REVIEW ARTICLE

Detecting Social Learning Using Networks: A Users Guide

WILLIAM HOPPITT* AND KEVIN N. LALAND
School of Biology and Centre for Social Learning and Cognitive Evolution, University of St. Andrews, Scotland, United Kingdom

Controversy over claims of cultures in nonhuman primates and other animals has led to a call for quantitative methods that are able to infer social learning from freely interacting groups of animals. Network-based diffusion analysis (*NBDA*) is such a method that infers social transmission of a behavioral trait when the pattern of acquisition follows the social network. As, relative to other animals, primates may be unusual in their heavy reliance on social learning, with learning frequently directed along pathways of association; in this study, we draw attention to the significance of this method for primatologists. We provide a "users guide" to *NBDA* methodology, discussing the choice of *NBDA* model and social network, and suggest model selection procedures. We also present the results of simulations that suggest that *NBDA* works well even when the assumptions of the underlying model are violated. Am. J. Primatol. 73:834–844, 2011.   © 2011 Wiley-Liss, Inc.

INTRODUCTION

"Social learning" is defined as learning that is influenced by observation of, or interaction with, a conspecific or its products. Most research is focused on social learning that results in social transmission: the spread of a behavioral trait from one individual to another. Such social learning may be particularly important in nonhuman primates, with a large number of cases being documented in which social learning and transmission is either directly observed or inferred to have occurred in the recent past on the basis of observed inter- or intraspecific variation [Dindo et al., 2009; Horner et al., 2010; Kendal et al., 2004, 2009, 2010a; Mcgrew et al., 1979; Perry & Manson, 2003; Sussmann, 1977; van Schaik et al., 2003; Whiten et al., 1999, 2005, 2007]. These include the famous case of the spread of potato and wheat washing through the Japanese macaques of Koshima Islet [Kawai, 1965], which led to the study of the spread, or "diffusion," of numerous other behavioral traits in natural populations of primates [Lefebvre, 1995].

In recent years, interest in social learning has grown dramatically, in part inspired by claims of cultures in apes [McGrew, 1998; van Schaik et al., 2003; Whiten et al., 1999] and monkeys [Perry & Manson, 2003]. In each case, local differences in behavior are thought to arise as a result of different behavioral innovations spreading through different populations. Laboratory studies have revealed that a capacity for social learning is widespread in primates [Whiten, 2000; Whiten & Mesoudi, 2008], as it is throughout both vertebrate and invertebrate taxa

[Galef & Laland, 2005; Hoppitt & Laland, 2008; Laland & Galef, 2009], so such explanations are highly plausible. However, in each case, this explanation remains controversial, since differences in behavior could be caused by genetic differences or ecological differences between populations [Laland & Hoppitt, 2003; Laland & Janik, 2006; Langergraber et al., 2010]. Consequently, researchers in the field of social learning are calling for quantitative methods that are able to infer social learning from freely interacting groups of animals, both in the wild and in captive but naturalistic settings [Kendal et al., 2010b].

One type of data that has been widely used to infer social transmission in groups of individuals is diffusion data, where the spread of a behavioral trait between individuals is monitored over time. The traditional method, diffusion curve analysis, plotted the cumulative number of individuals to have exhibited the behavioral trait against time, producing a "diffusion curve" [e.g. Lefebvre, 1995]. The

---

*Correspondence to: William Hoppitt, School of Biology, University of St. Andrews, Bute Building Queens Terrace, St. Andrews, Fife KY16 9TS, UK. E-mail: wjeh1@st-andrews.ac.uk

assumption was that if asocial learning alone were responsible for trait acquisition, this would occur at a constant rate and one would observe an r-shaped diffusion curve. Conversely, if social transmission were occurring, the rate of acquisition should increase with time as the number of informed individuals to learn from increased, resulting in an acceleratory curve, which will become S-shaped if the behavior spreads throughout the whole population being monitored. This method has been used extensively both in humans and nonhuman animals [Henrich, 2001; Reader, 2004; Rogers, 1995; Roper, 1986]. For example, Lefebvre [1995] used this method to analyze 21 diffusions of foraging innovations from the primate literature, including cases from Japanese macaques (e.g. fish-eating, [Watanabe, 1989]), vervet monkeys (acacia-pod dipping, [Hauser, 1988]), and chimpanzees (mango and lemon eating, [Takahata et al., 1986; Takasaki, 1983]. Lefevbre found an overall trend for accelerating learning rates, consistent with models of social transmission.

However, there are good reasons to doubt that an S-shaped diffusion curve is diagnostic of social learning [Hoppitt et al., 2010b; Reader, 2004]. First, a number of researchers have suggested that social learning might not necessarily result in an S-shaped curve, if, as is likely to be the case, the population is structured into subgroups. For instance, Kendal [2003] and Reader [2004] suggest that directed social learning can result in a step-shaped function, with acceleratory component parts, if the trait spreads more rapidly through closely connected subgroups, such as family units [e.g. Fritz et al., 2000]. Furthermore, differences in the rate of acquisition between different sub-sections of the population might act to obscure any underlying pattern; for example, a strong sex difference might result in a bimodal distribution of latencies to acquire the trait [Reader, 2004].

