debris flow that carried the still partially articulated bodies deeper into the cave, to deposit them along a subterranean stream (Fig. 3).

References and Notes

1. L. R. Berger et al., Science **328**, 195 (2010).
2. R. Broom, J. T. Robinson, Nature **164**, 322 (1949).
3. A. R. Hughes, P. V. Tobias, Nature **265**, 310 (1977).
4. R. Clarke, S. Afr. J. Sci. **94**, 460 (1998).
5. T. C. Partridge, in The Cenozoic of Southern Africa, T. C. Partridge, R. R. Maud, Eds. (Oxford Univ. Press, New York, 2000), pp. 100–125.
6. T. C. Partridge, D. E. Granger, M. W. Caffee, R. J. Clarke, Science **300**, 607 (2003).
7. C. K. Brain, in Swartkrans: A Cave's Chronicle of Early Man, C. K. Brain, Ed. (Monogr. 8, Transvaal Museum, Pretoria, ed. 2, 2004).
8. D. J. de Ruiter et al., J. Hum. Evol. **56**, 497 (2009).
9. T. C. Partridge, Nature **275**, 282 (1978).
10. C. K. Brain, in Palaeoclimate and Evolution, with Emphasis on Human Origins, E. S. Vrba, Ed. (New Haven, 1995), pp. 385–424.
11. R. Pickering et al., J. Hum. Evol. **53**, 602 (2007).
12. D. McB. Martin, C. W. Clendennin, B. Krapez, N. J. McNaughton, J. Geol. Soc. London **155**, 311 (1998).
13. P. G. Eriksson et al., in The Geology of South Africa, M. R. Johnson, C. R. Anhaeusser, R. J. Thomas, Eds. (Geological Society of South Africa, Johannesburg, 2006), pp. 237–260.
14. K. A. Eriksson, J. F. Truswell, Trans. Geol. Soc. S. Afr. **77**, 211 (1974).
15. C. K. Brain, The Hunters or the Hunted? An Introduction to African Cave Taphonomy (Univ. of Chicago Press, Chicago, 1981).
16. H. Blatt et al., Petrology: Igneous, Sedimentary and Metamorphic (Freeman, New York, ed. 3, 2006).
17. M. E. Tucker, Sedimentary Petrology (Blackwell, Oxford, ed. 2, 1991).
18. G. J. Retallack, Soils of the Past: An Introduction to Palaeopedeology (Harper Collins, London, 1990).
19. P. Coussot, M. Meunier, Earth Sci. Rev. **40**, 209 (1996).
20. L. R. Berger, R. Lacruz, D. J. De Ruiter, Am. J. Phys. Anthropol. **119**, 192 (2002).
21. J. P. Brugal, H. Roche, M. Kibunjia, C. R. Palevol. **2**, 675 (2003).
22. F. H. Brown, I. McDougall, I. Davies, R. Maier, in Ancestors: The Hard Evidence, E. Delson, Ed. (Liss, New York, 1985), pp. 82–90.
23. M. E. Lewis, L. Werdelin, in Hominin Environments in the East African Pliocene: An Assessment of the Faunal Evidence, R. Bobe, Z. Alemseged, A. K. Behrensmeyer, Eds. (Springer, Netherlands, 2007), pp. 77–105.
24. J. D. Woodhead et al., Quat. Geochronol. **1**, 208 (2006).
25. M. A. Lanphere, D. E. Champion, R. L. Christiansen, G. A. Izett, GSA Bull. **114**, 559 (2002).
26. J. G. Ogg, A. G. Smith, in A Geologic Time Scale 2004, F. Gradstein, J. G. Ogg, A. G. Smith, Eds. (Cambridge Univ. Press, Cambridge, UK, 2004), pp. 63–86.
27. K. Nishiizumi, M. W. Caffee, R. C. Finkel, G. Brimhall, T. Mote, Earth Planet. Sci. Lett. **237**, 499 (2005).
28. S. R. Hall, D. L. Farber, L. Audin, R. C. Finkel, A. S. Meriaux, Tectonophysics **459**, 186 (2008).
29. A. Matmon et al., Geol. Soc. Am. Bull. **121**, 688 (2009).
30. We thank the South African Heritage Resource Agency for allowing work on the Malapa site; the Nash family for granting access to the Malapa site and continued support; the Palaeontological Scientific Trust (PAST), the Institute for Human Evolution, University of the Witwatersrand and the School of Geosciences, University of the Witwatersrand, for funding; and the University of the Witwatersrand's Schools of Geosciences and Anatomical Sciences and the Bernard Price Institute for Palaeontology for support and facilities. Further support was provided by AfricaArray (P.H.G.M.D.), the France-SA, !Khure Africa project (G.C.P.K., D.L.F., A.-S.M.), and LLNL GEO-CAMS and the US-NSF EAR-0345895 (D.L.F.). R.P. is supported by a Swiss National Science Foundation Post-Doctoral Grant and gratefully acknowledges J. Woodhead, A. Grieg, and the team at Melbourne University. A.I.R.H. was funded through a UNSW Faculty of Medicine research grant and ARC Discovery Grant DP0877603 and acknowledges the support of the Liverpool University Geomagnetism Laboratory staff. Contributions to mapping and sample preparation by G. Charlesworth and Z. Jinnah are kindly acknowledged.

Why Copy Others? Insights from the Social Learning Strategies Tournament

L. Rendell,[1]* R. Boyd,[2] D. Cownden,[3] M. Enquist,[4,5] K. Eriksson,[5,6] M. W. Feldman,[7] L. Fogarty,[1] S. Ghirlanda,[5,8] T. Lillicrap,[9] K. N. Laland[1]*

Social learning (learning through observation or interaction with other individuals) is widespread in nature and is central to the remarkable success of humanity, yet it remains unclear why copying is profitable and how to copy most effectively. To address these questions, we organized a computer tournament in which entrants submitted strategies specifying how to use social learning and its asocial alternative (for example, trial-and-error learning) to acquire adaptive behavior in a complex environment. Most current theory predicts the emergence of mixed strategies that rely on some combination of the two types of learning. In the tournament, however, strategies that relied heavily on social learning were found to be remarkably successful, even when asocial information was no more costly than social information. Social learning proved advantageous because individuals frequently demonstrated the highest-payoff behavior in their repertoire, inadvertently filtering information for copiers. The winning strategy (discountmachine) relied nearly exclusively on social learning and weighted information according to the time since acquisition.

