Analytical reasoning task reveals limits of social learning in networks

Iyad Rahwan1,2, Dmytro Krasnoshtan1, Azim Shariff3 and Jean-François Bonnefon4,5

1Department of Electrical Engineering and Computer Science, Masdar Institute of Science and Technology, Abu Dhabi 54224, United Arab Emirates
2School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, UK
3Department of Psychology, 1277 University of Oregon, Eugene, OR 97403-1227, USA
4Centre National de la Recherche Scientifique (CNRS), Toulouse, France
5Unité 5263 Cognition, Langues, Language et Ergonomie, Université de Toulouse, Toulouse, France

Social learning—by observing and copying others—is a highly successful cultural mechanism for adaptation, outperforming individual information acquisition and experience. Here, we investigate social learning in the context of the uniquely human capacity for reflective, analytical reasoning. A hallmark of the human mind is its ability to engage analytical reasoning, and suppress false associative intuitions. Through a set of laboratory-based network experiments, we find that social learning fails to propagate this cognitive strategy. When people make false intuitive conclusions and are exposed to the analytic output of their peers, they recognize and adopt this correct output. But they fail to engage analytical reasoning in similar subsequent tasks. Thus, humans exhibit an ‘unreflective copying bias’, which limits their social learning to the output, rather than the process, of their peers’ reasoning—even when doing so requires minimal effort and no technical skill. In contrast to much recent work on observation-based social learning, which emphasizes the propagation of successful behaviour through copying, our findings identify a limit on the power of social networks in situations that require analytical reasoning.

1. Introduction

Social learning is a key cultural mechanism that improves the performance of individuals and groups, often outperforming individual trial-and-error learning [1,2]. Although the social imitation of successful behaviour is not uniquely human [3–6], it has been particularly important to human evolution. Sophisticated social learning mechanisms allowed humans to incorporate and accumulate the knowledge accrued by others, instead of solely relying, as most mammals do, on information that the individual can learn on its own, during its own lifetime. This ability to exploit what is called a ‘cultural niche’ [7] allowed humans to spread throughout radically different environments in time scales too slow to produce genetic adaptations. Social learning has retained its importance for the spreading of best practices [8,9], healthy habits [10], cooperation [11] or democratic participation [12]. Nevertheless, social learning has its limitations. For example, extensively copying the behaviour of successful or prestigious models is a low-cost way to acquire successful behaviours—but it often comes at the potential cost of not understanding the reasons why these behaviours were successful in the first place. In other words, social learners can be prone to blindly copying the behaviour of their models, without acquiring the causal knowledge or the reasoning processes that were responsible for this behaviour [7]. For instance, social learners may copy several behaviours of a successful fisher without knowing which behaviour is actually responsible for catching many fish, much less why that technique is successful.

This limitation is especially relevant in an age of increasing connectivity, facilitated by the Internet and social media [13]. While modern telecommunication technologies may impede some cognitive functions [14], many have suggested
that social media could make for better decisions [15,16]. Here, we investigate this claim by exploring whether and how social networks can be conducive to the uniquely human strategy of rational, analytic reasoning—whose engagement is critical for sound financial decisions, accurate risk assessments and many other demanding mental feats [17].

As social learning can, in theory, copy either a surface behaviour or the reasons behind this behaviour, networks can serve two purposes in relation to analytic reasoning. First, networks may propagate analytical reasoning processes. That is, individuals who witness rational decisions going against their intuition may be prompted to reflect on their initial intuition, recognize it as incorrect and spontaneously switch to a more analytic thinking style in subsequent, similar tasks. We refer to this phenomenon as the contagion of analytical processing (note that we use reasoning and processing interchangeably throughout). Another possibility is that networks propagate correct responses to analytic problems. That is, individuals who witness rational decisions going against their intuition may recognize their intuition as incorrect and adopt the correct decision, but do so without engaging analytic reasoning themselves. Thus, increased connectivity, by increasing the availability of diverse information sources, may enable individuals to obtain higher quality information and perspectives, without necessarily being able to generate similar insights independently. We refer to this phenomenon as the contagion of analytical output.

Not all networks may be able to propagate analytical processing or analytical output. Indeed, the effectiveness of social learning can depend on the topology of the network in which interactions occur. For example, in the context of complex exploration tasks, social learning is especially beneficial in networks with shorter average path length [18], through which information about good solutions propagates faster [19]. Network topology can also affect the effectiveness and efficiency of group coordination [20] and consensus-making [21]. To investigate whether, how and which social networks might propagate analytical reasoning, we ran a series of five laboratory-based network sessions, involving 20 subjects each. In each session, subjects sat at individual computer workstations and solved a series of analytic problems. Each subject was randomly assigned to a node in an underlying network, which determined the neighbours (in the sense of the network, rather than physical proximity) whose responses were visible to the subject. This general protocol was recently used to explore the effect of network structure on coordination [20], consensus [21,22], foraging [18] and behaviour diffusion [23].

