Social information can undermine individual performance in exploration-exploitation tasks

Kyanoush Seyed Yahosseini (yahosseini@mpib-berlin.mpg.de)[1], **Samuli Reijula** (samuli.reijula@helsinki.fi)[2],
Lucas Molleman (molleman@mpib-berlin.mpg.de)[1] and **Mehdi Moussaïd** (moussaid@mpib-berlin.mpg.de)[1]

[1] Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany
[2] TINT / Social and Moral Philosophy, University of Helsinki, Finland

Abstract

In many daily life situations, people face decisions involving a trade-off between exploring new options and exploiting known ones. In these situations, observing the decisions of others can influence people's decisions. Whereas social information often helps making better decisions, research has suggested that under certain conditions it can be detrimental. How precisely social information influences decision strategies and impacts performance is, however, disputed. Here we study how social information influences individuals' exploration-exploitation trade-off and show that this adaptation can undermine their performance. Using a minimal experimental paradigm, we find that participants tend to copy the solution of other individuals too rapidly, thus decreasing the likelihood of discovering a better solution. Approximating this behavior with a simple model suggests, that individuals' willingness to explore only depends on the value of known existing solutions. Our results allow for a better understanding of the interplay between social and individual factors in individual decision-making.
Keywords: Exploration-exploitation trade-off, social learning, decision-making

Introduction

Social information is crucial to help individuals and groups adapt to novel circumstances. Through social interaction and observation, people can collect up-to-date information about their environment and efficiently deal with its uncertainty (Hills et al., 2015). For example, by selectively copying successful others, individuals can readily improve their decisions, while avoiding the costs of trial-and-error learning (Mesoudi, 2011; Rendell et al., 2010; Wisdom, Song, & Goldstone, 2013). Moreover, the transmission of known solutions in populations can lead to the accumulation of knowledge, which can be built upon and refined over time (Boyd, Richerson, & Henrich, 2011; Derex & Boyd, 2015; Tomasello, 1999; Moussaïd & Seyed Yahosseini, 2016).

Yet, theoretical models from a range of disciplines suggest that social information can also be detrimental for the efficiency of decision-making (March, 1991; Mehlhorn et al., 2015). Several mechanisms can underlie this counter-intuitive effect. First, it might be due to the structure of the environment. When exploration has an opportunity cost, the availability of social information can motivate people to free-ride and wait for others to discover a profitable solution, thus reducing the group's exploration range (Bolton & Harris, 1999). This is particularly critical in dynamic environments that change in time and space, and where social information might be outdated or ill-fitted to one's own situation (Henrich & Boyd, 1998; Rogers, 1988). Second, the detrimental effects of social information can be caused by social factors: One example of such an effect are *information cascades*, where mutual reinforcement can lead groups to converge upon sub-optimal solutions, while leaving potentially superior solutions unexplored (Bala & Goyal, 1998; Giraldeau, Valone, & Templeton, 2002; Salganik & Watts, 2008). Finally, individual factors may impact how social information affects decision-making. Social influence is often modulated by individuals' perceived skill, experience and knowledge about the environment, which are not always accurately evaluated (Laland, 2004; Moussaïd et al., 2017; Rendell et al., 2010). Additionally, individuals' aspiration levels may change when observing the rewards of very successful or unsuccessful others (March, 2006).

Recently, empirical studies have delineated how environmental and social factors impact collective performance in decision problems. It has been shown that reducing the flow of social information among individuals – e.g., through sparsely connected social networks – can enhance performance at the group level (Fang, Lee, & Schilling, 2009; Derex & Boyd, 2016; Mason, Jones, & Goldstone, 2008). However, other studies suggest the opposite: they show that networks that facilitate the exchange of information between individuals tend to enhance group performance (Derex & Boyd, 2015; Mason & Watts, 2012). These conflicting conclusions suggest that a clear picture of the processes at play and their interaction between each other are currently not available. As most simulation studies assume simplified decision rules, while experimental approaches often involve a combined manipulation of environmental and social factors, making it difficult to understand the contributions of the different factors (Mehlhorn et al., 2015). How do people respond to social information when searching for a problem solution? How do they adapt their exploration and exploitation decisions when exposed to the behavior of a peer? Addressing these questions helps disentangle the complex interactions between these processes and facilitates understanding the resulting dynamics in its entirety.