Another concern is that S-shaped diffusion curves might arise as a result of purely asocial processes, which would cause false positives for social transmission when using diffusion curve analysis. Lefebvre [1995, p 326] suggests that "trial-and-error learning could be characterized by a cumulative S-shaped curve at the population level if individual variation in learning latency were normally distributed." However, in their mathematical analysis, Hoppitt et al. [2010b] find that S-shaped curves are unlikely to arise as a result of individual variation in learning ability. Nonetheless, Hoppitt et al. detect reasons to be concerned that S-shaped curves might arise through asocial learning. One such case arises if an individual must move through a number of stages in order to acquire a trait. For example, there may be a number of different steps required to solve a foraging task, such as defenses that need to be removed to access a fruit [Whiten, 1998]. If the time

to complete each step of the task is exponentially distributed, then we would expect the overall latency to solve the task to follow an approximately gamma distribution, causing the diffusion curve to become more and more S-shaped as the number of task steps increases [Hoppitt et al., 2010b]. Even if a task is relatively simple in structure, we might still expect a systematic increase in the per-capita rate at which individuals solve it, due to a reduction in neophobia of the task, e.g. utilization of a novel food source. This too is likely to result in an S-shaped diffusion curve [Hoppitt et al., 2010b]. In conclusion, recent theoretical analyses suggest that researchers cannot reliably infer social learning from the shape of the diffusion curve.

Fortunately, in recent years, a more promising method for analyzing diffusion data has emerged, which addresses many of these concerns. It is known as network-based diffusion analysis (*NBDA*). *NBDA*, first invented by Franz and Nunn [2009], infers social transmission if the pattern of spread of a behavioral trait, as measured by the time to acquire a trait, follows the patterns of association in a social network. As such, it inherently addresses the concern that the pattern of spread of a trait will be influenced by population structure. Hoppitt et al. [2010a] later extended *NBDA* to apply to order of acquisition data (now known as order of acquisition diffusion analysis(*OADA*)) as well as time of acquisition (now known as time of acquisition diffusion analysis (*TADA*)). They also extended the method to include individual-level variables that might influence the rate of acquisition, such as sex, age, and dominance, which can statistically control for the effects of these variables when testing for social transmission [Hoppitt et al., 2010a]. This is particularly important if such variables are correlated with the social network structure, since this can result in false positives for social transmission if such variables are not taken into account. Franz and Nunn's [2009] *NBDA*, like diffusion curve analysis, is susceptible to false positives if there is a systematic increase in the asocial rate of acquisition, due to, for instance, a multi-step task structure or a decrease in neophobia over time [Hoppitt et al., 2010b]. This problem can be alleviated in one of two ways: either by ignoring the exact times of acquisition and fitting the model to the order of acquisition data [Hoppitt et al., 2010a], or by fitting a model that allows for the fact that the asocial rate of acquisition might increase (or decrease) over time [Hoppitt et al., 2010b].

With these refinements, *NBDA* is able to address many of the problems inherent in diffusion curve analysis, and so provide a reliable method for analyzing diffusion data and inferring whether social learning has taken place. As such, *NBDA* is of potential utility to any researcher interested in inferring social learning in natural populations, or

in captive, but freely interacting groups, a common objective of primatologists. *NBDA* has already been used to analyze the diffusion of the solution to an artificial foraging task in wild lemurs (*Lemur catta*), though there was no evidence of social transmission following the social network in this particular case [Kendal et al., 2010a]. Nonetheless, we feel that *NBDA* is likely to prove especially useful for primatologists interested in social learning, as primate groups are often studied in sufficient detail that diffusions of naturally occurring traits are likely to be documented at the level of individuals, and social networks are now commonly quantified in primate groups. In addition, *NBDA* can be modified to provide a statistical test of hypotheses regarding the pathways of social transmission within a group (e.g. [Nahallage & Huffman, 2007]).

Our aim in the first part of this article is to provide a users' guide to *NBDA*, describing the variants of the method in a manner that is accessible to a nonmathematical reader, and providing guidelines for how to apply the method. All the methods described in this article can be implemented in the statistical package *R* [R Development Core Team, 2008], using the code provided on our website (http://lalandlab.st-andrews.ac.uk/freeware.html), where instructions for use are also available. We refer any reader wishing to know the underlying technical details to the primary literature on *NBDA* [Franz & Nunn, 2009; Hoppitt et al., 2010a,b].

Despite its advantages, *NBDA* does make a number of simplifying assumptions, most critically, that individuals move directly from a naïve to a fully informed (knowledgeable) state at the time they first perform the behavioral trait. In some circumstances, this assumption will be unrealistic, for instance, where the behavioral trait gradually becomes established in an individual's repertoire through reinforcement of repeated performances. However, at this stage, it is not clear whether violations of this assumption would lead to error in the use of *NBDA*; they may not. Accordingly, in the second part of the article, we present the results of simulation analyses that investigate the conditions under which current *NBDA* models are able reliably to detect social learning when the task requires repeated trials. We do this by using agent-based simulations, in which individuals learn to solve tasks using established learning rules, and different types of social cues. We then apply NBDA to these simulated datasets in order to assess the method's efficacy in assessing the evidence for social transmission.

