

Network-based diffusion analysis: a new method for detecting social learning

Mathias Franz^{1,*} and Charles L. Nunn^{1,2,3}

¹Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

²Department of Integrative Biology, University of California, Berkeley, CA 94720-3840, USA

³Department of Anthropology, Harvard University, Peabody Museum 11 Divinity Avenue, Cambridge, MA 02138, USA

Social learning has been documented in a wide diversity of animals. In free-living animals, however, it has been difficult to discern whether animals learn socially by observing other group members or asocially by acquiring a new behaviour independently. We addressed this challenge by developing network-based diffusion analysis (NBDA), which analyses the spread of traits through animal groups and takes into account that social network structure directs social learning opportunities. NBDA fits agent-based models of social and asocial learning to the observed data using maximum-likelihood estimation. The underlying learning mechanism can then be identified using model selection based on the Akaike information criterion. We tested our method with artificially created learning data that are based on a real-world co-feeding network of macaques. NBDA is better able to discriminate between social and asocial learning in comparison with diffusion curve analysis, the main method that was previously applied in this context. NBDA thus offers a new, more reliable statistical test of learning mechanisms. In addition, it can be used to address a wide range of questions related to social learning, such as identifying behavioural strategies used by animals when deciding whom to copy.

Keywords: animal cultures; social learning; diffusion of innovations; social network; agent-based model; maximum-likelihood estimation

1. INTRODUCTION

Social learning involves the acquisition of new behaviours from other group members, e.g. by directly observing and copying their behaviour (Heyes 1994). Evidence collected over the past decade suggests that a wide diversity of animals learn socially, including mammals, birds, fishes and invertebrates (Galef & Laland 2005; Leadbeater & Chittka 2007). Potato washing by a group of Japanese macaques (*Macaca fuscata*) provides one possible example of social learning. After a young female macaque started to wash sandy potatoes in a stream, three quarters of her group members acquired this new food processing technique within 9 years. Because the new behaviour spread most quickly between relatives and close associates, Kawai (1965) suggested that the animals imitated each other. However, whether potato washing was learned socially is still debated (Galef 1992, 2004; Lefebvre 1995). More generally, it has been difficult to demonstrate social learning in wild animals, despite much effort and interest in disentangling social and asocial learning mechanisms.

The study of social learning in animals, especially in primates, is often guided by an interest in the evolutionary roots of human culture and cognition (McGrew 1998; Tomasello 1999; Boesch 2003; Laland 2008). Investigating social learning in animals is also of broad relevance for understanding animal behaviour, cognition and evolution. By altering ecological or social environment, social learning dynamics can have a direct impact on

evolutionary dynamics in a process known as cultural niche construction (Laland *et al.* 2000) that can lead to coevolutionary processes between cultures and genes (Feldman & Laland 1996).

Assessing the importance of social learning for behavioural adaptation and evolution depends crucially on being able to distinguish social from asocial learning in free-living animals, as illustrated by the potato-washing example above. However, our ability to pursue this research is limited by the methods that are currently available to investigate learning dynamics in animal societies. Common approaches include laboratory experiments with captive animals and observational studies of variation in behavioural repertoires among populations of wild animals (Galef 2004). However, both approaches have been criticized for the inability to reflect social and ecological conditions in the wild (van Schaik *et al.* 2003), or to reliably identify social learning (Galef 2004; Laland & Janik 2006).

An alternative method to infer learning mechanisms is to investigate the spread (or diffusion) of new behavioural traits through a group of animals. The main method that is used to analyse these dynamics is diffusion curve analysis (DCA). In this approach, the investigator plots the cumulative number of animals with the trait over time and then fits mathematical curves to these data. The basis for this approach is that different learning mechanisms should produce different learning curves (figure 1). For example, simple mathematical models predict that asocial learning results in decelerating diffusion curves. This happens because the number of naive individuals decreases over time. In contrast to this pattern, social

* Author for correspondence (mathias.franz@eva.mpg.de).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2008.1824> or via <http://rsob.royalsocietypublishing.org>.

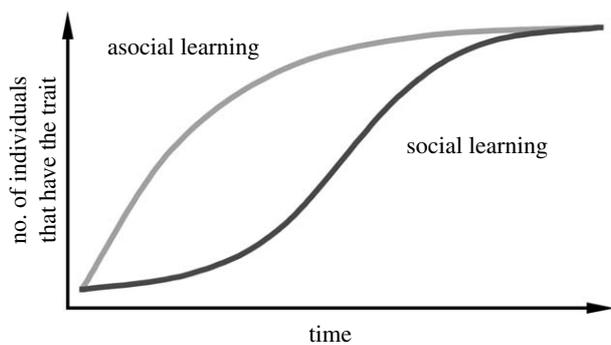


Figure 1. Shapes of diffusion curves consistent with two learning mechanisms. Simple mathematical models, which assume randomly interacting individuals, predict that asocial learning (e.g. by trial and error) produces decelerating curves, and social learning (e.g. by observing other group members) results in sigmoid-shaped curves.

learning causes sigmoid diffusion curves owing to a temporally increasing number of skilled individuals, from which naive individuals can learn. Often it is assumed that any accelerating curve (e.g. exponential or hyperbolic sine) is produced by social learning, and any non-accelerating curve (including linear curves) indicates asocial learning (Reader 2004).

