environment seemed to depend not only on its value with regard to reward, but also to its social nature. This idea has also been reinforced by data suggesting that, while areas such as the dIPFC adjust behavior according to both socially-based and non-sociallybased information, there may be regions within the human brain that learn preferentially (or perhaps exclusively) about outcomes coming from social agents, even when a non-social agent yields outcomes that are otherwise identical to those of the social agent (Stanley, 2015).

Prior research has also used economic tasks to examine participants' internal models of social behavior (Chang and Sanfey, 2011; Fareri, Chang, & Delgado, 2012; Xiang, Lohrenz, & Montague, 2013). In these studies, participants' earnings were based upon the decisions made by their partners, and participants' expectations regarding partner behavior were generated and updated using procedures based upon re-inforcement learning. While the aforementioned experiments are valuable for examining the behavioral and neural responses that occur when these expectations are violated, they share a crucial feature: their models generate expectations and values that are already in the domain of the reward-based outcome (i.e., the participant anticipates a certain number of points from his or her partner each round). Despite these

findings, it remains to be seen how the regions of the brain mentioned above select actions when reward-based information and social influence are present simultaneously, especially when that social influence cannot be readily transformed into an explicit value in the reward-based domain.

Here, we add to our understanding of the combined influence of different types of information through the use of a novel decision-making paradigm and an accompanying computational model. The experiment presented here examined the processing of reward-based and social stimuli by measuring behavioral and neural responses (the latter by using functional magnetic resonance imaging, fMRI) simultaneously from all members of a group of interacting participants over the course of multiple decisions. This study built on earlier work identifying a set of regions within the brain (most notably the insula, along with the dorsolateral prefrontal cortex, dorsal anterior cingulate, and parietal cortex) that are responsive to the degree of similarity between an individual's decisions and those of a group, and suggesting that these regions bias choices toward socially convergent behavior (Tomlin et al., 2013). However, these previous findings do not provide insight into how rewards, social information, and the history of an individual's choices are all used to produce decisions.

In this study, we sought to determine the individual impact of these influences, as well as the process and brain regions under which these influences are combined to make a decision. To accomplish the latter goal, we fit the behavioral data with a computational model that quantified these sources of information, as well as their effect upon the decision process. We then further refined this model through the analysis and interpretation of reaction times, location of gaze, and neuroimaging data.

Materials and methods

Participants

Groups of three participants carried out a social decisionmaking task, with groups separated into an eye-tracking cohort and a neuroimaging cohort. Data for the eye-tracking cohort were collected at Princeton University, New Jersey, while data for the neuroimaging cohort were collected at Baylor College of Medicine in Houston, Texas.

Participants in the eye-tracking cohort were recruited via Princeton's paid experiment website and word of mouth, and informed consent was obtained according to protocols approved by Princeton University's Institutional Review Board. Groups of three participants engaged in a series of decision-making tasks while data were recorded logging their responses, direction of gaze, and pupil diameter. Participants did not meet prior to the experiment, nor were they required to see one another afterward (though chance meetings in the hallway were possible after the experiment). However, behavioral suites were adjacent to one another so that incidental noise could be heard between suites. In addition, participants had to wait for all group members to be ready before the experiment could begin. As a result, no participant indicated doubt in the social nature of the experiment. The eye-tracking cohort consisted of 84 groups (n = 252 individuals; 148 female, 104 male; ages 18–45 years, with a mean age of 20 years).

Participants in the neuroimaging cohort were recruited at Baylor College of Medicine via email and word of mouth, and informed consent was obtained according to protocols approved by Baylor College of Medicine and Princeton University's Institutional Review Boards. Groups of three participants engaged in a series of decision-making tasks while fMRI data were acquired simultaneously (Montague et al., 2002), thus maximizing the amount of neuroimaging data collected from each group. Participants did not meet prior to the experiment, nor did they see one another afterward. However, the neuroimaging subject population at Baylor College of Medicine is accustomed to multi-agent experiments and, as above, participants had to wait for all group members to be ready before the experiment to begin. As with the eye-tracking cohort, no participant indicated doubt that the experiment involved a legitimate social group. The neuroimaging cohort consisted of 24 groups (n = 72 individuals; 47 female, 25 male; ages 18-57 years, with a mean age of 32 years). Of these 72 participants, eight were excluded for excessive movement during the scanning session (excessive movement being defined as more than 1 mm of translation from the participant's original position along any axis across the duration of the study) and an additional two were excluded for not responding during an entire block of the study, leaving a neuroimaging subset of 62 participants (40 female, 22 male; ages 18–53 years, with a mean age of 33 years).

Spatial decision-making tasks with social information

Groups of three participants carried out a decision-making game in which they each controlled a red marker (a dot) that they could move within the confines of a white rectangle. The game was presented via a monitor (eye-tracking cohort) or rear projection screen (neuroimaging cohort). Participants could move in one of four directions – up, down, left, or right – using a set of four buttons (keys on a standard keyboard for the eye-tracking cohort; two buttons on each of two optical button boxes for the neuroimaging cohort). Locations within the rectangle were discretized into ten possible positions along each dimension (horizontal and vertical); the rectangle therefore represented a 10×10 grid of 100 possible locations. Each decision by the participant changed his or her location by one increment in the selected direction, except those trials for which the participant did not make a choice and remained in

the same location for the next trial. If a participant was located along the edge of the rectangle and made a choice that would place him or her outside of the rectangle's border, the choice was registered, no other choices were permitted during that trial, and the participant remained in the current location for the next trial. Participants were instructed that the number of points earned on each trial depended upon the red dot's location (henceforth "his or her location") within the white rectangle on that trial, with some locations yielding more points than others. Participants were also instructed that the experiment would consist of four separate tasks, each consisting of 90 trials, and that they should try to earn as many points as they could since their total points would determine their compensation after the experiment. This paradigm placed decisions and their outcomes in a quantifiable domain (points), in contrast with other studies that have examined the impact of social information on preferences - a domain in which generalizable computational models are difficult to construct due to the idiosyncratic nature of the decisions and the inability to compare outcomes in an objective manner (e.g., Berns et al., 2005; Klucharev et al., 2009; Mason, Dyer, & Norton, 2009).

Each trial within a task consisted of a screen that reflected the participant's current location as well as the number of points he or she had earned based on that location (Fig. 1A). At the beginning of each trial, the screen was updated to reflect this information and the participant had 3 s to make a choice for the trial. A set of cartoon arrows confirmed the choice - the arrow corresponding to the chosen direction remained white, while the other arrows changed to gray. If the participant did not make a choice during the 3-s period, all the arrows became gray, and the red dot stayed in its current location for the next trial. Once the participant had either made a choice or allowed the 3-s period to elapse, the non-selected arrows remained gray and the participant could not register a new choice until the total trial length reached 6 s. After the total trial length reached 6 s, the screen was updated to reflect the participant's new location and most recent earnings, all the arrows became white once more, and a new trial began. The total duration of each trial was kept constant so that the onset of each trial could be synchronized across the three participants in the group. That is, trials occurred every 6 s, with the timing between a participant's individual choice and the onset of the next trial varying according to the time the participant took to make a choice.

The experiment consisted of four separate tasks, with each task defined by a combination of a "social condition" and a predefined "reward function." The onset of a new task, which always corresponded to a reward function and social condition the participant had not yet experienced, was preceded by a screen with the words "New Task," as well as text indicating the social condition for the new task. The social conditions defined what information was given to each member of the



Fig. 1 Timing and social conditions in the experiment. Groups of three participants performed a decision-making task under each of four "social conditions" that determined what information was available about other group members. Social conditions and reward functions were pseudorandomly paired across groups and randomly ordered at the beginning of each experiment. (A) In the task, trial onsets were synchronized and occurred every 6 s. Participants had up to 3 s to select a direction - if none was selected, the screen timed out and the corresponding marker remained in its location for the next trial (blue participant). If the participant made a selection, the screen timed out immediately and the participant's location was updated at the beginning of the next trial (red and green participants). At the beginning of the next trial, points earned for the current locations were updated and participants could make another response. Time-out periods were individually tailored so that the total trial length was 6 s, thus synchronizing trials across the group. (B) In the Isolated condition, participants saw only their own location (indicated by a red dot within a rectangle) and the number of points earned for being in that location on that trial. (C) In the Locations condition, participants saw the information present in the Isolated condition, plus two dots representing the locations of the other two group members. (D) In the Points condition, participants saw the information present in the Isolated condition, plus numbers in the upper right portion of the screen representing the earnings of the other two group members. (E) In the Both condition, participants saw all the information present in the Isolated, Locations, and Points conditions

group with regard to the other two participants. In the "Isolated" condition, participants had access only to their location within the rectangle and the number of points they had earned for the previous trial (Fig. 1B). In the "Locations" condition, participants had access to their own location and points earned on the last trial as in the Isolated condition; however, they could also see the current locations of the two other participants in the group (represented by blue and green markers), but no information about the points they had earned (Fig. 1C). In the "Points" condition, participants again could see their own location and recent earnings, and could also see the points each of the other group members had recently earned, but not their locations (Fig. 1D). Finally, in the "Both" condition participants had access to all the aforementioned information: their own location and recent points, as well as the locations and most recent points for the other two group members (Fig. 1E). This design allowed us to examine the separate and potentially combined influences of the two types of social information. Participants were informed that all members of the group would have access to the same kind of information for each task. For the purposes of the present article, we will focus on the impact of the participant's own points and the locations of the other group members, leaving consideration of the Points condition (and the influence of reward-based information in the Both condition) as a target for future presentation.

As mentioned above, each of the four tasks also employed a different predefined reward function that determined the number of points a participant earned for each location. Unbeknownst to the participants, these functions were actually one-dimensional - only one of the two dimensions of location within the grid (horizontal or vertical) was used to determine the points earned on a given trial. We refer to this dimension as the "relevant dimension" and to the other, which had no impact upon earnings, as the "irrelevant dimension." The orientation of the relevant dimension (vertical or horizontal) was determined randomly at the beginning of each task, was the same for all three members of the group, and remained constant across all 90 trials of that task. In order to keep this fact from being obvious to participants, noise was drawn from a zero-mean, uniform distribution and added to the earnings for each trial (otherwise, it would have been immediately obvious that points varied as the participant moved along one dimension, but not the other). In order to determine how much noise was appropriate, a separate pilot experiment was performed before the collection of the data discussed here. This pilot experiment was used to select an amount of noise high enough to mask the identity of the irrelevant dimension, but not so high that the maxima along the relevant dimension could not be determined over time. Of the 324 participants in this study, only one reported being aware of the irrelevant dimension. As demonstrated below, participants were sensitive to the different mean earnings provided by different locations, and spent more time at the maxima of the reward functions.