A USERS GUIDE TO *NBDA*

Randomization Method

A relatively simple way of testing whether the order of acquisition of a behavioral trait follows the pathways of association a social network is to use a randomization test [Manly, 2007]. One intuitive way to do this is to sum the network connections between each individual that acquires the trait, and the preceding individual to have solved the task, to yield a test statistic. The null distribution for the test statistic is then generated by repeatedly randomizing the order in which individuals solved the task (say, 999 times), and calculating the same test statistic in each case. The P value against the null hypothesis of no social transmission is given by the formula $(x+1)/(N+1)$, where N is the total number of randomizations and x is the number of randomizations for which the test statistic was greater than or equal to that observed for the real data. There are various other ways in which the test statistic can be calculated, but the general procedure remains the same.

The randomization approach has been used on fish (guppies) [Morrell et al., 2008] and birds (starlings) [Boogert et al., 2008], but failed to find a significant effect in both cases. In at least one of these cases, there are grounds for thinking this is a false negative, as the application of a more powerful *NBDA* found an effect of social learning [Hoppitt et al., 2010a]. Unfortunately, the randomization approach has some serious weaknesses. First, it does not estimate an effect size for social transmission, which is essential in making comparisons between species or contexts. Ideally, one would also require confidence intervals, which would give us an idea of the power of our data to detect social transmission. This is especially important since statistical power will not just depend on sample size but also network topography, with homogeneous networks yielding lower power [Hoppitt et al., 2010a]. Another severe limitation is that the randomization method does not allow for the inclusion of other variables that might influence the rate at which individuals acquire the trait. If a variable affecting asocial rate of acquisition is correlated with network structure, then this can result in false positives [Hoppitt et al., 2010a]. For example, dominant individuals might both associate with one another, and also be faster to acquire the trait than subordinates, making it appear like social transmission is occurring. Inclusion of variables such as social rank in the analysis allows us to statistically control for these effects, and even when such variables are not correlated with the social network, their inclusion can result in greater power to detect social transmission [Hoppitt et al., 2010a]. Finally, the randomization method does not enable us to incorporate information about the times of acquisition into the analysis, which can result in greater power, although, as we discuss below, it also has its weaknesses.

TADA AND *OADA*

The problems associated with the randomization method can be addressed by fitting a model in which

the rate of social transmission between individuals is assumed to be a positive function of the strength of the network connection between them. Franz and Nunn's [2009] *NBDA*, does just this: fitting a model to the time at which individuals acquire the trait, assuming that the rate of social transmission between an informed and a naïve individual is linearly proportional to the network connection between them. This effect is then scaled by a parameter representing the strength of social transmission, which is fitted to the data to assess the evidence for social transmission. In the software on our website this is a parameter (s) that varies from 0 (indicating all acquisition is asocial learning) to 1 (indicating that all acquisition is through asocial learning).

Hoppitt et al. [2010a] extended *NBDA* to include other, individual-level variables, such as sex or social rank, influencing the asocial rate of acquisition, allowing researchers to both estimate the effects of such variables and statistically control for them. They recognized two ways in which asocial learning and social transmission might interact, additively and multiplicatively. In the former case, each is an independent stochastic process, and either form of learning can result in acquisition, as would be the case if social transmission occurs as a direct result of observation, such as by imitation, or some other form of observational learning. Alternatively, the two acquisition processes might interact multiplicatively (see Fig. 1), which would be more appropriate if social transmission occurs indirectly through, for instance, local enhancement [see Hoppitt & Laland, 2008; Whiten & Ham, 1992 for discussion on direct and indirect routes to social learning]. It follows that if researchers are to include individual-level variables in their analysis, they must select between these models (see below). Hoppitt et al. [2010a] also note that *NBDA* can be extended simply to include multiple diffusions on different groups and/or of different traits.

Franz and Nunn [2009] and Hoppitt et al. [2010a] suggest slightly different methods for fitting the model to the data. Franz and Nunn [2009] recommend that time is split into a number a discrete units, and the researcher specifies in which unit, if any, each individual acquired the trait (this is the method used by Kendal et al. [2010a]). Hoppitt et al. [2010a] instead treat time as a continuous variable, and the times of acquisition are specified for each individual. In practice, the two methods will fit equivalent models if the number of time periods used in the discrete version is large. We find that the computation speed is much faster for the continuous version, and recommend that this be used if the exact times of acquisition are known. However, the discrete version remains useful if the exact times of acquisition are not known. For example, data might be collected as a series of scans to determine which
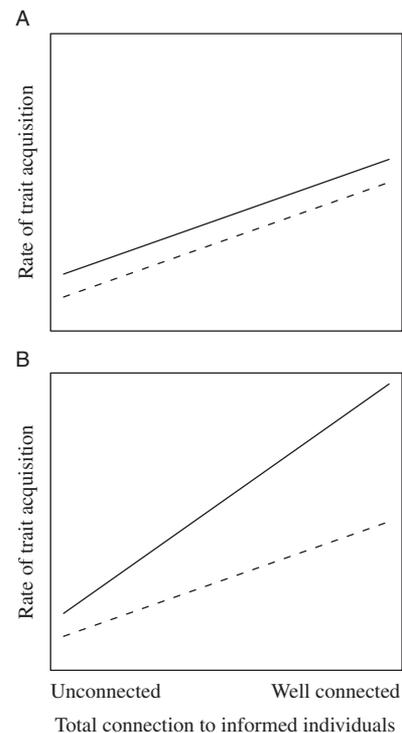

