Wise or mad crowds? The cognitive mechanisms underlying information cascades

Alan N. Tump¹ (tump@mpib-berlin.mpg.de), Tim Pleskac¹,², Ralf H. J. M. Kurvers¹

¹Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany
²University of Kansas, Department of Psychology, Jayhawk Blvd, Lawrence, KS 66045, USA

November 26, 2019

Abstract

Whether getting vaccinated, buying stocks, or crossing streets, people rarely make decisions alone. Rather, multiple people decide sequentially, setting the stage for information cascades whereby early-deciding individuals can influence others’ choices. To understand how information cascades through social systems, it is essential to capture the dynamics of the decision-making process. We introduce the social drift-diffusion model to capture these dynamics. We tested our model using a sequential choice task. The model was able to recover the dynamics of the social decision-making process, accurately capturing how individuals integrate personal and social information dynamically over time and when they timed their decisions. Our results show the importance of the interrelationships between accuracy, confidence, and response time in shaping the quality of information cascades. The model reveals the importance of capturing the dynamics of decision processes to understand how information cascades in social systems, paving the way for applications in other social systems.

Keywords: collective intelligence, information cascades, sequential decision-making, social information, confidence, herding, conformity, social impact, collective behaviour

Main

In many situations—be they financial investments, consumer choices, or simply crossing the street—one is generally not making a decision alone. Rather, there are multiple others present each making their own decisions. In such situations, decision-makers can observe the choices of others and use that information to inform their own decisions. Early-deciding individuals can thereby trigger information cascades, in which later-deciding individuals adopt earlier choices, potentially creating a situation where, in the extreme case, everyone does what everyone else is doing, even at the expense of abandoning their private information (Banerjee [1992], Anderson and Holt [1997], Bikhchandani et al. [1998], Gallup et al. [2012]).
Yet for a myriad of reasons—from limited time and computational resources to biases in the decision process—people’s choices do not always perfectly reflect the true state of the world. Information cascades can thus promote both positive and negative outcomes: in online environments, for example, both true and false news can spread quickly (Vosoughi et al., 2018); in offline environments, the behaviour of initial pedestrians crossing a road can amplify both safe and risky behaviours in other pedestrians (Faria et al., 2010; Pfeffer and Hunter, 2013). Understanding the conditions leading to positive and negative information cascades is crucial across many domains, including financial markets (Welch, 2000; Shiller, 2002), consumer preferences (Chen, 2008), political opinion formation (Battaglini, 2005), and opinion dynamics in social networks (Xiong and Liu, 2014).

To understand the conditions underlying positive and negative information cascades, we need to comprehend the timing of individual decisions as well as how individuals integrate personal and social information (i.e., other people’s decisions) dynamically over time. We, currently, however, lack a detailed understanding of the individual decision process in sequential choice paradigms. Many models of information cascades assume a random decision order and are thus ill-equipped to predict who will respond earlier and why (e.g., Anderson and Holt, 1997; Banerjee, 1992; Bikhchandani et al., 1998; Deneubourg et al., 1990; Mann, 2018; Sumpter and Pratt, 2008). When models of information cascades do refer to the timing of decisions, they do so from an optimal Bayesian perspective based on the quality of each individual’s private information (e.g., Chamley and Gale, 1994; Gul and Lundholm, 1995; Zhang, 1997; Ziegelmeier et al., 2005). Yet we know that people’s actual choice behaviour often deviates systematically from optimal Bayesian models (Hertwig et al., 2019; Pleskac and Busemeyer, 2010; Tversky and Kahneman, 1974).

To address these shortcomings, we developed a dynamic theory of social decision making by focusing on each individual’s decision process. As a basis, we took a well-established modeling framework of individual decision making that models decisions as a dynamic process in which information is accumulated as evidence over time until a threshold is reached (e.g., Edwards, 1965; Laming, 1968; Link and Heath, 1975; Ratcliff, 1978; Stone, 1960; Usher and McClelland, 2001). This evidence accumulation process has been successful in accounting for a wide range of decisions in domains including perception (Ratcliff and Smith, 2004), memory (Ratcliff, 1978), categorization (Nosofsky and Palmeri, 1997), preference (Busemeyer and Diederich, 2002; Busemeyer and Townsend, 1993), and inference (Pleskac and Busemeyer, 2010), and has successfully been applied to analyse the influence of static social information (Germar et al., 2014; Toelch et al., 2018).

We extended this evidence accumulation framework by showing how the choices of others are integrated with personal information and together accumulated as evidence. This approach provides a process-level account of the choices and response times of individuals in dynamic social systems. We tested the model in an empirical study. Findings showed that participants self-organize based on the quality of their personal information so that later deciders benefit from observing the choices of early deciders. Fitting the model to the data allowed us to test several hypotheses about how individuals simultaneously combine personal and social information, and how they time their decision in groups. In addition, we reveal mechanisms leading to the amplification of correct or incorrect cascading information.

The Social Drift-Diffusion Model

Models of the evidence accumulation process during decision-making include the drift-diffusion model (DDM; Ratcliff, 1978; Ratcliff and McKoon, 2008), the linear ballistic accumulator model (Brown and Heathcote, 2008), and the leaky competing accumulator model (Usher and McClelland, 2001). Most of these models can, in principle, be extended to model a social system. Here, we focus on the DDM, arguably the most successful framework for accounting for human choice behaviour, including...
some of the most basic aspects of the decision process, such as the speed–accuracy trade-off (Ratcliff and Smith, 2004; Voss et al., 2004), the construction of preferences (Busemeyer and Townsend, 1993), the formation of confidence judgements (Pleskac and Busemeyer, 2010), the emergence of response biases (Leite and Ratcliff, 2011; Pleskac et al., 2018), and how attention guides the evidence accumulation process (Diederich, 1997; Krajbich and Rangel, 2011).

According to the DDM, people faced with a choice between two options, A or B, base their choice on an internal level of evidence. Initially, people can have a bias and lean towards either option. This is modeled as an initial level of evidence. Over time, people extract further information about the options and accumulate this information as evidence. This accumulation gives rise to an evolving (latent) level of evidence, as depicted by the jagged line in Figure 1a. The jaggedness arises because each sample of evidence is noisy (i.e., the stimuli itself and the cognitive and neural processes introduce variability into the evidence accumulation). Once a choice threshold has been reached, a decision is made. If the accumulated evidence reaches the upper threshold, option A is selected; if it crosses the lower threshold, option B is selected. The time it takes for the evidence to reach either threshold is the predicted response time. In the social DDM, we modify this framework to cover multiple individuals accumulating evidence at the same time (Fig. 1a). In this case, the evidence comes from two sources: personal information, gathered from sampling the physical environment (e.g., for visual or auditory cues), and social information, gathered by observing the behaviour of others (Dall et al., 2005; Galef and Laland, 2005).