Human culture is widely thought to underlie the extraordinary demographic success of our species, manifest in virtually every terrestrial habitat (1, 2). Cultural processes facilitate the spread of adaptive knowledge, accumulated over generations, allowing individuals to acquire vital life skills. One of the foundations of culture is social learning, learning influenced by observation or interaction with other individuals (3), which occurs widely in various forms across the animal kingdom (4). Yet it remains something of a mystery why individuals profit by copying others and how best to do this.

At first sight, social learning appears advantageous because it allows individuals to avoid the costs, in terms of effort and risk, of trial-and-error learning. However, social learning can also cost time and effort, and theoretical work reveals that it can be error-prone, leading individuals to acquire inappropriate or outdated information in nonuniform and changing environments (5–11). Current theory suggests that to avoid these errors individuals should be selective in when and how they use social learning, so as to balance its advantages against the risks inherent in its indiscriminate use (9). Accordingly, natural selection is expected to have favored social learning strategies, psychological mechanisms that specify when individuals copy and from whom they learn (12, 13).

These issues lie at the interface of multiple academic fields, spanning the sciences, social sciences and humanities, from artificial intelligence to zoology (5, 14–18). Formal theoretical analyses [e.g., (2, 5–9, 11–13, 19)] and experimental studies (20, 21) have explored a small number of plausible learning strategies. Although insightful, this work has focused on simple rules that can be studied with analytical methods and can only explore a tiny subset of strategies. For a more authoritative understanding of when to acquire information from others and how best to do so, the relative merits of a large number of alternative social learning strategies must be assessed. To address this, we organized a computer tournament in which strategies competed in a complex and changing simulation environment. €10,000 was offered as first prize. The organization of similar tournaments by Robert Axelrod in the 1980s proved an extremely effective means for

[1]Centre for Social Learning and Cognitive Evolution, School of Biology, University of St. Andrews, Queen's Terrace, St. Andrews, Fife KY16 9TS, UK. [2]Department of Anthropology, University of California, Los Angeles, CA 90095, USA. [3]Department of Mathematics and Statistics, Queen's University, Jeffery Hall, University Avenue, Kingston, Ontario K7L 3N6, Canada. [4]Department of Zoology, Stockholm University, 11691 Stockholm, Sweden. [5]Centre for the Study of Cultural Evolution, Stockholm University, 11691 Stockholm, Sweden. [6]School of Education, Culture, and Communication, Mälardalen University, 72123 Västerås, Sweden. [7]Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA. [8]Department of Psychology, University of Bologna, 40127 Bologna, Italy. [9]Centre for Neuroscience Studies, Queen's University, 18 Stuart Street, Kingston, Ontario K7L 3N6, Canada.

*To whom correspondence should be addressed. E-mail: ler4@st-andrews.ac.uk (L.R.); knl1@st-andrews.ac.uk (K.N.L.)

investigating the evolution of cooperation and is widely credited with invigorating that field (*22*).

The tournament. The simulated environment for our tournament was a "multiarmed bandit" (*18*), analogous to the "one-armed bandit" slot machine but with multiple "arms." In the tournament, the bandit had 100 arms, each representing a different behavior and each with a distinct payoff drawn independently from an exponential distribution. Furthermore, we posited a temporally varying environment realized by changing the payoffs with a probability, p_c, per behavior per simulation round, with new payoffs drawn from the same distribution. The possibility of acquiring outdated information is seen as a crucial weakness of social learning [e.g., (*6*)].

Entered strategies had to specify how individual agents in a finite population choose between three possible moves in each round, namely Innovate, Observe, and Exploit. Innovate represented asocial learning, that is, individual learning stemming solely through direct interaction with the environment, for example, through trial and error. An Innovate move always returned accurate information about the payoff of a randomly selected behavior previously unknown to the agent. Observe represented any form of social learning or copying through which an agent could acquire a behavior performed by another individual, whether by observation of or interaction with that individual (*3*). An Observe move returned noisy information about the behavior and payoff currently being demonstrated in the population by one or more other agents playing Exploit. Playing Observe could return no behavior if none was demonstrated or if a behavior that was already in the agent's repertoire is observed and always occurred with error, such that the wrong behavior or wrong payoff could be acquired. The probabilities of these errors occurring and the number of agents observed were parameters we varied. Lastly, Exploit represented the performance of a behavior from the agent's repertoire, equivalent to pulling one of the multiarmed bandit's levers. Agents could only obtain a payoff by playing Exploit.

Evolutionary dynamics were realized by a death-birth process (*23*). Agents died with a constant probability of 1/50 per round and were replaced by the offspring of another agent. The probability that an agent was chosen to reproduce was proportional to its mean lifetime payoff, calculated as its summed payoff from playing Exploit divided by the number of simulation rounds that it had been alive. The obtained payoffs thus directly affected an agent's fitness. Offspring inherited their parent's strategy unless mutation occurred, in which case the offspring was given a strategy randomly chosen from the others playing in that simulation. We recorded the average frequency of each strategy in the population over the last 2500 rounds of each 10,000-round simulation and gave each strategy a score that was the mean of these values over the simulations in which it participated.

