DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Investigating the Impact of Observation Errors on the Statistical Performance of Network-based Diffusion Analysis

(Article begins on next page)

Investigating the impact of observation errors on the statistical performance of network-based diffusion analysis

Mathias Franz[1,2,3,*] and Charles L. Nunn[4]

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

[2] Integrative Primate Socio-Ecology Group, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

[3] Courant Research Center Evolution of Social Behavior, University of Göttingen, Kellnerweg 6, 37077 Göttingen, Germany

[4] Department of Human Evolutionary Biology, Harvard University, Peabody Museum 11 Divinity Avenue, Cambridge MA 02138, USA

* corresponding author's email address: mathias.franz@zentr.uni-goettingen.de

Suggested running head: The impact of observation errors on NBDA

Abstract

Experiments in captivity have provided evidence for social learning, yet it remains challenging to demonstrate social learning in the wild. Recently, we developed network-based diffusion analysis (NBDA) as a new approach to infer social learning from observational data. NBDA fits alternative models of asocial and social learning to the diffusion of a behavior through time, where the potential for social learning is related to a social network. Here, we investigate the performance of NBDA in relation to variation in group size, network heterogeneity, observer sampling errors and duration of trait diffusion. We find that observation errors, when severe enough, can lead to increased Type I error rates to detect social learning. However, elevated Type I error rates can be prevented by coding the observed times of trait acquisition into larger time units. Collectively, our results provide further guidance to applying NBDA and demonstrate that the method is more robust to sampling error than initially expected.

Key words

Introduction

In many animal species, individuals learn socially by observing the behavior of other individuals. While sophisticated experiments on captive animals have identified different learning mechanisms in animals (Galef & Giraldeau, 2001; Hoppitt & Laland, 2008), inferring the existence of social learning in wild animals remains an important and challenging task. Abandonment of the highly controlled experimental settings of captive studies is needed to investigate how social and ecological conditions in wild animals affect social learning dynamics and the emergence of traditions, and more generally for understanding the evolution of social learning and culture. However, the lack of experimental control also introduces new methodological problems. Important methods of inferring social learning in wild animals, such as the "ethnographic method" (Perry & Manson, 2003; Rendell & Whitehead, 2001; van Schaik, et al., 2003; Whiten, et al., 1999) and diffusion curve analysis (Reader, 2004), can have low statistical power to detect social learning and produce a high rate of false positives (Franz & Nunn, 2009; Galef, 2004; Laland & Galef, 2009; Laland & Hoppitt, 2003; Laland & Janik, 2006; Laland & Kendal, 2003; Reader, 2004). Thus, new methods are needed to investigate social learning.

Recently, we developed network-based diffusion analysis (NBDA) to overcome the limitations of previous approaches (Franz & Nunn, 2009). NBDA makes the reasonable assumption that social learning is more likely to take place among conspecifics that are relatively more closely linked in a social network (Coussi-Korbel & Fragaszy 1995). In the case of a food processing technique that is transmitted socially, for example, we expect that food related behaviors spread most quickly among individuals that often feed together, and thus have strong connections in a co-feeding network (for an example, see Figure 1). In the case of social learning,

but not asocial learning, we expect that the structure of a social network (i.e. the strength of connections among individuals) influences how a novel behavior spreads through a group of animals. NBDA aims to identify whether such an influence occurred in the observed diffusion of a novel behavior. For this purpose, alternative agent-based models of social and asocial learning are fitted to the observed diffusion of a novel behavior. These models provide a way to estimate the probabilities with which each individual learns through social or asocial learning during the different stages of a diffusion. By comparing these probabilities to the actual learning events, one can assess which learning mechanism was most likely to have caused the observed diffusion.

The statistical analysis of NBDA is based on maximum likelihood model fitting. The original NBDA version developed by Franz and Nunn (2009) is based on two alternative models: (1) pure asocial learning, in which each individual acquires the new behavior independently of others, and (2) a pure social learning model in which each individual (except the "inventor") acquires the new behavior by social learning. In the asocial learning model, learning is assumed to occur with a constant probability in each time step. In the social learning model, the probability of learning from other group members is assumed to be determined by the strength of the connections to others in a social network. Thus, the likelihood that an individual follows another in the diffusion of the novel behavior is dependent on the strength of its relationship to the knowledgeable individual. The identification of the best fitting model is based on the Akaike Information Criterion (Burnham & Anderson, 2002).