DCA has been criticized heavily for failing to reliably infer learning mechanisms in animals (Reader 2004). A major criticism is that the method assumes that animals interact randomly when they learn from one another. This assumption is frequently violated, as it has been shown that social learning dynamics can be influenced by factors such as age, sex, dominance and kinship (Tanaka 1998; Nicol & Pope 1999; Smith *et al.* 2002). Because these social learning biases can strongly influence the dynamics of trait spread through a group of animals (Voelkl & Noe 2008), the predicted shapes of the diffusion curves might be incorrect (Reader 2004).

(a) *Network-based diffusion analysis*

In light of these concerns, we offer network-based diffusion analysis (NBDA) as an alternative to DCA. NBDA takes into account the order and timing with which individuals in the group acquired the behavioural trait. These data are compared with a social network that contains information about potential social learning opportunities, e.g. in the context of grooming, feeding or simply the time that animals spend in close proximity. NBDA therefore makes use of the fact that socially learned traits will spread most quickly between animals that have strong connections in a social network (Coussi-Korbel & Fragaszy 1995).

To apply NBDA, we developed two agent-based models (ABMs) in which the trait is learned by either pure asocial or pure social learning. The social learning ABM assumes that social learning dynamics are explicitly linked to a social network, while asocial learning occurs probabilistically among individuals without regards to the social network (e.g. through independent trial and error learning). Both models are fit to the observed diffusion data using maximum-likelihood methods. The model that fits better indicates the learning mechanism that is more likely to have produced the observed data.

To test the performance of NBDA relative to the DCA, we generated large numbers of artificially created diffusion data and then applied both methods to these data. The artificial data were created with the ABMs that are fit in the NBDA. In this way, we had perfect knowledge of the underlying learning mechanism for each dataset. Throughout, we used an empirically derived social network (and variations of it) to simulate social learning dynamics. In this way, we assessed the statistical power of the method based on a real-world scenario, and we demonstrate that the method can be applied to the actual data. We also provide the necessary computer script in R (R Development Core Team 2007) for others to implement the method and to extend it for their own purposes (see the electronic supplementary material).

2. MATERIAL AND METHODS

(a) *Description of the ABMs*

In the following we describe two ABMs. Because both models are very similar in their structure, they are described in one framework. The model descriptions are based on the ODD protocol for describing individual- and ABMs (Grimm & Railsback 2005; Grimm *et al.* 2006).

(i) *Purpose*

The purpose of the models is twofold. Firstly, they are used to simulate how a new trait spreads through a group of animals by either asocial or social learning. Secondly, both models form the central element of the NBDA, in which they are fit to observed diffusion data (the pattern of trait spread through a group of animals).

(ii) *State variables and scales*

The main entities in both models are agents that can exist in two states: (i) naive, i.e. the agent did not yet learn the new trait and (ii) skilled, i.e. the agent already learned the new trait. In the first model, each agent possesses an asocial learning rate that is identical for all agents. This learning rate determines the probability that a naive agent will acquire the new trait through asocial learning that happens independently of other agents. In the second model, agents are connected through a user-defined social network. The strengths of the connections in this network determine the social learning rates at which naive agents learn from skilled group members.

(iii) *Process overview and scheduling*

The model dynamics proceed in discrete time steps. In each step, naive agents can learn asocially in the first model and socially in the second model. The simulation stops once all agents become skilled. In the first model, a naive agent will become skilled with a probability that equals its asocial learning rate. In the second model, the probability that a naive agent becomes skilled depends on the parameter τ and the sum s of the strengths of all connection with agents that were already skilled in the last time step. We assume that as s increases, the probability of social learning increases asymptotically to 1. The probability p that a naive individual i acquires the trait is given by

$$p(i) = 1 - \exp(-\tau \cdot s). \quad (2.1)$$

Note that equation (2.1) makes the assumptions that s is linearly related to the rate of learning and that for fixed s this learning rate is constant over time.