The reward functions themselves are depicted in Fig. 2. As can be seen from panels A–C, the mean number of points earned (z-axis) was constant for all locations along the irrelevant dimension (y-axis in this example), while they varied significantly across the relevant dimension (x-axis). Four functions were used, each designed to examine a different phenomenon. Figure 2A depicts the Valley task, which was characterized by a local maximum (far left), a global maximum (far right), and a deep minimum between them. While the initial location was randomized across the groups, all members of a given group were initialized at the same location. For the Valley task, this location was always at the local maximum along the relevant



Fig. 2 Reward functions in the experiment. Participants carried out four tasks, each of which provided different mean earnings for locations within the rectangle. Each task was defined by a relevant dimension (shown here on the x axis) and an irrelevant dimension (shown here on the y axis) – participants were rewarded based upon their locations along the former, but not the latter. Noise was added to the points earned for each trial. The surface height (z axis) indicates relative mean earnings at each location; actual points earned were determined by a scale factor as described in the text. (A) In the Valley task, participants began at the local maximum and

dimension (the group's initial location along the irrelevant dimension was randomly selected); for the three other tasks, initial locations were randomized along both the relevant and irrelevant dimensions. In order to maximize earnings in the Valley task, a participant had to move to the global maximum by persevering through a patch of low earnings. Because the points earned at the global maximum were approximately twice those at the local maximum, this function rewarded participants that thoroughly explored locations along the relevant dimension. The Single Peak function, depicted in Fig. 2B, was designed to produce the opposite incentive. The global maximum for this function was achieved by climbing the gradient toward an optimal location near the middle of the grid (along the relevant dimension). Thus, participants were not rewarded for exploring areas far from this obvious peak. The Double Peak function, shown in Fig. 2C, had two equally profitable maxima on average, a participant earned the same number of

only reached the best location if they persevered through a region in which earnings were low. (**B**) In the Single Peak task, there was a single maximum along the relevant dimension, allowing participants to climb the points gradient to reach the most profitable locations. (**C**) In the Double Peak task, two equally profitable maxima were spaced apart from one another. (**D**) In the Flat task, there effectively was no relevant dimension – mean earnings were the same no matter where participants were located

points if he or she was located at either peak. However, if a participant was drawn from one maximum to the other (e.g., in order to be closer to other group members – see Supplementary Fig. S1), his or her earnings would suffer by moving through the local minimum between the two peaks. Finally, Fig. 2D depicts the Flat task, in which the reward function was equal for all locations (i.e., neither dimension was relevant). This task was used to examine participants' decisions when there was no advantage to be gained by any decision, as a baseline against which to compare performance in the other tasks in which decisions could be meaningfully influenced by social and/or reward information (i.e., in which there were good and bad locations, and information about these could be inferred from observing others). In particular, we used the Flat task to determine if there were any systematic biases in how participants distributed themselves around the rectangle when they had no incentive to move to any particular locations. The mapping between reward function and social condition were pseudorandomly permuted across groups using a Latin squares design; the order of the four tasks was determined randomly at the beginning of the experiment.

Participants were informed prior to the experiment that all members of the group would be playing the same task at any given time – that is, locations that were good (or bad) for one participant would be good (or bad) for all participants – but that the earnings of other participants did not interact with their own earnings (i.e., a participant did not earn fewer points by virtue of other participants' locations or earnings). These instructions were designed to allow for the influence of social information without the influence of task-based incentives for competitive or cooperative behavior.

So that participants would not attribute special significance to particular locations within the white rectangle, the alignment of the reward functions to the grid was jittered along the relevant dimension across tasks. This was done by defining each function along a superset of twelve locations and choosing a set of ten contiguous locations within that set for each task, thus shifting the alignment of the reward function relative to the grid. The value of this shift was determined randomly at the beginning of the task. Furthermore, to minimize potential bias toward one side of the rectangle or the other, the orientation of the reward function along the relevant dimension was reversed with probability .5 between tasks. Each group experienced all four reward functions and all four social conditions exactly once, and the combinations of social condition and reward function were balanced across groups. The earnings during each task were also assigned a random scale factor between 50 and 99. Because this scale factor varied from task to task and was not revealed to participants, participants did not know if they had found the best location for a given task at any time (i.e., the points earned on one task could not be used to assess performance on the next).

Before each task, a screen was presented that lasted 8 s and indicated the "social condition" under which the next task would be performed. After the last task was completed, a screen was presented indicating (a) that the experiment was complete, and (b) how much the participant would be paid for the session.

Participants' compensation for the experiment were based on the points they earned in each of the four tasks. For each task, earnings were computed by summing the points earned on each trial and normalizing this sum by the aforementioned scale factor so that each task had the same potential payoff. By adding these normalized totals together and comparing the resulting value to the minimum and maximum possible earnings, the participant's compensation was determined with the range of US\$12 to US\$30 (eye-tracking cohort) or US\$30 to US\$50 (neuroimaging cohort). Participants were informed of these compensation procedures prior to the experiment.

Analysis of participant locations along the relevant dimension

In order to determine the distributions of participant locations for the tasks, location data for each participant were first adjusted to account for the orientation of the reward function (nominal or reversed; see above), as well as the shift used to jitter the reward functions along the relevant dimension (see above). This procedure realigned participants' locations so that they were relative to the global maxima and minima for each task, allowing for the pooling of data across orientations and shifts. Because mean earnings were uniform across the Flat task, there was no need for this procedure for that reward function. Similarly, no such adjustment was necessary when computing locations across the irrelevant dimensions of the other three tasks.

Analysis of inter-participant distances

For both dimensions, inter-participant distances were defined as the sum of the inter-participant distances across the three possible pairings of the group members:

Inter-participant relevant distance =	$ x_1-x_2 + x_2-x_3 + x_3-x_1 $
Inter-participant irrelevant distance =	$ y_1 - y_2 + y_2 - y_3 + y_3 - y_1 $

where x and y denote the participants' positions along the respective dimensions (in the case of the Flat task, both dimensions were considered to be irrelevant, and interparticipant distances was averaged across both dimensions). These values were computed for each of the 90 trials in the task, after which they were averaged and sorted according to the social condition under which the task was performed. The same method was applied to the analogous computations for simulated agents (see below).

Decomposition of actions in the computational model

The computational model applied to the decision-making data used a simplified action space in which choices were defined relative to the participant's previous choices (i.e., same choice, opposite choice, orthogonal choice or no choice) rather than the cardinal directions of the choice being made (up, down, left, or right). Because the identity of the relevant dimension and the location of the global maximum changed from one task to the next (and for the same task across groups), this was a more useful frame of reference for describing individual choices. On a given trial, participants could select a direction in which to move – this defined their *direction of travel* (Fig. 5A). If the participant chose to move in the same direction, this was labeled a *continue* action. If he or she returned to the previous location, this was labeled a *reverse* action. If he or she moved in either direction orthogonal to the direction of travel, this was labeled a *turn* action. Finally, if the participant did not make a choice during the allotted 3 s, the trial timed out and the participant remained in the same location – this was labeled a *stay* action. Because the direction of travel could only be defined once the participant had moved, trials preceding the first choice were excluded from analysis.

Action probabilities for reward-based choices

One consistent pattern in the data was participants' attraction to maxima in the reward functions (see Results, section *Sensitivity to reward-based information* and Fig. 3). To account for this trend, the first component of the decision-making model assumed that the probabilities of each of the four actions depended on a single variable summarizing the participant's recent earnings. This component will hereafter be referred to as "reward-based," as participants preferentially spent their time at maxima in the reward functions and were compensated according to the points they earned. We use the term "reward" in a narrow sense, referring to that aspect of the task concerned with task performance and the acquisition of secondary rewards (i.e., money). We later refer to the component of decision-making influenced

791

by other agents as "socially-based" in order to differentiate the two influences. However, these terms should not be taken to imply that social herding behavior might not itself represent a form of intrinsic reward.

In order to decompose reward-based actions, several metrics summarizing recent earnings were constructed and examined at the group level. These included the ratio of current earnings to earnings at the previous location, the ratio of current earnings to mean earnings across the block, and the ratio of current earnings to the maximum earnings experienced up to point in the block. In addition, a metric was tested that used the ratio of current earnings to previous earnings (i.e., the first of the aforementioned metrics), but combined these ratios across the history of the block using an exponential weighting. Group-level analysis suggested a logistic relationship between the reward ratio and the probability of each action. Analysis revealed that the best metric for predicting behavior was also one of the simplest (that is, the one requiring the lowest memory load): the ratio of earnings at the current location to earnings at the previous one - henceforth, the reward ratio (see Supplementary Fig. S2 for details regarding this analysis). Each choice was categorized according to action type and paired with the reward ratio corresponding to that trial. As



Fig. 3 Distribution of locations for participants and simulated agents. Blue bars indicate the fraction of trials (y-axis) that were spent by human participants at a given location (x-axis) along the relevant dimension in the Isolated condition; green bars indicate the fraction of trials spent by simulated agents at the same location (simulated without the influence of other group members' locations). Black dashed lines indicate the relative number of points yielded by each location. (A) In the Valley task, participants were concentrated at the local maximum (left), and the global maximum (right), while spending little time where earnings were low (middle). Simulated agents were concentrated at the local maximum

(left) and avoided regions where earnings were low (middle), but did not find the global maximum as easily as human participants. (**B**) In the Single Peak task, participants easily discovered the global maximum and spent the plurality of trials there. Simulated agents stayed near the global maximum like human participants, although they were not as tightly clustered near the maximum. (**C**) In the Double Peak task, participants spent nearly half of all trials on one of the two global maxima. Simulated agents spent the most time at the two global maxima, as well. (**D**) In the Flat task, participants and simulated agents were both uniformly distributed, with a slight preference for the middle of the rectangle

the participant's current earnings were always visible, this reward-based component was applied to all four social conditions. Thus, we used logistic functions to describe the relationship between the reward ratio on a trial and the probability that the participant would choose an action of that type. These functions were defined by four parameters:

$$P_{reward}(action) = \frac{A_1}{1 + e^{-A_2 * R_t + A_3}} + A_4$$

where A_1 determines the amplitude, A_2 determines the slope, A_3 determines bias (i.e., the location of the inflection point along the x-axis), A_4 determines the minimum probability of the action, and R_t is the reward ratio for trial *t*.