Fig. 1. A graphical depiction of (**A**) the additive *NBDA* and (**B**) multiplicative *NBDA*, showing the rate of trait acquisition for two individuals as a function of the total connection to informed individuals. At the extreme left of the range, individuals spend no time with any informed individuals, whereas at the extreme right, individuals are extremely well connected to individuals that have acquired the trait. For both (A) and (B), the asocial rate of acquisition for individual A (solid line) is double that for individual B (dashed line). In the additive model, the absolute difference in the rate of acquisition remains constant as the total connection increases, whereas in the multiplicative model, the ratio between the two remains constant.

individuals are informed at a number of discrete points in time. In this case, an individual observed performing the trait could have acquired it any time since the previous scan (assuming scans pick up all informed individuals), and so the discrete method is appropriate.

Even if data is collected continuously, in the wild it is unlikely that individuals can be observed all the time, meaning there will be observation error in the time at which each individual is recorded as acquiring the trait. Franz and Nunn [2010] found that observation errors can lead to inflated type-I error rates, but that this problem could be alleviated by using a discrete *TADA* in which the time units are long enough. They suggest a rule of thumb that there should be at least a 50% probability that an individual who has acquired the trait will be observed performing it within any given time unit. This can be checked by calculating the proportion of time units in which individuals are observed performing the trait after the time unit in which they are initially observed to do so.

As mentioned above, Hoppitt et al. [2010a] introduced a variant of *NBDA*, which is fitted to

data on the order in which individuals acquire a behavioral trait (*OADA*), rather than the *TADA*. They found that the *TADA* typically has greater power than *OADA*, but this comes at the cost of more assumptions. In the forms given by Franz and Nunn [2009] and Hoppitt et al. [2010a], *TADA* assumes that the baseline rate of acquisition (the rate of acquisition in the absence of social transmission) remains constant over time. This can result in false positives in the same circumstances as diffusion curve analysis: when the asocial rate of acquisition increases over time [Hoppitt et al., 2010a]. This could occur if the trait is the solution to a complex task that involves the completion of a number of steps. Alternatively, if the trait is behavior directed to a novel object, such as a novel food source, a decrease in neophobia over time would result in an apparent social transmission effect. *OADA* has the advantage that it is not sensitive to such effects. However, Hoppitt et al. [2010b] extend *TADA* such that it can accommodate an inhomogeneous (nonconstant) baseline rate of acquisition. They suggest that using a baseline function corresponding to a gamma distribution in the asocial time of acquisition will correct for an increasing baseline rate, and allow the *TADA* to be used in these circumstances, at the cost of an extra parameter being fitted to the data. However, *OADA* remains an attractive alternative if the baseline rate function is thought to take a form that is not easily modeled, e.g. fluctuating according to unmeasured environmental conditions.

In summary, the researcher has to decide whether to use (i) discrete *TADA*, continuous *TADA* or *OADA*; (ii) a multiplicative or additive model; and (iii) if using *TADA*, a constant or nonconstant baseline rate of acquisition. In Figure 2, we give a flow diagram to aid choice of (i), but choices (ii) and (iii) can be incorporated into the model selection framework, which we turn to now.

Model Selection Procedures

We will first cover the case where the main aim of the researcher is to test for social transmission through a network, either in a single diffusion, or where the parameter controlling the rate of transmission can be assumed to be the same for all diffusions being analyzed. (We will go on to consider the case multiple diffusions for which there may be different rates of transmission).

The goal of the researcher is to work out which statistical model provides the best fit to their diffusion data. In this article, although we assume the researcher is aiming to select the model best able to account for the data in order to assess the evidence for social transmission, in practice a model averaging approach [Burnham & Anderson, 2002] might also be applied). To this end, the researcher engages in a model selection procedure in which they must decide (i) which individual-level variables to include in the model; (ii) whether they should use an additive or a multiplicative model; (iii) whether they should assume a constant or nonconstant baseline rate of acquisition; and ultimately (iv) whether the inclusion of a parameter representing social transmission improves the model's fit to the data. If the latter is answered in the affirmative, then the researcher can conclude that social learning occurred.