NBDA thus uses a flexible statistical framework that compares the fit of different *a priori* models (Bolker, 2008; Burnham & Anderson, 2002). This flexibility is important because it provides a powerful platform to investigate more complex

4

learning processes than those included in the simple NBDA version of Franz and

Nunn (2009). Thus, different versions have been developed that include "mixed"

learning models with both social and asocial learning (Franz & Nunn, 2009; Hoppitt,

Boogert, & Laland, 2010; Hoppitt, Kandler, Kendal, & Laland, this issue), models in

which individual specific variables, such as dominance rank or neophobia, can impact

asocial learning dynamics (Hoppitt, et al., 2010) and models that include task

structure, i.e. learned behaviors that consist of multiple steps (Hoppitt, et al., this

issue).

In addition, Hoppitt, Boogert and Laland (2010) developed a variation of

NBDA in which model fitting only focuses on the order, rather than actual timing, in

which individuals acquired the new behavior. They refer to this method as *order of*

acquisition diffusion analysis (OADA), and to methods that use explicit information

on the timing of events as *time of acquisition diffusion analysis* (TADA). Included

under TADA is the original method of Franz and Nunn (2009) and extended versions

developed by Hoppitt et al. (2010). Focusing on the order of diffusion and ignoring

the exact timing of learning events is a fundamental difference to the NBDA method

developed by Franz and Nunn (2009). However, OADA is also a network-based

method that analyses diffusion dynamics by fitting alternative models of social and

asocial learning (in which social learning depends on network structure). Therefore,

Hoppitt, Boogert and Laland proposed using NBDA as an umbrella term for network-

based analyses of trait diffusion. In the following we will adopt this classification of

NBDA methods and therefore also refer to the methods developed by ourselves

(Franz and Nunn, 2009) as TADA.

Basic analysis of NBDA performance has revealed that the power to detect

social learning increases with increasing group size (Franz and Nunn, 2009, Hoppitt et

5

al., in press) and length of diffusion (Hoppitt, et al., 2010). Heterogeneity in weights of network edges seems to have a stronger impact on the performance of OADA compared to TADA (Hoppitt, et al., 2010). Furthermore, Hoppitt et al. (2010) showed that violating assumptions of specific NBDA implementations, such as ignoring effects of individual-level variables (e.g. dominance rank) on asocial learning, can lead to decreased power to detect social learning and increased Type I error rates.

These analyses give a first impression of the factors that impact the probability to correctly or erroneously infer social learning using NBDA. To effectively apply NBDA methods, however, we need a richer understanding of how basic factors, such as group size and network features, influence statistical performance, including interactions among the factors. Furthermore, we have very limited knowledge about the effects of observation errors on the statistical performance of NBDA. Observation errors can potentially increase Type I error rates and decrease power, resulting in Type II errors. Trait diffusion in wild animals is rarely (if ever) known with absolute certainty, especially when the trait is performed rarely and when observations do not cover the entire activity period of the group. Investigating the effects of observation errors thus is of particular importance for applying NBDA methods to wild animals, where unfettered observations of each individual across a given time period are typically impossible.

In this study, we addressed these problems by significantly extending our previous analyses of the TADA method that we (Franz and Nunn, 2009) referred to as "extended NBDA." This method fits a pure asocial learning model and a mixed learning model with both social and asocial learning. Specifically, we investigated the effects of group size, network heterogeneity, duration of trait diffusion and sampling errors on times of trait acquisition for each individual. We also investigated whether

6

re-coding these observations into discrete time steps (Figure 2) can offset negative effects of sampling errors. To systematically investigate these effects, we simulated artificial diffusion data and the "observation" and "coding" of these data using extensions of the agent-based models described by Franz and Nunn (2009). To simulate errors in observing trait diffusion and the analysis of diffusion data (including different coding regimes for diffusion times), we implemented a virtual "observer" and "analyst."

Methods

One part of our analyses focuses on assessing Type I error rates, i.e. the probability to erroneously infer social learning when only asocial learning caused the spread of a novel behavior. For this purpose, we simulated learning dynamics with pure asocial learning and then applied the "extended NBDA" method of Franz and Nunn (2009) to the diffusion data that is generated by these simulations. The other part of our analyses focused on assessing the statistical power to detect social learning when it actually occurred. Although we expect that in real animals social learning usually co-occurs with asocial learning, we performed simulations with pure social learning. By focusing on this extreme case our results reveal the upper limit of the power to detect social learning.

We used the "extended NBDA" method of Franz and Nunn (2009) because we aimed to assess the performance of a method that can be applied to wild animals for which we expect that a new trait spreads through a mixture of individual and social learning instead of pure social learning. The "extended NBDA" method is more suitable in this case because it fits two models: one involving pure asocial learning and another in which both social and asocial learning take place.