### 2. The experiments

#### 2.1. The networks

Different network topologies were used in the five sessions. The first session provided a Baseline condition in which subjects were not connected to any neighbour, and thus did not see any of the other participants’ responses. The other sessions spanned a wide range of possible structures (figure 1). At one extreme, in the Clustered topology, we have a collection of five fully connected cliques, with single connections between them. This structure provides minimal opportunities for long-range communication of ideas, while reinforcing local interaction. As such, it reduces the diversity of information sources. Moreover, this structure captures hierarchical situations: people who connect different cliques have privileged positions, just as the members of the central clique. The peripheral cliques may correspond to different departments in an organization, with a single member (manager) communicating with the central management clique. At the other extreme, we have a Full Connectivity topology wherein each individual is connected to every other individual, facilitating instant information propagation between every two individuals. In between, we have two topologies wherein connections are randomly determined. In the Erdős–Rényi topology, every two individuals have the same probability of being connected—as a consequence, all individuals in the final network have roughly the same number of connections [24]. By contrast, the Barabási–Albert topology is constructed in such a way that well-connected individuals are the most likely to acquire new connections—as a consequence, the network includes a few highly connected individuals who serve as communication hubs [25].

#### 2.2. The problems

Subjects were asked to solve a series of three questions known as the cognitive reflection test (CRT). These three questions have been used in hundreds of studies as a test of analytic reasoning [26]. All three require to engage analytic reasoning in order to overcome an incorrect intuition. It is important to recognize that no particular skill or knowledge is required to generate the correct answer—only the engagement of effortful, analytic reasoning process. Thus, there is no particular ‘trick’ which, once learned, can be used in subsequent tasks. The subject should simply recognize that initial intuition cannot be trusted, and a more reflective attitude is needed.

Table 1 displays the three questions, their incorrect intuitive response and their correct response. To measure the effect of social connectivity, each subject answered five trials for each of the three questions. In the first trial, subjects responded independently. In the subsequent trials 2–5, subjects could see the responses that their network neighbours (determined by the subjects’ network topology) entered during the previous rounds. No information was given about the accuracy of these responses. Subjects were informed that they would accumulate monetary rewards for every correct response they gave, on every trial. This set-up provides us with an ideal test-bed to pit analytical process contagion against analytical output contagion. Output contagion should improve performance from one trial to the next (within each question), but not from one question to the next. Processing contagion should improve performance from one question to the next, in addition to improving performance from one trial to the next.

### 3. Results

#### 3.1. Process contagion

Subjects’ performance appears in figure 2, trial by trial and question by question. Separate logistic regressions were conducted in each topology, in order to detect evidence of process contagion or output contagion. In order to detect process contagion, we tested whether the performance of subjects in each of our four topologies improved across questions, over and above the progression observed in the Baseline condition. For example, in the case of the Clustered topology, we conducted a logistic regression in which the predictors were the
question (first, second, third), the topology (Baseline, Clustered) and their interaction. The dependent measure was always the performance (correct or incorrect) during the first trial of each question. What counts as evidence for process contagion is a significant interaction between question and topology, showing that the increase in performance in the network group is greater than the increase in performance in the Baseline group. We detected no such significant interaction for any topology, all $z < 1.05$, all $p > 0.28$. It appears that whatever the topology, performance never improves significantly from one question to the next.

### 3.2. Output contagion

To detect output contagion, we tested whether the performance of subjects in each of our four topologies improved across trials within each question, over and above the progression observed in the Baseline condition. For example, in the case of the Clustered topology, we conducted a logistic regression in which the predictors were the trial (first, last), the topology (Baseline and Clustered in this case) and their interaction. What counts as evidence for process contagion is a significant interaction between trial and topology, showing that the increase in performance in the network group is greater than the increase in performance in the Baseline group.
than the increase in performance in the Baseline group. As shown in table 2, we obtained such evidence for all topologies except Clustered. In all other topologies, subjects’ performance largely improved across trials, as the correct response to each question spread in turn across the network.

### 3.3. The connectivity effect

The Clustered topology was an exception insofar as it seemed unable to improve performance over and above what was already observed in the Baseline group. One possible reason might be that connectivity in the Clustered network is insufficient to spread the correct, analytical response. To test whether the individual connectivity of a node was linked to the final performance of the subject in this node, we computed an index of connectivity (global distance to all other nodes, i.e. closeness centrality) and an index of final performance (average proportion of correct responses during the last trial of each question), for each node in each network. As expected, these two indices were significantly correlated, \( r(78) = 0.38, p < 0.0001 \).