Franz and Nunn [2009] suggest comparing models with (a) social transmission; (b) asocial learning; and (c) asocial learning and social transmission. We prefer to assume that asocial learning of the trait can always occur, even if it is only at a very low rate. Note that (a) can only be fitted if the diffusion starts with informed individuals in the population. There are a number of ways through which researchers can determine which model provides the best fit to their data (including information criteria or likelihood ratio tests (LRTs),
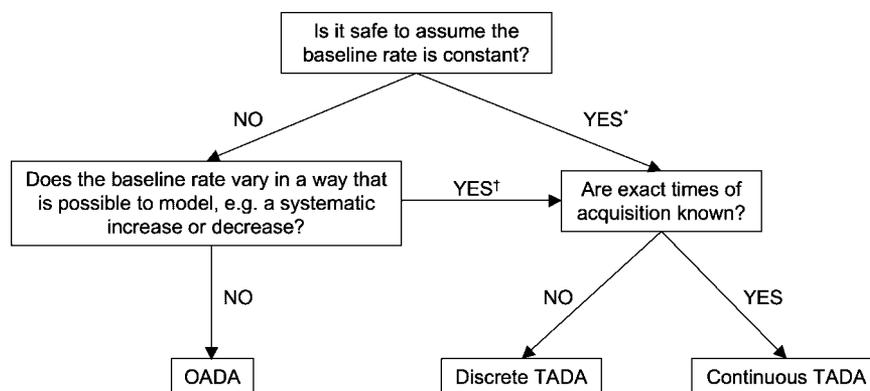

Fig. 2. Flowchart for selecting the appropriate *NBDA* model. *Researchers should be cautious in assuming the baseline rate of acquisition is constant, as a number of factors can cause increases in the baseline rate of acquisition [see Hoppitt et al., 2010b]. †In principle, any function can be used to model the baseline rate. However, the software provided on our website only allows for a systematic increase or decrease in the baseline rate. In cases, where environmental variables are thought to unpredictably influence the rate of acquisition, but to do so for all individuals in the same way, *TADA* becomes intractable, whereas *OADA* remains appropriate [see Hoppitt et al., 2010a].

forward selection, backward elimination, and so forth). In this article, we aim to provide some practical advice on how to proceed based on our experience and understanding of *NBDA*. In the ESM, we provide an example of the application of the strategy implemented below, using the code provided on our website. In this article, we suggest an approach to fitting a *TADA* model. For simplicity, in this article we assume that the baseline rate of acquisition is either constant, monotonically increasing, or monotonically decreasing. If this assumption is violated, the researcher should use *OADA* (Fig. 2). Note, the approaches we describe below can also be used for fitting an *OADA* model, but in this case one does not need to consider the baseline rate function.

The strategy involves first finding the best null model (model without social transmission) using Akaike's Information Criterion (AIC) or corrected AIC (AIC_c). Burnham and Anderson [2002] recommend using AIC_c when $N/k < 40$, where N is the sample size and k is the number of model parameters. We suggest taking N to be the number of acquisition events, summed across diffusions. When the number of individual-level variables being considered is small enough, we recommend fitting every possible combination of individual-level variables twice, the first time assuming a constant baseline rate function, and the second time assuming a nonconstant (gamma) baseline rate function. If the number of variables makes this infeasible, we suggest using both forward selection and backward elimination on the individual-level variables to find the best model for a constant and a nonconstant baseline. The best of these two models is selected as the null model, in comparison with which the best model containing social transmission must prove a superior fit to the data if social learning is to be inferred.

The researcher then repeats the exercise, this time seeking the best model containing social transmission. This time the procedure is carried out four times, assuming additive learning and multiplicative learning, each with constant and nonconstant baseline rates. The researcher chooses the best fitting model (smallest AIC or AIC_c) from among these four cases. Finally, the researcher compares the best model containing social transmission to the best model without social transmission ($s = 0$). The model that has the lowest AIC/AIC_c has most support, but the magnitude of the difference should also be assessed. Following the rules of thumb suggested by Burnham and Anderson [2002, p 70], we suggest that the evidence for social transmission be considered strong if AIC/AIC_c for the best social transmission model is > 4 greater than that for the best null model, and very strong if the difference is > 10. The full model-fitting procedure can be performed automatically using the AICTable function provided on our website, which returns a full table of models ordered by AIC or AIC_c.

A more stringent test for social transmission might be a procedure in which one simply adds social transmission to the null model, where one essentially tests for social transmission once everything else has been taken into account. However, one still has to choose between the additive and multiplicative models to compare with the null model.

When the researcher is analyzing multiple diffusions for which there may be different rates of transmission, we suggest an initial comparison between three models: (i) without social transmission; (ii) a common rate of social transmission for all diffusions; (iii) different rates of social transmission for cases where it is assumed to differ (e.g. different groups). This can be carried out using the strategy recommended above in each case. This will give us an initial idea of whether there is social transmission and whether it differs in strength between diffusions. If model (iii) is preferred, one can then test (i) which diffusions have significantly different rates of social transmission, by fitting models in which each pair are constrained to be the same, and using a LRT, and (ii) which are significantly different to zero, by fitting models in which each is constrained to be zero, and using a LRT (see ESM for an example of how this is performed).

Once the model has been fitted, it is important to obtain confidence intervals for the parameters of interest in the model. Even if there is no evidence for social transmission, we recommend getting confidence intervals for the social transmission parameter(s), as this can give us an idea of the statistical power to detect social transmission. If the confidence interval includes zero, but is also very wide, this suggests that the data did not have much power to detect an effect (either because of sample size, or homogeneous network structure). However, if the confidence intervals are narrow and include zero, this indicates that we have evidence that there was little or no social transmission involved in a diffusion (or more accurately, little or no social transmission following the social network provided).