Action probabilities for socially-based choices

A second behavioral pattern revealed by analysis of the data was participants' tendency to cluster together when the locations of other group members were visible (see Results, section *Sensitivity to social information* and Fig. 4). For tasks run under the Locations or Both conditions, the model therefore included a component quantifying the relationship between the locations of other group members and the probability of each of the four action types. As with the reward-based component, the socially-based component of the model assumed that the action probabilities depended on a single variable. This variable, θ , summarized the location of the group relative to the participant's current location and direction of travel. Specifically, θ was the absolute value of the angle between the vector describing the participant's current direction of travel and the vector pointing from the participant's location to the centroid of the group:

$$\theta = \arccos\left(\frac{V_p \cdot V_c}{||V_c||}\right)$$

where V_p is the unit vector describing the participant's current direction of travel (e.g., [1,0] if the participant



Fig. 4 Group clustering across social conditions. Inter-participant distance (y-axis) was defined as the sum of the absolute values of the distances between participants along the dimension of interest: lower distance values indicate tighter clustering within the group. P-values indicate the results of Tukey's tests between inter-participant distances for the Isolated condition and inter-participant distances for each of the other three conditions. (A) For the three tasks in which the relevant dimension was meaningful, there was a significant main effect of both social condition (F_{3,312} = 21.88, p < 10⁻¹², ω^2 = .13) and task (F_{2,312} = 38.12, p < 10⁻¹⁴, ω^2 = .16), as well as an interaction (F_{6,312} = 2.45, p < .05, ω^2 = .02). Tukey's tests revealed that inter-participant distances were significantly lower for the Locations and Both conditions than for the Isolated

condition; distances for the Points condition were marginally lower. (**B**) For the same three tasks along the irrelevant dimension, ANOVA indicated a significant main effect of social condition ($F_{3,312} = 404.4$, $p < 10^{-15}$, $\omega^2 = .20$) and task ($F_{2,312} = 33.13$, p < .05, $\omega^2 = .01$), but no interaction ($F_{6,312} = 31.54$, p = .37). Tukey's tests revealed that inter-participant distances were lower for the Locations and Both conditions than for the Isolated condition. (**C**) For the Flat task, which had no relevant dimension, there was a significant main effect of social condition ($F_{3,212} = 13.55$, $p < 10^{-7}$, $\omega^2 = .15$) and Tukey's tests between the Isolated condition and the other three social conditions revealed that inter-participant distances were lower only for the Locations condition

had most recently moved right) and V_c is the vector pointing from the participant's location to the group centroid. Under this formulation, θ ranged from 0° to 180° , with 0° corresponding to a centroid directly ahead of the participant (with *ahead* being defined by the participant's direction of travel), 90° corresponding to a centroid directly to one side or the other, and 180° corresponding to a centroid directly behind the participant (Fig. 5C). Each choice in the experiment was categorized according to its action type and paired with the value of θ corresponding to that trial. This allowed us to examine the effect of θ on the probability of each type of action. Because the input variable was angular, we used sine functions to describe the relationship between θ on a trial and the probability that the participant would choose an action of that type (see Supplementary Fig. S3 for a group-level analysis of the relationship between θ and the probability of each action). These functions were defined by four parameters:

$$P_{social}(action) = B_1 * \sin(B_2 * \theta_t + B_3) + B_4$$

а

where B_1 determines the amplitude, B_2 determines the period, B_3 determines the phase shift along the x-axis, and B_4 determines the mean probability of the action.

Action probabilities for perseveration-based choices

The data also revealed a third behavioral trend, in which participants often committed consecutive sequences of the same action (see Results, section *Computational modeling: source weighting and perseveration* and Fig. 6B). The final component of the model assessed this tendency to *perseverate* on the action. We defined this tendency as the probability of committing an action as a function of the number of times that action had been consecutively repeated on previous trials. For example, if the last two trials were both *continue* actions, the perseveration function would indicate the probability of making the same action (*continue*) as a function of the number of consecutive times it had been repeated in the immediately preceding trials (in this case, once). This function was constrained

Reverse

3.5

180

2.5 3



b

Fig. 5 A computational model for reward-based and socially-based decisions. (A) Imagine that on a given trial t-1, a participant chose to move to the right, thus defining the participant's direction of travel for the next trial. For trial t, actions could then be defined relative to the movement on the last trial. Four actions were possible: *continue, reverse, turn, and stay.* (B) Reward-based action probabilities were calculated for each of the four action types as a function of the *reward ratio*: the number of points earned on the current trial divided by the number of points earned at the previous location. Dark lines represent the mean value of each function across

participants, while lighter bands around each line represent the standard error of the mean (SEM). (C) For trials in which information was available regarding the locations of other group members, the angle to the group centroid relative to the participant's direction of travel, θ , was used as the input variable for the socially-based action probabilities. (D) Socially-based action probabilities are shown for each of the four action types as a function of θ . Dark lines represent the mean value of each function across participants, while lighter bands around each line represent the SEM

135

to have an inflection point of (0, 0) – that is, the probability of perseverating was defined to be zero when the action had not been previously repeated. This function had the form:

$$P_{perseverate}(action) = \frac{C_1}{1 + e^{-C_2 * N_t}} - \frac{C_1}{2}$$

where C_1 determined the maximum amplitude of the function, C_2 determined the slope of the function and

 N_t was the number of consecutive repeated actions of that type prior to trial *t*.

Fitting participants using the computational model

For each action type, the total probability of the action on a given trial was:

$$P_{total}(action) = P_{perseverate}(action) + (1 - P_{perseverate}(action)) * (W * P_{social}(action) + (1 - W) * P_{reward}(action))$$

where W was a parameter that determined the relative balance of the reward-based and socially-based sources.

As a single function described the probability of perseveration for all four actions, a parameter representing the







(RTs) across participants, sorted into six decision classes. ANOVA revealed a significant effect of both decision class ($F_{3,2365} = 7.03$, $p < 10^{-16}$, $\omega^2 = .03$) and location availability ($F_{1,2365} = 107.60$, $p < 10^{-23}$, $\omega^2 = .04$), but no interaction ($F_{3,2365} = 1.08$, p = .35). RTs when all sources favored the chosen action (left column) were significantly faster than when only one source (or no sources) favored the action. White bars correspond to RTs during conditions in which location information was available – decisions belonged to one of six classes. Gray bars correspond to RTs during conditions in which information other group members' locations were unavailable – without the socially-based source, so decisions belonged to one of four classes

balance between the perseveration-based source and the combination of the other two sources was subsumed by C_1 (see above) and was therefore unnecessary. Under this framework, the probabilities of the four actions always summed to one, regardless of whether any action had been previously repeated (i.e., whether $P_{perseverate}$ was greater than zero for any action). Parameter values were determined individually for each participant, and were assigned according to the combination of values that minimized the total likelihood error for that participant's measured behavior (that is, the set of parameters most likely to have produced participant's decisions under the model).

Standard function minimization, in which an algorithm assigns values to a set of continuous parameters in order to minimize an error function, was not a feasible method for fitting parameters for two reasons. First, such minimization algorithms are not guaranteed to find the global minimum in the error function, especially when the number of free parameters is large. In this case, fitting a participant's behavior, even without the perseveration component, involved 16 parameters for the reward-based component, 16 parameters for the socially-based component, and one parameter (W) that dictated the relative influence the two components. The second reason that standard function minimization was insufficient was that there were two independent constraints to satisfy: the minimization of likelihood error and the requirement that the functions corresponding to the four possible actions had to sum to one for each point in the range of input values ($0 \le R \le$ 3.5 for reward-based functions, 3.5 being the maximum reward ratio across the data set; $0^{\circ} \le \theta \le 180^{\circ}$ for socially-based functions). This fact also made genetic algorithms unfeasible, for the fitness functions of such algorithms would also depend on these two independent constraints.

To overcome this obstacle, an alternative method was devised. Rather than fit continuous parameters as in standard function minimization, a large set of enumerated combinations of parameter values was created for the reward-based, sociallybased and perseveration-based components. By enumerating the combinations a priori, our method ensured that the constraint on the sums of the reward-based and socially-based combinations was satisfied by assessing the likelihood error only for those combinations of parameters that satisfied the constraint on the sums (those that summed to $1 \pm .001$ for each point in the range of possible input values). Combinations were generated by first creating a set of candidate functions, then by exhaustively determining which sets of four of these functions summed to one for each point in the range of input values. The numbers of candidate functions for reward-based and sociallybased parameters were iteratively adjusted until the number of tested combinations was as equal as possible – greater numbers of combinations for one of the components would have artificially improved the fits for that component, potentially biasing the weighting parameter. W. Due to their asymptotic nature. logistic functions (corresponding to the reward-based component) were more likely to sum to one for all input values than sinusoid functions (corresponding to the socially-based component). Thus, 4,000 candidate logistic functions and 40,000 candidate sinusoid functions yielded 1.04 million and 1.11 million parameter combinations, respectively (1 million was chosen as a target value in order to balance the accuracy of the fits with the computational requirements of testing such a large number of combinations). There were, in turn, 24 possible mappings between a given combination of functions and the four action types. Because perseveration was a function with only two parameters (rather than having a set of parameters dedicated to each action type), only 10,000 parameter combinations were needed to cover the possible mappings between the number of consecutive actions and the likelihood of perseverating.

By definition, the perseveration-based component exerted no influence on behavior when an action had not been repeated previously. Therefore, reward-based and socially-based components could be assessed independently of perseveration by examining trials on which there was no repeated action. Trials for which the *continue* action was impossible due to the boundaries of the rectangle (other actions were always possible regardless of location), or for which the reward ratio was undefined (the first trial of each block) or the group centroid angle were undefined (due to co-localization with the participant's own marker) were similarly excluded for the purposes of fitting. After filtering trials accordingly, parameter fitting was accomplished using a hierarchical procedure that focused first on fitting the reward-based and socially-based components (A₁-A₄ and B₁-B₄, respectively) and their relative weightings (W), then moving on to the perseveration-based component (C_1 and C_2). For each participant, the combination of reward-based functions, socially-based functions, and W that minimized the likelihood error was determined by exhaustively calculating the error for each possible combination; W was tested in the range of [0, 1] in increments of .05. Once these parameters had been determined, the trials for which there were previously repeated actions were reintroduced and the combination of perseveration parameters that minimized the likelihood across all of the participant's behavior was selected. Once all parameters were fit for the individual, the probability of each action type could be estimated for each decision.