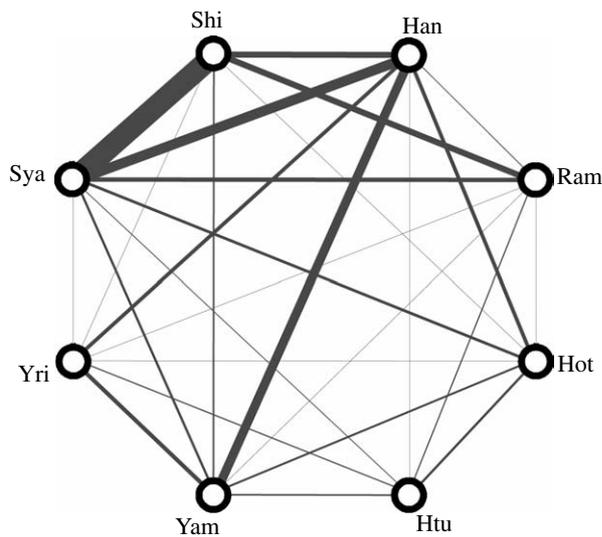


Figure 2. Schematic of the social network that was used as model input to generate artificial diffusion data. Circles represent individuals, and lines indicate social connections between the individuals. Line widths are proportional to the strengths of social connections. This network is based on an empirically observed co-feeding network of eight Japanese macaques (Ventura *et al.* 2006). See §2 for details.

(iv) *Initialization*

Both models are always initialized in the same way, with all but one agent set to be naive, and one agent set to be skilled. This skilled agent is called the inventor.

(b) *Generation of artificial diffusion data*

Artificial diffusion data were generated using both models assuming a group size of eight agents. The asocial learning rate in the first model was set to 0.125. The social network for the second model was based on an empirically observed co-feeding network of a group of eight Japanese macaques (figure 2; Ventura *et al.* 2006). The co-feeding indices calculated by Ventura *et al.* (2006) were obtained by focal animal observations, which resulted in a non-symmetric interaction matrix (i.e. the co-feeding index of dyad A-B is not necessarily the same as for dyad B-A). To use these data as model input we created a symmetric matrix by calculating for each dyad the mean of the two values in the asymmetric matrix. The parameter τ was set to 0.2. (Note that NBDA can also use non-symmetric interaction matrices as input. However, we transformed the network data because we suspected that the asymmetries in the interactions in this example data set were artefacts that emerged from the data collection.)

Artificial diffusion data were created for each model separately. Simulations for one model were performed for eight different initial conditions, each with a different individual as inventor. Simulations were repeated 10 000 times for every initial condition, resulting in 160 000 sets of artificial diffusion data in which the learning mechanism was known (10 000 diffusion datasets \times 8 individuals as inventor \times 2 models).

(c) *Diffusion curve analysis*

This analysis requires as input the cumulative number of skilled individuals for each time step. Nonlinear least-squares fitting was used to fit a decelerating function

$$f(x) = 1 - a/\exp(b \cdot x), \quad (2.2)$$

and a sigmoid function

$$f(x) = 1/(1 + \exp(a - b \cdot x)), \quad (2.3)$$

to the data. Based on the residual sum of squares, likelihoods and the Akaike information criterion (AIC) were calculated for each function (Burnham & Anderson 2002). A smaller AIC value indicates a better fit of the corresponding function, and we assumed that a difference in AIC values of more than two indicates a better fit of the model with the lower AIC value. We also calculated Akaike weights (Burnham & Anderson 2002) for each model, which can be interpreted as the probability that a specific model produced the given data (in comparison with the other model).

(d) *Network-based diffusion analysis*

The required input for this analysis is the number of individuals in the group, the corresponding social network (as a matrix) and the times at which the individuals became skilled. Using maximum likelihood, the unspecified parameters are fit in the two models; these involve the asocial learning rate in the first model and τ in the second model. For the identification of the parameter value that maximizes the log likelihood we used the *optimize* function that is provided in R (R Development Core Team 2007).

The log-likelihood values for a specific model parametrization were directly calculated from the given model parameters and diffusion data without performing any simulations of the model itself. This is possible because the likelihood for a specific model is completely determined by the corresponding probabilities of successful and unsuccessful learning events (i.e. for each time step the probabilities that naive individuals learned or failed to learn). To calculate the overall log likelihood, we calculated the log likelihoods for single events separately (i.e. in each time step for each individual) and then summed them. If it was observed that an agent learned in a specific time step, then the corresponding log likelihood is given by the natural logarithm of the related learning probability. In the first model, this probability is given directly by the asocial learning rate. The learning probabilities in the second model had to be calculated separately for each time step because they depend on the value of τ and on the number and identity of individuals that were already skilled in the previous time step (as described in §2(a)(iii)). If an agent did not learn, then the log likelihood equals the natural logarithm of one minus the corresponding learning probability. The log likelihood is zero if the individual was already skilled in the previous time step.

Based on the log likelihoods, we calculated AIC values for each model (Burnham & Anderson 2002). As for DCA, we assumed that one model fits better if the difference in AIC values is larger than two (for an analysis with a different assumption see §3(d)), and we calculated Akaike weights for each model.

