G Model
YLMOT-1377; No. of Pages 6

ARTICLE IN PRESS

Learning and Motivation xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Learning and Motivation

journal homepage: www.elsevier.com/locate/l&m

Mental time travel, memory and the social learning strategies tournament

L. Fogarty*, L. Rendell, K.N. Laland

Centre for Social Learning and Cognitive Evolution, School of Biology, University of St Andrews, Bute Building, Westburn Lane, St Andrews, Fife KY16 9TS, United Kingdom

ARTICLE INFO

Available online xxx

Keywords:
Social learning strategies tournament
Social learning
Memory
Strategy
Discounting
Mental time travel

ABSTRACT

The social learning strategies tournament was an open computer-based tournament investigating the best way to learn in a changing environment. Here we present an analysis of the impact of memory on the ability of strategies entered into the social learning strategies tournament (Rendell, Boyd, et al., 2010) to modify their own behavior to suit a changing environment. The tournament showed that a strategy's ability to remember the past and to predict the future were both key to its success. The possibility that a strategy needs to engage in an approximation of 'mental time travel' to succeed in the tournament strongly implies that investment in randomly timed social learning is not enough to guarantee success. A strategy must use social learning strategically with reference to both predicted future environmental states and past environmental states. We examine the two most successful strategies (DiscountMachine and Intergeneration) in terms of their use of memory and discuss the impact of their complex memory use on their ability to time learning moves strategically and track environmental change. The tournament suggests that the human capacity for mental time travel may have improved the efficiency of social learning and allowed humans to invest in more sophisticated social learning than is seen elsewhere in the animal kingdom.

Learning and memory are two clearly related concepts, with the ability to learn resting on the ability to form memories. Learning is generally defined as an extended and long-term process whereby individuals can alter their behavior and state of knowledge, based, in part, on their previous experiences. As the neural substrate for learning, memory can be considered to be a description of how changes in knowledge state, motor abilities or behavioral repertoire are encoded in the brain and later retrieved to form the basis of behavioral changes (Richter, 1966). It is therefore not a huge leap in imagination to glean information about memory use from models of learning. Here we do that, paying special attention to the role of memory in the learning exhibited in the social learning strategies tournament (Rendell, Boyd, et al., 2010; Rendell et al., 2011).

Social learning is learning that is facilitated by observation of, or interaction with other individuals or their products (Heyes, 1994). The idea that social learning was a cheap and efficient form of learning in which individuals need not encounter the dangers or time-consuming costs associated with individual learning was generally accepted until, in 1988, an anthropologist, Alan Rogers, proposed what came to be known as Rogers' paradox. Rogers developed a simple mathematical model which established that agents in a population who engaged in unbiased (random) social learning were, at equilibrium, no fitter than agents who engaged in asocial learning (Rogers, 1988; Enquist, Eriksson, & Ghirlanda, 2007; Rendell, Fogarty,

* Corresponding author.
E-mail address: lf31@st-andrews.ac.uk (L. Fogarty).

0023-9690/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.lmot.2012.05.009

Please cite this article in press as: Fogarty, L., et al. Mental time travel, memory and the social learning strategies tournament. *Learning and Motivation* (2012), http://dx.doi.org/10.1016/j.lmot.2012.05.009

& Laland, 2010). Though not strictly speaking a paradox (rather, a counter-intuitive consequence of frequency-dependent selection), this result was viewed as surprising (hence 'Rogers' paradox'), since social learning underpins cultural learning, and culture, in turn, is widely thought to have increased human fitness substantially.

The notion that social learning cannot generally be random if it is to be adaptive led to interest in the strategic use of social learning: carefully choosing when to learn and from whom (Boyd & Richerson, 1985; Laland, 2004). Social learning strategies involving a mixture of social and asocial learning were widely discussed in cultural evolution (Boyd & Richerson, 1985; Enquist et al., 2007; Feldman, Aoki, & Kumm, 1996; Henrich & McElreath, 2003; Rogers, 1988) and animal social learning (Galef & Laland, 2005; Kendal, Coolen, van Bergen, & Laland, 2005; Laland, 2004) literatures. However, the sheer number of possible strategies posed a challenge to the field and only one or two strategies could be examined at a time (e.g. Enquist et al., 2007; Rendell, Fogarty, et al., 2010).