Although this procedure was computationally tractable, it is worth noting that our method treated the influence of each source as independent. Under this framework, the three components of the model could combine to influence behavior but, for example, the current value of θ could not alter the functions used by the reward-based component (more concretely, the current location of the group could not change how the participant related current reward to behavior). While allowing for such an effect is an interesting possibility, it would dramatically increase the complexity of the model and the amount of computation required to fit parameters to behavior; we consequently have deferred such a modification to later work.

Correlation analysis of individual difference measures and task-related behavior

In addition to behavioral responses, gaze location, and brain activity, we also collected data on participants' age and personality, with the latter assessed using the NEO personality inventory (also known as the Big Five; Costa & McCrae, 1985). Correlations were calculated between these six measures and three task-related measures: W (the weighting between reward-based and socially-based components of our model), earnings during the Isolated and Points conditions (i.e., those conditions for which location information for the other group members was unavailable), and earnings during the Locations and Both conditions (i.e., those conditions for which locations were visible). Because there were significant correlations between certain personality metrics, as well as between personality metrics and demographic variables, multiple comparison correction via the Bonferroni method was not appropriate (this method, while simple, assumes independence between tests). Instead, multiple comparison correction was achieved via a Monte Carlo permutation test.

Comparison with alternative models

The complexity of the model detailed above begged an important question: was the three-component model really the best explanation of the observed data? We used the Akaike Information Criterion (AIC; Akaike, 1974), which weighs the explanatory power of a model against the number of free parameters that the model employs, to compare the quality of the three-component model to several alternatives. One of the key characteristics of the three-component model is its sensitivity to the context of the decision: the current values of the reward ratio R and the centroid angle θ , as well as the number of repeated actions the participant had made, on any given trial. The first alternative model removed the dependency on context while retaining all three components: reward-based and socially-based actions were modeled as static probabilities for each of the four action types, rather than functions of R and θ ; perseveration-based actions were modeled as a static probability of choosing a previously repeated action. The full, threecomponent model was also compared to three other alternative models, each one corresponding to a single component within the model. These single-component models were, like the three-component model but unlike the static probability model, sensitive to context. These comparisons therefore asked whether multiple components were actually necessary in explaining participants' behavior, or if a single component could sufficiently explain behavior, instead. The AIC was calculated for each participant under the three-component model and the alternative models, and these values were compared across models.

Simulations of agents using the computational model

In order to assess whether the model could recapitulate key aspects of observed behavior, a simulation was run that used the action probabilities described above to generate choices. Groups of three simulated agents were placed on one of the four reward functions using the same initialization procedure used for human participants. They were then forced to make one move in a random direction (generated independently for each agent) in order to define their direction of travel. The composition of the groups of simulated agents was analogous to the composition of the human groups. That is, each group of human participants was simulated according to the parameters fit to each member of that group, thereby recapitulating potentially crucial individual differences within the group. Each group was simulated 1,000 times with 90 trials per simulation, and the locations within the rectangle and the inter-agent distances were averaged across simulations. Two environments were considered for each of the four reward functions: one in which only the reward-based and perseveration functions were used (the analog of the Isolated condition), and another condition in which the socially-based functions were added (the analog of the Locations condition). The first environment was used to assess simulated agents' preferences for locations along the relevant dimension of the reward functions (see Fig. 3), while the second was used to assess simulated agents' degree of clustering when location information was available (see Fig. 6C).

Classification of individual trials using the computational model

While our computational model was able to capture several key aspects of group behavior (see Results, section Computational modeling: simulations of participant behavior and Figs. 3 and 6C), the model alone could not indicate how the three sources of information (reward-based, sociallybased, and perseveration-based) combined to affect decisionmaking. In order to address this question, individual trials were classified according to which source favored the action chosen on that trial (or, when the chosen action was most likely according to multiple sources, how many sources could be used to explain it). Prior to this classification, however, certain trials were excluded from analysis. These included all trials preceding the participant's first movement (prior to which the direction of travel was undefined), the last trial (which had no external consequences), trials on which continuing would move the participant's marker outside the rectangle, and any trial for which the group centroid was the same as the participant's location (in which case θ was undefined, as a value of zero would have indicated a group centroid ahead of the participant's direction of travel). For the Locations and Both conditions, in which the locations of other group members were visible, each of the remaining trials was categorized as belonging to one of six classes. These six classes were defined according to whether the chosen action was favored by:

- 1) all (three) sources
- 2) two of the three sources
- 3) only the reward-based source
- 4) only the socially-based source
- 5) only the perseveration-based source
- 6) none of the sources.

For the Isolated and Points conditions, the locations of other group members were not visible. Thus, the sociallybased source was unavailable, leaving only two sources to influence decision-making. Trials were therefore categorized as belonging to one of four classes according to whether the chosen action was favored by:

- 1) all (two) sources
- 2) only the reward-based source
- 3) only the perseveration-based source
- 4) none of the sources.

Eye-tracking analysis

For the eye-tracking cohort, gaze information was synced with each trial and recorded at 60 Hz using an Eye-Trac 6000 eyetracking system from Applied Science Laboratories with a video head tracker. Participants with eyeglasses, or for whom the pupil was not reliably observed, were excluded from the eye-tracking data. For each sample, gaze location was calculated and classified according to whether the participants were looking at the number of points they had earned on the current trial, the points earned by other participants (available under the Points and Both conditions), their own marker, the markers of other participants (available under the Locations and Both conditions), the arrows indicating the participant's choice, or none of these features. It is worth noting that a classification of the participant's own marker or the markers of other group members was subject to a conservative criterion, requiring the participant's location of gaze to fall within the region of the 10 x 10 grid corresponding to the marker's location. Each grid location therefore occupied approximately 2.7° of horizontal visual angle and 1.5° of vertical visual angle; the values are approximate because, while eye-tracking was calibrated at a distance of 24 in. between the eye and the tracking camera, participants' heads were not fixed in place for the duration of the experiment. This requirement may have underestimated the number of samples during which participants were looking at their own or others' markers. However, it is reasonable to expect that such underestimation occurred consistently across all the classes of decisions mentioned above, and so did not bias the location of gaze analysis depicted in Fig. 7A. Because trials were of variable length and gaze location could not always be reliably calculated (due to blinking or loss of the pupil), gaze times were normalized to the amount of data contained within the trial and expressed as the fraction of the trial spent on that particular location.

Neuroimaging data acquisition and preprocessing

Functional magnetic resonance imaging data (Kwong et al., 1992; Ogawa et al., 1990) were collected using three Siemens 3.0 Tesla Trio scanners. Each session included a high-resolution, T1-weighted scan (MP-RAGE; Siemens). Whole-brain imaging was collected during the tasks in a single session using echo-planar imaging with a repetition time (TR) of 2,000 ms, echo time (TE) of 40 ms, and a flip angle of 90°. The images were acquired as matrices of 64 x 64 x 26 voxels aligned to the anterior and posterior commissures of the corpus callosum, resulting in voxels with a resolution of $3.4 \times 3.4 \times 4.0$ mm.

Image preprocessing was performed with SPM8 (http:// www.fil.ion.ucl.ac.uk/spm; Frackowiack et al., 1997; Friston et al., 1995). Slice-timing correction was followed by realignment to the first functional scan using a six-parameter rigidbody transformation, with the six parameters corresponding to three axes of rotation and three directions of translation. The mean of the realigned images was co-registered to the T1weighted structural image using a twelve-parameter affine transformation. Tissue segmentation was determined for each structural image, and the gray matter designated by this classification was used for spatial normalization by applying a twelve-parameter affine transformation. The functional images were then normalized and smoothed with an 8 mm FWHM Gaussian kernel for inter-subject analyses.

Neuroimaging data analysis

We used SPM8 (http://www.fil.ion.ucl.ac.uk/spm; Frackowiack et al., 1997; Friston et al., 1995) to implement a standard general linear model (GLM) for each participant. The design matrix for each participant included delta functions for the "New Task" screen preceding each task, the onset of each trial, and the end of each trial (whether the trial ended with a decision by the participant or timed out). Trials were grouped according to whether (a) location information regarding other group members was available (i.e., the Locations and Both conditions) or not



Fig. 7 Additional evidence for the *integration hypothesis* of decisionmaking. (**A**) The fraction of each trial that participants looked at their points was greater than zero for all three single-source decision classes ($p < 10^{-27}$ for all). ANOVA revealed a significant effect of location availability ($F_{1,686} = 36.20$, $p < 10^{-8}$, $\omega^2 = .05$), but no effect of decision class ($F_{1,686} = .70$, p = .40) and no interaction between the two factors ($F_{1,686} =$.003, p = .96). Because the locations of other group members were unavailable under the Isolated and Points conditions, there was no "Social Only" class of decisions under these conditions. The fraction of each trial that participants in the eye-tracking cohort looked at the markers of other group members was greater than zero for all three types of single-source decision ($p < 10^{-31}$ for all three), but did not differ from one another ($F_{2,513} = .27$, p = .76). As others' markers were not present in the Isolated and Points conditions. Error bars represent standard error

of the mean. (**B**) Reaction times did not differ between "switch" and "non-switch" trials for any source, regardless of whether location-based information regarding other group members was available or not. P-values for paired t-tests are indicated above each group; error bars represent standard error of the mean. (**C**) General linear model analysis revealed brain regions whose activity was linearly proportional to the number of decision sources that disagreed with the chosen action (corrected p < .001, threshold of five contiguous voxels). The top row of images corresponds to analyses for the Isolated and Points conditions, in which location-based information for other group members was unavailable (and the socially-based component consequently did not contribute to decision-making). The bottom row corresponds to analyses for the Locations and Both conditions, in which location-based information was available and all three components contributed to decision-making

(i.e., the Isolated and Points conditions), and (b) the participant moved his or her marker on the trial (i.e., a *continue*, *reverse*, or *turn* trial) or not (i.e., a *stay* trial). The former grouping accounted for the fact that the socially-based component of the model depended on location information, and was only defined during the Locations and Both conditions; the latter grouping accounted for the fact that *stay* trials did not involve a button press and had a uniform duration, while other actions involved a button press and a variable reaction time. In addition, button presses that occurred after the participant's choice for the trial had been registered were included as delta functions in the design matrix, even though they did not affect the content or timing of events within the experiment.