3. RESULTS

(a) *Simulations of social learning*

For artificial diffusion data created by pure social learning, DCA was able to correctly infer social learning in 66.2 per cent of all analyses. In 26.2 per cent of the tests, however, DCA erroneously inferred asocial learning, with the remaining 7.6 per cent of cases being undecided because the differences in the AIC values were smaller than two. By contrast, the NBDA had much higher power

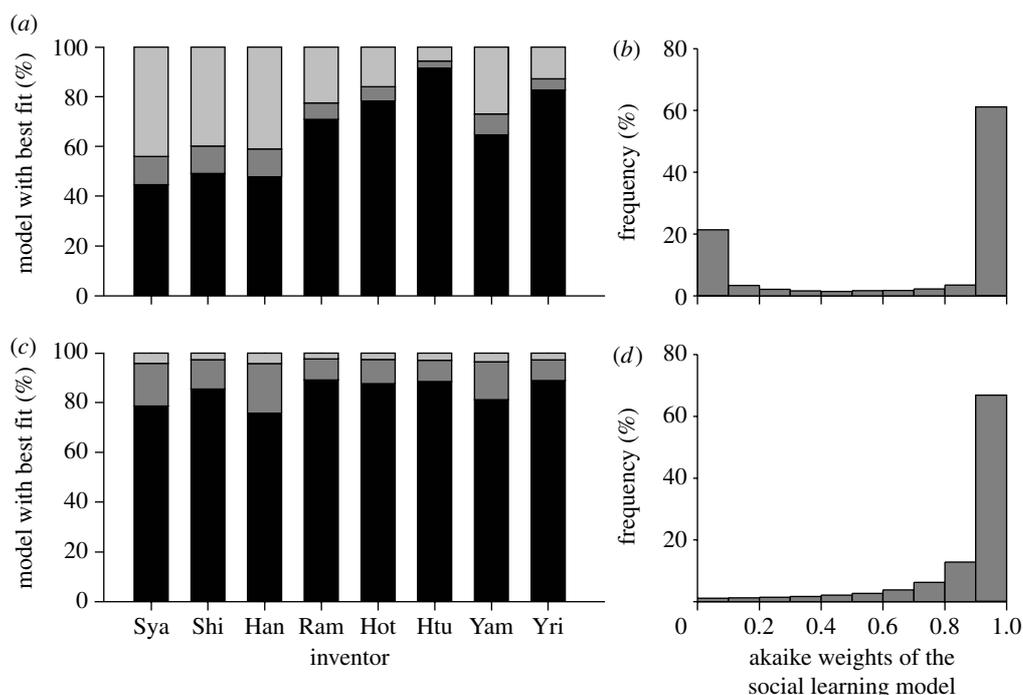


Figure 3. Results for artificial diffusion data that were created by pure social learning. The analysed data consisted of 80 000 sets of artificial diffusion data, 10 000 for each inventor. (a,b) Results for DCA; and (c,d) results for NBDA. (a,c) The proportions of analyses in which DCA and NBDA inferred social learning (black) and asocial learning (light grey) are shown for each inventor separately. Inferences of the learning mechanisms were based on the differences in AIC values of the two alternative models, which were fit to the diffusion data. We assumed that a difference of at least two indicates strong support for one model. The ‘either’ (dark grey) category contains cases where neither model was supported on this basis. (b,d) The distribution of the Akaike weights for the social learning model was calculated from the whole data for each method. The social network, which was used for creating artificial data and for NBDA, is shown in [figure 2](#).

for detecting social learning. This method correctly inferred social learning in 84.3 per cent of the datasets and asocial learning in only 3.3 per cent of the datasets.

Taking a closer look at these results revealed that the performance of DCA depended strongly on the identity of the inventor ([figure 3a](#)). Specifically, DCA was more likely to erroneously classify the learning mechanism when the inventor belonged to a subgroup of very strongly connected individuals (i.e. individuals Sya, Shi and Han in [figure 2](#)). This makes intuitive sense. If an individual has strong connections with few individuals, but these individuals have mostly weak connections with the remaining group members, then the new trait is likely to first spread very quickly in the strongly connected subgroup. The weakness of connections outside the subgroup produces lower learning probabilities, which slows the subsequent spread of the trait and favours fitting of decelerating curves (i.e. evidence for asocial learning if DCA is used). By contrast, the NBDA takes the network structure into account and thus had much higher power to correctly infer social learning.

(b) Simulations of asocial learning

Results for diffusion data produced by asocial learning revealed similar differences between the two methods. Again, the DCA performed poorly. Asocial learning was correctly inferred in only 57.6 per cent of the simulated datasets, while social learning was incorrectly inferred in 30.2 per cent. NBDA provided stronger inferences of the learning mechanism, as this method found evidence for asocial learning in 80.1 per cent of simulated datasets, and for social learning in only 6.4 per cent.

In these analyses, the DCA was less dependent on the identity of the inventor ([figure 4a](#)). This outcome was expected because asocial learning rates were identical for all agents and they were not influenced by other group members. The reason why social learning was erroneously inferred in approximately 30 per cent of the tests reflects stochastic effects in small groups, which cause asocial learning to ‘accidentally’ produce sigmoid-shaped curves. Again, the NBDA has an advantage in these cases because this method only finds strong support for social learning if the spread of the new trait corresponds to the structure of the social network.