In this paper we show that social information directly affect individuals' decision to explore or exploit, that is the used decision strategy, and as a result hamper performance. We examine how social information affects individuals' tendency to *exploit* their own best solution, to *copy* the solution of a peer, and to *explore* their environment in search for superior alternatives. To this end, we designed a experiment in which we tightly control the value of social information. We aim to eliminate several common confounding factors stemming from temporal and spatial heterogeneity in the environment, endogeneity of information propagation in the network, and pre-existing individual variation of knowledge and skills (e.g., Mason and Watts (2012); Jayles et al. (2017); Jönsson, Hahn, and Olsson (2015). This approach allows us to draw a clear and simple picture of the isolated effects of social information on how individuals solve an exploration-exploitation task.

Methods

We designed a sequential decision-making task with unidirectional information flow in which participants played the role of farmers trying to maximize their cumulative payoff over 30 rounds. In every round, each participant could choose between three options: (1) plant a new unknown crop (i.e. explore a new solution), (2) plant the best crop he or she found so far (i.e. exploit the best known solution), and (3) plant the crop with the highest value that one other participant had discovered while taking the same task before (i.e. copy social information). Figure 1A shows the experimental interface of the experiment.

Each crop was associated with a fixed payoff ranging from 1 to 100 points. The crop payoffs were randomly drawn from an exponential distribution with $\lambda = 0.05$, capped at 100. That is, many crops were associated with a low payoff, and a few of them had a high payoff, with a maximal possible value of 100 (Figure 1B).

Each round, participants could either (1) draw a new value from the exponential distribution and receive the associated payoff (explore), (2) receive the payoff X_e associated to the highest value that was drawn so far (exploit), or (3) receive the highest payoff X_c that another participant had discovered (copy). In each round, participants could see the payoffs X_e and X_c while the payoff for explore was hidden.

Experimental treatment. To systematically examine the effect of social information on individual decision strategies and performance, we implemented seven experimental treatments and a control condition in a between-subjects design. In the control condition, no social information was available, and participants could only choose to *explore* or *exploit*. We gathered the highest payoff found by each participant in the control condition, that is X_e after 30 rounds of independent search. To systematically assess the effect of these values on decision strategies, we selected a subset of them to display as social information X_c in the experimental treatments. We selected the following values: 16, 21, 26, 31, 36, 46, and 56. This setup allows the participants in the experimental conditions to observe actual social information generated by other participants, and it ensures a uniform sample size over a wide range of different values of social information. For each participant, the value X_c was constant over the 30 farming rounds.

Procedure and participants. We ran two sessions of the experiment on Amazon Mechanical Turk. In the first session all participants were assigned to the control condition. In the second session participants were randomly assigned to one of the seven experimental treatments, where a specific X_c was

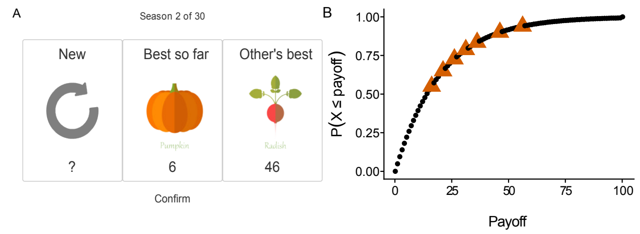

Figure 1: Experimental design. (A) The experimental interface as shown to the participants. In every round (*season*), participants could choose to plant a new unknown crop (left button), to collect the payoff associated to their best discovered crop (middle button) or to collect the payoff associated to the best crop found by another participant (right button). (B) The cumulative distribution function of the exponential random variable used to generate the payoffs every time a participant decided to plant a new crop. The values of social information used in our experimental conditions are marked by red triangles.

selected from the first session.

Participant were informed that the payoffs could range from 1 to 100 and that most crops were associated to a low number of points. They where also briefed about how the social information was acquired. Participants were instructed to maximize their cumulative payoff over 30 rounds and informed that at the end of the experiment, their cumulative payoff would be converted into real money (2,000 points = US$1). Before starting the experiment, participants completed a short interactive tutorial and had to pass a comprehension check. At the end of each round, participants were informed about their payoff from that round and their cumulative payoff.