Several regressors were also included, with each regressor corresponding to delta functions whose amplitude varied parametrically with the variable's value on each trial. Variables

that could be calculated at the onset of each trial, such as the absolute magnitude of the reward earned (i.e., the number of points displayed in the lower left portion of the participant's screen), the reward ratio R, the centroid angle θ , and the number of consecutive repeated actions N, were associated with trial onset times. In addition, two variables associated with the decisions made by the participant were associated with the end of each trial. The first of these was the reaction time for the trial (because *stay* trials were of uniform length, this regressor was not included for the stay trial group). The second regressor associated with the end of each trial indicated the number of sources not recommending the chosen action – this was used to examine the impact of disagreement between the three different sources of information. Regressors were orthogonalized prior to inclusion. Trials for which model inputs were undefined (e.g., the first trial of each task, in which the reward

ratio R could not be calculated, or trials in which the group centroid was identical to the participant's current location, thus preventing θ from being calculated) constituted a separate group. In order to avoid spurious activations due to movement by the participant, the six regressors for head motion were constructed from the motion values computed by the six-parameter rigid body transformation used for spatial realignment.

In a parallel analysis designed to test the *mixture hypothesis* detailed below, trials associated with a single source (see Materials and methods: Classification of individual trials using the computational model) were additionally classified according to whether they were "switch trials" or "non-switch trials"; analysis parameters and the assignment of regressors were otherwise the same (with the exception of the disagreement regressor described above, which was excluded). For all analyses, the GLM for each participant was fit voxel-wise to the BOLD data for that participant, and a random effects analysis was performed across participants for each regressor. Multiple comparisons correction was performed by a nonparametric permutation test computed via SPM's SnPM toolbox (http://warwick.ack.uk/snpm; Nichols and Holmes, 2001) and all resulting statistical maps were thresholded at p < .001 with an extent threshold of five voxels. All statistical maps found in this article and supplementary material can be found on NeuroVault: http://neurovault.org/collections/2490/ (Gorgolewski et al., 2015).

For the purpose of control analysis, a region of interest (the vmPFC) was designated anatomically with a center of (X = 0 mm, Y = 51 mm, Z = -8 mm; MNI coordinates) and a radius of 8 mm. BOLD signals were extracted from the corresponding voxels and the mean signal was calculated across all voxels in the ROI for each acquired image, with an offset of 5 s to account for the lag in the hemodynamic response. Because decision times did not necessarily align with scan boundaries (i.e., the times at which images were acquired), linear interpolation between the two bounding data points was used to estimate the magnitude of the BOLD response at the time of the trial's end, with the end defined either by the participant's first response after trial onset, or the elapse of the 3-s deadline. Thus, the mean response was calculated for every trial in the experiment and categorized according to the trial's class (see above).

Results

Sensitivity to reward-based information

As shown in Fig. 3, participants' behavior in the Isolated condition exhibited a strong preference for locations along the relevant dimension that yielded higher mean earnings. As shown for the Valley task in panel 3A, the majority of trials were spent close to the local maximum (the line of locations at which groups were initialized), with relatively few trials spent at the global minimum. However, those participants (56%) who managed to make it past the minimum and find the global maximum spent a large proportion of trials there (see Supplementary Fig. S4).

Similarly, panel 3B demonstrates that the global maximum in the Single Peak task was easily found, and that participants spent nearly 40% of their trials at that location along the relevant dimension. The same was true for the Double Peak task – nearly 50% of trials were spent on one of the two global maxima. Finally, participants playing the Flat task exhibited a fairly uniform distribution across all locations, with a slight tendency to cluster near the middle of the rectangle. Taken together, these data suggest that participants were sensitive to reward-based information along the relevant dimension. Participants distributed themselves nearly uniformly along the irrelevant dimensions of each task although, as in the Flat task, there was a slight tendency to concentrate on the middle of the rectangle (see Supplementary Fig. S5).

Sensitivity to social information

In addition to being influenced by earnings, participants were also biased by information about other group members. Figure 4 shows that clustering among groups, as defined by the summed distances between the three participants (see Materials and methods: Analysis of inter-participant distances), varied significantly as a function of the type of social information available. Panel 4A depicts the degree of clustering along the relevant dimension, averaged across the three tasks in which mean points varied meaningfully across locations: the Valley, Single Peak, and Double Peak tasks. Repeated measures analysis of variance (ANOVA) revealed significant main effects of social condition ($F_{3,312} = 21.88$, p < 10^{-12} , $\omega^2 = .13$) and task (F_{2.312} = 38.12, p < 10^{-14} , $\omega^2 = .16$), as well as a significant interaction (F_{3.312} = 2.45, p < .05, ω^2 = .02). As compared to the Isolated condition, participants exhibited tighter clustering (i.e., lower inter-participant distance) in the Locations and Both conditions, under which participants had information about the locations of other group members (see Supplementary Information, Table S1 for additional statistical data). There was also a marginally significant effect for the Points condition (p = .08), in which the earnings of other group members were visible. Although clustering was not significantly different between the Isolated and Points conditions for any individual task, the comparison was closest to significance (p = .11) for the Single Peak task, in which participants' locations could be uniquely inferred from their earnings if they were located at the global maximum (i.e., high earnings by other participants were an indicator of location, even when location was not explicitly presented).

The effects described above allow the possibility that social information had no independent influence, but simply

reinforced the reward-based information already available to the participant (i.e., the preference for more rewarding locations could have been strengthened by witnessing that others also had such a preference). However, panel 4B shows that groups also clustered more tightly along the irrelevant dimension in the Location and Both conditions than in the Isolated condition, an effect that cannot be accounted for by earnings alone (see Supplementary Information, Table S1 for additional statistical data). Repeated measures ANOVA revealed significant main effects of social condition ($F_{3,312} = 404.4$, p < 10^{-15} , ω^2 = .20) and task (F_{2,312} = 33.13, p < .05, ω^2 = .01), but no significant interaction ($F_{6.312} = 31.54$, p = .37). There was no significant difference between the Isolated and Points conditions, as locations along the irrelevant dimension could not be inferred based on earnings as they could for the relevant dimension of the Single Peak task. A similar phenomenon was observed along both dimensions of the Flat task (Fig. 4C), with a significant main effect of social condition ($F_{2,312}$ = 13.55, $p < 10^{-7}$, $\omega^2 = .15$). As compared to the Isolated condition, clustering was not significantly greater for either the Points or Both conditions. However, inter-participant distances were lower in the Locations condition (see Supplementary Information, Table S1 for additional statistical data). That is, when (a) participants could see both the locations and earnings of other group members, and (b) earnings depended meaningfully on decisions (as they did in the Valley, Single Peak, and Double Peak tasks) there was more clustering, even along the dimension that had no impact on their performance. However, when the same information was present but the contingency between decisions and outcomes was removed (as was the case in the Flat task) there was not significant clustering, suggesting that the combination of other group members' locations and earnings allowed at least some participants to discover the lack of utility in clustering. Thus, clustering was not an inevitable group behavior, but rather could be avoided when outcomes were independent of behavior and group members possessed full information.

Computational modeling: Reward-based influences

To quantify and better understand the influence that participants' earnings had on their behavior, we fit a simple decisionmaking model to the behavioral data. In the model, the action taken on each trial was chosen based upon the location previously occupied by the participant and the points earned at that previous location (see Fig. 5A and *Materials and methods*: *Action probabilities for reward-based choices* and *Fitting participants using the computational model*). For reward-based choices, the probability of each of the four actions was fit to each individual as a function of the individual's recent earnings. Figure 5B depicts the probabilities of each action type as a logistic function of the ratio between the participants' current earnings and their earnings at their previous location. These functions conformed to reasonable expectations about how decision-makers should behave. For example, when the reward ratio was greater than one (i.e., the participant had earned more points at the current location than the previous one), continuing in the same direction was estimated to be the most likely action. When the reward ratio was low (earnings had decreased), the participant was half as likely to continue. A similar effect was seen for staying in the same location: when earnings were increasing, participants were more likely to remain in that location than if they were decreasing. Conversely, the probability of reversing was very low when earnings went up, but much higher when they went down. The probability of turning varied to a lesser degree, with that action slightly less likely for higher reward ratios than for lower ones. Each function suggests a tendency toward gradient-climbing.

Computational modeling: Social influences

We also used the model to examine the influence of social information on behavior by fitting a second component of the model under which the action taken on each trial was chosen based upon θ , the angle to the centroid of the other group members (see Fig. 5C and Materials and methods: Action probabilities for socially-based choices and Fitting participants using the computational model). For sociallybased choices, the probability of each of the four actions was fit to each individual as a sinusoidal function of θ : Fig. 5D depicts the mean decision functions corresponding to each action type. These functions also conformed to expectations of human behavior in the tasks. Participants were more likely to continue in the same direction when the group centroid lay in the sector ahead of them ($\theta < 45^{\circ}$) than when it was directly behind them ($\theta = 180^\circ$). Conversely, the choice to reverse was at its maximum when the group centroid lay behind them ($\theta > 135^{\circ}$). The choice to *turn* was both at its maximum and the most likely of the four actions when the group centroid was between 90° and 165°. The choice to stay did not exhibit as much variation as the other three choices, but staying was most likely when the group centroid was behind the participant. These functions suggest a tendency toward approaching other group members.

Computational modeling: Source weighting and perseveration

In addition to fitting the action probabilities for each participant, our analysis assessed *the degree to which* choices depended on social information as compared to rewardbased information. The weighting between these two influences, summarized by the parameter W, could range from 0 (entirely reward-based) to 1 (entirely socially-based) and varied from one participant to the next. The distribution of values for W is shown in Fig. 6A. This distribution had a mean of .53, indicating that, on average, participants were slightly more reliant on social information than their own earnings. While participants' values for W varied across the entire interval, there was a notable concentration near the boundary at .95. This subset represents participants whose behavior during the Isolated and Points conditions depended on the reward ratio R, but whose behavior during the Locations and Both conditions was governed almost entirely by the locations of other group members (participants for which W = 1 used the average values of their socially-based functions to govern their behavior during the Isolated and Points conditions).