(c) Akaike weights

Analyses of Akaike weights revealed further problems with DCA and advantages of NBDA. The results for the DCA showed bimodal distributions in the simulation of social and asocial learning ([figures 3b](#) and [4b](#)). In most cases, the model with the correct learning mechanism had very high Akaike weights, which indicate strong support for this model. In a considerable proportion of all analyses, however, the model with the correct learning mechanism was supported very poorly, thus indicating that the model with the incorrect learning mechanism was supported very strongly. This shows that the identification of the wrong learning mechanism is an inherent problem of the DCA and cannot be solved by performing more conservative analyses (e.g. by setting a higher threshold for the difference in AIC values when judging one model relative to the other; the results of a corresponding analysis are shown in ‘Additional analyses.doc’ in the electronic supplementary material).

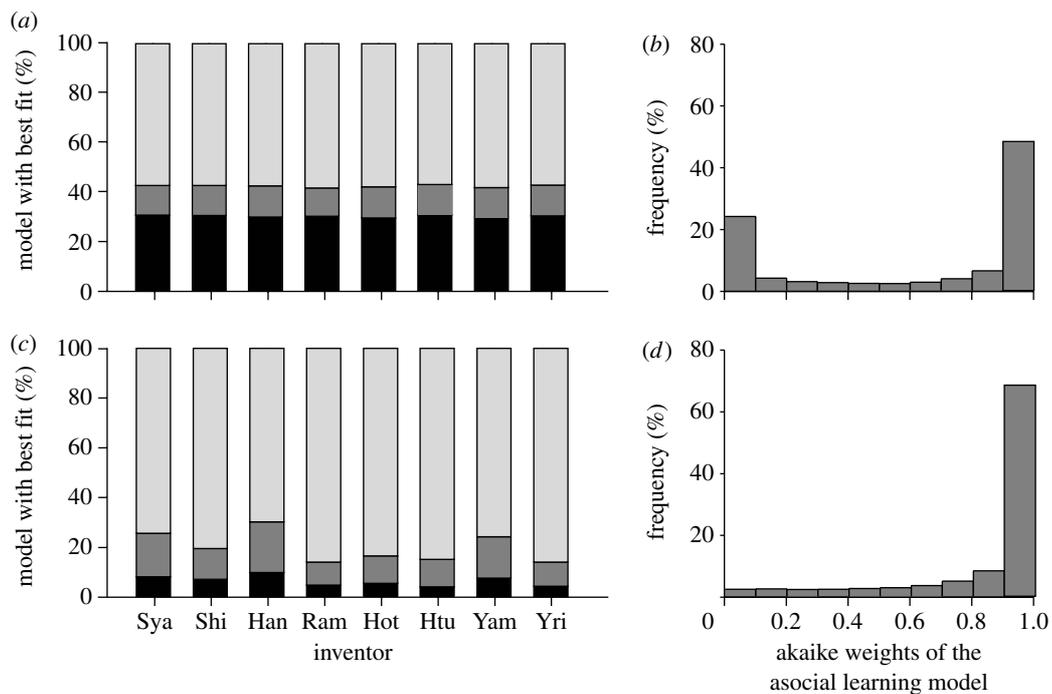


Figure 4. Results for artificial diffusion data that were created by pure asocial learning. The analysed data consisted of 80 000 sets of artificial diffusion data, 10 000 for each inventor. (a,b) Results for DCA; and (c,d) results for NBDA. (a,c) The proportions of analyses in which DCA and NBDA inferred social learning (black) and asocial learning (light grey) are shown for each inventor separately. Inferences of the learning mechanisms were based on the differences in AIC values of the two alternative models, which were fit to the diffusion data. We assumed that a difference of at least two indicates strong support for one model. The ‘either’ (dark grey) category contains cases where neither model was supported on this basis. (b,d) The distribution of the Akaike weights for the asocial learning model was calculated from the whole data for each method. The social network, which was used for NBDA, is shown in [figure 2](#).

By contrast, we found unimodal distributions of Akaike weights for the NBDA for simulations of social and asocial learning, and in most cases, the model with the correct learning mechanism was very well supported ([figures 3d](#) and [4d](#)). This shows that the better performance of the NBDA is a robust result.

(d) Additional analyses

The electronic supplementary materials provide additional analyses to test the performance of DCA and NBDA under the conditions of (i) varied structure of the social network, (ii) different group sizes, and (iii) changing the threshold in AIC differences, which is used in model selection. Our results showed that the performance of NBDA and DCA is not strongly impacted by variation in network structure. By contrast, an increase in group size led to an improvement in the performance of NBDA and DCA. Increasing the threshold in AIC differences from two to four increased the cases in which DCA and NBDA reveal inconclusive results (i.e. that none of both models fits better). While this effect strongly reduced the frequency with which NBDA identifies the wrong learning mechanism, the performance of DCA was only weakly affected and the frequency with which DCA identified the wrong learning mechanism remains high.