7

To generate artificial diffusion data we used the social and asocial learning agent based models (ABMs) described by Franz and Nunn (2009). These models assume that initially one individual (the inventor) is already skilled, i.e. it has already learned the new behavior. All other individuals are assumed to be naïve, i.e. they have not yet learned the new behavior. In the asocial learning model each naïve individual can acquire the new trait with a fixed probability in every time step independently of others in the simulated population. In the social learning model, naïve individuals learn from skilled individuals with whom they have social contact. Specifically, learning probabilities depend on the strength of social network connections to skilled group members and a parameter τ, which determines how connection strengths are translated into social learning probabilities. Increasing the value of τ generally results in higher probabilities of social learning among connected individuals and thus shorter diffusion durations (see Franz and Nunn 2009). Therefore, by varying this parameter we were able to investigate the effects of diffusion duration on statistical performance of NBDA.

We extended these models by explicitly simulating an *observer* and an *analyst*. The observer records the first performance of the new trait for each individual. By using an observer, we were able to simulate observation and sampling errors by varying the probability that the observer identifies when an individual first performed the new trait. For this purpose we assumed that in each time step the observer has a fixed probability of detecting a performed behavior. Thus, the observer can detect the behavior when it first occurred or some time thereafter (e.g., Figure 2). The analyst takes the data from the observer and codes the recorded times of acquisition into time units that might be larger than the time units that were used by

8

the observer (i.e. the simulated time steps, see Figure 2). These "coded" data were

then used as input to TADA.

NBDA methods require as input a social network that is assumed to reflect

social learning opportunities for the observed trait. Here we assumed that the analyst

has a perfect knowledge about the social network. Therefore, the social network used

to simulate social learning dynamics was also used as input to TADA.

We used Latin hypercube sampling to investigate how multiple parameters

influence the performance of TADA (Rushton, Lurz, Gurnell, & Fuller, 2000;

Seaholm, Ackerman, & Wu, 1988). Latin hypercube sampling is a type of stratified

Monte Carlo sampling that has been used in epidemiological modeling and more

efficiently explores parameter space than random sampling procedures. We varied (1)

group size, (2) network heterogeneity, (3) learning speed, (4) probability that the

observer records a performed behavior and (5) the length of time units that the analyst

uses to code the observed data (Figure 2). Table 1 gives ranges in which these

parameters were varied. We used Latin hypercube sampling to generate 1,000,000

parameter sets separately for social and asocial learning simulations. For each of these

parameter sets we simulated a single diffusion, in which a randomly chosen individual

was initially set to be the skilled "inventor".

To generate social networks with different degrees of heterogeneity we draw

all edge weights (or connection strengths between individuals) from differently

skewed distributions. The weight w for an edge between two individuals i and j was

calculated based on a random number r from the interval between 0 and 1 and a

heterogeneity coefficient h:

9

$$w_{i,j} = \frac{r^h}{\frac{1}{N}\sum_{all\ i,j} w_{i,j}}$$

where *N* is the number of individuals in a group. Setting *h* to zero results in a homogeneous network. A value of one results in edge weights that are uniformly distributed. Values larger than one result in more skewed distributions with many weak and few strong edges. Dividing by the mean sum of edge weights from one individual to all others ensures that diffusion duration is largely independent of group size and network heterogeneity. In real animals, group size and network heterogeneity might have an effect on the duration of diffusion, which for instance could lead to an indirect influence of group size on performance of TADA. In our analysis we tried to eliminate such indirect effects because we wanted to infer the direct effects of group size, network heterogeneity and diffusion duration on TADA performance.

To analyze the performance of TADA, we coded results in which the method inferred social learning as a one and otherwise as zero. We then used these codes as the dependent variable in the statistical analysis, focusing in particular on the probability of detecting social learning when data were generated under social learning (a correct inference) or under asocial learning (an incorrect inference). In addition, we calculated the duration of the observed diffusion after the analyst coded the observed data into new time units. To identify the important predictor variables and possible interactions among them we performed classification tree analysis (De'ath & Fabricius, 2000) using the package "tree" in the statistical software R (R Development Core Team, 2007). Classification tree analysis hierarchically splits the data set into subcategories in a way that minimizes the variance of the response variable in the emerging subsets. Graphical output based on this analysis can reveal

and help to visualize complex interactions among predictor variables. In our analysis we used the five predictor variables: group size, degree of network heterogeneity (*h*), diffusion duration (measured in the number of coded time units, see also Figure 2), length of coded time unit (measured in simulated time steps, see also Figure 2), and observation probability per coded time unit (which reflects the probability that the observer records the behavior performed by an individual in one time unit that was coded by the analyst.) In the electronic supplementary materials we provide an R script with the performed simulations and analyses.

Results

We first examined the results involving *pure asocial learning*. We found that the probability to incorrectly infer social learning depended mainly on two factors: observation probability and diffusion duration. Note that observation probability and diffusion duration are measured in coded time units used by the analyst (rather than simulated time steps). Group size, network heterogeneity and length of coded time unit had no major influence on the probability of erroneously detecting social learning.