Formally, we denote the cumulative evidence at time point \( t \) as \( L(t) \). At the start, individuals may favour one option over the other, described by their start point \( L(t=0) = \beta \). Here, their start point is based on previously collected personal information and is estimated from confidence ratings provided during the initial stage of the decision task. However, the start point can also represent initial biases towards either option (e.g., Voss et al., 2004). At each time step \( \Delta t \), the current state of evidence \( L(t) \) is updated by sampling new evidence until a decision is made (i.e., until the level of evidence reaches the choice threshold \( \theta \)):

\[
L(t + \Delta t) = L(t) + [\delta_p + \delta_s(t)] \times \Delta t + \sqrt{\Delta t} \times \epsilon, \tag{1}
\]

where \( \epsilon \) is Gaussian white noise (i.e., the diffusion process) with a mean of 0 and a variance of 1. The parameters \( \delta_p \) and \( \delta_s(t) \) correspond to the strength of the personal and social information uptake, respectively. Personal information uptake describes the integration of information extracted directly from the physical environment, as well as the evaluation of information from memory. Social information is defined as the size of the majority of individuals \( M(t) \) who already decided at time point \( t \) (see also Bikhchandani et al., 1998):

\[
M(t) = N_A(t) - N_B(t), \tag{2}
\]

where \( N_A(t) \) and \( N_B(t) \) are the number of individuals who have already decided for option A or B, respectively. The impact of majority size on the social drift rate is described by a power function, analog to Latané (1981):

\[
\delta_s(t) = s \times M(t)^q. \tag{3}
\]

The parameter \( s \) is a scaling factor that influences the strength of the social drift; \( q \) governs the shape of the power function. When \( q = 1 \), each additional choice for the majority option has the same influence on the social drift rate (i.e., a linear effect); when \( q > 1 \) (\( q < 1 \)), each additional choice for the majority option has an increasingly stronger (weaker) impact on the social drift rate. Note that, in contrast to the individual drift rate, the social drift rate can vary over time (indicated by the changing direction of the arrows in Fig. 1a). By incorporating a social drift into the classical
Table 1: Description of the parameters of the social DDM

<table>
<thead>
<tr>
<th>Model feature</th>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nondecision time</td>
<td>$\tau$</td>
<td>Response latency (e.g., motor response time). The parameter $\tau$ describes the time relative to the individual’s fastest response.</td>
</tr>
<tr>
<td>Start point</td>
<td>$\beta = \frac{1}{1 + e^{-(C-b)}}$</td>
<td>The start point is a function of the confidence in the personal choice $C$, which ranges from highly confident but incorrect to highly confident and correct (Fig. 4b). The parameter $a$ determines how sensitive the start point is to changes in confidence; $b$ captures other factors besides confidence in the personal decision that impact the start point.</td>
</tr>
<tr>
<td>Personal drift rate</td>
<td>$\delta_p$</td>
<td>The average rate of evidence accumulation supporting the personal choice (Fig. 4c).</td>
</tr>
<tr>
<td>Social drift rate</td>
<td>$\delta_s = s \times M(t)^q$</td>
<td>The social drift rate describes the impact of social information, with $s$ being a scaling parameter that influences the strength of the social drift rate, and $q$ being a parameter that shapes the power function describing the relationship of majority size $M(t)$ and social drift rate (Fig. 4d).</td>
</tr>
<tr>
<td>Choice threshold</td>
<td>$\theta$</td>
<td>The amount of evidence an individual has to accumulate to make a decision; $\theta (-\theta)$ reflects the correct (incorrect) choice threshold (Fig. 4e).</td>
</tr>
</tbody>
</table>

DDM, the social DDM can account for individuals being emitters and receivers of social information and thereby capture the dynamic information exchange among group members.

In sum, the social DDM characterizes (i) how individuals incorporate personal information with the parameters $\beta$ and $\delta_p$, (ii) how individuals incorporate social information depending on the majority size via the parameters $s$ and $q$, and (iii) individuals’ willingness to wait for social information with the parameter $\theta$ (see Table 1 for all parameter descriptions).

The predator detection task

We tested the social DDM in an empirical study (see Fig. 1b; see Methods for full details). In brief, participants were divided into groups of varying sizes (‘small’, ‘medium’, or ‘large’). Each group of participants was seated together in a single room, facing a large screen. Participants were asked to imagine being a fish in a school facing a choice between two alternatives—namely whether to
Figure 1: Illustration of the social DDM and the experimental paradigm. (a) A generic example of the social DDM with five individuals, each represented by a jagged line. The start point of each individual indicates the personal evidence accumulated up to that point. At the start, no individual exceeds the choice threshold and social information is absent, implying no social drift (as indicated by the horizontal arrow). Individuals who begin close to either of the thresholds (red lines) are likely to choose early, providing social information for undecided individuals. This social information impacts the rate of evidence accumulation, with the drift rate shifting towards the choice threshold favoured by the majority (as indicated by the arrow pointing upwards). (b) The stages of the predator detection task. During the personal phase, individuals briefly observe a grid of ‘sharks’ and ‘tuna.’ They then make a personal decision whether to ‘Stay’ or ‘Escape’ and report their confidence in that decision. In the subsequent social phase, they are asked to make a second decision on whether to ‘Stay’ or ‘Escape,’ but now they can freely time their decisions and simultaneously observe the choices of others before doing so. Finally, the correct answer is displayed, and the next trial begins (with 40 trials in total).

escape or not—depending on the presence of predators, in this case, sharks. They were instructed to escape when five or more sharks were present and to stay when four or fewer sharks were present. At each trial, participants were shown—for 2 seconds—a grid with a varying number (3, 4, 6, or 7) of sharks hidden among harmless fish. Participants first made a personal choice on whether to ‘Stay’ or ‘Escape’ and then reported their confidence in that choice on a scale from 50% to 100%.
They then entered the social phase, in which they had a maximum of 20 seconds to make a second decision on whether to ‘Stay’ or ‘Escape’, but without seeing the grid again. Instead, the display showed a count of the number of choices for each option. Participants were free to enter their choice at any point in time; they could thus respond early (thereby providing social information) or wait to observe the decisions of others. However, they could only decide once. Finally, we provided feedback on the correct choice.

Results

Empirical results: groups show beneficial self-organization according to information quality

Participants achieved an accuracy of 74% in their personal choice (Fig. 2a), and participants reporting higher confidence in their personal choice were also more accurate (Fig. 2b; $\beta = 3.82$, CI $= [3.35, 4.28]$). Participants were thus—at least partly—aware of the quality of their personal information.

We fitted a Two-Stage Dynamic Signal Detection model (2DSD; Pleskac and Busemeyer 2010) to the choice, RT, and confidence data from the personal phase (see Supplementary Results and Discussion). The close correspondence between the model and the data suggests that a drift–diffusion process is a good description of the decision process during this stage of the experiment.

With an average accuracy of 79%, participants’ choices during the social phase, where they had
Figure 3: **Empirical results and predictions of the social DDM.** Participants reporting higher confidence in their personal choice (a) improved less and (b) responded earlier during the social choice. (c) The larger the majority favouring the opposing option, the more likely participants were to change their decision. (d) The choices of participants who responded later in the social choice were less accurate in the personal choice (declining blue dots) but improved more in the social choice (indicated by the increasing difference between blue and yellow dots at later RTs). For visualization purpose, RTs are binned by rounding to the closest integer. RTs greater than 13 seconds (less than 1%) were assigned to the 12 seconds bin. (a–d) The dashed lines show the choices and RTs predicted by the social DDM, accurately capturing all relationships. For frequency distributions, see Supplementary Figure S2. (e) Participants improved most when more confident individuals were more accurate (yellow dots) and responded earlier. Numbers indicate the number of trials. For all panels, the points and error bars depict the mean and the 95% credible intervals of the posterior distribution of the Bayesian regression model.

the opportunity to wait for social information before choosing again, were more accurate (Fig. 2: $\beta = 0.3, CI = [0.20, 0.39]$). The reported level of confidence in their personal choice predicted their likelihood to improve (Fig. 3: $\beta = -4.27, CI = [-4.88, -3.68]$): participants reporting the lowest confidence level improved in more than 15% of trials; whereas the most confident, in less than 1% of trials. Why do unconfident participants achieve such higher gains from the social process? There are two mechanisms underlying this. First, participants reporting lower confidence waited longer before making a decision during the social phase (Fig. 3: $\beta = -4.86, CI = [-5.22, -4.5]$). Second, participants partly adopted the decisions of others (Fig. 3: $\beta = 0.62, CI = [0.57, 0.67]$): the
larger the majority for the opposing option, the more likely participants were to change their decision. Individuals rarely changed their minds if the majority agreed with their personal decision. As Supplementary Figure S1 shows, participants followed both correct and incorrect majorities, highlighting the importance of the accuracy of early-deciding participants for triggering positive/negative information cascades. Figure 3d shows the consequences of these patterns: participants whose personal choices were accurate (and confident) tended to respond early in the social phase, whereas those whose choices were inaccurate (and unconfident) tended to wait longer, as illustrated by the downward trend of the blue dots (slope: $\beta = -0.16$, CI = [$-0.18, -0.14$]). The latter participants increased their accuracy during the social phase through social influence, as illustrated by the higher yellow dots compared to the corresponding blue dots at higher RTs (interaction: $\beta = 0.11$, CI = [0.09, 0.13]).