The 'social learning strategies tournament' was a way to examine the relative merits of a large number of strategies in one standard simulation environment. The organisers invited researchers to take part in a competition where they would submit social learning strategies, and these strategies would compete against each other in a simulation environment. The tournament attracted 104 entries from 14 countries and a myriad of academic disciplines (Rendell, Boyd, et al., 2010). The contests involved 100 agents who could learn about a possible 100 acts, each with a payoff drawn from an exponential distribution, which changed with the environment at a rate p_c. Although the opportunity for advancement in the study of social learning was the original focus of the tournament, it has become clear that the large number of strategies submitted to the tournament coupled with the strictly controlled simulation environment has yielded information on a variety of topics, including optimal use of social learning, optimal timing of social learning moves, the type of culture strategies can produce, and the effect of social learning on the persistence of knowledge and culture (Rendell, Boyd, et al., 2010; Rendell et al., 2011).

The submitted strategies varied substantially in their performance, affected by a number of factors, the majority of which we will not consider here (but see Rendell, Boyd, et al., 2010). However, the winning strategy, called DiscountMachine and submitted by Dan Cownden and Tim Lillicrap, appeared to enhance its performance through a simple form of 'mental time travel', as did several other successful strategies. Here we discuss what this use of 'mental time travel' by a number of the strategies submitted to the tournament might imply about memory in a social learning context.

Typical definitions of mental time travel involve 'episodic memory' of the past, consideration of the future and an understanding of how these relate to the self (Dudai & Carruthers, 2005; Suddendorf & Corballis, 1997; Tulving, 1983). Thus, mental time travel involves subjective reconstruction or construction of past or future events. In the case of humans, researchers can clearly see and demonstrate the presence of episodic memory and future planning. Conversely, in the case of animals, who are unable to verbalize their experiences of memory, researchers must rely on their actions to draw conclusions about the content of their memories and the mechanisms by which they access that content. This has led to the use of the terms 'episodic-like memory', 'future planning' or '*what, where, when*' (*www*) memory in discussions of animal mental time travel (Clayton & Dickinson, 1998; Clayton, Bussey, Emery, & Dickinson, 2003; Raby, Alexis, Dickinson, & Clayton, 2007).

The importance of mental time travel and its specificity to humans has been hotly debated for some years (Clayton et al., 2003; Suddendorf & Busby, 2003; Suddendorf & Corballis, 1997). Even in the midst of this debate, it is useful and interesting to examine the effects of memory, and future projection, on the success of individuals in a changing environment. Here we take a general definition of 'mental time travel' and apply it to the tournament strategies. Our intention is to elucidate the effect of memory use on the success, or otherwise, of individuals using these strategies. The tournament provides us with a unique opportunity to examine the effects of different memory use capabilities on evolutionary success in a standardized and simplified environment.

When discussing memory in computer models such as the tournament, we encounter a series of definitional problems that need addressing before we proceed. Each agent in the tournament had full access to its past moves and the results of those moves. In essence, our agents had access to what Tulving (1983) called 'memory as a warehouse'. Therefore, if they chose to, they could remember every move they made from the moment of their birth to their last simulation round. However our agents did not have access to the computer memory containing information about other agents' histories or environmental parameters (Rendell, Boyd, et al., 2010 S.O.M.).

The agents in our tournament are incapable of the type of complex mental task, vividly reliving the past and imagining the future, described above, making it challenging to define the type of memory to which agents in our model have access. What we are seeing when we look at our computer agents is how the information encoded in their simple memories can be used (for instance, by weighting more recent learning more heavily than older knowledge, or making predictions into the future about the likely success of cultural behavior), and what effect the extent to which they access this information can have on the agent's success. We are narrowly focussing on personal *www* memory, but since there is no spatial context in our simulations, the 'where' aspect is ignored. The difference between mental time travel and *www* memory is really a difference in subjective experience, say the difference between remembering the time, date and location of your birth and being able to mentally relive the event itself (Suddendorf & Busby, 2003). Thus we are interested in the content of memories rather than the subjective experience of them. This is true for many computer models of learning, which focus primarily on what information is retained and the origins of the information – what individuals learn from whom. As a result, the models are generally agnostic as to the exact memory mechanisms used to encode the information. It is therefore possible for us to discuss the effects of learning and memory without defining the memory mechanisms in detail.