Choosing a Social Network

So far we have not discussed the type of social network that should be used in an *NBDA*. There are numerous types of data that can be used to construct social networks [Croft et al., 2008; Whitehead, 2008], some result in binary networks, where a connection between individuals is either present (1) or absent (0), whereas others result in weighted networks, where the strength of a network connection is also specified. Networks can also be directed, where the connection from individual i to j is not necessarily the same as that from j to i. *NBDA* has been developed primarily for application to a weighted network, but there is no reason *NBDA* cannot be used for a binary network too. In such instances, the

model assumes that social transmission occurs at the same rate between all connected individuals, and not at all between nonconnected individuals.

In general, *NBDA* assumes that the rate of social transmission between two individuals is proportional to the network connection between them. Consequently, we suggest that a network is chosen that reflects the frequency of opportunities for social learning between each pair of individuals. For instance, if the trait is an arbitrary behavior pattern that can be performed at any time, one could use a measure of the proportion of time two individuals spend together, or within observation distance of one another. Alternatively, if the trait is a skill used in foraging, one could use a measure of the proportion of time individuals feed together. Although such reasoning is defensible, we are conscious that further work is required to investigate which network measures best correspond to the proportion of time individuals spend observing one another. However, while the preceding discussion assumes that a naïve individual must observe an informed individual performing a trait for social transmission to occur, this is not necessarily true. There are a number of cases where it has been shown that social transmission can occur when a naïve individual is exposed to the products of an informed individual's behavior [e.g. Campbell & Heyes, 2002; Mitchell et al., 1999; Sherry & Galef, 1984; Terkel, 1995]. In this case, the most appropriate social network would be one that reflects how often an individual is likely to encounter the products of another's behavior.

An alternative way to approach *NBDA* is to use a number of different types of social network, to investigate which one best explains the pattern of diffusion, and so make inferences about the manner in which social transmission occurs [Franz & Nunn, 2009]. One way to do this would be to formulate different social networks that correspond to different hypotheses about how social transmission occurs, such as specific theories regarding "directed social learning" [Coussi-Korbel & Fragaszy, 1995]. For instance, a hypothesis that all social transmission is vertical (from parents to offspring) could correspond to a directed binary social network which has 1 weightings from adults to their offspring, and 0 weightings elsewhere. Similar networks could be constructed to allow oblique transmission (older generations to younger generations) and/or horizontal transmission (between peers within a generation), and the fit of each model compared using AIC. In other cases, hypotheses might link social transmission to social rank. Although *NBDA* has not yet been used in this way, it represents a highly promising application of the method.

NBDA WITH A MORE REALISTIC LEARNING MODEL

NBDA fits a relatively simple model of the learning process. The model assumes that individuals move directly from a naïve state (before their first performance of the behavioral trait) to an informed state, after which they can transmit the behavioral trait to others at a rate that equals all other informed individuals. This is perhaps unrealistic, since in many circumstances individuals are not likely to be immediately proficient in performing the behavioral trait. More plausibly, individuals may increase their rate of performance of the trait as they are reinforced for doing so, thus transmitting the trait at a higher rate to others. The fact that *NBDA* makes a simplifying assumption about the learning process is not inevitably problematic, since all mathematical and statistical models make simplifying assumptions, and sometimes these do not really affect their utility. However, in this case, it is important to test whether *NBDA* works accurately when individuals learning in a more biologically plausible manner. To this end, we conducted simulations that modeled diffusions occurring according to a widely recognized learning rule, and then applied *NBDA* to the simulated data sets, to test whether *NBDA* could still detect social transmission in such circumstances.

In all simulations, naïve individuals spontaneously performed the target trait at a low baseline rate: the first individual to do so in the population would be the innovator [Reader & Laland, 2003]. When an individual performed the trait, it would be reinforced for doing so, following the Rescorla–Wagner learning rule, meaning that the association would converge on its maximum value (1) as the individual performed the trait more and more times (see Fig. 3). Therefore, the rate of trait performance, $\lambda_i(t)$, by individual i at time t was given as follows:

$$\lambda_i(t) = (\mu + \omega V_i(t)), \qquad (1)$$

where μ is the baseline rate of trait performance, $0 \leq V_i(t) \leq 1$ is individual i's association of the trait with reward at time t, and ω gives the maximum effect learning can have on the rate of trait
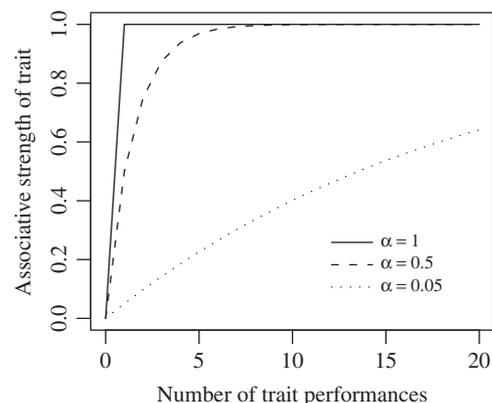

Fig. 3. Rescorla–Wagner learning rule used in the simulation models. The association of the trait with reward is shown as a function of the number of trait performances for different rates of learning (α).

performance. The quantity of interest here is ω/μ, which gives a measure of the rate at which experienced informed individuals perform the trait relative to naïve individuals, i.e. is a measure of how hard the trait is to invent asocially. Each time an individual solved the trait, its association of the trait with reward was incremented as follows:

$$\Delta V_i = \alpha(1 - V_i), \qquad (2)$$

where $\alpha \geq 0$, controls how quickly individuals reach the maximum level of learning. When $\alpha = 1$, this model corresponds to *NBDA*: after an individual's first performance of the trait, it is fully informed, and performs the trait at maximal rate. Our interest lies in testing whether *NBDA* is still appropriate when $\alpha < 1$.