Axelrod's cooperation tournaments were based on a widely accepted theoretical framework for

the study of cooperation: the Prisoner's Dilemma. Although there is no such currently established framework for social learning research, multiarmed bandits have been widely deployed to study learning across biology, economics, artificial intelligence research, and computer science [e.g., (*18*, *24–28*)] because they mimic a common problem faced by individuals who must make decisions about how to allocate their time in order to maximize their payoffs. Multiarmed bandits capture the essence of many difficult problems in the real world, for instance, where there are many possible actions, only a few of which yield a high payoff; where it is possible to learn asocially or through observation of others; where copying error occurs; and where the environment changes. When the payoffs of a multiarmed bandit change over time as in our tournament, the bandit is

termed "restless," and the framework has the advantage of proving extremely difficult, perhaps impossible, to optimize analytically [e.g., (*29*)]. Thus, we could be confident that our tournament would be a genuine challenge for all entrants.

In all other respects, we attempted to keep the model structure as simple as possible to maintain breadth of applicability and ease of understanding and to attract the maximum number of participants. We balanced this simplicity with the inclusion of three features that we considered critical, namely, individual memories (to facilitate learning); a degree of error associated with social learning (the existence of which nearly all the current literature agrees on); and replicator dynamics with mutation, to allow an evolutionary process. We used a common currency for costs (time) and made each possible move cost the same to
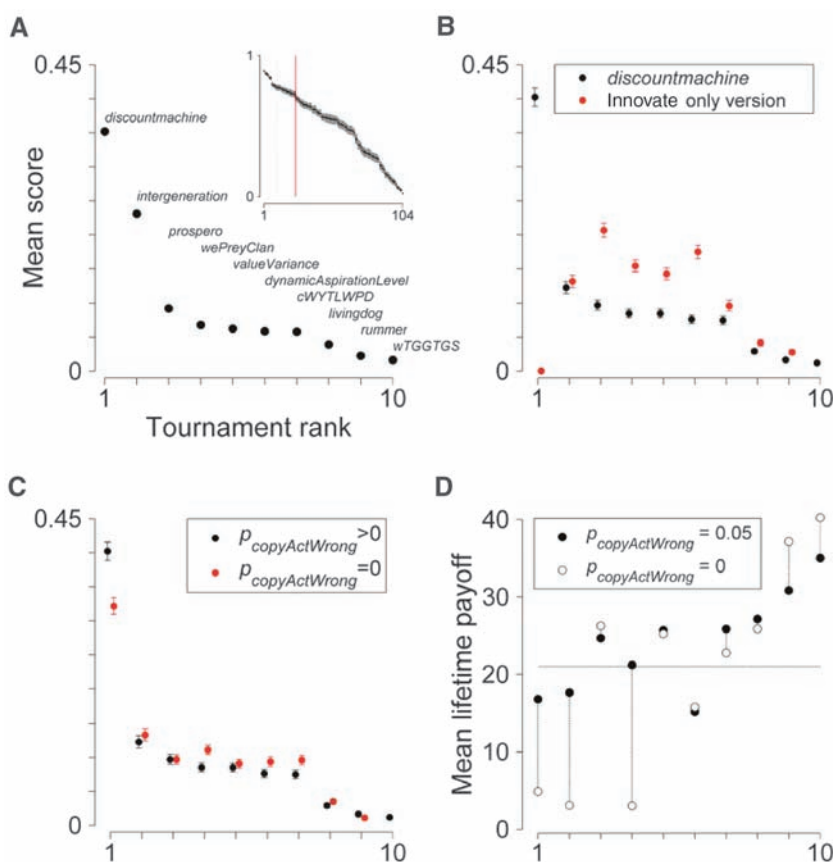

Fig. 1. Performance of entered strategies. (**A**) Ranked overall strategy scores in the final stage of the tournament (cWYTLWPD indicates copyWhenYoungThenLearnWhen-PayoffsDrop; wTGGTGS, whenTheGoingGetsToughGetScrounging). Scores are averaged over all final stage simulations. (Inset) Scores for all 104 entered strategies. Dotted black line indicates the 10 highest scoring strategies; solid red line indicates the 24 strategies entered into further pairwise conditions. Error bars are ± SEM but mostly not visible because all SEMs < 0.004. (**B**) Ranked scores from those final-stage simulations in which conditions were chosen at random (*33*), and under the same conditions but with the tournament winner, discountmachine, recoded to learn only with Innovate and never Observe (red). (**C**) As in (B) but comparing original results with $p_{copyActWrong}$ fixed at 0 (red). (**D**) Average individual fitness, measured as mean lifetime payoff, in populations containing only single strategies for each of the final-stage contestants, ranked by tournament placing. Data are average values from the last quarter of single simulations, which were run under the same conditions as the first stage of the tournament and also under the same conditions except with $p_{copyActWrong} = 0$. The horizontal dashed line represents the mean lifetime payoff of individuals when all strategies are played together under the same conditions. Strategies relying exclusively on social learning are those ranked 1, 2, and 4.

minimize structural assumptions about learning costs. The agents in our simulations could not identify or communicate directly with each other, an assumption that precluded the deployment of some model-based strategies present in the cultural evolution literature [e.g., prestige bias (30)]. Nonetheless, we reasoned that the simplicity, accessibility, and generality of the proposed tournament structure outweighed the benefits of further complexity.

Analyses. We received 104 entries, most, although not all (31), from academics across a wide range of disciplines and from all over the world. The tournament was run in two stages. Strategies first competed in pairwise round-robin contests, taking turns to invade or to resist invasion by another strategy under a single set of conditions (32). The 10 best performers progressed to a second stage, where all 10 strategies competed simultaneously in melee contests over a range of simulation conditions (33). Scores in the first stage ranged from 0.02 to 0.89 (with a theoretical maximum of 1), indicating considerable variation in strategy effectiveness (Fig. 1A).

Statistical analysis indicates that much of this variation is explained by the extent to which strategies used social learning, more social learning being associated with higher payoffs. We examined the factors that made strategies successful by using linear multiple regression and model selection by using Akaike's information criterion (AIC) (33). The best-fit model contained five predictors (Table 1). Two predictors had effect sizes more than twice the magnitude of the others: the proportion of those learning moves that were Observe and the variance in the number of rounds before a strategy first played Exploit. The proportion of learning moves dedicated to Observe had a strong positive effect on a strategy's score (Fig. 2A). Although Innovate cost no more than Observe, the best strategies relied almost entirely on social learning; that is, when learning, they almost exclusively chose Observe rather than Innovate. The proportion of moves that involved learning of any kind had a negative effect, indicating that it was detrimental to invest too much time in learning be-

cause payoffs came only through Exploit. The data reveal an unexpectedly low optimum proportion of time spent learning (Fig. 2C).