In total 322 participants (145 Female, mean age = 36.9 years, SD = 10.9 years) completed the experiment with an average of 40.2 participants (SD = 4.7) per experimental condition. Participants were rewarded by US$0.75 plus a monetary bonus based on their final score in the experiment (mean bonus=US$0.62, SD=US$0.24). The average completion time was 8 minutes. The self-reported understanding of the experimental task was very high, as 97% of the participants reported ≥ 6 points on a 7 point Likert scale.

Models

To gain a deeper understanding of how social information shapes individual decisions in our experiment, we first introduce two simple models involving different decision strategies: the *benchmark model* and the *threshold model*. In both models, agents make a decision between (1) *explore*, (2) *exploit*, and (3) *copy* at any given round t. The payoff $\rho(t)$ that the agent receives at round t depends on the chosen option as follows:

1. If the agent decides to explore, $\rho(t) = x$, where x is a randomly drawn value from the exponential distribution shown in figure 1B.
2. If the agent decides to exploit, $\rho(t) = X_e$, where X_e is the highest value that has been drawn by the agent since the beginning of the task.

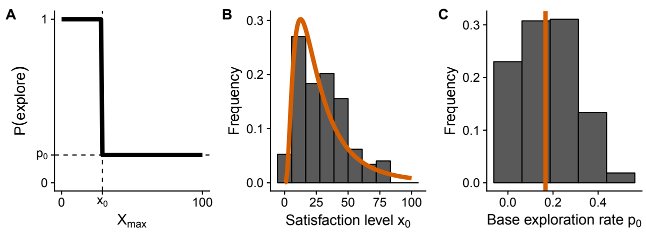

Figure 2: The building blocks of the threshold model. (A) Graphical representation of the step function used for the threshold model. The probability $P(\text{explore})$ equals 1 as long as the payoff of the best known solution X_{max} is lower than the satisfaction level x_0. Above that level, some exploration is maintained with a probability p_0, the base exploration rate. (B) Distribution of the estimated threshold values x_0 for all participants. The the red curve indicates the best log-normal fit (meanlog = 3.13, sdlog = .78, McFadden's $r^2 = .58$). (C) Distribution of the estimated values of p_0 for all the participants. The median value $p_0 = .17$ is indicated by the red line.

3. If the agent decides to copy, $\rho(t) = X_c$, where X_c is the payoff associated with the crop provided by social information.

For both models we measure performance as the average payoff that an agent achieves across the 30 rounds. We ran 610,000 repetitions of each model while systematically varying X_c between 0 and 60.

Benchmark Model. In the benchmark model, we assume that the agents explore their environment during the first τ rounds (i.e. the exploration phase) of the experiment, and then choose the most rewarding option between exploit and copy, during all the remaining rounds (i.e. the capitalization phase) (Rapoport & Tversky, 1970; March, 1991). The unique parameter of this normative model is the time τ at which the agents switch from exploration to capitalization. We call the optimal switching point τ^* the value of τ that yields the best mean performance for a given value of X_c. We estimate τ^* by systematically varying τ between 0 and 30 and pick the value that yielded the best mean performance.

Threshold Model. For the threshold model, we assume a simple decision strategy to facilitate the comparison with the experimental results. The threshold model assumes that the probability to choose the explore option $P(\text{explore})$ in a given round is solely dependent on the maximum payoff $X_{max} = max(X_e, X_c)$ that the agent can get by exploiting or copying. Thus, in each round the agent will either explore with a probability $P(\text{explore})$ or choose the exploit or copy option that will yield the higher payoff with a probability $1 - P(\text{explore})$.

The probability $P(\text{explore})$ is specified by a simple step function:

$$P(\text{explore}) = \begin{cases} 1, & \text{if } X_{max} < x_0 \\ p_0, & \text{otherwise} \end{cases}$$

Where x_0 is a threshold *aspiration level* based on X_{max}, and p_0 is the base exploration rate (see figure 2A).