In each simulation, each time one hypothetical individual performed the trait, there was a probability each other individual in the group would observe it. To determine who observed, for each simulated group, we randomly generated a social network, which was generated using the methods described in Hoppitt et al. [2010a]. This method generates a symmetrical network with values between 0 and 1, which here we take as the proportion of time individuals spend within observation distance of one another. The network was pruned such that connections less than 0.8 were set to 0 to give the network what we assumed to be a more realistic structure (the effect of this parameter is investigated in more detail in [Hoppitt et al., 2010a]). We assumed that that probability of observation was a linear function of network connection, corresponding to an observation probability of 0.3 for individuals that were within observation distance.

In the first set of simulations, social learning was modeled as occurring in an analogous manner to the asocial learning described above. For individuals who observed trait performance, their association of the trait with reward was incremented as follows:

$$\Delta V_i = s(1 - V_i), \qquad (3)$$

where $s \geq 0$ determines the strength of social transmission. We simulated 1,000 diffusions for each combination of parameter values (see ESM for technical details), each time recording the time each simulated individual first performed the trait, and taking this as the time of acquisition, as one would for real data. We then applied *OADA* and *TADA* (with constant and monotonically changing baseline rates) to the data, with no individual-level variables, recording whether social transmission was detected at the 5% significance level using an LRT. The proportion of cases in which social transmission was detected gives us an estimate of statistical power in each cases, or Type I error when $s = 0$.

Note that in applying *NBDA*, we used the same social network used to generate the data, and so assume that the measured social network reflects,

without error, the "true" social network determining the path of social transmission. Although this assumption is unrealistic, our main concern here is to compare the performance of *NBDA* when the assumptions of the model are and are not upheld, and so this assumption is not a concern.

We ran simulations for a group size of 20, for a social transmission effect $s = 0.025$, varying the relative effect of learning ω/μ, and the rate of learning α. The results are shown in Figure 4. We find that as the relative effect of learning, ω/μ, increases, the power to detect social transmission increases. This is not surprising. When ω/μ is low, individuals easily invent the behavioral trait for themselves, and so we would not expect the times of acquisition to follow the social network. Unsurprisingly, that power is highest when $\alpha = 1$, since the process underlying the data effectively follows the *NBDA* model. However, we note that though power does drop off as α decreases, this drop is surprisingly small even when α is as low as 0.05. This suggests that *NBDA* remains capable of detecting social transmission even when there is not a sharp transition from a naïve to an informed state. When we set $s = 0$, we found that the Type-I error rate still remained appropriate (~5%) for our more realistic learning model.

The first set of simulations assumed that social learning operated in an analogous manner to asocial learning, where observation directly altered an individual's association of the trait with food. However, there are numerous social learning processes that do not operate in this way. Some social learning processes transiently alter an observer's behavior in a way that leads to learning by otherwise asocial means [Hoppitt & Laland, 2008; Whiten & Ham, 1992]. We wanted to know whether *NBDA* was still effective in detecting social transmission that operated in this way. To model the process, we assumed that the observation of the trait transiently increased an observer's rate of trait performance (but did not result directly in learning), and that this effect faded away exponentially to baseline levels. This could taken as a model of local enhancement [Thorpe, 1956] where the observer is attracted to a location at which the trait is more likely to be performed, e.g. the solution to a foraging task can only be performed at the task itself [e.g. Whiten et al., 2005]. Alternatively, it could be a model of response facilitation [Byrne, 1994], where observation of the trait itself makes performance of the trait more likely.