The final component of the decision-making model accounted for participants' tendency to *perseverate* on an action. The function for *perseverative choices*, shown in panel 6B, also exhibited reasonable behavior. An action was more likely to occur on the current trial if it had been repeated in the past. For example, once an action had been repeated six times, the likelihood of choosing the same action again was over 50% (thus, the probability of choosing the action recommended by the reward-based and socially-based sources was less than 50%). This was true for all action types, but the effect was most apparent for *stay* and *continue* actions (see Supplementary Fig. S6).

Analysis of individual difference measures, modeling parameters, and task-related behavior

Participant data exogenous to the task (i.e., personality metrics assessed via the NEO inventory, as well as the sex and age of the participants) were analyzed with regard to task-based data (values of W and earnings during blocks with and without location-based social information). Initial analyses suggested a positive correlation between age and W (r = .14, p < .01), a positive correlation between Conscientiousness and earnings during block without location-based social information (r = .16, p < .005), and a significant correlation between sex and W (in this case, males had a higher average value for W than females, p < .05). However, these effects did not survive multiple comparison correction (corrected p = .49, .29, and .18, respectively).

Comparison with alternative models

In order to assess the quality of the three-component model described in Figs. 5 and 6, it was compared to four alternative models. These alternative models were designed so as to determine: (a) the impact of sensitivity to the context of each decision (i.e., the current reward ratio, the direction of the group centroid, and the number of times an action had been repeated), and (b) the importance of each separate component of the model. Analysis using the AIC indicated that the full, three-component model was better than each of the alternative models ($p < 10^{-22}$ in all cases; see Supplementary Fig. S7 for

detailed statistical comparisons). With regard to absolute (rather than relative) explanatory power, the model yielded a mean McFadden's adjusted pseudo- R^2 of .14 with a standard deviation of .12 (McFadden, 1974), as compared against a null model with no free parameters, under which the four actions had an equal probability of being selected on each trial; this "adjusted" pseudo- R^2 , like the AIC, penalizes free parameters.

Computational modeling: Simulations of participant behavior

In addition to using the model to characterize participants' behavior, we tested whether the mechanisms in the model provide a reasonable account for the decision processes used by the participants, by examining whether the model could generate patterns of behavior similar to those observed for the participants. These simulations were without social information (the analog of the Isolated condition), and the results are depicted in Fig. 3 alongside the data for human participants in the Isolated condition. For each of the four reward functions, simulated agents spent proportionally more time at those locations that provided more points. They were biased toward the local maximum on the Valley task, and toward the global maxima on the Single Peak and Double Peak tasks. They also were uniformly distributed across locations in the Flat task, with a slight preference for the middle of the rectangle. There were some noticeable differences, however. Simulated agents in the Valley task did not reach the global maximum as frequently as human participants, and they were not as tightly distributed near maxima as human participants. However, it is worth noting that the simplicity of the model precluded simulated agents from either remembering the locations of maxima encountered earlier than the previous trial or remembering which locations had been previously visited. That is, they could neither compute the mapping between locations and reward-based value, nor could they compute a plan for efficiently exploring the space and exploiting locations discovered to be valuable - capacities advantageous for an agent with more sophisticated cognition.

As shown in Fig. 4, inter-participant distances were lower along both *relevant* and *irrelevant* dimensions when the locations of other group members were provided to human participants (i.e., in the Locations and Both conditions). Additional simulations examined the behavior of groups of three interacting agents as they performed the Valley, Single Peak, and Double Peak tasks (the Flat task being excluded because it did not possess a relevant dimension), this time including location-based information for other agents. These simulations yielded the location of each agent over time, and these locations were used to calculate the analog of inter-participant distances. Figure 6C shows that simulated agents exhibited a pattern for the inter-participant distances that was similar to

those for human participants (see also Fig. 4): providing simulated agents with information regarding the locations of other group members resulted in tighter clustering along both the relevant and irrelevant dimensions. The model was therefore able to recapitulate several key behaviors demonstrated in the experiment: attraction toward reward maxima (see Fig. 3), greater clustering along the relevant dimension when location information was available, and greater clustering along the irrelevant dimension when location information was available. As with the purely reward-based simulations, there was a small discrepancy: the inter-participant distances for agents simulated under the Locations condition were slightly larger than those for the human data. This phenomenon is most likely attributable to the fact that, as in Fig. 3, simulated agents did not adhere as tightly to maxima, resulting in greater spread around these locations.

Behavioral and neural evidence in favor of information integration during decision-making

While the findings reported above indicate that participants were responsive to three sources of information — previous behavior, current earnings, and the behavior of others — they do not directly address how this information was combined to influence behavior. There are at least two possible hypotheses for how this might have occurred. All three sources of information could have been combined on every trial (as in the model described above), weighted by their relative influence; we will refer to this as the *integration* hypothesis. Alternatively, one of the three sources of information could have determined decision-making on any given trial, with the determining source changing from one trial to the next (and the relative influences of each source represented by the corresponding transition probabilities); we will refer to this as the *mixture* hypothesis.

The integration and mixture hypotheses make different predictions regarding participants' reaction times, allocation of gaze, and neural responses. For example, each source of information may have taken a different amount of time to process. Nevertheless, the *integration* hypothesis predicts that responses will be fastest when different sources of information agree and slowest when they disagree. The mixture hypothesis, in contrast, predicts that reaction times will not depend on disagreement between sources because only one source is used on a given trial, and that actions for which the source is ambiguous will be comprised of a mixture of the reaction times corresponding to the possible sources used for that choice. That is, while the exact proportions of trials corresponding to each source is unknown, their mean must necessarily lie within the range of reaction times for each source. However, an individual employing only one source for a given trial would be expected to incur a "switch cost" when the identity of the source changed from one trial to the next (Monsell, 2003). This cost would be associated with increased reaction times when consecutive trials used different sources (as compared to when they used the same source), as well as increased activity in frontal cortex during such "switch trials" (Dove et al., 2000; Rushworth et al., 2002).

To test these predictions, decisions were classified according to whether the three sources of information - current earnings, social, and prior behavior - favored the same action and, if they did not, which was the most likely cause of the action (see Materials and methods: Classification of individual trials using the computational model). Mean reaction times were extracted for each decision class (excluding "stay" actions, which were achieved by allowing the trial to time out) and sorted according to whether the locations of other group members were available (Location and Both conditions) or not (Isolated and Points conditions). The results of this analysis are shown in Fig. 6D. Repeated measures ANOVA revealed a significant effect of decision class ($F_{3,2365}$ = 7.03, p $< 10^{-16}$, $\omega^2 = .03$) and location availability (F_{1,2365} = 107.60, p $< 10^{-23}$, $\omega^2 = .04$), but no interaction between the two factors $(F_{3,2365} = 1.08, p = .35)$. Responses when all three sources of information recommended the same action ("all sources") were significantly faster than when only one source of information recommended the action (p < .05 in all cases; see)Supplementary Table S2 for additional statistical data) - this finding supports the *integration* hypothesis. Furthermore, the fact that the mean reaction time for the "all sources" and "two of three sources" trials lies outside the range of reaction times for the one-source trials suggests that the mixture hypothesis is unlikely to be correct.

With regard to participants' allocation of gaze, the mixture and integration hypotheses again make different predictions. Under the *integration* hypothesis, all sources of information are processed, and the visual stimuli representing these sources must each be processed on every trial. In contrast, the *mixture* hypothesis predicts that only one source is used for any given trial, and that only the visual stimulus representing that source requires attentional resources. While such visual stimuli existed for reward-based decisions and socially-based decisions, it is important to note that no such stimulus existed for perseveration-based decisions (i.e., there was no visible counter for the number of consecutive actions that the participant had made). Because the source being used under the *mixture* hypothesis was ambiguous when multiple decision-making sources agreed, decisions that were associated with a single source represent the most parsimonious route to addressing these different predictions.

We analyzed data from the eye-tracking cohort and calculated the portion of each trial that participants spent looking at their earned points (necessary for reward-based decisions; see Fig. 1) and the locations of the markers for the other group members (necessary for socially-based decisions; see Fig. 1B and D). Figure 7A depicts these results. Repeated measures

ANOVA revealed that participants spent more time looking at their recent earnings when the locations of other group members were not available than when they were available ($F_{1.686}$ = 36.20, p < 10^{-8} , ω^2 = .05). However, there was no effect of decision class, regardless of whether the data were combined $(F_{1.686} = .70, p = .40)$ or analyzed separately for choices when location information was available ($F_{2.513} = .31$, p = .73; Fig. 7A, white bars) and when it was not $(F_{1,341} = .23, p =$.63; Fig. 7B, gray bars). In addition, there was no significant interaction between decision class and location availability $(F_{1.686} = .003, p = .96; see Supplementary Table S3 for$ additional statistical data). In all cases, the mean fraction of each trial spent attending to the participant's points was greater than zero ($p < 10^{-27}$ for all combinations of decision class and location availability). Figure 7A also depicts the repeated measures ANOVA for the time participants spent looking at the markers of other group members. As these markers were only visible under the Locations and Both conditions, analysis was restricted to this subset of the data and averaged across the two conditions. The mean proportion of each trial in which participants directed their gaze at others' markers was greater than zero for all three decision classes ($p < 10^{-31}$), but there was no effect of decision class on gaze duration ($F_{2.513} = .27$, p = .76; see Supplementary Table S4 for additional statistical data). Thus, participants spent significant time attending to both sources of information, but the allocation of their attention did not vary with the identity of the source. Again, these results support the *integration* hypothesis of decision-making.

Analysis revealed no evidence of "switch costs" in the reaction times of participants. By examining single-source trials that were themselves preceded by single-source trials, reaction times were grouped according to whether the source for the preceding trial differed from that of the current trial ("switch trials") or matched that of the current trial ("non-switch trials"). The results of this analysis are shown in Fig. 7B. There were no significant differences between switch and non-switch trials for any given source, regardless of whether location-based information was available or not.