We fit the step function to the behavioral data of each participant individually. For that we use the step function with $p_0 = 0$ as a binary classifier to predict if a participant explores. We define exploration as a positive and exploitation as a negative outcome. We then calculate x_0 such that the predictive accuracy of the classifier is maximized. Accuracy is defined as the $(\text{number of true positives} + \text{number of true negatives})/30$. For the most accurate x_0 we set p_0 to the number of false negatives$/30$. That is, p_0 is the probability that exploration was wrongly classified as exploitation or copy. This procedure provides an estimation of the parameters x_0 and p_0 for each participant. The distributions of the two parameters are shown in figure 2B-C. For the simulations, we generate agents by picking x_0 from the log-normal distribution fitted to the estimated values of the participants and p_0 by the median of those.

Results

We now look at our experimental results and compare them to the predictions from the two models.

Performance. The presence of social information has a strong and non-monotonic influence on the participants' performance (Figure 3). We observe a decay in performance in the conditions where participants received social information of low value, specifically when X_c ranges between 21 and 31, but not in case of $X_c = 16$. Overall, participants who received no social information performed equally well or better than those who received social information of value X_c lower than 46. For higher values of X_c social information was beneficial and allowed participants to improve their performances as compared to the control condition.

Furthermore, our simulations show that the threshold model predicts a similar decay in performance around $X_c = 26$, whereas the benchmark model – which was not calibrated on the data but for maximizing performance – suggests that participants could have reached a much better score with a different decision strategy, such as exploring early to be able to assess the relative value of X_c.

Decision strategies. To explain the decrease in performance observed in our experimental results for values of X_c surrounding 26, and predicted by the threshold model, we analyzed the underlying decision strategy of the participants. Specifically, we looked at how frequently participants chose to explore, to exploit and to copy in each treatment (figure 3B). We anticipated that participants would explore less in the presence of social information, and indeed participants explored on average 10 rounds (SD = 5.5) in the control condition, while exploring 8.1 (SD = 5.2) rounds when social information was available (W = 6761.5, p = .04). Also we predicted that the frequency of exploration would depend on the value of social information. Surprisingly the data shows that the exploration rate changes only marginally between the experimental conditions. For $X_c = 16$ participants explored on average 7.8 rounds (SD = 4), very close to the 7.5 rounds (SD = 5.4) of exploration

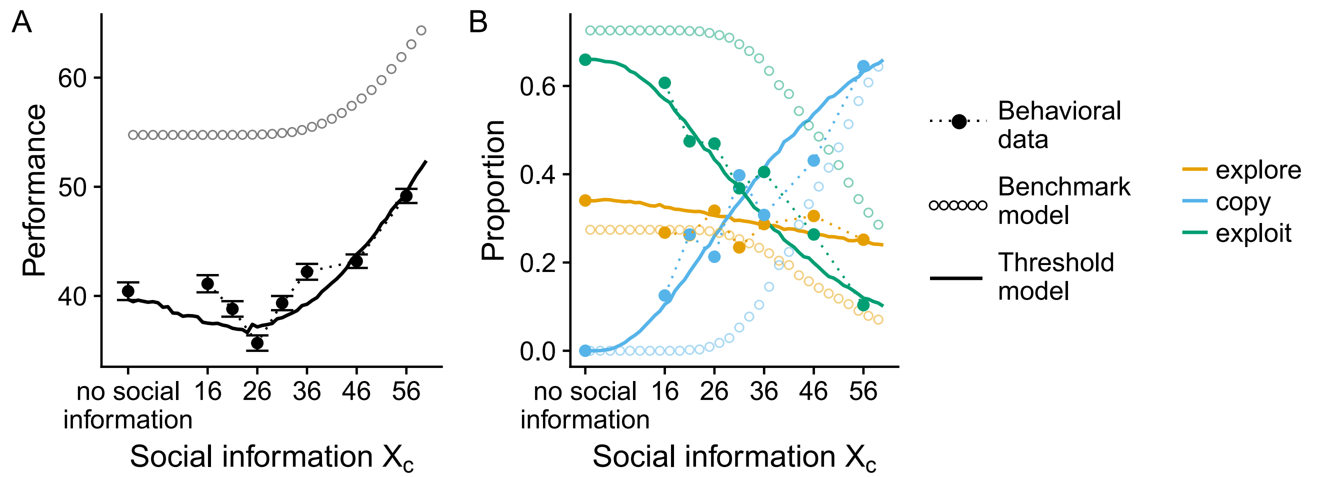

Figure 3: Experimental results and model predictions. (A) Average individual performance as a function of social information X_c as measured in the experiment and predicted by the two models. Average individual performance is calculated as the average payoff per round, $\sum(\rho(t))/30$. Error bars indicate the standard error. (B) Proportion of explore, copy, and exploit decisions as a function of the social information X_c.