The timing of (either form of) learning also emerged as a crucial factor. Strategies with a high variance in the number of rounds spent learning before the agent first played Exploit, caused by occasionally waiting too long before beginning to exploit, tended to do poorly (Fig. 2B). Conversely, strategies that engaged in longer bouts of exploiting between learning moves tended to do significantly better (Fig. 2D). Successful strategies were able to target their learning to coincide with periods when average population payoffs dropped, indicating a change in the environment that had rendered a behavior less profitable (Fig. 2E). This pattern was observable statistically as the lagged correlation between the time series of average payoff and the proportion of learning moves in the population. We calculated Pearson correlation coefficients between the average payoff at simulation round t and the proportion of learning moves at round $t + \Delta$, with $0 < \Delta < 10,000$. Accurate targeting of learning to periods where payoffs are dropping produces large negative correlation coefficients for small Δ. We compared the correlations for populations containing the 10 strategies that progressed to the final stage with the correlations from simulations run with strategies ranked 78 to 88 in the first stage of the tournament (i.e., markedly less-successful strategies). For the final-stage strategies, the strongest negative correlations were always found with lags of less than three ($\Delta < 3$) and were significantly stronger than the strongest correlations found for the less-successful strategies [two-sample t test, $P < 0.0001$ (fig. S9)]. Successful strategies

Table 1. Parameters of the AIC best-fit model predicting strategy scores in the first, pairwise, tournament stage. Adjusted $R^2 = 0.76$. Dash entry indicates not applicable.

Predictor	Effect size (β weight)	β	SE	t	$p(>\|t\|)$
(Intercept)	–	0.32	0.06	5.43	<0.0001
Proportion of learning that is Observe	0.42	0.43	0.06	7.15	<0.0001
Variance in rounds to first Exploit*	−0.42	−0.06	0.01	−6.62	<0.0001
Proportion of learning moves	−0.17	−0.34	0.12	−2.79	0.0063
Average rounds between learning moves	0.16	0.01	<0.01	3.09	0.0026
Estimate p_c? (yes = 1, no = 0)	−0.07	−0.04	0.03	−1.47	0.1452

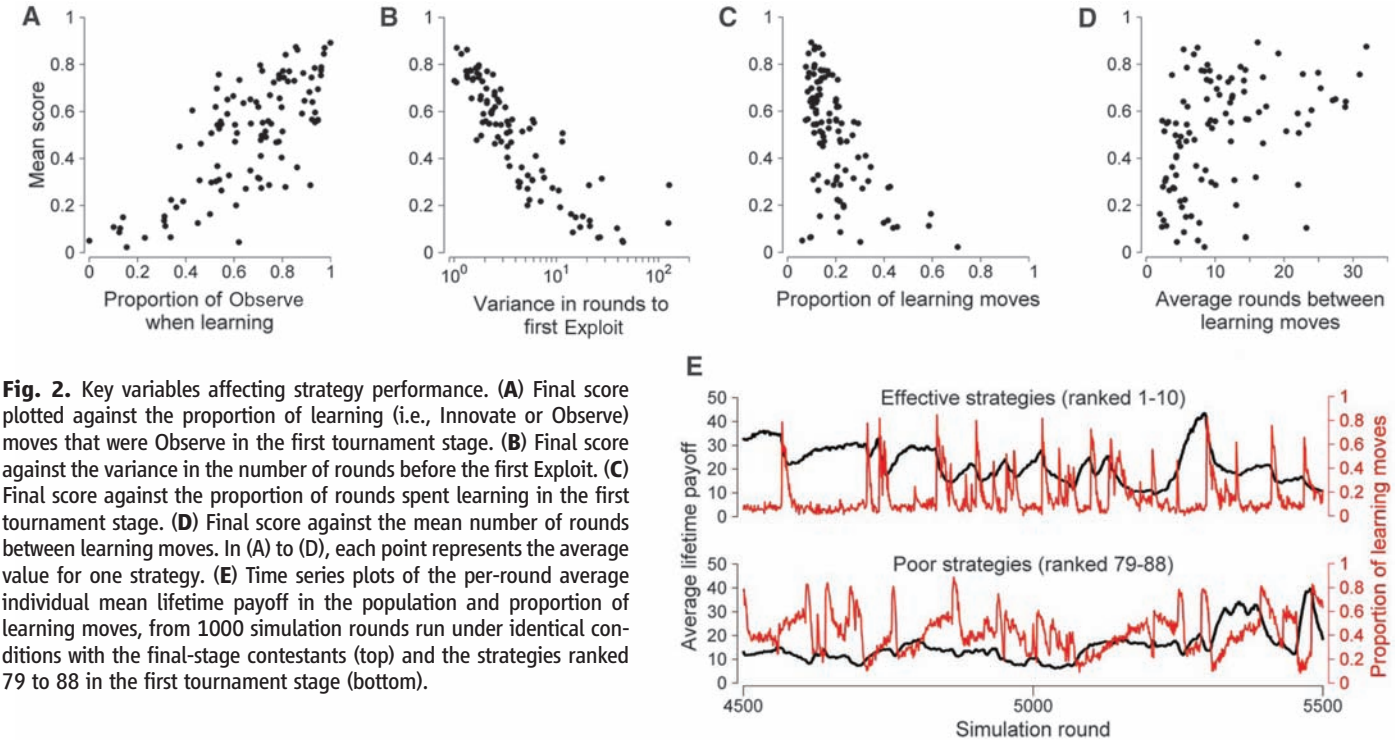*We used the natural log of this predictor to give a better linear relationship.