The electronic supplementary material also includes an analysis of the performance of NBDA when using a network with disturbed structure. This analysis showed that the results obtained by NBDA are robust to small disturbances in the structure of the social network, which for instance might be caused by observation errors.

Finally, the electronic supplementary material includes an extended version of NBDA, in which a model of social and asocial learning is fitted to the diffusion data. The performance of this method in detecting social learning was similar to that described above for NBDA.

4. DISCUSSION

Our results show that heterogeneous social networks can alter the shape of the diffusion curve, and this can lead to erroneous assessments of the learning mechanism in DCA ([Reader 2004](#)). In addition, the simulations revealed that stochastic effects in small groups can alter the shape of the diffusion curve, thus explaining why the DCA so often failed to correctly infer asocial learning in our tests. By using additional information about the social network and the order in which individuals acquired the behavioural trait, NBDA overcomes these problems. Thus, NBDA provides greater power to infer social learning. NBDA is also more conservative than DCA because it is much less susceptible to incorrectly inferring social learning, when only asocial learning processes were occurring. Furthermore, our additional analyses showed that NBDA is robust, including to errors in estimating network structure and alternative methods of analysis.

The strength of NBDA lies in its integration of different kinds of information about the diffusion process and the social network into a comprehensive analysis that is based on an information-theoretic framework ([Burnham & Anderson 2002](#)). In contrast to NBDA, most other approaches that incorporated information on social

interactions to identify social learning used only limited information about the spread of the new trait to infer learning mechanisms. For instance, [Perry *et al.* \(2003\)](#) and [Bonnie & de Waal \(2006\)](#) focused solely on the outcome of a diffusion process by comparing the social network structure with the distribution of a dyadic-interaction behaviour. While [Perry *et al.* \(2003\)](#) performed a simple test to infer whether dyads that performed a new behaviour also have stronger connections in a social network, [Bonnie & de Waal \(2006\)](#) conducted a more complex analysis in which they tested whether the match between the social network and the distribution of the behaviour also could have emerged by chance. [Boogert *et al.* \(2008\)](#) and [Morrell *et al.* \(2008\)](#) applied conceptually similar approaches to [Bonnie & de Waal \(2006\)](#), but they used more information about the diffusion process by integrating information on the order in which a new foraging behaviour spreads through a group. The timing of learning events was, however, neglected in these analyses.

Other than NBDA, the approach of [Kendal *et al.* \(2007\)](#) is the only one that we are aware of that incorporates information about the order and the timing of learning events. Some elements of their approach are similar to NBDA, as they also estimate parameters that describe social and asocial learning processes from the observed diffusion data. The main advantages of NBDA to the approach of [Kendal *et al.* \(2007\)](#) are that we used maximum likelihood to parametrize our models, and it is easier to collect the required information to implement NBDA, which is especially important for studies on wild animals.

(a) *Applying network-based diffusion analysis*

Network-based diffusion analysis offers the identification of social learning in groups of captive animals as well as in wild populations. In contrast to the experiments with captive animals, in field studies it is likely to be difficult to ensure high-quality data that accurately record the first expression of the behaviour by individuals, especially if the spreading behaviour is practiced only rarely. We expect that poor data quality will mainly impact the power to detect social learning, because the fit of the corresponding model should be most sensitive to erroneous information on the time of learning events or the social network.

Here, we focused on only one empirically observed social network. This approach is used as a proof of concept for NBDA, and to explore the effects of heterogeneous networks on the performance of DCA. In an additional analysis, we have shown that our findings are robust to variation in the structure of this network. Nevertheless, it is still possible that the network size and topology have an effect on the performance of one or both methods. As demonstrated by our analyses, the ABMs can be used to investigate the expected performance of the methods in a specific group of animals (i.e. of a particular size or with a different social network). Even in the case when data are unavailable on the social network for a specific group, social networks from other groups of the same species might be used to estimate the method's performance (although of course it would always be preferable to use networks derived from the group of interest). For example, our results indicate that DCA should not be used for

studies of Japanese macaques. Results of such analyses that were already conducted for this species (e.g. [Galef 1992](#); [Lefebvre 1995](#)) should thus be treated cautiously.

(b) *Extensions of network-based diffusion analysis*

In our description of the NBDA, we assumed a static group whose composition does not change during the spread of a new trait. This assumption might not always be well justified because diffusion processes can take multiple years to complete, as shown by the example of potato washing in macaques ([Kawai 1965](#)). In these cases it is possible that births, deaths and migrations alter the number of skilled and naive individuals, which might crucially influence the dynamics of the diffusion process. In contrast to DCA, NBDA can easily control for demographic changes by incorporating these changes in the corresponding ABMs. To implement this extension, the input for the described models would simply require time series of group sizes and social networks.