The classification tree is shown in Figure 3. At the base of this tree is the observation probability, which illustrates the importance of observation errors. In general higher observation probabilities resulted in lower Type I error rates. While observation probabilities lower than 0.37 resulted in 18 percent probability to incorrectly infer social learning (leaf A), observation probabilities between 0.37 and 0.75 led to an incorrect inference of social learning in only six percent of simulations (leaf B). When observation probabilities were at least 0.75 incorrect inferences of social learning were predicted to occur in no more than two percent of simulated data

sets (leaves C and D). Under these conditions, very short diffusion durations additionally decreased Type I error rates (leaf C).

We then examined data generated under *pure social learning* to assess statistical power. The classification tree analysis revealed that the probability to correctly infer social learning also depended mainly on two factors: the diffusion duration and group size. Observation probability, network heterogeneity and the length of coded time units had no major influence on the probability to correctly detect social learning.

Figure 4 shows the classification tree from the analysis of traits spread through social learning. The probability of correctly inferring social learning was very low when the diffusion duration was shorter than three time units (leaves A and B). In these cases TADA could not infer social learning because information from the spread of the new behavior was insufficient to allow a better fit of the combined asocial and social learning model relative to the pure asocial learning model. Higher probabilities for inferring social learning emerged when the diffusion duration was equal or longer than three coded time units (leaves C, D and E), but it was limited to 50 percent when the diffusion duration did not exceed four (leaf C). Note that for leaves A, B and C the statistical power was completely determined by diffusion duration and not strongly affected by another predictor variable.

When the duration of coded diffusion was equal to or larger than five, group size had a strong effect on the power to detect social learning. In groups with at least 17 individuals, social learning was correctly inferred in 87 percent of the cases (leaf F). For smaller group sizes (the simulated minimum was five), social learning was inferred in 60 percent of simulations (leaf D).

Discussion

We investigated how group size, network heterogeneity, diffusion duration, observation probability and length of coded time units impact the statistical performance of NBDA. In our analysis we focused on the TADA that Franz and Nunn (2009) referred to as "extended NBDA," which involves the comparison of a mixed model of social and asocial learning with one of pure asocial learning. Our most significant finding was that observation errors, when severe enough, can lead to increased Type I error rates. Elevated Type I error rates can be prevented by coding the observed diffusion times into larger units before applying TADA.

Observation errors might occur because animals cannot be observed continuously, for example when focal sampling is used (Altmann, 1974). In our simulations observation errors resulted in delayed recording of the acquisition times of the diffusing behavior (Figure 2). Importantly, observation errors of this type can also result in errors in the *order* of trait diffusion. Indeed, we found that in 89 percent of our simulated data sets, observation errors created sequence differences between the simulated and "observed" diffusion (i.e., changes in order of acquisition). Observation errors will generally become more pronounced when sampling effort per individual decreases and when the new behavior is performed less frequently.

To assess the power to detect social learning and Type I error rates we used simulated diffusion dynamics created by pure social and asocial learning as input to TADA. The statistical analysis of TADA results was performed using classification trees. This method hierarchically splits the data set into subcategories in a way that minimizes the variance of the response variable in the emerging subsets. Graphical output from the classification tree analysis can reveal and help to visualize complex interactions among predictor variables. However, resulting numbers indicating sharp

differences on the tree should not be viewed as thresholds that involve sudden jumps in power or Type I errors. While they may in fact represent thresholds, they may also reflect an underlying gradual effect of one or more predictors.

We found that observation errors can lead to elevated Type I error rates (Figure 3). To our surprise, however, TADA was more robust to observation errors than we anticipated. Previously we suggested that the length of coded time units should be long enough to ensure that each acquisition of the new behavior occurred with high certainty in the coded time step (Franz & Nunn, 2009). In the example displayed in figure 2, this would require us to choose the coding possibility 2. The results of the current study lead us to modify this requirement. We now offer a rule of thumb that having at least about a 50 percent probability that the acquisition of the behavior occurred in the coded time step is sufficient to prevent elevated Type I error rates (i.e., a bit more conservative than the >37% shown on Figure 3, since the predicted Type I error rate as 37% is 0.06). In figure 2, this would allow us to choose the coding possibility 1 (because in this case two out of three individuals acquired the new trait in the same coded time step in which they were first observed to perform this trait). In cases with a lower observation probability, appropriate coding of time steps can be used to keep the probability of Type I errors low.