Fig. 2. Key variables affecting strategy performance. (**A**) Final score plotted against the proportion of learning (i.e., Innovate or Observe) moves that were Observe in the first tournament stage. (**B**) Final score against the variance in the number of rounds before the first Exploit. (**C**) Final score against the proportion of rounds spent learning in the first tournament stage. (**D**) Final score against the mean number of rounds between learning moves. In (A) to (D), each point represents the average value for one strategy. (**E**) Time series plots of the per-round average individual mean lifetime payoff in the population and proportion of learning moves, from 1000 simulation rounds run under identical conditions with the final-stage contestants (top) and the strategies ranked 79 to 88 in the first tournament stage (bottom).

targeted learning to periods when it was likely to be most valuable (i.e., when the environment changed) but otherwise minimized learning, allowing them both to improve their payoffs through learning and to maintain high rates of exploiting (Table 1). The issue of when to break off exploiting current knowledge in order to invest in further knowledge gain, the exploitation/exploration tradeoff, had not been incorporated into previous theory in this field, and our tournament introduces this dimension into the domain of understanding social learning.

The strategy discountmachine (34) emerged as a convincing winner (Fig. 1A) in the second stage of the tournament, which pitted the 10 best performers in the first stage against each other in simultaneous competition under a range of conditions (it was also the winner of the pairwise phase). Strikingly, both discountmachine and the runner-up, intergeneration, relied nearly exclusively on Observe as their means to learn (Fig. 3, C and D), and at least 50% of the learning of all of the second-stage strategies was Observe. Although all second-stage strategies increased their amount of learning as the rate of environmental change increased, the best performers capped the level of learning to a max-

imum to maintain payoffs (Fig. 3A). The winning strategy stood out by spreading learning more evenly across agent life spans than any other second-stage strategy (Fig. 3B). It did this by, uniquely among the finalists, using a proxy of geometric discounting to estimate expected future payoffs from either learning or playing Exploit.

Winning strategies also relied more heavily on recently acquired than older information. The top two strategies shared the following expression for estimating the expected payoff (w_{expected}) of a known behavior:

$$w_{\text{exp}} = w(1 - p_{\text{est}})^i + \overline{w}_{\text{est}}\left(1 - (1 - p_{\text{est}})^i\right) \quad (1)$$

where w is the current payoff held in the agent's memory and acquired i rounds ago, $\overline{w}_{\text{est}}$ is the estimated mean payoff for all behavior, and p_{est} is an estimate of p_c, the probability of payoff change. This expression weighs expected payoffs increasingly toward an estimated mean as the time since information was last obtained increases. Given the uncertain and potentially conflicting nature of information obtained through social learning, the winning strategy used a further weighing based on its estimate of p_c, dis-

counting older social information more severely in more variable environments than in relatively constant ones. No other strategies in the melee round evaluated payoffs in this way.

In the melee round, simulations were run to explore the effects of varying the rate of environmental change (p_c); the probability and scale of errors associated with social learning; and the relative costs of the two forms of learning, the last achieved by increasing the number of other agents sampled when playing Observe (social learning being cheap when multiple individuals are observed). We found the tournament results to be unexpectedly robust to variation in these factors (Fig. 4). The first- and second-place strategies switched rank in some conditions, namely when the environment was more stable (Fig. 4A) and when social learning was cheap relative to asocial learning [i.e., the number of agents sampled by Observe was high (Fig. 4D)]. Increasing the probability and the magnitude of errors associated with social learning made nearly no difference to the strategy rankings (Fig. 4, B and C); even at extreme values, strategies heavily reliant on social learning thrived (fig. S11). This implies that social learning is of widespread utility even when it provides no in-
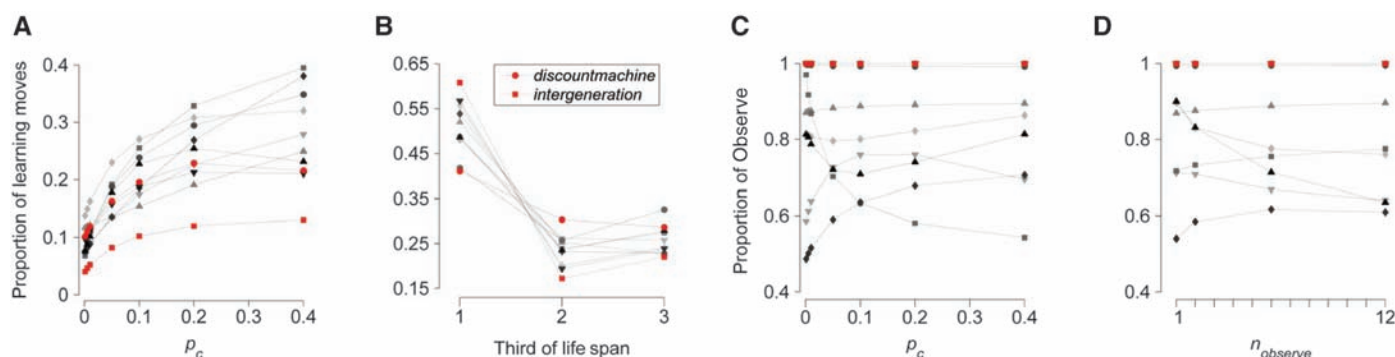