We also examined neural evidence to distinguish between the *mixture* and *integration* hypotheses. This was motivated by the prior observation that mechanisms in dorsolateral prefrontal cortex (dIPFC) and dorsal anterior cingulate cortex dACC) are engaged by information integration and processing conflict (Botvinick et al., 2001; Carter et al., 2000; Carter & van Veen, 2007). Accordingly, the *integration* hypothesis predicts that these systems should show differential activity based on the presence of conflict between the three sources. In contrast, the *mixture* hypothesis — which assumes that only one source impacts decision-making on a given trial — predicts that "switch trials" (i.e., those on which a participant shifts from one source to another for the current decision) should be associated with increased frontal activity relative to nonswitch trials. Data from the neuroimaging cohort were thus used not only to localize neural activity, but to determine which of these theories was more plausible (for other uses of neuroimaging data beyond the localization of function; see Mather, Cacioppo, & Kanwisher, 2013).

As shown in Fig. 7C, there were significant differences in BOLD activity according to the level of disagreement between the sources. For this analysis, the level of disagreement was quantified as the number of sources not recommending the chosen action; these values were included in the GLM analysis as a parametric regressor - see Materials and methods. Figure 7C depicts the results of this analysis for two non-overlapping subsets of the data: conditions for which location-based information was unavailable (Isolated and Points), and conditions for which location-based information was available (Locations and Both). The patterns of activity seen across across the two analyses share a common set of regions that were more active when the sources disagreed. These regions include the dorsal anterior cingulate cortex (dACC), bilateral dorsolateral prefrontal cortex (dlPFC), bilateral parietal cortex, bilateral inferior frontal gyrus, and the left insula (see Supplementary Table S5 for details regarding the neuroimaging results). Thus, there was strong agreement between the two analyses: BOLD responses increased with the number of disagreeing sources. In contrast, there was no significant increase in BOLD response on switch trials relative to non-switch trials.

Additional neuroimaging analysis

While the results detailed above suggest that conflict between different types of information occurred at the level of individual decisions, three important questions remained. First, did the relative weighting of reward-based and socially-based information (represented by the parameter W) correspond to differential patterns of brain activity across subjects? That is, did this trait-level parameter have an effect on how information was processed during the task? Correlation analysis revealed no relationship between W and the magnitudes of responses to screens presented at the beginning or end of trials (i.e., when new information was first presented and when the participant's decision was made). In addition, analysis of differential responses to these screens when other group members' locations were available versus unavailable, corresponding to trials when the socially-based component was defined and undefined, similarly did not reveal a relationship with W. In sum, there was no indication that that participants' emphasis on reward-based or socially-based information altered their neural activity in a global fashion.

There remained a second important question: are the inputs themselves (i.e., reward ratio, angle to the group centroid, and number of previously repeated actions) also processed in a colocalized fashion, or do these inputs have separable neural substrates that are only brought together at a later stage?

Analysis suggested that that latter was the case (see Supplementary Fig. S8 and Supplementary Tables S7-S9). Trial-by-trial values of the reward ratio R were positively correlated with activity in visual cortex, the midbrain, and the putamen - the latter two of which have previously been shown to process comparisons between rewards (comparisons between alternatives in the case of the midbrain - see O'Doherty et al. 2006, and comparison between received and expected rewards in the case of the putamen - see McClure et al., 2003). The group centroid angle on a given trial was significantly correlated with activity in the supplementary motor area (SMA) and prefrontal cortices. This pattern of activity is more difficult to interpret, as angular orientations are typically processed in the visual cortex, but may reflect anticipated changes in the participant's direction of travel (see Materials and methods), as larger centroid angles (and thus higher activity) were associated with such changes. Values for the perseveration-based input to the model (number of consecutively repeated actions) were positively correlated with activity in the vmPFC. Experimental evidence suggests that this region has a role in automatic, context-independent behaviors (Ashby et al., 2010), and lesioning the analogous area in rats (infralimbic prefrontal cortex; Coutureau & Killcross, 2003) causes a transition from perseverative behavior to more goal-directed behavior.

The other remaining question focused on the specificity of the results depicted in Fig. 7. That is, were other regions in the brain also sensitive to disagreement between sources in the proposed model? To answer this question, a control regionof-interest (i.e., one whose activity was not correlated with source conflict like those in Fig. 7) was selected: the vmPFC. This region was not revealed by our source conflict analysis, but has been implicated by previous studies to be involved in the prediction of outcomes based upon integration of multiple sources of information (Behrens et al., 2008; Boorman et al., 2013; Chung et al., 2015). Repeated measures ANOVA examined the relationship between activity in the vmPFC and the number of disagreeing sources according to the model above: the results of this analysis are shown in Supplementary Fig. S9. A joint analysis pooling data for which locations of other group members were available and unavailable revealed no effect of source disagreement, location availability, or interaction (respectively, $F_{2,339} = 2.15$, p = .12; $F_{1,339} = .07$, p = .79; $F_{2,339} = .79$, p = .46). There was also no effect of source disagreement in a separate analysis of trials for which location-based information was available ($F_{3,225}$ = 5.72, p = .31). There was an effect of source disagreement on trials for which location-based information was unavailable $(F_{2,175} = 5.72, p < .005, \omega^2 = .05)$; however, the pattern of activity exhibited by the vmPFC was the opposite to that shown in Fig. 7: vmPFC activity was lower for trials in which there was more disagreement between sources, thereby suggesting a different form of processing in this region.

Discussion

In this study, we used a spatial foraging task performed in a group setting to verify, quantify, and computationally model the influence of reward-based, socially-based, and perseveration-based information. Groups of three participants explored a two-dimensional space in which their location within the space determined their earnings, under conditions that varied the availability and type of information they had about other participants' behavior. Participants' earnings were contingent upon their locations along only one of the two dimensions, so that locations along the other dimension revealed the influence of social information independent of the effects of earnings. Analysis of behavior revealed two important findings. First, participants were sensitive to earnings along the relevant dimension: they preferred profitable locations and avoided unprofitable ones. Second, participants clustered more tightly when information was provided regarding other group members' locations. This occurred along the irrelevant as well as the relevant dimension, and was thus independent of the effects of earnings. From these data, we developed a simple computational model that accounted for these behaviors by combining and weighting summary metrics of recent earnings, recent choices, and the current locations of other group members.

Our findings also indicate that different sources of information were integrated for each decision, rather than one source having exclusive influence on a given decision, with the source varying from decision to decision. This support came in multiple forms: reaction time data, duration of gaze, and neuroimaging analysis. In the latter case, regions such as the parietal and dorsolateral prefrontal cortices were more active when the sources of information favored different actions than when they favored the same action. Prior experimental evidence has indicated that parietal cortex exhibits increased activity when the response favored by social information conflicts with that favored by the participant's own experience (Berns et al., 2005). In addition, the dlPFC has been shown to be critical for exerting top-down control in a variety of domains (Dosenbach et al., 2008; MacDonald et al., 2000; Miller & Cohen, 2001), including social decision-making (Knoch et al., 2006; Sanfey et al., 2003; Soutschek, Sauter, & Schubert, 2015).

The dorsal anterior cingulate cortex (dACC) also exhibited this pattern of activity. A recent paper by Chung et al. also found increased activity in the dACC when the choices of other group members were incongruent with the participant's own choice (Chung et al., 2015): this would be analogous to competing influences by the reward-based and socially-based components in our model. Here, we augment our understanding of the dACC's role in processing conflict by adding a third conflicting component: the participant's own choice history. Another recent paper by Suzuki et al. investigates this phenomenon in the context of consensus decision-making (Suzuki et al., 2015). In this case, experimental evidence suggests that the dACC may be the site of integration for social decision-making.

One region noticeably absent from the results of this analysis was the vmPFC - a region previously proposed as an integration site (Behrens et al., 2008; Hackel et al., 2015; Suzuki et al., 2012). Our analysis indicates that processing in the vmPFC was fundamentally different from that in the aforementioned areas (in the Isolated and Points conditions, activity actually *decreased* as conflict increased). There are two important distinctions between the models employed in this literature and that which we propose here. First, the vmPFC in these studies was reported to have been integrating the values of expected outcomes (or discrepancies between pieces of information indicative of these values), rather than integrating potentially conflicting actions recommended by separable components. In many previously used paradigms, this difference is less notable: decisions should correspond to (or at least be biased toward) high-value outcomes, and social influence can be transformed easily into quantities that exist in the same domain as the outcome. This is the basis of the second important distinction. The experiment presented here was designed to explicitly examine the integration of fundamentally different sources of information that could not be easily transformed into a single expected value: quantifiable reward, the behavior of other group members, and the history of the individual's choices. Both of these factors (integration of outcome value vs. integration of separate decision components, and the immiscible nature of the sources of information) indicate potentially different roles for the vmPFC and the regions revealed by our analysis.

Prior work has elucidated the possibility that separate populations of neurons within the midbrain and striatum may allow these structures not only to evaluate reward-based value, but also the information-based value of a stimulus or decision (Bromburg-Martin, 2010). Recent studies have indeed provided support for such a distinction in the human striatum (e.g., see Tricomi & Fiez, 2012; Iigaya, 2016; Smith, 2016). Although we did not model the value of information accrued across trials in the experiment (and constructing a principled quantification of such value would present a complex inference problem in and of itself), there is certainly variance left to account for in the behavioral data. An information-based source could provide an important, fourth component to the model presented here, and provides an interesting avenue for future work.

In our computational model, the parameter W was used to represent the balance between reward-based and sociallybased information. Analysis revealed that participants emphasized social information slightly more heavily than reward-based information. However, this balance varied significantly across participants, and values of W fit to single individuals could be used to predict the severity of certain maladaptive behaviors. On one end of the spectrum, insensitivity to social cues and/or a lack of prosocial behaviors (analogous to low values of W) might be characteristic of autism spectrum disorder (Baron-Cohen, 2001; Frith, 2001). On the other end of the spectrum, overdependence on social information when selecting actions (analogous to high values of W) might be used to assess and/or predict an individual's susceptibility to peer pressure, particularly in populations that are heavily biased by the behavior of others (Gardner & Steinberg, 2005). It is also worth noting that our computational model placed a uniform weighting on information coming from each group member (i.e., the group centroid was equidistant between the markers of other group members). While unequal weightings of social information based on the degree of social closeness have indeed been shown to affect reward-based decision-making (Fareri et al., 2012; Fareri, Chang, & Delgado, 2015), the experiment presented here was conducted in the absence of personal information regarding other group members. Therefore, this potentially intriguing modification of the computational model must be set aside for future work.