Previously, we advised users to keep coded time steps short enough to avoid the possibility that two individuals learned in the same coded time step because we expected that this would reduce the power to detect social learning (Franz & Nunn, 2009). Our current results suggest that this advice also can be relaxed. Increasing the length of coded time steps only became a problem in our analyses when the resulting duration of the diffusion dropped below five. Therefore, we advise users of TADA who are uncertain about the extent of observation errors on diffusion times to code

14

diffusion times in rather long time units if this does not lead to a very short duration of the coded diffusion. Our analysis indicates that such a procedure minimizes Type I errors while not strongly affecting the power to detect social learning.

In contrast to our method, the TADA methods developed by Hoppitt at al. (2010) and Hoppitt at al. (this issue) are time continuous. In other words, they allow users to directly include observation times, rather than coding observed data into discrete time steps. Nevertheless, it can be expected that time continuous TADA methods are as susceptible to observation errors as the time discrete TADA method that we used in this study. Thus, if observation errors are strong enough, coding of time steps would also be needed when using these other methods.

The order of acquisition (OADA) method proposed by Hoppitt et al. (2010) may, however, be more robust than TADA against observation errors in the times of behavior acquisition. OADA does not use information about the time of behavior acquisition but only about the corresponding order. As long as observation errors do not change the order of the diffusion, the performance of OADA should not be impacted. However, it is difficult to predict how strongly the performance of OADA is impacted in cases in which observation errors lead to changes in the order of the observed diffusion.

Our finding that group size can strongly increase the power to detect social learning while network heterogeneity has little effect is in agreement with previous studies of the performance of TADA by Franz and Nunn (2009) and Hoppitt et al. (2010). Note that this result does not imply that including the network into the analysis of diffusion dynamics is not necessary. Franz and Nunn (2009) showed that diffusion curve analysis (Reader, 2004), which ignores social network structure, can strongly reduce the power to detect social learning and increase the probability to

erroneously infer social learning. The results of our current analyses do not contradict

these findings. Nevertheless, intuitively one might expect that the power to detect

social learning should increase with increasing network heterogeneity because the

more heterogeneous a social network is the stronger the likelihood that it determines

the pathway that a diffusion can take through a group of individuals. This argument

makes sense and also explains why the power of OADA to detect social learning

increases with increasing network heterogeneity dynamics (Hoppitt, et al., 2010).

However, to distinguish between alternative learning models TADA takes into

account both order of the diffusion and information about the timing of individual

learning events (or at least the first observed production of the behavior). This

information becomes increasingly important for distinguishing among alternative

learning models in more homogeneous social networks. In completely homogeneous

networks, TADA exclusively uses data on the timing of learning events and is

therefore equivalent to diffusion curve analysis (Reader, 2004). The timing of

learning events can be used to distinguish between social and asocial learning because

social learning, but not asocial learning, can be expected to result in a positive

feedback of the number of skilled individuals and learning rate (but see Hoppitt et al.,

this issue). The greater the number of individuals that have acquired a novel behavior,

the more opportunities exist for others to observe this behavior and learn it

themselves. The influence of this feedback mechanism is weak in heterogeneous

social networks because learning dynamics are strongly determined by the structure of

the network. However, with increasing network homogeneity the influence of the

feedback mechanism on diffusion dynamics increases. Thus, in more homogeneous

social networks, TADA makes greater use of information on the timing of learning

events to distinguish between alternative learning models. This explains why

decreasing network heterogeneity does not lead to a decrease in power to detect social learning.

In this context, we additionally emphasize that we explored the effect of network heterogeneity only based on the distribution of connection strengths in randomly created networks. Thus, we did not explicitly explore possible effects of other network characteristics such as community modularity, which can vary substantially in real world networks (Kasper & Voelkl, 2009). Furthermore, we did not consider in our analysis from which individual in the network the diffusion starts. As indicated by the results of Franz and Nunn (2009), this can also affect Type I error rates and power.

Other factors that can strongly affect the statistical performance of TADA which we did not consider in our analysis include (1) that a new trait spreads through combined effects of social and asocial learning, which would reduce the power to detect social learning (Hoppitt, et al., 2010), (2) that observations might begin long before the innovation and spread of the new behavior, which would increase the power to detect social learning, (3) observation errors in the structure of the social network, and (4) errors associated with incorrect identification of individuals in the group, such that behaviors are assigned to individuals who may not have learned how to perform the behavior. Results from Franz and Nunn (2009) indicate that TADA is robust against small disturbances in network structure. To investigate observation errors in network structure in more detail it might be suitable to follow the approach we have taken in this study and explicitly simulate observations that are used to construct social networks.