Fig. 3. Why the winner won. Error bars are ± SEM but mostly not visible because all SEMs < 0.003. (**A**) Proportion and (**B**) timing of learning moves in the final tournament stage. First and second place strategies are highlighted; the rank of the other strategies is indicated by shading, with darker shading indicating higher rank. (**C** and **D**) Variation in the proportion of learning moves that were Observe with (**C**) variation in the rate of environmental change (p_c) and (**D**) the number of agents sampled when playing Observe (n_{observe}), in the final tournament stage.
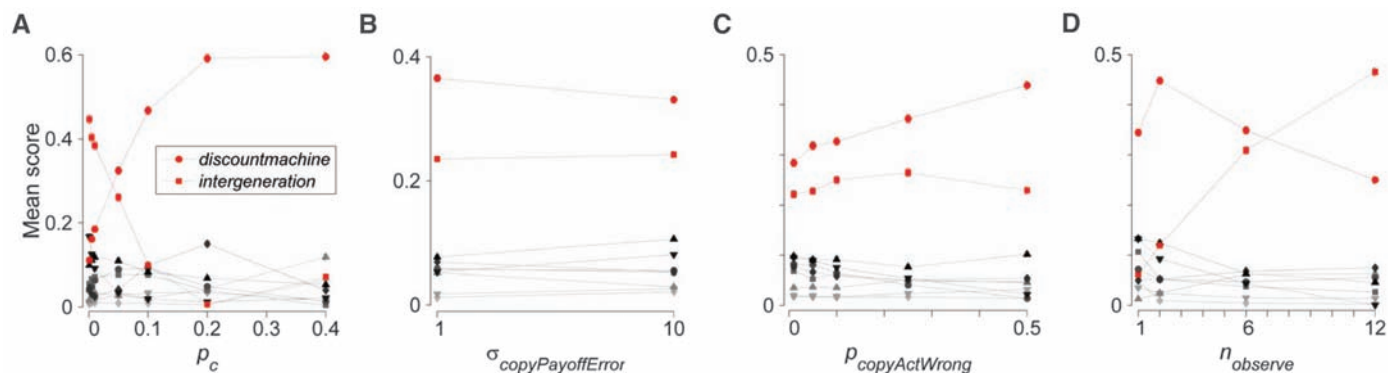

Fig. 4. Social learning dominates irrespective of cost across a broad range of conditions. Plots show mean strategy scores (±variance) across systematic melee conditions with respect to (**A**) variation in the rate of environmental change (p_c), (**B**) $\sigma_{\text{copyPayoffError}}$, the standard deviation of a normally distributed error applied to payoffs returned by Observe, (**C**) $p_{\text{copyActWrong}}$, the probability that Observe returned a behavior selected, at random from those not actually observed, and (**D**) the number of other agents sampled when playing Observe (n_{observe}). First and second place strategies are highlighted; the rank of the other strategies is indicated by shading with darker shading indicating higher rank. Error bars are ± SEM but mostly not visible because all SEMs < 0.01.

Fig. 5. Results of a series of simulations in which the tournament winner played against a version of itself altered to learn only by Innovate. The rate of environmental change (p_c) was systematically varied. Five simulations were run at each level of p_c. Other parameters were fixed at $n_{observe} = 1$, $p_{copyActWrong} = 0.05$, and $\sigma_{payoffError} = 1$.
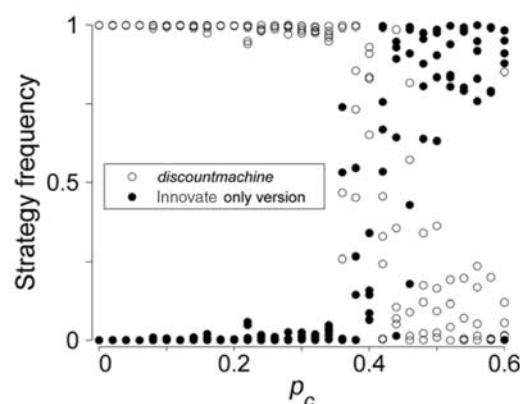
formation about payoffs. Nor does this utility rest on our assumption that copying errors can introduce new behaviors (fig. S13). These are surprising results, given that the error-prone nature of social learning is widely thought to be a weakness of this form of learning, whereas the ability to copy multiple models rapidly or preferentially copy high-payoff behavior are regarded as strengths (*1*). Strategies relying heavily on social learning did best irrespective of the number of individuals sampled by Observe (Fig. 4D). These findings are particularly unexpected in the light of previous theoretical analyses (*5–8, 10, 11, 13*), virtually all of which have posited some structural cost to asocial learning and errors in social learning.

Previous theory also suggests that reliance on social learning should not necessarily raise the average fitness of individuals in a population (*6, 7, 10*) and may even depress it (*35*). However, this was not the case for the strategies successful enough to make the second stage; in this second round, average individual fitness in mixed-strategy populations was positively correlated with the proportion of learning in the population that was social [$r = 0.16$, $P = 0.02$ (fig. S9)]. In contrast, for poorly performing strategies the relationship between average individual fitness and the rate of social learning was strongly negative ($r = -0.71$, $P < 0.001$; fig. S9). This highlights the importance of the strategic use of social learning in raising the average fitness in a population (*5, 12, 19*).

Strategies that did well were not, however, those that maximized average individual fitness when fixed in a population. Instead, we found a strong inverse relationship between the mean fitness of individuals in populations containing only one strategy and that strategy's performance in the tournament (Fig. 1D). Furthermore, the mean lifetime payoff in the population when all strategies competed together under the same conditions was lower than the levels achieved by lower-ranking strategies when playing alone. This illustrates the parasitic effect of strategies that rely heavily on Observe (e.g., discountmachine, intergeneration, wePreyClan, and dynamicAspirationLevel; ranked 1, 2, 4, and 6; all played Observe on at least 95% of learning moves). From this we can conclude

that strategies using a mixture of social and asocial learning are vulnerable to invasion by those using social learning alone, which may result in a population with lower mean fitness. An established rule in ecology specifies that, among competitors for a resource, the dominant competitor will be the species that can persist at the lowest resource level (*36*). Recent theory suggests an equivalent rule may apply when alternative social learning strategies compete in a population: The strategy that eventually dominates will be the one that can persist with the lowest frequency of asocial learning (*13*). Our findings are consistent with this hypothesis.