Participants in groups thus self-organized according to information quality, with confident and accurate participants deciding early, thereby providing high-quality information for the less confident and less accurate participants, who decided later. This beneficial self-organization depended on two crucial aspects: (i) a positive relationship between confidence and accuracy of personal choice across group members, and (ii) a negative relationship between confidence and RT during the social choice phase. As Figure 3d illustrates, groups showed the highest improvement when both conditions were met, and this occurred in the majority of trials. Improvement was credibly lower for all other conditions (Supplementary Table S1).

Model results: the cognitive mechanisms driving self-organization

To understand the processes leading to the self-organization of groups, we need to understand the cognitive mechanisms underlying individuals’ dynamic integration of personal and social information over time. To this end, we developed the social DDM (Fig. 1a; Table 1), which allowed us to test competing hypotheses on how participants integrate personal and social information over time. We examined three model features: (A) Individuals base their start point on their personal decision and reported confidence. (B) When participants start drifting, they drift towards the correct option, their initially chosen option, or neither of the two. (C) When social information becomes available, participants drift towards the option favoured by the majority. We tested several candidate models composed of various combinations of these three features, and used the deviance information criterion (DIC; Spiegelhalter et al. (2002)) to compare their performance. Figure 4a shows the models’ DIC values relative to that of the best model (see also Supplementary Table S2). In the following, we present the results of the model with the lowest DIC (see Supplementary Table S3 for parameter estimates). Finally, to test how the cognitive mechanisms were affected by group size, we compared the different group sizes (Supplementary Table S4).

Individuals incorporate personal information via start point and drift rate

Participants incorporated their personal information (i.e., personal choice and confidence) during the social decision process in two distinct ways. First, consistent with current models of choice and confidence judgements (Moran et al. 2015; Pleskac and Busemeyer 2010; Yu et al. 2015), they shifted their start point towards their initially chosen option: Individuals who reported higher confidence in the [in]correct option started closer to the threshold of the [in]correct option (Fig. 4b: small: $a = 4.20$, CI = [3.11, 5.35]; medium: $a = 3.42$, CI = [2.81, 4.07]; large: $a = 3.90$, CI = [3.46, 4.37]). This implies that individuals with high confidence in their personal choice were more likely to decide in favour of this option and to do so fast. Second, participants drifted towards the threshold of their initially chosen option (Fig. 4c; small: $\delta_p = 0.65$, CI = [0.45, 0.86]; medium: $\delta_p = 0.62$, CI
Figure 4: Model comparison and individual- and group-level fittings of the social DDM for different group sizes. (a) The deviance information criterion (DIC) values of all models relative to the model with the lowest DIC. The model with the lowest DIC (i.e., preferred model) features a (i) confidence-dependent start point, (ii) drift towards the initially chosen option, and (iii) social drift. (b) Participants reporting higher confidence in the correct/incorrect choice started closer to the correct/incorrect decision threshold at y-value 1/0. (c) Evidence tended to drift towards the choice threshold of the option chosen during the personal phase. (d) The larger the majority favouring an option, the more strongly participants drifted towards the choice threshold favoured by the majority. Participants in smaller groups had a stronger drift given the same majority size. (e) The choice threshold $\theta$, reflecting a participant’s willingness to wait for social information, did not differ between group sizes. Grey lines/dots represent individual-level fittings; coloured lines/dots, the estimates on a group size-level.

$$= [0.50, 0.75]; \text{large: } \delta_p = 0.53, \text{CI} = [0.47, 0.59]).$$ Both processes were independent of group size (Supplementary Table S4). To sum up, across all group sizes, highly confident participants started close to the choice threshold of their initially chosen option and, on top of that, drifted towards that option, whereas participants with low confidence started out unbiased (i.e., in the middle between the thresholds).

Individuals incorporate social information via drift rate

We found that the drift rates were credibly influenced by the majority (Fig. 4d). The larger the majority favouring an option, the more strongly participants drifted towards that option. The shape
of the relationship between majority size and social drift rate (the $q$ parameter) differed between group sizes (small vs. medium: $q = 0.82$, CI = [0.22, 1.44]; medium vs. large: $q = 0.27$, CI = [0.14, 0.41]). In small groups, the drift rate increased exponentially with increasing majority size. In larger groups, each additional individual voting for the majority had less impact than the preceding one, and this function followed a concave shape. Accordingly, the influence of a single individual was larger in small groups than in large groups. Comparing the strength of the personal drift rate (i.e., towards the choice threshold of the initially chosen option) to the social drift rate (i.e., towards the option favoured by the majority) showed that a majority of approximately two is required to counteract an individual’s tendency to drift towards the choice threshold reflecting their initial choice. This highlights participants’ tendency to give personal information more weight than social information. Corroborating this finding, Figure 3c shows that a majority of approximately four participants in favour of the opposing option is required to induce a 50% likelihood of changing a participant’s decision. Finally, we found that participants’ willingness to wait for social information, captured by the threshold parameter $\theta$, did not differ between group sizes (Fig. 4e).

Model predictions: the social DDM captures the self-organizing dynamics

Importantly, the model described above was able to recover all the key features of the dynamics of the social decision-making process. The dashed lines in Figure 3 show the model predictions of the social DDM. In line with the empirical data, the social DDM predicts that unconfident participants wait longer before making a decision (Fig. 3b), that individuals are increasingly likely to follow the majority as the size of that majority increases (Fig. 3c), and that participants whose personal choices were inaccurate wait longer and improve more during the social phase (Fig. 3d). As a result, participants with low confidence in their personal choice improved most (Fig. 3a). We investigated the validity of the model with a parameter recovery analysis (see Supplementary Information). For all parameters, the generating and recovered parameters were highly correlated, implying that each parameter describes a distinct mechanism. Further, all recovered parameter estimates were close to the generating parameters, affirming the validity of the magnitude of the parameter estimates as captured by the social DDM (Supplementary Fig. S3).

Discussion

We have shown that the behaviour of individuals in a social sequential decision-making task can be described by an evidence accumulation process whereby personal and social information is integrated until a decision is made, formalized by the social DDM. The model accurately predicts decision time and choice by taking personal information, social information, and the willingness to wait for social information into account. It successfully captured all the interrelationships of the key behavioural results of the social phase, thereby revealing the cognitive underpinnings of the group-level self-organization according to information quality. Measuring how individuals process personal and social information affords a deeper understanding of how individuals in a social environment cope with the complex problem of evaluating personal information, how they time their decision, and incorporate social information.

During the social decision-making process, individuals incorporated personal information in two ways: at the start of the process, they adjusted their subjective level of evidence to their confidence (i.e., they adjusted their start point), and during the process, they reinforced their ‘belief’ in their original choice over time (i.e., they drifted towards the decision threshold of their personal choice). We also found evidence for such ‘belief reinforcement’ over time in the personal phase (see 2DSD model analysis in the Supplementary Information). The reinforcement of initial beliefs can
potentially have a large influence in real-world social choices. Because individuals generally gather personal information before receiving social information, reinforcement of initial beliefs can lead to situations where even strong counterfactual social information may no longer prove persuasive (i.e., confirmation bias; Klayman, 1995; Kloriat et al., 1980; Nickerson, 1998). Many studies have found that individuals indeed weight personal information more strongly than social information, a phenomenon called egocentric discounting (e.g., Jayles et al., 2017; Larrick and Soll, 2006; Noes Tump et al., 2018; Yaniv and Kleinberger, 2000). In almost all of these studies, participants made a personal judgement before receiving social information. When the order was reversed, the influence of social information indeed increased (Koehler and Beauregard, 2006). Our finding of belief reinforcement provides a compelling explanation for egocentric discounting, simply by providing personal information first. Future studies could test whether increasing the length of the delay between personal choice and provision of social information reduces the influence of social information, as predicted by the social DDM.