The Tournament

Competitors entering the tournament were asked to specify the circumstances under which individual agents should learn asocially (INNOVATE), learn socially (OBSERVE), or perform an act from their repertoire (EXPLOIT). These rules were subsequently translated into computer code.

INNOVATE returned accurate information about the payoff of a randomly selected behavior not in the agent's repertoire. (While in reality which novel behavior individuals adopt may be chosen non-randomly, our assumption that novel behavior was acquired at random offered the advantages of simplicity and tractability, which were important to us in establishing an accessible and appealing game). OBSERVE returned noisy information about the behavior and payoff demonstrated in the population by $n_{observe}$ other agents, selected at random from those playing EXPLOIT. Playing EXPLOIT performed an act from the individual's repertoire, chosen by the strategy and the agent received the associated payoff. The simulation model was organized into a series of iterations, or rounds. Each round a new entry was made in the memory matrix (which we termed '*myHistory*') of each agent in the simulation, whose first row contained the round number or age of the agent, second row contained the previous moves (OBSERVE = 0, INNOVATE = −1, EXPLOIT > 0), third row contained the acts learned or exploited, and whose final row contained the payoffs associated with those acts. For example the agent with the following *myHistory*matrix,

$$myHistory\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 2 & 2 \\ 3 & 33 & 5 & 8 \\ 6 & 9 & 3 & 1 \end{bmatrix},$$

is four rounds old, played OBSERVE in the first round, learning act 3 with payoff 6, etc.

The evolutionary dynamics of the tournament simulation were modeled as a death–birth process with each individual having a fixed probability of dying per iteration. After each death, individuals were selected from the survivors to reproduce in proportion to their mean lifetime payoff gained from EXPLOIT moves, and their offspring replaced dying individuals. Offspring usually inherited their parent's strategy, but could mutate with a low probability, which allowed new strategies to invade the population.

The tournament was run in two stages, although we discuss only the results of the first stage in the current paper. The first stage was a pair-wise round-robin tournament, and the second was a melee that included the top 10 strategies from the first stage. Each pair-wise contest in the first stage consisted of 10 simulations in which agents with one strategy were introduced to a population of agents with another, and 10 simulations in which the first strategy dominated the population with the second invading. The mean frequency of a strategy over the last 2500 simulation rounds was its score for that simulation. These scores were then averaged over 20 simulations, and this average recorded as the overall score for that strategy in that contest. Strategies were ranked according to average score across all pairwise contests.

Memory in the Tournament: Definitions and Difficulties

Using submitted prose descriptions as well as the computer code submitted with or generated for each strategy, we can divide the strategies entered to the tournament into a number of memory-use categories (Table 1). These categories by necessity neglect aspects of mental time travel (like theory of mind) that apply only to humans (and perhaps a few

Table 1
Loose memory categories in the tournament strategies.

	Memory type	Example	Example strategies
0	Minimal use of declarative memory	E.g. I know one act, that's what I'll do	exploitOneInnovation, genderedStrategy, piRounds
1	Used memory as a guide for their next action in terms of agent's age only.	E.g. If I am 7 rounds old, do this	aHandfulOfSkill, innovateAndObserve, observeNoThanks, keepUp
2	Used memory as a guide for their next action in terms of last action only.	E.g. if I did this in the last round, do that in the next	anyRandGambit,
3	Used memory to generate an estimate of temporally local environmental conditions.	E.g. my payoff dropped in the last round, so I have undergone environmental change, better do this	copyIfBetter, balancedCopyWhenPayoffsDecrease, infantJuvinileMature
4	Used memory to estimate environmental parameters and use these to predict the probability of certain environmental changes in the future or discounting (see below)	E.g. I have estimated that there is a 95% chance of an environmental change in the next round, better do this next	W00t, discountMachine, prospero, whenTheGoingGetsThoughGetScrounging

non-human animals) and instead concentrate on the use of memory by our computer agents. Thus we can account for their 'understanding' of environmental changes and motivational states but must remain agnostic as to the emotional or representational content of these memories.