In summary, our results confirm that TADA can be a statistically powerful tool to infer social learning in wild animals. While the method seems to be robust to a

moderate degree of observation errors on recorded times of behavior acquisition, the coding of diffusion time steps allows it to cope with strong observation errors. Our results suggest that the power to detect social learning will not be strongly affected by observation errors. In cases in which observation errors cannot be precisely estimated, we therefore recommend relying on a conservative approach by using long time units for coding diffusion times.

Using NBDA to study social learning in wild animals might enable us to gain important new insights about 'culture' in animals and the evolution of cultural capacities in humans. While many social learning studies have focused on identifying the existence of social learning mechanisms in different species (Galef & Giraldeau, 2001; Hoppitt & Laland, 2008), NBDA methods provide a way to study and compare dynamics that are created by and influence social learning. Understanding which learning biases, or 'social learning strategies' (Laland, 2004) drive social learning dynamics in groups of animals and humans has received increasing attention in empirical and theoretical studies (Boyd & Richerson, 1985; Enquist, Eriksson, & Ghirlanda, 2007; Henrich & Boyd, 1998; J. Kendal, Giraldeau, & Laland, 2009; J. R. Kendal, Rendell, Pike, & Laland, 2009; R. L. Kendal, Coolen, & Laland, 2004; Laland, 2004; McElreath, et al., 2008; McElreath, et al., 2005; Mesoudi & Lycett, 2009). However, most empirical studies that investigated related questions (reviews: Kendal, Coolen, van Bergen & Laland, 2005; Kendal, Coolen & Laland, 2009) were performed with captive animals. Extensions of current NBDA approaches offer a way to approach these questions in the wild.

References

Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour, 49*(3-4), 227-267.

Bolker, B. (2008). *Ecological models and data in R*. Princeton, NJ: Princeton University Press.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: Univ. of Chicago Press.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York: Springer.

De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology, 81*(11), 3178-3192.

Enquist, M., Eriksson, K., & Ghirlanda, S. (2007). Critical social learning: A solution to Rogers's paradox of nonadaptive culture. *American Anthropologist, 109*(4), 727-734.

Franz, M., & Nunn, C. L. (2009). Network-based diffusion analysis: a new method for detecting social learning. *Proceedings of the Royal Society B-Biological Sciences, 276*(1663), 1829-1836.

Galef, B. G. (2004). Approaches to the study of traditional behaviors of free-living animals. [Article]. *Learning & Behavior, 32*(1), 53-61.

Galef, B. G., & Giraldeau, L. A. (2001). Social influences on foraging in vertebrates: causal mechanisms and adaptive functions. *Animal Behaviour, 61*, 3-15.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior, 19*(4), 215-241.

Hoppitt, W., Boogert, N. J., & Laland, K. N. (2010). Detecting social transmission in networks *Journal of Theoretical Biology*.

Hoppitt, W., Kandler, A., Kendal, J. R., & Laland, K. N. (this issue). The Effect of Task Structure on Diffusion Dynamics: Implications for Diffusion Curve and Network-Based Analyses. *Learning & Behavior*.

Hoppitt, W., & Laland, K. N. (2008). Social processes influencing learning in animals: A review of the evidence. *Advances in the Study of Behavior, Vol 38, 38*, 105-165.

Kasper, C., & Voelkl, B. (2009). A social network analysis of primate groups. *Primates, 50*(4), 343-356.

Kendal, J., Giraldeau, L.-A., & Laland, K. (2009). The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. *J Theor Biol, 260*(2), 210-219.

Kendal, J. R., Rendell, L., Pike, T. W., & Laland, K. N. (2009). Nine-spined sticklebacks deploy a hill-climbing social learning strategy. *Behavioral Ecology, 20*(2), 238-244.

Kendal, R. L., Coolen, I., & Laland, K. N. (2004). The role of conformity in foraging when personal and social information conflict. *Behavioral Ecology, 15*(2), 269-277.

Laland, K. N. (2004). Social learning strategies. [Article]. *Learning & Behavior, 32*(1), 4-14.

Laland, K. N., & Galef, B. G. (2009). *The Question of Animal Culture*. London: Harvard University Press.

Laland, K. N., & Hoppitt, W. (2003). Do animals have culture? *Evolutionary Anthropology, 12*(3), 150-159.

Laland, K. N., & Janik, V. M. (2006). The animal cultures debate. [Review]. *Trends in Ecology & Evolution, 21*(10), 542-547.

Laland, K. N., & Kendal, J. R. (2003). What the models say about social learning. In D. M. Fragaszy & S. Perry (Eds.), *The Biology of Traditions: Models and Evidence* (pp. 33 - 55). Cambridge: Cambridge University Press.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B-Biological Sciences, 363*(1509), 3515-3528.

McElreath, R., Lubell, M., Richerson, P. J., Waring, T. M., Baum, W., Edsten, E., et al. (2005). Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior, 26*(6), 483-508.