When looking at how social information entered the evidence accumulation process, we found that individuals incorporated social information by drifting towards the decision threshold favoured by the majority. The larger the majority size, the more strongly individuals drifted towards that majority choice. For medium- and large-sized groups, the relationship followed a concave power function, where each additional individual voting for the majority choice had less additional impact on the drift rate. Such saturating influence is consistent with the findings of earlier studies (Asch and Guetzkow, 1951; Bond, 2005; Latane, 1981; Milgram et al., 1969). In groups of three, the relationship followed an exponential function. Weighting single choices less with increasing group size is probably an adaptive strategy: In larger groups, waiting for further decisions avoids confirming fast, but wrong, choices, as others can still correct initial mistakes. In small groups, fast but wrong choices will also occur, but since there are few others to correct those choices, there is little point in delaying a response via a reduced social drift rate.

The social DDM can also characterize other features of the dynamics of the social decision-making process. Beyond capturing how social information impacts the accumulation of evidence, it also captures an individual’s willingness to wait for social information via the threshold parameter $\theta$. Thus, the model is able to distinguish, for instance, between individuals who are sensitive to majorities but unwilling to wait for social information and individuals who may be interested in observing the decisions of others but put more weight on their own personal information. The capacity to unify these different facets of social decision-making within a single theoretical framework is a long-standing goal of social decision-making in the areas of collective animal behaviour (Deneubourg et al., 1990; Sumpter and Pratt, 2008) and social psychology (Latane, 1981). Future studies could investigate the interrelationships between the different parameters, and potential links to established personality measures.

Previous studies have provided evidence for both positive information cascades, such as knowledgeable individuals leading others to resources or safety (Dyer et al., 2008; Kurvers et al., 2015; Stroeymeyt et al., 2011; Watts et al., 2016), and for negative ones, such as the spread of fake news, mobbing, or stampedes (Giraldeau et al., 2002; Bikchandani et al., 1998; Raafat et al., 2009). Here, we have shown the importance of two key aspects promoting positive information cascades. First, a positive confidence/accuracy relationship across group members. In many contexts, confidence is a valid cue for accuracy (Freund and Kasten, 2012; Hertwig, 2012; Bahrami et al., 2012). The strongest association of confidence and accuracy across group members arises when all individuals are more confident when they are more accurate and when their confidence scales are well aligned (i.e., a given level of confidence implies the same level of accuracy across individuals; see also Marshall et al., 2017; Bang and Frith, 2017). The second key aspect promoting positive information cascades is a negative relationship between confidence and RT, meaning that more confident individuals respond faster.
Several mechanisms in the social DDM can influence this relationship—for example, how individuals adjust their start point depending on their confidence. If confident individuals do not start closer to a decision threshold, they are not expected to respond earlier. Also, interindividual differences in model parameters such as choice thresholds or personal and social drift rate can negatively impact the confidence–RT relationship.

The quality of information cascades is shaped by the relationship between accuracy and response time, whereby it is crucial for positive information cascades that accurate individuals respond faster than inaccurate individuals. The social DDM framework allows us to predict the quality of information cascades on the basis of individual or task characteristics. For example, if individuals differ in their ability to solve a task (e.g., individual differences in drift rates), those with higher ability are expected to make faster, more accurate decisions than the less competent ones, triggering positive information cascades. In contrast, when individuals differ systematically in their speed–accuracy tradeoff (e.g., differences in threshold separation; Chittka et al., 2009; Ratcliff et al., 2016), and groups harbour both fast, but inaccurate individuals and slow, but accurate individuals, we expect relatively many fast errors, triggering negative information cascades.

Because the DDM has been successful in accounting for behavioural phenomena across a wide range of tasks, our extension to social environments opens up new possibilities for studying a range of social and collective phenomena. It makes it possible to measure how individuals combine personal and social information and time their decisions whenever decisions are made sequentially and the choices are—at least partially—observable by others. We hope future work will apply and extend the social DDM to areas such as dynamics in consumer preferences (Chen, 2008), emergency evacuations (Moussaïd et al., 2016), and social media (Vosoughi et al., 2018), or to areas of animal social and collective behaviour such as predator detection and mate choice (Danchin et al., 2004).

Methods

Experimental procedure

Participants were 141 students from Wageningen University (the Netherlands) and the University of Bielefeld (Germany). Participants were divided into 16 groups, with group size ranging from small (3 individuals; 5 groups) to medium (7–10 individuals; 6 groups), to large (15–17 individuals, 5 groups; see also Supplementary Table S5). Prior to participation, each participant signed an informed consent form. Each group of individuals was seated on chairs facing a large screen. They were confronted with the following binary decision task: individuals briefly (for 2 seconds) observed an image of a shoal of 72 stylized fish (tuna and sharks aligned in an 8 x 9 grid; see Fig. 1b). Participants were instructed to choose “Escape” if there were five or more sharks and “Stay” if there were four or fewer. The number of sharks present was three, four, six, or seven, and each number was repeated ten times, resulting in 40 trials. Treatment order was randomized. After observing a shoal of fish, individuals had five seconds to report their personal decision and an additional five seconds to report their confidence in their personal decision. Participants were instructed to use confidence as the subjective probability of being correct on a scale from 50% to 100%. In the subsequent social phase, participants made a second decision on the same image. During this phase, they received social information in the form of the number of group members who had already decided on a particular option, displayed on the screen. The social information was first updated after three seconds and then iteratively every two seconds (i.e., at sec 3, 5, 7, 9, . . . 19). The social phase lasted 20 seconds. A countdown timer on the screen indicated the remaining choice time. Participants made all decisions using a wireless keypad. Afterwards, we provided feedback on the correct choice. Participants received 0 points for an incorrect decision and 100 points for a correct decision. To
avoid a scenario in which all participants waited until the last second for social information, we introduced a small cost of one point per second for correct decisions during the social phase. The members of each group with the highest payoff got a small reward in kind. Prior to the 40 study trials, participants completed two test trials to familiarize themselves with the procedure. These results were excluded from the analyses.

**Statistical analysis**

We used Bayesian hierarchical generalized linear models with the “brms” package (Bürkner et al., 2017) to analyse the empirical data in R (R Core Team, 2019). The parameter estimates were generated by running five Markov Chain Monte Carlo (MCMC) simulations in parallel with 5,000 iterations, of which the first 2,500 were discarded as burn-in to reduce autocorrelations. To analyse the difference in the accuracy of personal and social choices (Fig. 2a), we fitted choice correct (yes/no) as a binary response variable and type of choice (personal/social) as a population-level effect (i.e., fixed effect). In this model (and all following models, unless stated otherwise), we included individual and group identity as group-level effects (i.e., mixed effects). We ran separate models to investigate how confidence related to (i) personal accuracy (Fig. 2b), (ii) likelihood to improve (Fig. 3a), and (iii) RT during social choice (Fig. 3b). ‘Personal accuracy’ (correct/incorrect) and ‘likelihood to improve’ (yes/no) were fitted as binomial response variables and ‘RT during social choice’ as an exponentially modified Gaussian (ex-Gaussian) distributed response variable. Confidence was included as a population-level effect in all three models. To investigate whether the majority size affected the likelihood of an individual changing its decision (Fig. 3c), we fitted the likelihood to change the decision as a binary response variable (yes/no) and majority size favouring the opposing option as a population-level effect. To analyse the relationship between RT in the social phase and accuracy of personal and social choices (Fig. 3d), we used decision correct (yes/no) as a binary response variable and type of choice (personal/social) in interaction with RT as a population-level effect.