**Figure 2.** Proportion of correct responses for each of the three CRT questions, across trials, in the five topologies.

**Table 2.** Results of the logistic regressions testing for output contagion. A significant interaction effect means that the rate of correct responses increases more in the tested topology than in the Baseline condition. This is the case in all but the Clustered topology.
4. Discussion

In sum, our data show that networks can help to solve analytic problems—with two important caveats. First, networks do not propagate the analytic reasoning style required to independently arrive at correct answers. They can only propagate the selection of the correct response to analytic problems, one at a time. Second, low-connectivity networks will not do, and the least connected individuals within a network will not obtain full benefits. Of these two results, the failure of networks to propagate analytical processing is especially striking. Consider that it is possible to prime analytical processing using very subtle cues—such as an evocative image of Rodin’s Thinker [27] or listing questions using a challenging font [28]. How can we explain, then, that repeated exposure to the analytic output of peers in a network, and even the subsequent recognition and adoption of their correct answer, all fail to prime analytic reasoning in subsequent tasks?

Social learning is a low-cost phenomenon because learners evaluate behaviours, not on the basis of an understanding of what makes a behaviour successful, but merely on the characteristics of others who perform those behaviours. The trade-off for minimizing those costs, though, is that without that deep understanding, learners can be inaccurate in what they choose to copy [7]. This propensity may explain why subjects persist in copying only analytical responses in our tasks, while copying analytical processing would be fairly easy, cost-less and financially rewarding. The current data therefore reveal what we call an unreflective copying bias—the tendency to copy what others do as a result of successful analytic processing, without engaging analytic processing oneself.

This observation suggests that there are limits to the efficacy of social learning in propagating successful reasoning strategies. As ‘cultural learning can increase average fitness only if it increases the ability of the population to create adaptive information’ [7], our results exemplify imitation as a form of free riding that ultimately may not improve society’s capacity to innovate through analytical reasoning.

The discovery of the unreflective copying bias also sheds new light on the ongoing debate about the promises and perils of social media and the Internet. Some have suggested that the Internet is ‘making us stupid’ [29] by encouraging rapid, unthoughtful sampling of small bits of information from many sources, thus limiting our capacity for concentration, contemplation and reflection [30], and eliminating the healthy diversity of opinions [31]. Yet, others have argued that these technologies significantly expand opportunities for learning, problem solving and informed decision-making [16]. Intriguingly, our results suggest that both these views might be correct, in their own time. On the one hand, the unreflective copying bias can facilitate the rapid propagation of analytical responses over social networks, fulfilling their promise of improved decision-making. But on the other hand, the bias may very well decrease the frequency of analytical processing, by making it easy and commonplace for people to reach analytical responses without engaging analytical processing. In sum, the unreflective copying bias alone can explain why increased connectivity may eventually make us stupid by making us smarter first.

Our results complement the large literature on dual-process accounts of reasoning, which has been recently popularized by a number of authors [17,32–34]. In particular, our results suggest that while people’s common bias in favour of intuition can lead to problematic decisions, social learning fixes this problem, but only superficially. In other words, social learning does not seem to help individuals bypass their bias in favour of intuition, but rather helps society as a whole thrive despite this bias.

Several limitations should be noted. Cultural anthropologists make a distinction between copying, which has low accuracy but also low cost, and teaching, which has higher accuracy, but is less accessible because of difficulty of finding willing teachers [35]. Most social learning happens using copying [7], and this is the form of social learning we focus on.

Secondly, our results do not entirely rule out the possibility of contagion of analytical processing. There is always a risk when drawing conclusions from a null result, because null effects can result from insufficient statistical power. Note though that our statistical power was largely sufficient to detect contagion of analytical output, which suggests that there was no contagion of analytical processing to detect. A possible response, of course, is that contagion of analytical processing may simply require a longer period of time to take place compared with contagion of analytical output. Further experiments using a larger collection of questions may help increase confidence in our findings.

Thirdly, one limitation of our study is that the order of the three CRT questions was kept constant across sessions, which prevented us from controlling for the relative difficulty of the questions. The literature is inconclusive about this relative difficulty, though, as it appears to vary across samples. Note that the data collected in the Baseline condition suggest that, at least in our sample, participants found the three questions to be equally difficult.

Finally, a possible objection to our result is that failure to propagate analytical reasoning may be due to a qualitative difference in the skills required to solve each of the CRT questions, and therefore may not indicate an absence of social learning. We believe that this objection is unlikely to hold. We specifically used the three standard CRT questions [26] because they have been used in numerous studies to test analytic reasoning and ability to suppress intuition. This vast literature has never hinted at the possibility that the three questions might use significantly different, domain-specific skills. Detailed studies of the psychological structure of the CRT suggested that it indeed measures a single construct [36].