Mesoudi, A., & Lycett, S. J. (2009). Random copying, frequency-dependent copying and culture change. *Evolution and Human Behavior, 30*(1), 41-48.

Perry, S., & Manson, J. H. (2003). Traditions in monkeys. [Article]. *Evolutionary Anthropology, 12*(2), 71-81.

R Development Core Team (2007). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Reader, S. M. (2004). Distinguishing social and asocial learning using diffusion dynamics. [Review]. *Learning & Behavior, 32*(1), 90-104.

Rendell, L., & Whitehead, H. (2001). Culture in whales and dolphins. *Behavioral and Brain Sciences, 24*(2), 309-382.

Rushton, S. P., Lurz, P. W. W., Gurnell, J., & Fuller, R. (2000). Modelling the spatial dynamics of parapoxvirus disease in red and grey squirrels: a possible cause of the decline in the red squirrel in the UK? *Journal of Applied Ecology, 37*(6), 997-1012.

Seaholm, S. K., Ackerman, E., & Wu, S. C. (1988). Latin hypercube sampling and the sensitivity analysis of a Monte-Carlo epidemic model. *International Journal of Bio-Medical Computing, 23*(1-2), 97-112.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., et al. (2003). Orangutan cultures and the evolution of material culture. [Article]. *Science, 299*(5603), 102-105.

Ventura, R., Majolo, B., Koyama, N. F., Hardie, S., & Schino, G. (2006). Reciprocation and interchange in wild Japanese macaques: Grooming, cofeeding, and agonistic support. *American Journal of Primatology, 68*(12), 1138-1149.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (1999). Cultures in chimpanzees. [Article]. *Nature, 399*(6737), 682-685.

Figures and Tables

Table 1: Ranges in which varied parameters were sampled.

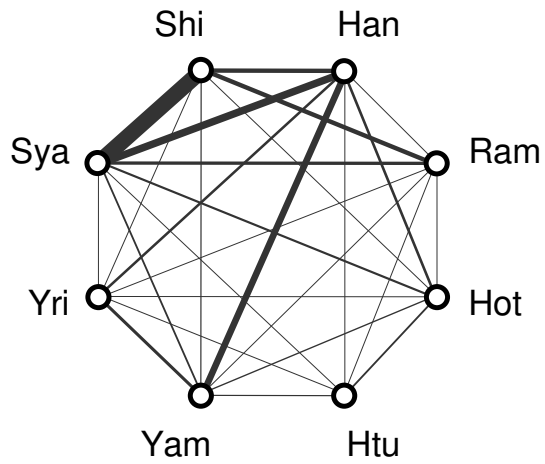| Parameter	Range
Group size	5 - 50
Network heterogeneity (h)	0 - 10
Observation probability	0.02 - 1
Length of coded time unit	1 - 50
Asocial learning rate (only in the asocial learning model)	0.01 - 0.1
τ (determines social learning rates in the social learning model)	0.05 - 0.5