Discussion. The most important outcome of the tournament is the remarkable success of strategies that rely heavily on copying when learning in spite of the absence of a structural cost to asocial learning, an observation evocative of human culture. This outcome was not anticipated by the tournament organizers, nor by the committee of experts established to oversee the tournament, nor, judging by the high variance in reliance on social learning (Fig. 2A), by most of the tournament entrants. Although the outcome is in some respects consistent with models that used simpler environmental conditions and in which individual learning is inherently costly relative to social learning (*5*), in our tournament the environment was complex and there was no inherent fitness cost to asocial learning. Indeed, there turned out to be a considerable cost to social learning because it failed to introduce new behavior into an agent's repertoire in 53% of all the Observe moves in the first tournament phase, overwhelmingly because agents observed behaviors they already knew. Nonetheless, social learning proved advantageous because other agents were rational in demonstrating the behavior in their repertoire with the highest payoff, thereby making adaptive information available for others to copy. This is confirmed by modified simulations wherein social learners could not benefit from this filtering process and in which social learning performed poorly (fig. S12). Under any random payoff distribution, if one observes an agent using the best of several behaviors that it knows about, then the expected payoff of this behavior is much higher than the

average payoff of all behaviors, which is the expected return for innovating. Previous theory has proposed that individuals should critically evaluate which form of learning to adopt in order to ensure that social learning is only used adaptively (*11*), but a conclusion from our tournament is that this may not be necessary. Provided the copied individuals themselves have selected the best behavior to perform from at least two possible options, social learning will be adaptive. We suspect that this is the reason why copying is widespread in the animal kingdom.

That social learning was critical to the success of the winning strategy is shown by the results of running the random conditions portion of the second tournament stage with a version of discountmachine recoded to learn only by Innovate: it placed last (Fig. 1B). We also found that discountmachine dominated its recoded cousin across a large portion of the plausible parameters space with respect to environmental change (Fig. 5), with payoffs needing to change with 50% probability per round before the Innovate-only version could gain a foothold. This is another way that our tournament challenges existing theory, which predicts that evolution will inevitably lead to a stable equilibrium where both social and asocial learning persist in a population [e.g., (*6*)].

It is important to note that, although our tournament may offer greater realism than past analytical theory, the simulation framework remains a simplification of the real world where, for instance, model-based biases and direct interactions between individuals (*15*) operate. It remains to be established to what extent our results will hold if these are introduced in future tournaments, where the specific strategies that prospered here may not do so well. Nonetheless, the basic generality of the multiarmed bandit problem we posed lends confidence that the insights derived from the tournament may be quite general.

The tournament also draws attention to the importance of social learning errors as a source of adaptive behavioral diversity. In our tournament, there was a probability, $p_{copyActWrong}$, that a social learner acquired a randomly selected behavior rather than the observed behavior. Modeling social learning errors in this way means new behavior can enter the population without explicit innovation. The importance of these errors is illustrated by the fact that strategies relying exclusively on social learning were unable to maintain high individual fitness when $p_{copyActWrong} = 0$ (Fig. 1D). This does not mean that the success of the winning strategy depended on the condition $p_{copyActWrong} > 0$; in the presence of other strategies providing the necessary innovations, discountmachine and intergeneration maintain their respective first and second places when $p_{copyActWrong} = 0$ (Fig. 1C). Other models have highlighted copying errors as potentially important in human cultural evolution (*37*), but the extent to which adaptive innovations actually come about through such

errors is an important empirical question ripe for investigation.

The ability to evaluate current information on the basis of its age and to judge how valuable that information might be in the future, given knowledge of rates of environmental change, is also highlighted by the tournament. There is limited empirical evidence that animals are able to discount information on the basis of the time since it was acquired (38), but little doubt that humans are capable of such computation. Our tournament suggests that the adaptive use of social learning could be critically linked to such cognitive abilities. There are obvious parallels with the largely open question of mental time travel, the ability to project current conditions into the future, in nonhumans (39), raising the hypothesis that this cognitive ability could be one factor behind the gulf between human culture and any nonhuman counterpart. A critical next step will be to evaluate experimentally to what extent human behavior mirrors that of the tournament strategies [e.g., (40)]. By drawing attention to the importance of adaptive filtering by the copied individual and temporal discounting by the copier, the tournament helps to explain both why social learning is common in nature and why human beings happen to be so good at it.