5. Material and methods

5.1. Subjects and sessions

Five experimental sessions were conducted during the spring and summer of 2013. Each session involved 20 subjects, totalling 100 subjects. Subjects were students at the Department of Psychology from the University of Oregon. The maximum of age was 26, the minimum was 18 with the average age of 19.65 (s.d. 1.68).

Participation in the experiment was voluntarily and monetary compensation was provided to participants. Each student attempted 7 questions \( \times 5 \) attempts each = 35 total attempts. Each correct attempt earns the student $0.25, which leads to a maximum earning of 7 questions \( \times 5 \) attempts \( \times $0.25 = $8.75 \). This is in addition to the subject pool credit or show-up remuneration the subjects have already received.

5.2. Procedure

Each session lasted approximately 45 min. Participants were randomly assigned to one of the five experimental sessions (one
control and four networked conditions). In the control condition, each participant worked completely independently of others. For each of the games in the four networked conditions, a computer program randomly assigned each participant to one of the 20 available nodes in the network and hid the identities of the participants. Note that subjects were unaware of the full structure of the network to which they were assigned.

On arrival to the laboratory, participants had time to familiarize themselves with the system by going through a step-by-step Web-based tutorial. They were able to explore what the system interface would look like, what type of questions would be asked, how many questions and iterations will be in the game, and how to submit answers, how to read information about the answers of their neighbours in the network (except in the control condition) and how the payment is calculated (see the electronic supplementary information for screen shots).

During the tutorial, subjects could raise their hands to ask the experimenter any question about the game flow. Before starting the experiment, each subject took a mandatory quiz to ensure that s/he had read and understood the tutorial. The quiz tested their understanding of key concepts of the game: how many times will you see each question, does your reward depend on the responses of other players (it does not) and how the reward is calculated.

Once a participant successfully finishes the quiz, s/he enters the experiment and waits for other participants. The experiment begins as soon as all 20 participants are ready. Participants were asked not to communicate with each other. Each participant had a paper and a pen in case they were needed.

At the beginning, the subject sees a question and has 60 s to read it. After 60 s, the screen changes to another one with a text field for submitting the answer. On the first attempt of each question, the participant answers the question completely independently. The subject has 30 s to submit the answer. A countdown timer on the screen indicates the amount of time left. If the subject was unable to submit the answer during the allocated time, his/her answer would not be recorded and s/he does not receive credit for that attempt. But s/he will still be able to move further. If the subject submits his/her answer before the timer reaches zero, s/he receives the confirmation and waits until the timeout, to ensure all participants are synchronized.

When the timer reaches zero, the field for submitting the answer becomes unavailable, and the participants move to another screen. On this screen, the subject sees an invitation to review a list of answers by his/her network neighbours. The subject does not know who these people are or where they sit. The subject is aware that it is possible not to see all of their neighbours’ answers (this happens if those neighbours were unable to submit their answers on time).

The subject has 15 s to review these neighbours’ answers and to consider whether to change his/her answer in the next iteration. At the end of the 15 s, the subject moves to a screen similar to the previous one. The only difference is that if the person submitted the answer on previous iteration, that answer appears in the field by default. If the subject does not wish to alter his/her answer, s/he still needs to click ‘Submit’ in order to receive credit for the attempt. Also the right upper table is populated with answers of the participant’s neighbours in the network (only from round 2 to 5). The user has 30 s to resubmit the answer or to change to a new answer.

The above process iterates five times, after which the subjects are moved to the next question (a completely new question). Again, subjects always have 60 s on their first attempt at a question, which they answer completely independently—without viewing network neighbours’ answers.

When a participant finishes the last (fifth) iteration of the last question, s/he is redirected to a summary page with the results of the experiment. On this page, s/he sees the final payment and information about his/her own answers in each attempt, as well as the correct answer to each question. The participant receives credit (money) for each correct answer they give, i.e. every attempt on every question. This means that they have incentive to submit correct answers every time, including in their first (independent) attempt.

All subjects played in the same controlled environment. During all experiments, the laboratory, equipment, material, descriptions, tutorial, list and sequence of questions, time to read, answer the question and review the answer of neighbours remained the same.

Acknowledgements. The authors thank Bethany Lasseter and Cassandra M. Brandes for help in running the laboratory experiments. Iyad Rahwan is grateful for discussions with Jacob Crandall. I.R., J.F.B. designed the research; D.K. implemented the platform; A.S. conducted the experiments; I.R., D.K., J.F.B., A.S. analysed the data; I.R., D.K., J.F.B., A.S. wrote the paper.

Data accessibility. Data used in the article are uploaded as online electronic supplementary material.

References