Individual differences may exist not only in the relative emphasis on reward-based and socially-based information, but also in the specific mapping between information and actions. In our model, this mapping was captured by four probability functions that varied according to a single variable (reward ratio for reward-based actions, the angle between direction of travel and the group centroid for socially-based actions). Important individual differences could include the likelihood of continuing or reversing in the face of repeatedly decreasing earnings. If such differences were stable, they could be used not only to predict performance in this task (such as the likelihood of reaching the global maximum on the valley task), but to predict more general attributes, such as the willingness to explore unknown options. Additional experiments are needed to determine the generalizability of each participant's parameters.

We also assumed that the mapping between earnings and action probabilities was constant across social conditions, when in fact reward-based action probabilities may have changed subtly once the participants knew that their performance was on display. Perhaps individuals played more conservatively, increasing the probability of "reverse" actions when earnings decreased or staying in the same location regardless of current earnings. While simultaneously fitting the reward-based action probabilities and socially-based action probabilities for each participant in each separate social condition is difficult, it is not impossible *in principle*, and these efforts are in progress.

We have also not modeled the behavior exhibited during the Points condition. Preliminary analyses of behavior and direction of gaze indicate that participants attend to, and are affected by, this information. Future work will focus on this condition (and the presence of other participants' earnings during the Both condition) in an attempt to more comprehensively model the integration of reward-based and socially-based information.

Taken together, the results reported here suggest that humans respond to both earnings and social information in a systematic fashion, and that the impact of such information on decision-making behavior can be captured by computational modeling. Furthermore, the results of our model may be useful in constraining the space of possible psychological and neural processes that support the integration of these different streams of information. By achieving a better understanding of the interplay between social and non-social influences, we can endeavor to avoid circumstances in which those influences are overemphasized (or ignored) to our detriment.

Acknowledgments This work was supported by the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-07-1-0528, Multidisciplinary University Research Initiative. The authors thank N. Apple and the Hyperscan Development Team at Baylor College of Medicine for assistance in the implementation of the experiment in the Network Experiment Management Object software (NEMO; http://labs. vtc.vt.edu/hnl/nemo/index.html), and the staff of the Human Neuroimaging Laboratory at Baylor College of Medicine for assistance in the collection of these data: K. Bartley, C. Symank, C. Howard, M. Ross, and J. White.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. S. Guetzkow (Ed.), *Groups Leadership and Men*. Pittsburgh: Carnegie Press.
- Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences*, 14(5), 208–215.
- Baddeley, M. (2010). Herding, social influence and economic decisionmaking: socio-psychological and neuroscientific analyses. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 365*(1538), 281–290.
- Baron-Cohen, S. (2001). Theory of mind and autism: a review. International Review of Research in Mental Retardation: Autism, 23, 169–184.
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, 456(729), 245– 249.
- Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *The Journal of Neuroscience*, 21(8), 2793–2798.
- Berns, G. S., Chappelow, J., Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., & Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biological Psychiatry*, 58(3), 245–253.
- Bikhchandani, S., & Sharma, S. (2000). Herd behavior in financial markets. *IMF Staff Papers*, 279-310.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models

of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), 700–764.

- Boorman, E. D., O'Doherty, J. P., Adolphs, R., & Rangel, A. (2013). The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron*, 80(6), 1558–1571.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624–652.
- Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, 30(2), 619–639.
- Bromburg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, 68(5), 815–834.
- Camus, M., Halelamien, N., Plassmann, H., Shimojo, S., O'Doherty, J., Camerer, C., & Rangel, A. (2009). Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex decreases valuations during food choices. *European Journal of Neuroscience*, 30(10), 1980–1988.
- Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., & Cohen, J. D. (2000). Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 97(4), 1944–1948.
- Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience,* 7(4), 367–379.
- Chang, L. J., & Sanfey, A. G. (2011). Great expectations: neural computations underlying the use of social norms in decision-making. *Social Cognitive and Affective Neuroscience*, nsr094.
- Chung, D., Christopoulos, G. I., King-Casas, B., Ball, S. B., & Chiu, P. H. (2015). Social signals of safety and risk confer utility and have asymmetric effects on observers' choices. *Nature Neuroscience*, 18(6), 912–916.
- Costa, P. T., & McCrae, R. R. (1985). Manual for the NEO personality inventory. Odessa: Psychological Assessment Resources.
- Coutureau, E., & Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behavioral Brain Research*, 146(1), 167–174.
- Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4p1), 377–383.
- Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cognitive, Affective, & Behavioral Neuroscience, 3*(1), 27–38.
- Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12(3), 99–105.
- Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an eventrelated fMRI study. *Cognitive Brain Research*, 9(1), 103–109.
- Fareri, D. S., Chang, L. J., & Delgado, M. R. (2012). Effects of direct social experience on trust decisions and neural reward circuitry. *Frontiers in Neuroscience*, 6, 148.
- Fareri, D. S., Niznikiewicz, M. A., Lee, V. K., & Delgado, M. R. (2012). Social network modulation of reward-related signals. *The Journal of Neuroscience*, 32(26), 9045–9052.
- Fareri, D. S., Chang, L. J., & Delgado, M. R. (2015). Computational substrates of social value in interpersonal collaboration. *The Journal of Neuroscience*, 35(21), 8170–8180.
- Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., & Mazziotta, J. C. (1997). *Human Brain Function*. San Diego: Academic Press.
- Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiack, R. S. J., & Turner, R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, 2(1), 45–53.

- Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, 32(6), 969–979.
- Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental Psychology*, 41(4), 625–635.
- Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwartz, Y., Ghosh, S. S., Maumet, C., Sochat, V. V., Nichols, T. E., Poldrack, R. A., Poline, J.-B., Yarkoni, T., & Margulies, D. S. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in Neuroinformatics*, 9, 8.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105–2108.
- Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. *Nature Neuroscience*, 18(9), 1233–1240.
- Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decisionmaking involves modulation of the vmPFC valuation system. *Science*, 324(5927), 646–648.
- Herrnstein, R. J. (1961). Relative and absolute strength of responses as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behaviour*, 4(3), 267–272.
- Ligaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J., & Dayan, P. (2016). The modulation of savouring by prediction error and its effects on choice. *eLife*, 5, e13747.
- Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., & Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, 61(1), 140–151.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, 314(5800), 829–832.
- Knutson, B., Wimmer, G. E., Kuhnen, C. M., & Winkielman, P. (2008). Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *NeuroReport*, 19(5), 509–513.
- Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., & Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences*, 89(12), 5675–5679.
- Latane, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10(3), 215.
- MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472), 1835–1838.
- Mason, M. F., Dyer, R., & Norton, M. I. (2009). Neural mechanisms of social influence. Organizational Behavior and Human Decision Processes, 110(2), 152–159.
- Mason, W., & Watts, D. J. (2012). Collaborative learning in networks. Proceedings of the National Academy of Sciences, 109(3), 764–769.
- Mather, M., Cacioppo, J. T., & Kanwisher, N. (2013). How fMRI can inform cognitive theories. *Perspectives on Psychological Science*, 8(1), 108–113.
- McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, 38(2), 339–346.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics (New York*. NY: Academic Press.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312), 75–78.
- Milgram, S. (1963). Behavioral Study of Obedience. *The Journal of Abnormal and Social Psychology*, 67(4), 371–378.

- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. Annual Review of Neuroscience, 24(1), 167–202.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, 16(5), 1936–1947.
- Montague, P. R., Berns, G. S., Cohen, J. D., McClure, S. M., Pagnoni, G., Dhamala, M., & Fisher, R. E. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *NeuroImage*, 16(4), 1159– 1164.
- Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36(2), 265–284.
- Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, 16(1), 72–80.
- Nichols, T. E., & Holmes, A. P. (2001). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, 15(1), 1–25.
- O'Doherty, J. P., Buchanan, T. W., Seymour, B., & Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*, 49(1), 157–166.
- Ogawa, S., Lee, T. M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance* in *Medicine*, 14(1), 68–78.
- Raafat, R. M., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in Cognitive Sciences*, 13(10), 420–428.
- Rangel, A., & Hare, T. (2010). Neural computations associated with goaldirected choice. *Current Opinion in Neurobiology*, 20(2), 262–270.
- Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *The Journal of Neuroscience*, 28(51), 13775–13785.
- Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, 15(8), 549–562.
- Rushworth, M. F. S., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *Journal of Neurophysiology*, 87(5), 2577–2592.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626), 1755–1758.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36(4), 393–414.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.
- Smith, D. V., Rigney, A. E., & Delgado, M. R. (2016). Distinct reward properties are encoded via corticostriatal interactions. *Scientific Reports*, 6, 20093.
- Soutschek, A., Sauter, M., & Schubert, T. (2015). The importance of the lateral prefrontal cortex for strategic decision making in the prisoner's dilemma. *Cognitive, Affective, & Behavioral Neuroscience*, 15(4), 854–860.
- Stanley, D. A., & Adolphs, R. (2013). Toward a neural basis for social behavior. *Neuron*, 80(3), 816–826.
- Stanley, D. A. (2015). Getting to know you: general and specific neural computations for learning about people. *Social Cognitive and Affective Neuroscience*, 11(4), 525–536.
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: neural currencies for

valuation and decision making. *Nature Reviews Neuroscience*, 6(5), 363–375.

Surowiecki, J. (2005). The Wisdom of Crowds. Random House LLC.

- Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., Cheng, K., & Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron*, 74(6), 1125–1137.
- Suzuki, S., Adachi, R., Dunne, S., Bossaerts, P., & O'Doherty, J. P. (2015). Neural mechanisms underlying human consensus decision-making. *Neuron*, 86(2), 591–602.
- Toelch, U., & Dolan, R. J. (2015). Informational and normative influences in conformity from a neurocomputational perspective. *Trends* in Cognitive Sciences, 19(10), 579–589.
- Tomlin, D., Nedic, A., Prentice, D. A., Holmes, P., & Cohen, J. D. (2013). The neural substrates of social influence on decision making. *PloS One*, 8(1), e52630.
- Tricomi, E., & Fiez, J. A. (2012). Information content and reward processing in the human striatum during performance of a declarative memory task. *Cognitive Affective and Behavioral Neuroscience*, 12(2), 361–372.
- Xiang, T., Lohrenz, T., & Montague, P. R. (2013). Computational substrates of norms and their violations during social exchange. *The Journal of Neuroscience*, 33(3), 1099–1108.