Figure 1: Schematic representation of a co-feeding network of eight Japanese macaques (Ventura, Majolo, Koyama, Hardie, & Schino, 2006). Circles represent individuals, and lines indicate social connections between individuals. Line widths are proportional to time spent feeding together in close proximity.
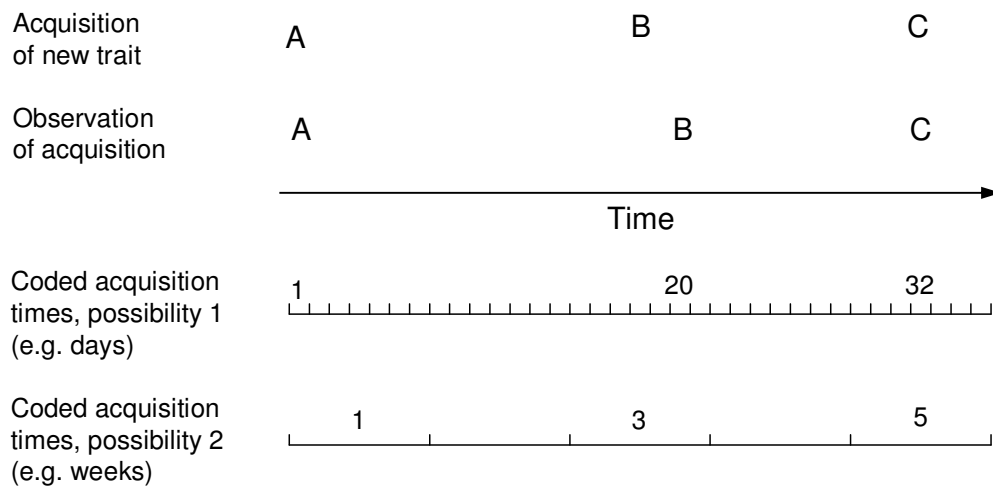
Figure 2: Example that illustrates observation errors and coding of diffusion times. The position of A, B and C mark times when three different individuals acquired a new trait and when an observer first recorded the new trait in each of these individuals. In addition, two possible ways of coding these data into discrete time steps of different length are shown. To illustrate how "length of coded time step" and "diffusion duration" are calculated in our analysis, we can assume that the depicted learning dynamics were simulated with a model in which one time step represents one day. The length of coded time steps in possibility 1 would then be one and in possibility 2 it would be seven. The diffusion duration using possibility 1 would be 32 and using possibility 2 it would be five.
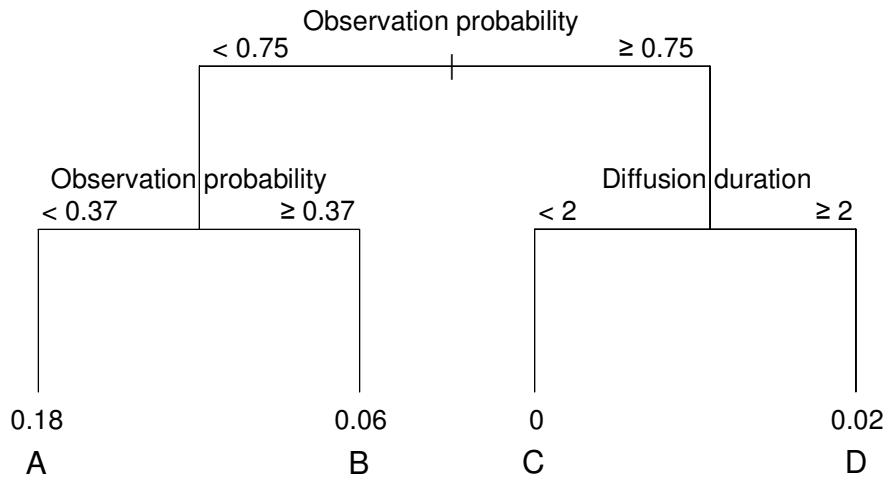
Figure 3: Classification tree of results from simulations of pure asocial learning. Numerical values of the leaves (A-D) indicate proportions of data sets in which social learning was erroneously inferred (representing Type I errors). Leaves are labeled to facilitate discussion of results that are depicted in the tree.
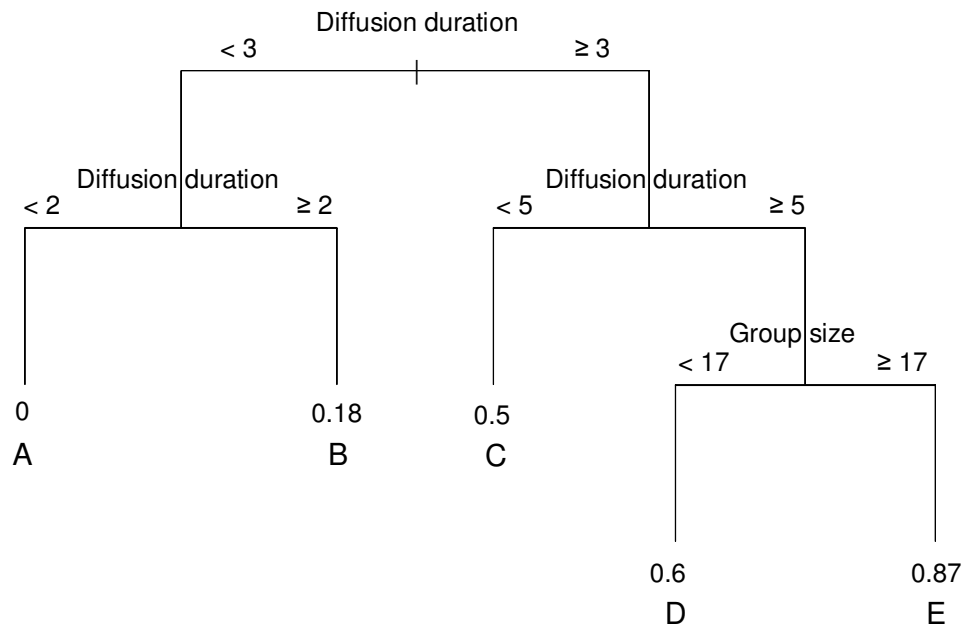
Figure 4: Classification tree of results from simulations of pure social learning. Numerical values of the leaves (A-E) indicate proportions of data sets in which social learning was correctly inferred. Leaves are labeled to facilitate discussion of results that are depicted in the tree.