Similarly, it has been argued that individual differences in asocial learning abilities can strongly impact diffusion dynamics, and under such conditions S-shaped diffusion curves can emerge from pure asocial learning ([Laland & Kendal 2003](#); [Reader 2004](#)). Although this is not necessarily true ([Henrich 2001](#)), our models offer the possibility to include individual variations, not only for the ability to learn asocially but also for social learning.

The ABMs that are used in the NBDA allow only one learning mechanism to take place. We extended this approach to allow both social and asocial learning to occur simultaneously (in the electronic supplementary material we provide a performance analysis of this extension and an R script to implement it). Incorporating such a model in NBDA provides an opportunity to test whether both learning mechanisms influenced the spread of a new trait. Furthermore, models that are restricted to a single learning mechanism are nested in the model that includes both mechanisms. This structure allows for likelihood ratio testing, which could be used to infer learning mechanisms based on *p*-values. Furthermore, a model that combines social and asocial learning is needed to analyse the data that include time periods prior to the first occurrence of a new trait (such data can be analysed using the extended version of NBDA that is in the electronic supplementary material). Using such additional information will increase the power to detect social learning. However, fitting a pure social learning model would not be reasonable in this case, because at least the first individual that acquired the trait must have learned asocially.

In our social learning model, we assumed that the probability of learning from a skilled individual was influenced only by spatial proximity of skilled and unskilled individuals during feeding. While this assumption seems to be reasonable, real-world scenarios might be more complex, because individual characteristics such as age, kinship or rank of either individual might influence learning probabilities ([Boyd & Richerson 1985](#); [Coussi-Korbel & Fragaszy 1995](#)). Although such influences are assumed to be likely because they should increase adaptability of social learning, only a few attempts have been made to test these hypotheses with empirical data (see [Laland \(2004\)](#) for a review). NBDA offers a framework to identify which social and individual

characteristics impact social learning processes. This could be achieved by identifying different networks that represent dyadic relationships according to the specific social interactions or individual characteristics, such as time spent grooming or age differences. Functional relationships between the values of dyadic relations and social learning probabilities can be formulated similar to the example used in this paper (see equation (2.1)). Fitting alternative models, in which social learning probabilities are determined by different networks or combinations of networks, could be used to identify which factors most likely influenced social learning dynamics in the observed diffusion of a new trait. Of course, it is also possible to extend the asocial learning model in a similar way to test for effects of individual characteristics such as age or rank on asocial learning rates.

NBDA is a flexible tool that allows more extensions than those described above. Additional influences or alternative assumptions about the dynamics of social and asocial learning can be included in NBDA, provided that they allow unambiguous calculation of learning probabilities for each individual in each time step. However, it should be noted that model selection based on AIC values favours models with fewer parameters if the alternative models explain the observed data equally well. The potential that a more complex model explains the data better than a simple model, in general, increases with increasing amount of observed data (i.e. number of individuals that have acquired the behaviour). Therefore, identifying complex learning strategies that require more parameter-rich models will require relatively large sample sizes.

We thank Richard McElreath, Peter Richerson, Mark Lubell and other members of the 'culture group' at UC Davis, Luke Matthews, Richard Wrangham and other members of the 'primate group' at Harvard University and two anonymous reviewers for their helpful suggestions and discussion. We also thank Damien Caillaud and Roger Mundry for valuable discussion of the maximum-likelihood method. This research was supported by the Max Planck Society.