To investigate how the interrelationships between confidence, accuracy, and RT affected improvement (Fig. 3e), we first calculated—for each group and trial—the Spearman’s correlation coefficients of confidence and accuracy as well as of confidence and RT. We converted these coefficients into dichotomous variables, with the correlation coefficient being either 0 and above or below 0. We excluded trials in which all individuals reported identical choices or confidences, because it was impossible to calculate correlation coefficients for these. We treated all four possible combinations of correlations as different levels of a single factor. We included the factor as a population-level effect and improvement as response variable. In this model, group identity was the only group-level effect. As statistical summary, we report the mean of the posterior distributions and the 95% credible intervals (CI). See Supplementary Table S1 for the results of the regression models. To visualize the results (Fig. 2 and Fig. 3), while accounting for the hierarchical structure of the data, we re-ran the regression models, treating the continuous variables as categorical data. Unless stated otherwise, the points and error bars reflect the mean and the 95% CI of the posterior distribution. Visual inspection of the Markov chains and the Gelman Rubin statistic ($\hat{R}$) indicated that all Markov chains converged.

**Social DDM: Model parameter estimation**

To understand the dynamics of the social phase, we developed the social DDM (Fig. 1a, Table 1). The model features decisions with variable drift rates, in order to obtain choice and RT predictions. We calculated the probability density function of RTs and associated choice probabilities of the
drift-diffusion process by implementing an extended version of a Markov chain approach (Diederich, 1997; Diederich and Busemeyer, 2003) in R (R Core Team, 2019). A detailed description of how to implement the Markov chain approach can be found in Diederich and Busemeyer (2003).

The model assumes that the state space of the decision-maker’s evidence \( L \) is ranging from the lower decision threshold \(-\theta\) (reflecting the wrong decision) to the upper threshold \( \theta \) (reflecting the correct decision) with a step size of \( \Delta \) and \( k \) being the number of steps to reach the decision threshold from a neutral start point:

\[
L = [-k\Delta, -(k-1)\Delta, \ldots, -\Delta, 0, \Delta, \ldots, (k-1)\Delta, k\Delta];
\] (4)

where \( \theta = k\Delta \).

Each time step \( h \) the evidence states change with probabilities given by a \( m \times m \) transition probability matrix \( P \), with \( m = \frac{2\theta}{\Delta} + 1 \). The elements \( p_{1,1} = 1 \) and \( p_{m,m} = 1 \) are the two absorbing states and reflect the decision thresholds. The other elements of \( P \) with \( 1 < i < m \) are:

\[
p_{i,j} = \begin{cases} 
\frac{1}{2\alpha} \left(1 - \frac{u}{\sigma^2} \sqrt{h}\right) & \text{if } j = i-1 \\
\frac{1}{2\alpha} \left(1 + \frac{u}{\sigma^2} \sqrt{h}\right) & \text{if } j = i+1 \\
1 - \frac{1}{\alpha} & \text{if } j = i \\
0 & \text{otherwise}
\end{cases}
\] (5)

with \( \sigma^2 \) being the diffusion coefficient and \( u = \delta_p + \delta_s \) the total drift rate, whereby \( \delta_p \) and \( \delta_s \) are drift rates reflecting the accumulation of personal and social information, respectively (see Table 1). The parameter \( \alpha > 1 \) improves the approximation of the continuous time process. We set \( \alpha = 1.3, \sigma^2 = 1 \) and \( h = 0.005 \). The transition probability matrix in its canonical form:

\[
P = \begin{bmatrix} 
P_I & 0 \\
R & Q
\end{bmatrix} = \begin{bmatrix}
1 & m & 0 & 0 & \ldots & 0 & 0 \\
0 & 1 & 0 & 0 & \ldots & 0 & 0 \\
p_{1,2} & 0 & 0 & 0 & \ldots & 0 & 0 \\
p_{2,3} & 0 & 0 & 0 & \ldots & 0 & 0 \\
p_{3,4} & 0 & 0 & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
p_{m-3,m-2} & 0 & 0 & 0 & \ldots & 0 & 0 \\
p_{m-2,m-1} & 0 & 0 & 0 & \ldots & 0 & 0 \\
p_{m-1,m} & 0 & 0 & 0 & \ldots & 0 & 0
\end{bmatrix}
\] (6)

With \( P_I \) being a \( 2 \times 2 \) matrix with the two absorbing states and \( R \) a \( (m-2) \times 2 \) matrix containing the transition probabilities that eventually lead to the absorbing states in a single transition. \( Q \) is a \( (m-2) \times (m-2) \) matrix including the remaining transition probabilities. The initial process is represented by \( Z \) an \( m-2 \) vector containing the initial probability distribution. The initial start point \( \beta \) is a function of confidence and choice and relative to the upper (correct) threshold:

\[
\beta = \frac{1}{1 + e^{-a(c-b)}};
\] (7)
with $a$ and $b$ being free parameters. $C$ is the reported confidence in the correct choice and is scaled from zero (i.e., highly confident and wrong) to one (i.e., highly confident and correct). We set the distribution of the initial evidence states by $Z_{2\beta} = 1$, with $\beta^* = \beta(m - 3) + 1$. Because $\beta^*$ is not always an integer, we avoid rounding errors by giving most probability mass $1 - (\beta^* - \text{round}(\beta^*))$ to $Z_{\text{round}(\beta^*)}$ and the rest $\beta^* - \text{round}(\beta^*)$ the closest integer of $\beta^*$. For example, if the process starts unbiased (i.e., $\beta = 0.5$) and $m = 7$ then $\beta^* = 3$ and $Z = [0, 0, 1, 0]$. However, if $\beta = 0.55$, then $\beta^* = 3.2$ and $\beta^* - \text{round}(\beta^*) = 0.2$ and therefore $Z = [0, 0, 0.8, 0.2, 0]$. We account for variable drift rates by updating the transition probabilities of $Q$ at $t = (3, 5, 7, 9, \ldots 19)$ seconds, reflecting the iterative updated social information (see Experimental procedure). With $Q_n$ containing the transition probabilities at time point $t = nh$, we can calculate the probability of choosing the correct or wrong option after $n$ time steps:

$$[Pr(\text{correct}|n), P(\text{wrong}|n)] = Z \times Q_1 \times Q_2 \times Q_3 \ldots Q_n \times R - \tau \times t_{\text{min}},$$

with $\tau$ being the non-decision time relative to the fastest response of the individual $t_{\text{min}}$. By varying the transition probabilities of $Q_n$ with changing $\delta$, we are able to account for varying social information over time.

Integrating the social DDM into a Bayesian estimation technique, namely a Differential-Evolution-MCMC algorithm, enables us to sample posterior probability densities of the model parameters (see Table 1). The Differential-Evolution-MCMC is an extension of the Metropolis-Hastings algorithm where proposals are generated by taking the Markov states of parallel computed chains into account (Ter Braak 2006; Turner et al. 2013). To estimate the effect of group size while controlling for individual differences, we used a hierarchical framework. Each parameter was fitted on an individual level but was simultaneously informed by a higher order group-level prior, a normal distribution described by two hyper parameters (i.e., mean and variance), which were informed by the individual fittings. To estimate the posterior probability densities we ran 24 chains in parallel, each with a chain length of 20,000 including a burn-in period of 10,000 and a thinning factor of 10 to reduce autocorrelations. The tuning parameter ($\gamma$) was set to $= 2.38/\sqrt{d}$, with $d$ being the dimensionality of the posterior, which was $d = 2$ for the hyper parameters and $d = 7$ for the individual parameters (see Ter Braak 2006; Turner et al. 2013). To further improve the mixing of the parallel chains, we included deterministic and probabilistic (i.e., relying on the Metropolis–Hastings probability) migration steps where chain-states are swapped across parallel chains (Turner et al. 2013). We performed the deterministic migration step with a probability of 5% where we first determine a random number of $n = 2, 3, \ldots, 24$ chains and then sample $n$ chains without replacement. We then swap the parameter set in a cyclic fashion where the set of the first sampled chain moves to the second, the second to the third and so on, until the last set moves to the first set. A deterministic migration step strongly improves the mixing behaviour of chains but does not resolve the frequent problem of Differential-Evolution-MCMC algorithms that outlier chains hardly converge. We, therefore, additionally implemented a probabilistic migration step which was carried out with a probability of 10%. For the probabilistic version we swapped proposal states instead of accepted states between chains which therefore still relied on the Metropolis–Hastings probability to be accepted. Thereby, we sampled two parallel chains and interchanged a single random parameter state.