for $X_c = 56$. That is, a reduction of exploration cannot explain the decay in performance. However, as expected, the overall tendency to copy increases with larger values of X_c, while the amount of exploitation decreases at the same time.

Because the average performance of individuals in the control condition – where no social information was available – is approximately 40 points, copying a crop of value X_c lower than 40 would necessarily be counter-productive. This is indeed confirmed by the prescriptive simulations of our benchmark model (blank green and blue circles in figure 3B). Yet, our results show that for values $X_c \leq 36$, people unknowingly copy too frequently – thus capitalizing on a sub-optimal option. Our threshold model predicts a similar trend.

Temporal changes. To draw a more accurate picture of the decision strategies, we looked at the temporal changes of behavior across the 30 rounds (figure 4). One striking observation is that the normative assumption that people would explore first and copy or exploit later, as implemented in the benchmark model and suggested by optimal stopping theory (Rapoport & Tversky, 1970), clearly does not capture the participants' decisions. On the contrary, copying is the predominant behavior during the first few rounds independent of the value of X_c. Only later on, if a better X_e is found, copying rates diminish.

Participants tend to alternate between explore and copy at the beginning of the experiment and, unsurprisingly, disregard X_c, if X_c is not sufficiently good compared to what has been sampled in the meantime. This dynamic is captured by the threshold model. Thus the reduced performance around $X_c = 26$, as shown in figure 3, can be explained by the fact that participants tend to copy social information too frequently during the first few rounds without being able to assess the relative value of X_c. Whenever the value of social information is lower than what they would have sampled independently (i.e. $X_c < 40$), individual performance is undermined. Due to

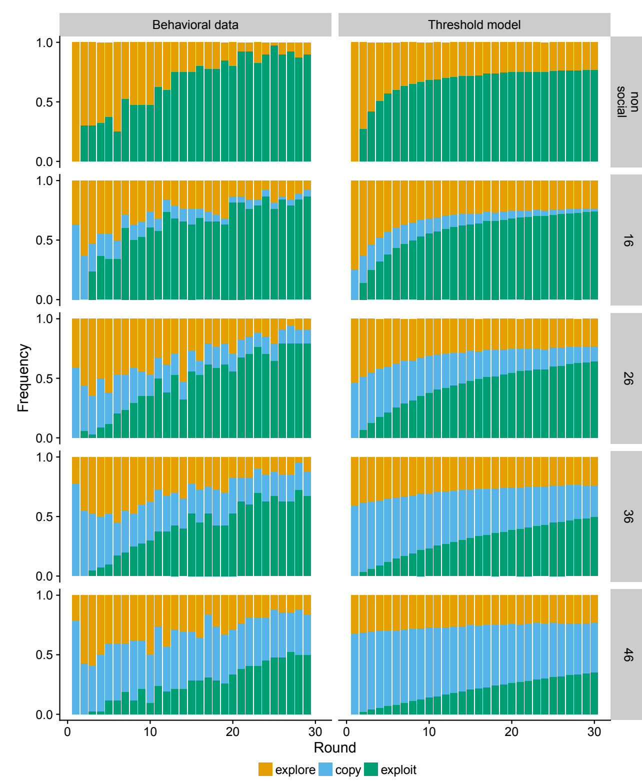

Figure 4: Proportions of explore, copy, and exploit as a function of the round in our experimental data (left column) and as predicted by the threshold model (right column). Each row refers to a different value of social information X_c (as indicated in the gray rectangles on the right).

the non-zero base exploration rate of the participants (modeled by $p_0 > 0$) a better solution is eventually discovered and exploited. But this discovery is delayed as compared to the control condition, decreasing the number of rounds when X_e can be exploited, and hence reducing overall individual performance. For $X_c = 16$ a better solution for X_e can be discovered very easily and thus the overall performance is not undermined.