References and Notes

1. P. J. Richerson, R. Boyd, *Not by Genes Alone* (Univ. of Chicago Press, Chicago, 2005).
2. L. L. Cavalli-Sforza, M. W. Feldman, *Cultural Transmission and Evolution: A Quantitative Approach* (Princeton Univ. Press, Princeton, NJ, 1981).
3. C. M. Heyes, *Biol. Rev. Camb. Philos. Soc.* **69**, 207 (1994).
4. W. Hoppitt, K. N. Laland, *Adv. Stud. Behav.* **38**, 105 (2008).
5. R. Boyd, P. J. Richerson, *Culture and the Evolutionary Process* (Univ. of Chicago Press, Chicago, 1985).
6. A. Rogers, *Am. Anthropol.* **90**, 819 (1988).
7. M. W. Feldman, K. Aoki, J. Kumm, *Anthropol. Sci.* **104**, 209 (1996).
8. T. Kameda, D. Nakanishi, *Evol. Hum. Behav.* **23**, 373 (2002).
9. L.-A. Giraldeau, T. J. Valone, J. J. Templeton, *Philos. Trans. R. Soc. London Ser. B* **357**, 1559 (2002).
10. J. Y. Wakano, K. Aoki, M. W. Feldman, *Theor. Popul. Biol.* **66**, 249 (2004).
11. M. Enquist, K. Eriksson, S. Ghirlanda, *Am. Anthropol.* **109**, 727 (2007).
12. K. N. Laland, *Learn. Behav.* **32**, 4 (2004).
13. J. Kendal, L.-A. Giraldeau, K. Laland, *J. Theor. Biol.* **260**, 210 (2009).
14. E. Danchin, L.-A. Giraldeau, T. J. Valone, R. H. Wagner, *Science* **305**, 487 (2004).
15. J. Apesteguia, S. Huck, J. Oechssler, *J. Econ. Theory* **136**, 217 (2007).
16. H. Whitehead, L. Rendell, R. W. Osborne, B. Würsig, *Biol. Conserv.* **120**, 427 (2004).
17. K. Dautenhahn, C. L. Nehaniv, Eds., *Imitation in Animals and Artifacts* (MIT Press, London, 2002).
18. K. H. Schlag, *J. Econ. Theory* **78**, 130 (1998).
19. J. Henrich, R. McElreath, *Evol. Anthropol.* **12**, 123 (2003).
20. J. R. Kendal, L. Rendell, T. W. Pike, K. N. Laland, *Behav. Ecol.* **20**, 238 (2009).
21. B. G. Galef Jr., *Adv. Stud. Behav.* **39**, 117 (2009).
22. R. Axelrod, W. D. Hamilton, *Science* **211**, 1390 (1981).
23. M. A. Nowak, *Evolutionary Dynamics: Exploring the Equations of Life* (Harvard Univ. Press, Cambridge, MA, 2006).
24. D. E. Koulouriotis, A. Xanthopoulos, *Appl. Math. Comput.* **196**, 913 (2008).
25. R. Gross et al., *J. R. Soc. Interface* **5**, 1193 (2008).
26. D. Bergemann, J. Valimaki, *Econometrica* **64**, 1125 (1996).
27. J. Niño-Mora, *Top (Madr.)* **15**, 161 (2007).
28. P. Auer, N. Cesa-Bianchi, P. Fischer, *Mach. Learn.* **47**, 235 (2002).
29. C. H. Papadimitriou, J. N. Tsitsiklis, *Math. Oper. Res.* **24**, 293 (1999).
30. J. Henrich, F. J. Gil-White, *Evol. Hum. Behav.* **22**, 165 (2001).
31. Three strategies were entered by high school students, and one of these (whenTheGoingGetsToughGetScrounging, submitted by Ralph Barton and Joshua Borin of Westminster School in the United Kingdom) achieved notable success by ranking 10th overall.
32. These conditions were $p_c = 0.01$, $n_{observe} = 1$ (the number of agents sampled when playing Observe), $p_{copyActWrong} = 0.05$ (the probability that Observe returned a behavior selected, at random, from those not actually observed), $\sigma_{payoffError} = 1$ (the standard deviation of a normally distributed error applied to observed payoffs); and we also ran more pairwise contests under several other conditions with the top 24 performing strategies to ensure that progression to the second stage was not solely dependent on these particular parameter values.
33. A full description of the simulation procedures and statistical analyses is available on *Science* Online.
34. This strategy was entered by Daniel Cownden and Timothy Lillicrap, who were subsequently invited to be authors on this paper.
35. L. Rendell, L. Fogarty, K. N. Laland, *Evolution* **64**, 534 (2010).
36. D. Tilman, *Resource Competition and Community Structure* (Princeton Univ. Press, Princeton, NJ, 1982).
37. J. Henrich, R. Boyd, *J. Cogn. Cult.* **2**, 87 (2002).
38. Y. van Bergen, I. Coolen, K. N. Laland, *Proc. Biol. Sci.* **271**, 957 (2004).
39. T. Suddendorf, M. C. Corballis, *Behav. Brain Sci.* **30**, 299, discussion 313 (2007).
40. M. J. Salganik, P. S. Dodds, D. J. Watts, *Science* **311**, 854 (2006).
41. The authors acknowledge the use of the U.K. National Grid Service (www.grid-support.ac.uk) in carrying out this work. We thank all those who entered the tournament for contributing to its success. We are also very grateful to R. Axelrod for providing advice and support with regard to the tournament design. This research was supported by the CULTAPTATION project (European Commission contract FP6–2004-NESTPATH-043434).

REPORTS

Two-Dimensional Phonon Transport in Supported Graphene

Jae Hun Seol,[1] Insun Jo,[2] Arden L. Moore,[1] Lucas Lindsay,[3,4] Zachary H. Aitken,[5] Michael T. Pettes,[1] Xuesong Li,[1,6] Zhen Yao,[2] Rui Huang,[5] David Broido,[3] Natalio Mingo,[7] Rodney S. Ruoff,[1,6] Li Shi[1,6]*

The reported thermal conductivity (κ) of suspended graphene, 3000 to 5000 watts per meter per kelvin, exceeds that of diamond and graphite. Thus, graphene can be useful in solving heat dissipation problems such as those in nanoelectronics. However, contact with a substrate could affect the thermal transport properties of graphene. Here, we show experimentally that κ of monolayer graphene exfoliated on a silicon dioxide support is still as high as about 600 watts per meter per kelvin near room temperature, exceeding those of metals such as copper. It is lower than that of suspended graphene because of phonons leaking across the graphene-support interface and strong interface-scattering of flexural modes, which make a large contribution to κ in suspended graphene according to a theoretical calculation.

Since graphene was first exfoliated from graphite and studied on dielectric substrates in 2004 (1), the monatomic layer of carbon atoms has attracted great interest for electronic applications because of superior charge mobility (2) and mechanical strength (3), as well as its compatibility with existing planar silicon devices. Other carbon allotropes, including diamond (4), graphite (5), and carbon nanotubes (CNTs) (6–8), have the highest thermal conductivity (κ) values reported because the strong bonding of the light carbon atoms results in a large phonon contribution to κ despite a much smaller electronic component. For similar reasons, graphene is

[1]Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712, USA. [2]Department of Physics, The University of Texas at Austin, Austin, TX 78712, USA. [3]Department of Physics, Boston College, Chestnut Hill, MA 02467, USA. [4]Department of Physics, Computer Science, and Engineering, Christopher Newport University, Newport News, VA 23606, USA. [5]Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX 78712, USA. [6]Texas Materials Institute, The University of Texas at Austin, Austin, TX 78712, USA. [7]Laboratoire d'Innovation pour les Technologies des Energies Nouvelles et les Nanomatériaux, Commissariat à l'Énergie Atomique Grenoble, 17 rue des Martyrs, 38054 Grenoble, France.

*To whom correspondence should be addressed. E-mail: lishi@mail.utexas.edu