Category 0 indicates that the strategy made no use of *myHistory*. Ignoring their age, past moves, and the results of those moves, these strategies relied on a range of other methods (in one not quite serious entry, the digits of pi) to determine their next move. Categories 1 and 2 describe those strategies that were relatively inflexible, deciding on future moves based on the agent's age or previous moves. This type of strategy partially used its memory, gleaning from it aspects of the agent's history, but neglecting the actual outcomes, in terms of payoff or environmental information that could be used to plan for future environmental changes. Category 3 describes strategies that pay close attention to their histories and use the information in *myHistory* to ascertain when a drastic environmental change has occurred. They can then adjust their behavior accordingly. Finally category 4 strategies use all the information encoded in *myHistory*. Similar to category 3, they are capable of detecting environmental change but they can also use past information to predict the likelihood of future environmental changes (i.e. the exact value of p_c, or other error rates) and to act accordingly. Another important feature of category 4 strategies is their ability to discount information based on the time since acquisition of the information *and* the likelihood of environmental change occurring in the intervening time. In order to do this, they typically catalogue past instances in which a particular behavior was used and keep track of instances in which that behavior did not return the expected payoff – perhaps the closest thing to episodic memory our agents achieved.

This categorization allows us to examine the effect of the complexity of memory use and to examine the link between the past and future aspects of mental time travel. Dudai and Carruthers (2005) suggest that there is a strong link, in the human mind at least, between memories of the past and prediction of the future. It is common sense to assume that prediction of the future without access to information about the past is little more than guessing, but equally it is possible that using memory in the very short term, noticing sudden changes, for example, without using that information to generate predictions about the future is almost equally fruitless. It is therefore possible that the selective advantage of episodic memory lies in its application to future events (Suddendorf & Corballis, 1997; Tulving, 1983). If this were the case, we would expect to see that the average scores of strategies in the pairwise contest of the tournament were similar for strategies using memory alone (category 3) and higher for those using both memory and future planning (category 4).

Results

We analyzed the memory categories in terms of median score using a Kruskal–Wallis test. The memory categories (0, 1, 2, 3, 4) were significantly different from each other ($p < 0.001$) at the 95% confidence level. Category 4, incorporating both use of memory, discounting and prediction of future environmental changes, had the highest median score (Fig. 1) and was significantly higher than categories 0, 1, 2 and 3. Both the eventual winner of the tournament, DiscountMachine, and the second place strategy, Intergeneration, were in category 4, and of the top twenty strategies in the first round of the tournament, 10 were from category 4 and 4 were from category 3.

The original analysis of the tournament strategies also stressed the importance of timing learning moves, whether social or asocial (Fig. 2, Rendell, Boyd, et al., 2010). The analysis showed that the ability to time learning moves to coincide with environmental changes was crucial to success in the tournament. Effective timing of learning combines elements of both

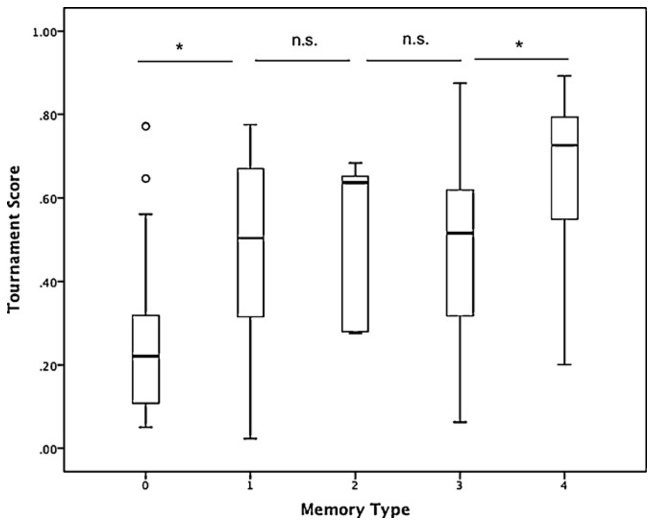

Fig. 1. Plot shows memory category (0, 1, 2, 3, 4) against tournament score for all 104 strategies in stage 1 of the tournament. Median score shown (± interquartile range and maximum and minimum values). $^*p < 0.05$, n.s implies non-significance.