Discussion

In this paper we investigated how the presence of social information can influence individual strategy and performance. We studied this question by means of a simple exploration-exploitation task. In contrast to existing research, we only implemented the most basic parts of the exploration-exploitation paradigm, thus focusing on how social information affects decisions. On that account, we ignored the structure of communication networks and considered unidirectional information flow between two individuals only (Mason & Watts, 2012; Toyokawa, Kim, & Kameda, 2014; Wisdom et al., 2013). Furthermore, we fixed the social information to one static value that does not change over time (in contrast to Mason and Watts (2012); Mesoudi (2011); Toyokawa et al. (2014)), and we eliminated the spatial correlations between the payoffs (in contrast to Mason et al. (2008); Wu, Schulz, Speekenbrink, Nelson, and Meder (2017)

With this design, we discovered that social information can undermine not only the collective, but also the individual performance (cf. Wisdom et al. (2013)). In line with a recent simulation study conducted by Barkoczi and Galesic (2016), our results show that the detrimental impact of social information can depend on the decision strategies employed and not exclusively on the network structure connecting people or the structure of the environment.

The rationale of early copying. The harmful effect of social information is caused by the participants' tendency to copy social information too early in the experiment, without knowing its relative value compared to what can be discovered by individual exploration. One common assumption is that people start exploring their environment and with the help of the gathered data evaluate the relative value of the social information, and only then capitalize on the best solution (cf. our benchmark model and also March (1991)). An alternative strategy, assuming an already exhaustive search by the previous participant, could be to rely completely on the social information provided and copy it all the time. Such a strategy would minimize the cost for exploration. However our results show that participants do not follow either of those strategies, but rather prefer to copy in the very early phases of the experiment and only then start to explore. But copying does not provide any information about the environment, so why do people adopt this seemingly irrational strategy? Early copying can actually be reasonable if the payoff of the copied solution is sufficiently good. In fact, the best crop found by participants in the control condition had an average payoff of 56.2, which is above the average of the payoff distribution. Thus, copying early could possibly have been beneficial if the social information given to the participants was representative of the performance in the absence of social information, which was not the case in our experiment. It is reasonable to assume that participants expected the social information to be representative of the underlying distribution of payoffs, which could have then justified their early copying.

Sequential versus simultaneous treatment. Compared to other studies, one specificity of our design is that the participants completed the experiment sequentially rather than simultaneously (Mason & Watts, 2012; Toyokawa et al., 2014). Each participant in the experimental condition was exposed to the best value found by another participant in the control condition during independent search. Other experimental designs implemented simultaneous interactions, in which participants could see what others have found at the end of every round. In this case, the dynamics might be different, as participants could not reasonably assume that their peer had extensively explored the environment beforehand. In this context, the rationale of the early copying strategy would vanish, and we would expect people to copy only later in the experiment. In agreement with this interpretation, late copying has been reported in experiments involving simultaneous interactions (Mason & Watts, 2012).

Accuracy of the threshold model. Despite its simplicity, the threshold model reproduces our experimental observations quite closely, suggesting that it has captured key aspects of the decision strategy. The model assumes that the probability to search for a new solution only depends on the payoff of the best available solution. The model therefore ignores most of the available information, such as the payoff distribution of previously explored solutions and the remaining number of rounds. The fact that the model captures key dynamics in the experiment without explicitly accounting for temporal dimension is surprising, since intuitively, people are likely to explore less as the end of the experiment approaches. Nevertheless, given the current quality of the model's predictions, adding a temporal component might only yield a marginal improvement of the predictions. Hence, time appears as a cue that has – if anything – a minor role on the decision strategy. Finally, whereas the model currently implements a fixed threshold value, it is also possible to consider an adaptive threshold that would vary with the observed sample of payoffs.