We used the social DDM to compare competing hypotheses on how individuals integrate personal and social information. More specifically, we examined three model features: (A) Individuals base their start point on their personal choice and reported confidence. (B) Individuals drift towards the correct option, their initially chosen option, or neither of the two. (C) Individuals drift towards the option favoured by the majority. We compared the performance of models composed of the various combinations of these three features using the deviance information criterion (DIC: Spiegelhalter et al. 2002). To investigate the effect of group size on the collective dynamics, we categorized groups...
as small (3 individuals), medium (7–10), or large (15–17), and fitted the parameters separately for each group size. As a statistical summary, we report the mean of the posterior distributions and the 95% CI. We excluded all observations for which personal choice, social choice, confidence, or RT of social choice were missing (~8%) and if the RT was below 0.1 second (~6%).

**Social DDM: predictions**

To analyse the predictions (i.e., choices and RTs) of the social DDM, we generated decisions by sampling from the probability density functions produced by the model using the mean of the individual-level posterior distribution as model estimates. The probability density function was computed for each individual and trial by taking into account the individual model estimates, the personal choice, the reported confidence, and the social information observed by the individual at a given trial. We then sampled 10 choices and RTs to account for stochasticity. The model predictions are shown as dashed lines in Figure 3.

**Acknowledgement**

We thank Max Wolf for helpful feedback, Susannah Goss for editing the manuscript, Oliver Krüger for hosting the experiments, and all tutors of the 2017 ‘Basismodul Biologie’ at the University of Bielefeld, Germany. R.H.J.M.K. acknowledges funding from the German Research Foundation (DFG, grant number: KU 3369/1-1). The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

**Author contributions**


**Competing interests**

The authors declare no competing interests.

**References**


Supplementary Information

Supplementary Results

Using the Two-stage Dynamic Signal Detection (2DSD) model to model the personal phase: To model the personal choice phase, we used the two-stage dynamic signal detection (2DSD) model. The 2DSD model is an evidence accumulation model that can account for choice and response time (RT) in the personal choice and the associated confidence judgement. In so doing, it can identify cognitive mechanisms potentially governing the interrelationships of these behavioral measures [Pleskac and Busemeyer 2010]. Like other evidence accumulation models, it assumes that individuals gather evidence over time until the amount of evidence surpasses a threshold. The two key assumptions of the 2DSD model are that evidence accumulation continues after the decision is made and that reported confidence depends on the evidence accumulated at the time of the confidence judgement. Thereby, the evidence state is mapped into confidence judgements using response criteria that serve as thresholds indicating the next higher confidence judgements (e.g., from 50 to 60, or 60 to 70). See Pleskac & Busemeyer (2010) for a detailed description of the 2DSD model.

We fitted the model in the hierarchical Bayesian framework, implemented with RStan in R [R Core Team 2019; Stan Development Team 2018], with five parallel chains with 10,000 iterations each and a thinning factor of 10. The first half of the iterations were discarded as burn-in. Descriptions of the main parameters are given in Supplementary Table S6. For the Wiener diffusion process, we included boundary separation $\alpha$, predecisional drift rate $\delta_{\text{pre}}$, relative start point $z$, and nondecision time $NDT$, which was calculated relative to the fastest response. Some trials were expected to be more difficult than others, because the number of sharks could be closer to (i.e., 4 and 6) or further away from (i.e., 3 and 7) the threshold number of sharks (5). We accounted for variations in difficulty by varying the predecisional drift rate $\delta_{\text{pre}}$, depending on trial difficulty:

$$\delta_{\text{pre}} = \begin{cases} \delta_{\text{difficult}}, & \text{if 4 or 6 sharks present} \\ \delta_{\text{difficult}} + \Delta_{\text{easy}}, & \text{if 3 or 7 sharks present} \end{cases}$$

with $\Delta_{\text{easy}}$ describing the additional effect of easy trials on the drift rate. For the postchoice process, we fitted confidence criteria and the postdecisional drift rate ($\delta_{\text{post}}$). The postdecisional drift rate is influenced by the predecisional drift rate, with the parameter $w$ controlling its strength, and $\delta_{\text{choice}}$ describing the influence of the choice on the subsequent drift:

$$\delta_{\text{post}} = \begin{cases} w \times \delta_{\text{pre}} + \delta_{\text{choice}}, & \text{if correct} \\ w \times \delta_{\text{pre}} - \delta_{\text{choice}}, & \text{if incorrect} \end{cases}$$

The evidence distribution at the time point when confidence is reported $L_{\text{conf}}$ is a combination of the evidence accumulated at the time point of choice and the evidence accumulated between choice and confidence judgement. It is normally distributed with a mean of

$$E[L_{\text{conf}}] = \begin{cases} \alpha + \delta_{\text{post}} \times IJT, & \text{if correct} \\ 0 + \delta_{\text{post}} \times IJT, & \text{if incorrect} \end{cases}$$

and a variance of

$$\text{var}[L_{\text{conf}}] = \sigma^2 IJT$$

with $IJT$ being the interjudgement time (i.e., the time between choice and confidence reporting).

Each decision maker has confidence criteria $c_j$ to map the evidence state $L_c$ into six possible confidence judgements $conf_j$ with $j = 0, 1, 2, ... 5$, corresponding to the confidence levels 50 to 100. The
probability of reporting $conf_j$ is given by the normal cumulative distribution $\sim N \left( E[L_c], \text{var}[L_c] \right)$ with:

$$P(c_j < L_c < c_{j+1})$$

where $c_0$ is equal to $-\infty$ and $c_6$ to $\infty$. The five remaining confidence criteria are fitted by the model. We assume the locations of the confidence criteria for correct and incorrect choices to be symmetrical. Hence, we set the locations relative to the choice thresholds with $alpha + c_j$ and $0 - c_j$ for correct and incorrect choices, respectively. For the fitting process, we excluded all choices with RTs below 0.1 sec. To compare the predictions of the model with the empirical data, we generated choices, RTs, and confidence judgements using the participant’s mean posterior parameter estimate. The confidence judgements were generated by sampling from the evidence distribution at the time point of the judgement and mapping this evidence state to a confidence judgement. We thus obtained confidence judgements given the individual’s choice, RT, and interjudgement time. To account for stochasticity generated by the sampling process, we sampled 100 confidence judgements, choices, and RTs per individual and trial.