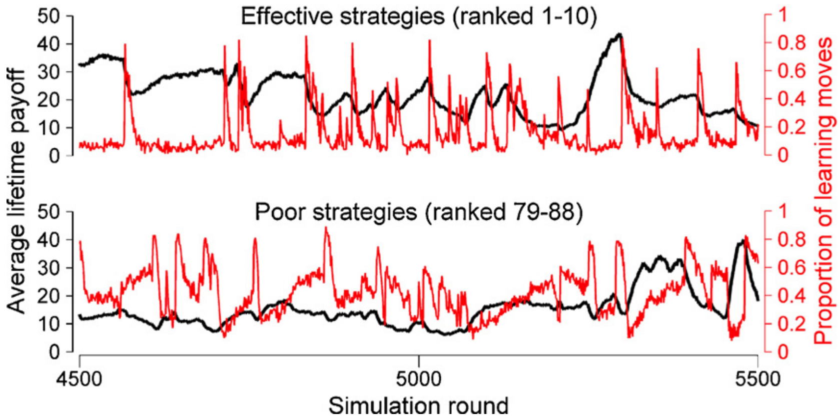

Fig. 2. Time series plots of the per-round average individual mean lifetime payoff in the population and proportion of social learning moves, from 1000 simulation rounds run under identical conditions with the final-stage contestants (top) and the strategies ranked 79–88 in the first tournament stage (bottom).
From Rendell, Boyd, et al. (2010), Rendell, Fogarty, et al. (2010).

category 3 and 4 memory use. The positive relationship between the number of learning moves that were social, and a good strategy's success in the tournament, coupled with a negative relationship between social learning and success in the poorer performing strategies, implies that social learning is adaptive only when used well. This suggests that natural selection could have selected for more efficient use of social learning, and our analysis here suggests that investment in mental time travel may be one means of increasing this efficiency.

Discussion

It is of course difficult to discuss aspects of the strategies submitted to the tournament in isolation since, as the original analysis of the tournament results showed, there were a number of factors that contributed to the success or otherwise of each strategy. The most important factors that emerged from that analysis were the proportion of learning moves that were social, and the timing of those learning moves (Rendell, Boyd, et al., 2010). It is easy however to see that there might be a significant link between the ability to time social learning moves correctly and the strategy's use of mental time travel.

The tournament winner, DiscountMachine, was a complex (category 4) strategy in terms of mental time travel. The most important and robust features of the strategy were (1) its overwhelming propensity to engage in social learning at the expense of individual learning (the strategy could only INNOVATE in one circumstance, when it was in the founding generation of a new simulation), (2) the timing of its social learning moves, which coincided optimally with environmental changes, and (3) its ability to discount information based on the age of the information and an estimate of the rate of environmental change.

DiscountMachine and the strategy that came in second overall in the tournament, Intergeneration, used the same formula to discount information based on its age:

$$w_{\text{exp}} = w(1 - p_{est})^i + \bar{w}_{\text{exp}}(1 - (1 - p_{est})^i),$$

where w was the payoff held in the agent's repertoire, i was the time since learning the information, $\bar{w}_{\text{exp}}$ was the estimated mean payoff for all behavior and p_{est} was an estimate of p_c, the rate of environmental change. p_{est} was estimated by dividing the number of changes in payoffs associated with each behavior by the number of rounds that behavior was known. The formula discounts the value of information towards the estimated mean for all behaviors as the time since acquisition increases. The success of the strategies using this kind of flexible discounting suggests the possibility that natural selection could have shaped memory in a similar way. The strategies suggest that the ability to discount information based on experience of the past and prediction of the future is an important part of survival in changeable environments. The winning and runner-up strategies contrast with a number of less successful strategies in their ability to discount the value of information. Less successful strategies, although also capable of discounting, did so in a fixed and unresponsive way. For example, both DynamicAspirationLevel (3rd in round 1) and SpyNWork (34th in round 1) discounted information according to different rules. SpyNWork behaved in a rational way, exploiting only its best act but only considering acts updated in the last 25 rounds. This amounts to generating a fixed estimate of the rate of environmental change to be $p_c = 0.04$, or one change every 25 rounds. DynamicAspirationLevel was similar but decayed the value of all behavior by a set amount each round. Again this made an assumption about the rate of environmental change without specific reference to what the agents actually knew about the simulation environment and the probability of change. The kind of discounting that these strategies engaged in is arguably similar to the kind of discounting observed in stickleback fish. In the face of conflicting social information, these fish were shown to value their asocial information less as time passed since they collected the information (van Bergen, Coolen, & Laland, 2004). It may be reasonable to assume that the costly calculations involved in constantly updating predictions about the world may not be worthwhile in all circumstances, and that in the absence of extreme variability in environmental conditions, natural selection may fashion some useful 'rules of thumb' about information discounting.