Future research directions To examine basic aspects of how social information affects decision strategies of exploration and exploitation, we deliberately started from a simple task. In future research we will gradually increase the complexity of the decision-making setting – up to the point where a complete, realistic situation can be described. This will consist of implementing simultaneous interactions, varying group size, manipulating the communication network, changing the payoff distribution and injecting private information. These additions would also allow us to determine the predictive power of the model, by testing it in different environments.

Acknowledgments

We thank the ARC research group and especially Wouter van den Bos for fruitful discussions. Samuli Reijula's research is supported by the Academy of Finland.

References

Bala, V., & Goyal, S. (1998). Learning from Neighbours. *The Review of Economic Studies*, *65*(3), 595–621.

Barkoczi, D., & Galesic, M. (2016). Social learning strategies modify the effect of network structure on group performance. *Nature Communications*, *7*, 13109.

Bolton, P., & Harris, C. (1999). Strategic Experimentation. *Econometrica*, *67*(2), 349–374.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*(Supplement 2), 10918.

Derex, M., & Boyd, R. (2015). The foundations of the human cultural niche. *Nature Communications*, *6*, 8398.

Derex, M., & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences*, *113*(11), 2982.

Fang, C., Lee, J., & Schilling, M. A. (2009). Balancing Exploration and Exploitation Through Structural Design: The Isolation of Subgroups and Organizational Learning. *Organization Science*, *21*(3), 625–642.

Giraldeau, L., Valone, T. J., & Templeton, J. J. (2002). Potential disadvantages of using socially acquired information. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1427), 1559.

Henrich, J., & Boyd, R. (1998). The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior*, *19*(4), 215–241.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*(1), 46–54.

Jayles, B., Kim, H.-r., Escobedo, R., Cezera, S., Blanchet, A., Kameda, T., ... Theraulaz, G. (2017). How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences*.

Jönsson, M. L., Hahn, U., & Olsson, E. J. (2015). The kind of group you want to belong to: Effects of group structure on group accuracy. *Cognition*, *142*, 191–204.

Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior*, *32*(1), 4–14.

Lazer, D., & Friedman, A. (2007). The Network Structure of Exploration and Exploitation. *Administrative Science Quarterly*, *52*(4), 667–694.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87.

March, J. G. (2006). Rationality, foolishness, and adaptive intelligence. *Strategic Management Journal*, *27*(3), 201–214.

Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, *137*(3), 422–433.

Mason, W. A., & Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, *109*(3), 764.

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., ... Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, *2*(3), 191–215.

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: Payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334–342.

Moussaïd, M., Herzog, S. M., Kämmer, J. E., & Hertwig, R. (2017). Reach and speed of judgment propagation in the laboratory. *Proceedings of the National Academy of Sciences*.

Moussaïd, M., & Seyed Yahosseini, K. (2016). Can Simple Transmission Chains Foster Collective Intelligence in Binary-Choice Tasks? *PloS one*, *11*(11), e0167223.

Rapoport, A., & Tversky, A. (1970). Choice behavior in an optional stopping task. *Organizational Behavior and Human Performance*, *5*(2), 105-120.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., ... Laland, K. N. (2010). Why Copy Others? Insights from the Social Learning Strategies Tournament. *Science*, *328*(5975), 208.

Rendell, L., Fogarty, L., Hoppitt, W. J., Morgan, T. J., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, *15*(2), 68–76.

Rogers, A. R. (1988). Does Biology Constrain Culture? *American Anthropologist*, *90*(4), 819–831.

Salganik, M. J., & Watts, D. J. (2008). Leading the Herd Astray: An Experimental Study of Self-fulfilling Prophecies in an Artificial Cultural Market. *Social Psychology Quarterly*, *71*(4), 338-355.

Tomasello, M. (1999). The Human Adaptation for Culture. *Annual Review of Anthropology*, *28*(1), 509–529.

Toyokawa, W., Kim, H.-r., & Kameda, T. (2014). Human Collective Intelligence under Dual Exploration-Exploitation Dilemmas. *PLOS ONE*, *9*(4), e95789.

Wisdom, T. N., Song, X., & Goldstone, R. L. (2013). Social Learning Strategies in Networked Groups. *Cognitive Science*, *37*(8), 1383–1425.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2017). Mapping the unknown: The spatially correlated multi-armed bandit. *bioRxiv*.