2DSD model results: Participants drifted towards the correct choice threshold ($\delta_{\text{difficult}} = 0.37$, CI = [0.33, 0.41]). Trials with three or seven sharks were easier than trials with four or six sharks, as indicated by a stronger drift towards the correct option in the former ($\Delta_{\text{easy}} = 0.05$, CI = [+0.00, 0.10]). Varying drift rates depending on difficulty were not included in the social DDM analysis, as the effect was comparatively small. After making a choice, participants continued accumulating evidence and, on average, kept gathering correct evidence ($w = 0.72$, CI = [0.62, 0.83]). Hence, participants who made an incorrect decision gathered more evidence over time contradicting their initial choice (resulting in lower confidence), whereas the evidence of those who made a correct choice was strengthened (resulting in higher confidences). This process explains the increasing difference in confidence ratings between correct and incorrect choices as interjudgement time increases (Supplementary Fig. S4a). Additionally, there was a choice effect on the postdecision drift, whereby participants accumulated evidence in favour of their already chosen option ($\delta_{\text{choice}} = 1.64$, CI = [1.47, 1.80]). As a result, longer interjudgement times are predicted to lead to higher confidence judgements (Supplementary Fig. S4b). Figure S4c shows that the 2DSD recreates the well-established relationship between confidence and accuracy, which is partly determined by the postdecisional processing evident in Figures S4a and b. In both the 2DSD and the social DDM analysis, we thus found that confidence is linked to the evidence state and that participants drifted in the direction of their chosen option (i.e., reinforced their ‘belief’ in their original choice). Figure S4c shows RT distributions for the personal choice. Overall, the empirical data (solid lines) correspond closely with the predictions of the 2DSD model (dashed lines), indicating that the personal phase can be described by a drift diffusion process. One distribution characteristic the model cannot recover is the higher average RTs for incorrect choices. This is a well-known property of the drift–diffusion model, and can be addressed by adding trial-by-trial variability to the drift rates (Ratcliff and Rouder, 1998). For simplicity, we have not included trial-by-trial variability.
Figure S1: **Improvement during the social phase depended on the quality of social information.** Participants’ choices were increasingly likely to improve/worsen as the size of the majority for the correct/incorrect option increased. Dots represent the mean; error bars represent twice the standard error. The dashed line shows the prediction of the social DDM.
Figure S2: **Distributions of key behavioural measures.** (a) The proportion of reported confidence scores for correct and incorrect choices. The higher the confidence score, the larger the proportion of correct choices, resulting in a positive confidence–accuracy relationship (see also Supplementary Fig. S4c). (b) The proportion of choices made in the presence of different majority sizes. In the social phase, most choices (≈60%) were made in the absence of a majority, and participants who experienced a majority were more likely to observe a confirming majority (i.e., negative values) than an opposing majority. Participants facing an opposing majority were more likely to change their choice the larger the size of this opposing majority. (c) Observed RT distributions during the social phase as a function of reported confidence. Participants reporting the highest level of confidence overwhelmingly responded within 4 seconds, whereas the distribution of participants reporting the lowest confidence level peaked after 4 seconds. (d) Observed RT distributions during the social phase for correct and incorrect choices. Given that unconfident participants are more likely to be wrong and waited longer, it follows that individuals who were wrong, on average, wait longer to respond. (e, f) RT distributions as predicted by the social DDM. The model recovers not only the relationship of RT with confidence and accuracy, but also the shape of the RT distributions. The RT distributions are multimodal because social information was first updated after 3 seconds and then every 2 seconds. The updating events often resulted in larger majorities which increased the likelihood of a response by an increase in the drift rate. (c–f) Dashed vertical lines represent the mean RTs.
Figure S3: **Model recovery.** The x-axis shows the actual (input) parameters; the y-axis shows the recovered parameters. The figure shows the results of a parameter recovery analysis conducted to ensure that the parameters of the social DDM are interpretable and capture distinct cognitive mechanisms. We repeatedly generated data with random input parameters and recovered them with the same hierarchical social DDM used to analyse the empirical data. The input parameters were sampled with a quasirandom number generator (using the sobol sequence), ensuring an even distribution across a large multidimensional parameter space. Using these input parameters, we generated social choices by computing probability density functions while taking into account the personal choice, reported confidence, and the social information observed by the participant at a given trial. The generated data thus have the same hierarchical structure as the empirical data, with 141 participants and varying group size. Again, we report the mean of the posterior distributions and the 95% CI of the higher order group-level estimate for each group size. To measure the relationship of input and recovered parameters, we calculated Spearman’s correlation coefficient $r$ for all parameters (except nondecision time, which is relative to a participant’s fastest response and thus meaningless on a group level). For all parameters, there was a strong positive correlation between the generated and the recovered parameters. The estimates provided by the social DDM thus describe separate identifiable features and are interpretable in their magnitude.
Figure S4: Results of the 2DSD model. (a) The longer the time between the personal choice and the confidence judgement (interjudgement time), the larger the difference in confidence between participants whose choices were correct vs. incorrect. Dots represent the average confidence judgements for correct choices minus the average confidence judgements for incorrect choices for different interjudgement times. (b) The longer the interjudgement time, the higher the reported confidence judgements. Dots represent the mean; error bars represent twice the standard error. (a–b) For visualization purposes, interjudgement times are binned by rounding to the closest integer. (c) Participants reporting higher confidence were more likely to be correct. Dots and error bars represent mean and 95% CI of the posterior distribution. (d) The solid lines represent the observed RT distribution of the personal choice for correct (blue) and incorrect (red) choices. (a–d) The dashed lines represent the predictions of the 2DSD model.
## Supplementary Tables

Table S1: Bayesian linear regression results

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Est.Error</th>
<th>l−95% CI</th>
<th>u−95% CI</th>
<th>Eff.Sample</th>
<th>Rhat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>Intercept (personal choice)</td>
<td>1.1</td>
<td>0.07</td>
<td>0.97</td>
<td>1.23</td>
<td>9657.34</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Social choice</td>
<td>0.3</td>
<td>0.05</td>
<td>0.2</td>
<td>0.39</td>
<td>32162.95</td>
<td>1</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>Intercept</td>
<td>-1.58</td>
<td>0.17</td>
<td>-1.91</td>
<td>-1.25</td>
<td>20461.1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Confidence</td>
<td>3.82</td>
<td>0.24</td>
<td>3.35</td>
<td>4.28</td>
<td>20270.94</td>
<td>1</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>Intercept (personal choice)</td>
<td>1.65</td>
<td>0.08</td>
<td>1.48</td>
<td>1.81</td>
<td>7563.13</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>RT</td>
<td>-0.16</td>
<td>0.01</td>
<td>-0.18</td>
<td>-0.14</td>
<td>15170.53</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>RT: social choice</td>
<td>0.11</td>
<td>0.01</td>
<td>0.09</td>
<td>0.13</td>
<td>18797</td>
<td>1</td>
</tr>
<tr>
<td><strong>Likelihood to change</strong></td>
<td>Intercept</td>
<td>-3.6</td>
<td>0.18</td>
<td>-3.96</td>
<td>-3.26</td>
<td>7735.1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Size of opposing majority</td>
<td>0.62</td>
<td>0.03</td>
<td>0.57</td>
<td>0.67</td>
<td>12889.8</td>
<td>1</td>
</tr>
<tr>
<td><strong>RT</strong></td>
<td>Intercept</td>
<td>6.96</td>
<td>0.23</td>
<td>6.49</td>
<td>7.41</td>
<td>4644.56</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Confidence</td>
<td>-4.86</td>
<td>0.18</td>
<td>-5.22</td>
<td>-4.5</td>
<td>9740.24</td>
<td>1</td>
</tr>
<tr>
<td><strong>Improvement</strong></td>
<td>Intercept</td>
<td>1.17</td>
<td>0.2</td>
<td>0.78</td>
<td>1.56</td>
<td>22984.38</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Confidence</td>
<td>-4.27</td>
<td>0.31</td>
<td>-4.88</td>
<td>-3.68</td>
<td>21332.94</td>
<td>1</td>
</tr>
<tr>
<td><strong>Improvement</strong></td>
<td>Intercept (earlier; more accurate)</td>
<td>0.09</td>
<td>0.01</td>
<td>0.08</td>
<td>0.11</td>
<td>3830.29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Earlier; less accurate</td>
<td>-0.09</td>
<td>0</td>
<td>-0.09</td>
<td>-0.08</td>
<td>19689.95</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Later; more accurate</td>
<td>-0.04</td>
<td>0.01</td>
<td>-0.05</td>
<td>-0.03</td>
<td>16925.55</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Later; less accurate</td>
<td>-0.04</td>
<td>0.01</td>
<td>-0.06</td>
<td>-0.02</td>
<td>17547.86</td>
<td>1</td>
</tr>
</tbody>
</table>
Table S2: Deviance information criterions (DIC) for different versions of the social DDM. The version with the lowest DIC is indicated in bold.

<table>
<thead>
<tr>
<th>No further drift</th>
<th>Drift towards initial choice</th>
<th>Drift towards correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither</td>
<td>79493</td>
<td>76026</td>
</tr>
<tr>
<td>Varying start point</td>
<td>76364</td>
<td>74183</td>
</tr>
<tr>
<td>Social drift</td>
<td>78058</td>
<td>74275</td>
</tr>
<tr>
<td>Both</td>
<td>74200</td>
<td><strong>71865</strong></td>
</tr>
</tbody>
</table>

Table S3: Mean parameter estimates and 95% credible intervals of the social DDM for different group sizes.