An analysis of the tournament strategies from the perspective of their memory use may go some way towards explaining why humans are particularly good at social learning – humans, perhaps alone, are capable of the kind of complex mental time travel required to maximize the benefits of social learning (through for example, rational discounting or cumulative culture, Vale et al., this issue) and ensure that any investment in social learning is strategic and low-risk. For instance, DiscountMachine computed whether investment in further learning would likely reap greater dividends in the future than relying on current behavior given its estimate of environmental change and the age of the information. We suspect that only humans are capable of this kind of calculation and that deployment of this kind of reasoning greatly enhances the efficiency of learning. Results from the tournament suggest strategic use of memory must be combined with strategic forgetting or discounting in order for a strategy to be successful.

Acknowledgements

The authors would like to thank R. Boyd, M. Enquist, K. Eriksson, M.W. Feldman, S. Ghirlanda and all tournament entrants, including the winners D. Cownden and T. Lillicrap, ERC FP6 'Cultaptation' and ERC Advanced Grant to KNL and a BBSRC studentship to LF.

References

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: Chicago University Press.
Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub-jays. *Nature*, *395*, 272–274. http://dx.doi.org/10.1038/26216
Clayton, N. S., Bussey, T. J., Emery, N. J., & Dickinson, A. (2003). Prometheus to Proust: The case for behavioral criteria for 'mental time travel'. *Trends in Cognitive Sciences*, *7*, 436–437. http://dx.doi.org/10.1016/j.tics.2003.08.003
Dudai, Y., & Carruthers, M. (2005). The Janus face of mnemosyne. *Nature*, *434*, 823–824.
Enquist, M., Eriksson, K., & Ghirlanda, S. (2007). Critical social learning: A solution to Rogers' paradox of non-adaptive culture. *American Anthropologist*, *109*, 727–734.
Feldman, M. W., Aoki, K., & Kumm, J. (1996). Individual versus social learning: Evolutionary analysis in a fluctuating environment. *Anthropological Science*, *104*, 209–231.
Galef, B. G., Jr., & Laland, K. N. (2005). Social learning in animals, empirical studies and theoretical models. *Biosciences*, *55*, 489–499.
Henrich, J., & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology*, *12*, 123–135.
Heyes, C. M. (1994). Social learning in animals: Categories and mechanisms. *Biological Review*, *69*, 207–231.
Kendal, R. L., Coolen, I., van Bergen, Y., & Laland, K. N. (2005). Trade-offs in the adaptive use of social and asocial learning. *Advanced Studies of Behavior*, *35*, 333–379.
Laland, K. N. (2004). Social learning strategies. *Learning & Behavior*, *32*, 4–14.
Raby, C. R., Alexis, D. M., Dickinson, A., & Clayton, N. S. (2007). Planning for the future by western scrub-jays. *Nature*, *445*, 919–921.
Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., et al. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, *328*, 208–213. http://dx.doi.org/10.1126/science.1184719
Rendell, L., Fogarty, L., & Laland, K. N. (2010). Rogers' paradox recast and resolved: Population structure and the evolution of social learning strategies. *Evolution*, *64*, 534–548.
Rendell, L., Boyd, R., Enquist, M., Feldman, M. W., Fogarty, L., & Laland, K. N. (2011). How copying affects the amount, evenness and persistence of cultural knowledge: Insights from the social learning strategies tournament. *Philosophical Transactions of the Royal Society B*, *366*, 1118–1128.
Richter, D. (1966). *Aspects of Learning and Memory*. London: Heinemann.
Rogers, A. (1988). Does biology constrain culture? *American Anthropologist*, *90*, 819–831.
Suddendorf, T., & Busby, J. (2003). Mental time travel in animals? *Trends in Cognitive Sciences*, *7*, 391–396 (http://dx.doi.org/10.1016/S1364-6613(03)00187-6)399, 682-685, (http://dx.doi.org/10.1038/21415)
Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, *123*, 133–167.
Tulving, E. (1983). *Elements of episodic memory*. Oxford, England: Clarendon Press.
Vale, G. L., Flynn, E. G., & Kendal, R.L. Cumulative culture and future thinking: Is mental time travel a prerequisite to cumulative cultural evolution? *Learning and Motivation*, http://dx.doi.org/10.1016/j.lmot.2012.05.010, this issue.
van Bergen, Y., Coolen, I., & Laland, K. N. (2004). Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proceedings of the Royal Society of London B*, *271*, 957–962.