REFERENCES

- Boesch, C. 2003 Is culture a golden barrier between human and chimpanzee? *Evol. Anthropol.* **12**, 82–91. (doi:10.1002/evan.10106)
- Bonnie, K. E. & de Waal, F. B. M. 2006 Affiliation promotes the transmission of a social custom: handclasp grooming among captive chimpanzees. *Primates* **47**, 27–34. (doi:10.1007/s10329-005-0141-0)
- Boogert, N. J., Reader, S. M., Hoppitt, W. & Laland, K. N. 2008 The origin and spread of innovations in starlings. *Anim. Behav.* **75**, 1509–1518. (doi:10.1016/j.anbehav.2007.09.033)
- Boyd, R. & Richerson, P. J. 1985 *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Burnham, K. P. & Anderson, D. R. 2002 *Model selection and multimodel inference: a practical information-theoretic approach*. New York, NY: Springer.
- Coussi-Korbel, S. & Fragaszy, D. M. 1995 On the relation between social dynamics and social learning. *Anim. Behav.* **50**, 1441–1453. (doi:10.1016/0003-3472(95)80001-8)
- Feldman, M. W. & Laland, K. N. 1996 Gene-culture coevolutionary theory. *Trends Ecol. Evol.* **11**, 453–457. (doi:10.1016/0169-5347(96)10052-5)
- Galef, B. G. 1992 The question of animal culture. *Hum. Nat.* **3**, 157–178. (doi:10.1007/BF02692251)
- Galef, B. G. 2004 Approaches to the study of traditional behaviors of free-living animals. *Learn. Behav.* **32**, 53–61.
- Galef, B. G. & Laland, K. N. 2005 Social learning in animals: empirical studies and theoretical models. *Bioscience* **55**, 489–499. (doi:10.1641/0006-3568(2005)055[0489:SLIAES]2.0.CO;2)
- Grimm, V. & Railsback, S. F. 2005 *Individual-based modeling and ecology*. Princeton, NJ: Princeton University Press.
- Grimm, V. *et al.* 2006 A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* **198**, 115–126. (doi:10.1016/j.ecolmodel.2006.04.023)
- Henrich, J. 2001 Cultural transmission and the diffusion of innovations: adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change. *Am. Anthropol.* **103**, 992–1013. (doi:10.1525/aa.2001.103.4.992)
- Heyes, C. M. 1994 Social-learning in animals—categories and mechanisms. *Biol. Rev. Camb. Philos. Soc.* **69**, 207–231. (doi:10.1111/j.1469-185X.1994.tb01506.x)
- Kawai, M. 1965 Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima Islet. *Primates* **6**, 1–30. (doi:10.1007/BF01794457)
- Kendal, J. R., Kendal, R. L. & Laland, K. N. 2007 Quantifying and modelling social learning processes in monkey populations. *Int. J. Psychol. Psychol. Ther.* **7**, 16.
- Laland, K. N. 2004 Social learning strategies. *Learn. Behav.* **32**, 4–14.
- Laland, K. N. 2008 Animal cultures. *Curr. Biol.* **18**, R366–R370. (doi:10.1016/j.cub.2008.02.049)
- Laland, K. N. & Kendal, J. R. 2003 What the models say about social learning. In *The biology of traditions: models and evidence* (eds D. M. Fragaszy & S. Perry), pp. 33–55. Cambridge, UK: Cambridge University Press.
- Laland, K. N. & Janik, V. M. 2006 The animal cultures debate. *Trends Ecol. Evol.* **21**, 542–547. (doi:10.1016/j.tree.2006.06.005)
- Laland, K. N., Odling-Smee, J. & Feldman, M. W. 2000 Niche construction, biological evolution, and cultural change. *Behav. Brain Sci.* **23**, 131–175. (doi:10.1017/S0140525X00002417)
- Leadbeater, E. & Chittka, L. 2007 Social learning in insects—from miniature brains to consensus building. *Curr. Biol.* **17**, R703–R713. (doi:10.1016/j.cub.2007.06.012)
- Lefebvre, L. 1995 Culturally-transmitted feeding-behavior in primates—evidence for accelerating learning rates. *Primates* **36**, 227–239. (doi:10.1007/BF02381348)
- McGrew, W. C. 1998 Culture in nonhuman primates? *Annu. Rev. Anthropol.* **27**, 301–328. (doi:10.1146/annurev.anthro.27.1.301)
- Morrell, L. J., Croft, D. P., Dyer, J. R. G., Chapman, B. B., Kelley, J. L., Laland, K. N. & Krause, J. 2008 Association patterns and foraging behaviour in natural and artificial guppy shoals. *Anim. Behav.* **76**, 855–864. (doi:10.1016/j.anbehav.2008.02.015)
- Nicol, C. J. & Pope, S. J. 1999 The effects of demonstrator social status and prior foraging success on social learning in laying hens. *Anim. Behav.* **57**, 163–171. (doi:10.1006/anbe.1998.0920)
- Perry, S. *et al.* 2003 Social conventions in wild white-faced capuchin monkeys—evidence for traditions in a neotropical primate. *Curr. Anthropol.* **44**, 241–268. (doi:10.1086/345825)
- R Development Core Team, 2007 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reader, S. M. 2004 Distinguishing social and asocial learning using diffusion dynamics. *Learn. Behav.* **32**, 90–104.

- Smith, V. A., King, A. P. & West, M. J. 2002 The context of social learning: association patterns in a captive flock of brown-headed cowbirds. *Anim. Behav.* **63**, 23–35. (doi:10.1006/anbe.2001.1886)
- Tanaka, I. 1998 Social diffusion of modified louse egg-handling techniques during grooming in free-ranging Japanese macaques. *Anim. Behav.* **56**, 1229–1236. (doi:10.1006/anbe.1998.0891)
- Tomasello, M. 1999 *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., Suzuki, A., Utami, S. S. & Merrill, M. 2003 Orangutan cultures and the evolution of material culture. *Science* **299**, 102–105. (doi:10.1126/science.1078004)
- Ventura, R., Majolo, B., Koyama, N. F., Hardie, S. & Schino, G. 2006 Reciprocation and interchange in wild Japanese macaques: grooming, cofeeding, and agonistic support. *Am. J. Primatol.* **68**, 1138–1149. (doi:10.1002/ajp.20314)
- Voelkl, B. & Noe, R. 2008 The influence of social structure on the propagation of social information in artificial primate groups: a graph-based simulation approach. *J. Theor. Biol.* **252**, 77–86. (doi:10.1016/j.jtbi.2008.02.002)