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDT ($T_s$)</td>
<td>0.4 [0.23, 0.56]</td>
<td>0.33 [0.25, 0.41]</td>
<td>0.31 [0.26, 0.37]</td>
</tr>
<tr>
<td>Start point ($a$)</td>
<td>4.2 [3.11, 5.35]</td>
<td>3.43 [2.81, 4.07]</td>
<td>3.9 [3.46, 4.37]</td>
</tr>
<tr>
<td>Start point ($b$)</td>
<td>0.5 [0.45, 0.54]</td>
<td>0.48 [0.45, 0.51]</td>
<td>0.5 [0.48, 0.52]</td>
</tr>
<tr>
<td>Personal drift</td>
<td>0.65 [0.45, 0.86]</td>
<td>0.62 [0.5, 0.75]</td>
<td>0.53 [0.47, 0.59]</td>
</tr>
<tr>
<td>Social drift ($s$)</td>
<td>0.51 [0.23, 0.82]</td>
<td>0.31 [0.24, 0.38]</td>
<td>0.36 [0.3, 0.41]</td>
</tr>
<tr>
<td>Social drift ($q$)</td>
<td>1.75 [1.16, 2.36]</td>
<td>0.93 [0.81, 1.05]</td>
<td>0.66 [0.6, 0.72]</td>
</tr>
</tbody>
</table>

Table S4: Differences between parameter estimates for different group sizes. Shown are the mean and the 95% credible intervals.

<table>
<thead>
<tr>
<th></th>
<th>Small – Medium</th>
<th>Small – Large</th>
<th>Medium – Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDT ($T_s$)</td>
<td>0.06 [-0.12, 0.25]</td>
<td>0.08 [-0.09, 0.26]</td>
<td>0.02 [-0.08, 0.11]</td>
</tr>
<tr>
<td>Start point ($a$)</td>
<td>0.77 [-0.48, 2.08]</td>
<td>0.3 [-0.88, 1.51]</td>
<td>-0.47 [-1.25, 0.31]</td>
</tr>
<tr>
<td>Start point ($b$)</td>
<td>0.01 [-0.04, 0.07]</td>
<td>-0.01 [-0.06, 0.04]</td>
<td>-0.02 [-0.06, 0.02]</td>
</tr>
<tr>
<td>Personal drift</td>
<td>0.03 [-0.2, 0.27]</td>
<td>0.12 [-0.08, 0.33]</td>
<td>0.09 [-0.04, 0.23]</td>
</tr>
<tr>
<td>Social drift ($s$)</td>
<td>0.2 [-0.09, 0.51]</td>
<td>0.15 [-0.13, 0.46]</td>
<td>-0.05 [-0.13, 0.04]</td>
</tr>
<tr>
<td>Social drift ($q$)</td>
<td>0.82 [0.22, 1.44]</td>
<td>1.1 [0.5, 1.71]</td>
<td>0.27 [0.14, 0.41]</td>
</tr>
<tr>
<td>Choice threshold</td>
<td>-0.21 [-0.93, 0.54]</td>
<td>-0.08 [-0.77, 0.65]</td>
<td>0.13 [-0.3, 0.56]</td>
</tr>
</tbody>
</table>

Table S5: The number of groups per group size.

<table>
<thead>
<tr>
<th>Group size</th>
<th>Number of groups</th>
<th>Number of participants</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5</td>
<td>15</td>
<td>Small</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>21</td>
<td>Medium</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>8</td>
<td>Medium</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>9</td>
<td>Medium</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>10</td>
<td>Medium</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>45</td>
<td>Large</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>16</td>
<td>Large</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>17</td>
<td>Large</td>
</tr>
<tr>
<td>Total:</td>
<td>16</td>
<td>141</td>
<td></td>
</tr>
</tbody>
</table>

29
Table S6: Description of the parameters of the 2DSD.

<table>
<thead>
<tr>
<th>Model feature</th>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nondecision time</td>
<td>$NDT$</td>
<td>A parameter between 0 and 1 accounting for nondecision time (e.g., motor response time), parameterized as the time relative to an individual’s fastest response.</td>
</tr>
<tr>
<td>Relative start point</td>
<td>$z$</td>
<td>Describes the initial evidence state before the evidence sampling process begins.</td>
</tr>
<tr>
<td>Predecisional drift rate</td>
<td>$\delta_{\text{pre}} = \begin{cases} \delta_{\text{difficult}}, &amp; \text{if difficult} \ \delta_{\text{difficult}} + \Delta_{\text{easy}}, &amp; \text{if easy} \end{cases}$</td>
<td>The baseline predecisional drift rate for difficult (i.e., 4 or 6 sharks) and easy (i.e., 3 or 7 sharks) trials.</td>
</tr>
<tr>
<td>Boundary separation</td>
<td>$\alpha$</td>
<td>The boundary separation determines how much evidence an individual has to accumulate to make a decision.</td>
</tr>
<tr>
<td>Carryover effect</td>
<td>$w$</td>
<td>A parameter controlling how strongly the predecisional drift rate carries over to the postdecisional drift rate.</td>
</tr>
<tr>
<td>Self-confirmation bias</td>
<td>$\delta_{\text{choice}}$</td>
<td>A parameter describing the influence of the choice (i.e., being correct or incorrect) on the subsequent drift rate.</td>
</tr>
<tr>
<td>Confidence criteria</td>
<td>$c_j$</td>
<td>Thresholds that divide the evidence space into confidence judgements.</td>
</tr>
</tbody>
</table>

Table S7: 2DSD parameter results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>l−95% CI</th>
<th>u−95% CI</th>
<th>Eff.Sample</th>
<th>Rhat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nondecision time</td>
<td>0.63</td>
<td>0.56</td>
<td>0.74</td>
<td>2362.88</td>
<td>1</td>
</tr>
<tr>
<td>Relative start point</td>
<td>0.53</td>
<td>0.52</td>
<td>0.54</td>
<td>2085.7</td>
<td>1</td>
</tr>
<tr>
<td>Predecisional drift rate (intercept, difficult)</td>
<td>0.37</td>
<td>0.33</td>
<td>0.41</td>
<td>1991.81</td>
<td>1</td>
</tr>
<tr>
<td>Predecisional drift rate (effect of easy)</td>
<td>0.05</td>
<td>0</td>
<td>0.1</td>
<td>2256.05</td>
<td>1</td>
</tr>
<tr>
<td>Boundary separation</td>
<td>2.5</td>
<td>2.45</td>
<td>2.56</td>
<td>2268.78</td>
<td>1</td>
</tr>
<tr>
<td>Carryover effect</td>
<td>0.72</td>
<td>0.62</td>
<td>0.83</td>
<td>2354.22</td>
<td>1</td>
</tr>
<tr>
<td>Self-confirmation bias</td>
<td>1.64</td>
<td>1.47</td>
<td>1.8</td>
<td>1962.74</td>
<td>1</td>
</tr>
<tr>
<td>Confidence criteria 5</td>
<td>3.27</td>
<td>2.46</td>
<td>4.09</td>
<td>1875.05</td>
<td>1</td>
</tr>
<tr>
<td>Confidence criteria 4</td>
<td>4.99</td>
<td>4.54</td>
<td>5.44</td>
<td>2444.76</td>
<td>1</td>
</tr>
<tr>
<td>Confidence criteria 3</td>
<td>3.83</td>
<td>3.5</td>
<td>4.17</td>
<td>2160.56</td>
<td>1</td>
</tr>
<tr>
<td>Confidence criteria 2</td>
<td>3.04</td>
<td>2.74</td>
<td>3.35</td>
<td>2499.86</td>
<td>1</td>
</tr>
<tr>
<td>Confidence criteria 1</td>
<td>2.4</td>
<td>2.11</td>
<td>2.72</td>
<td>2605.37</td>
<td>1</td>
</tr>
</tbody>
</table>