We set the strength and duration of the transient effect to approximate that we estimated for local enhancement in populations of meerkats solving a foraging task [Hoppitt et al., in preparation]. The strength of the effect was set such that the rate of trait performance immediately after observation was approximately 2,000× that of
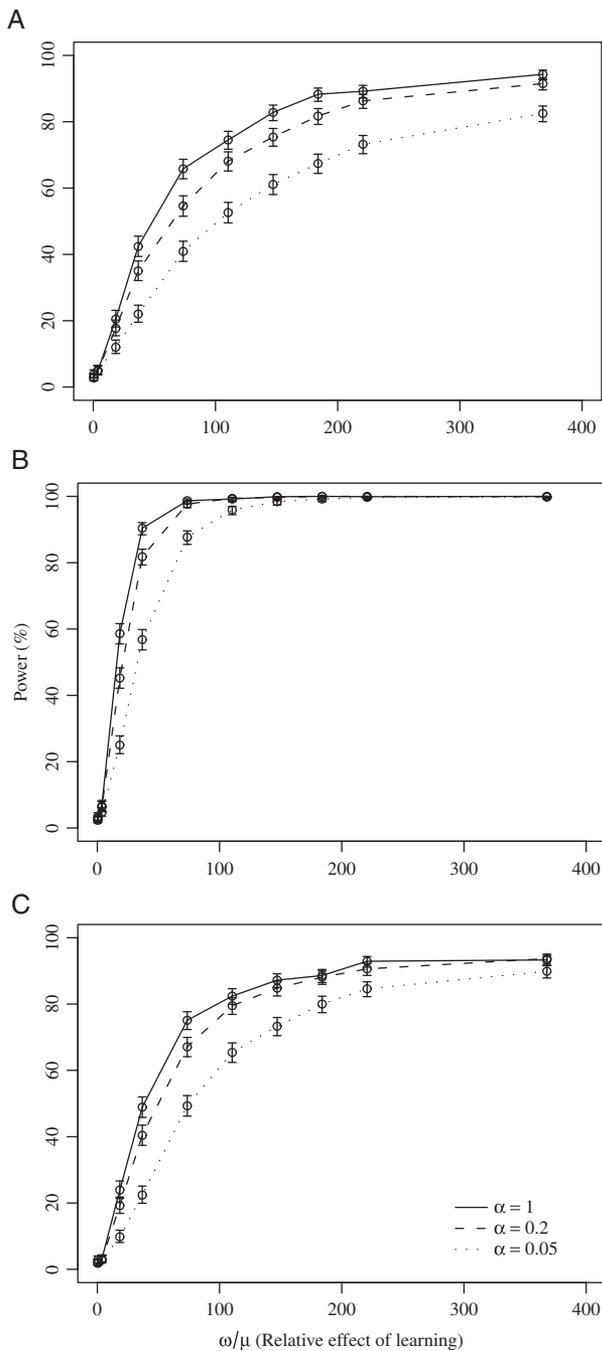
Fig. 4. Estimated power of different *NBDA* methods to detect direct social transmission ($s = 0.025$) as a function of the relative effect of learning (relative to the naïve rate of performance) and the rate of learning (α). Power at the 5% significance level was calculated using 1,000 simulations of groups of 20 individuals (see main text for details). Error bars show Wilson's 95% confidence intervals.

baseline levels, and this faded away with a half-life of 21 sec. We then ran sets of simulations for different values of α, ω/μ, and group sizes. The results are shown in Figure 5. The power was very high for the strength of transient effect that was modeled, even
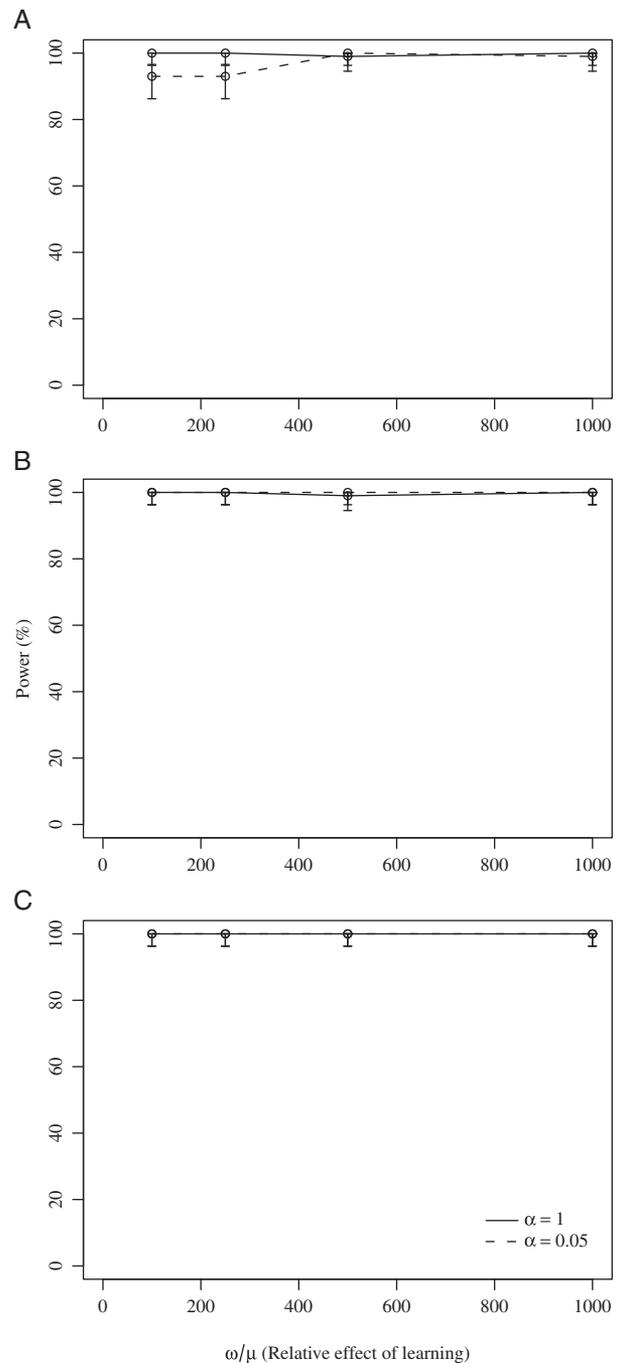
Fig. 5. Estimated power of different *NBDA* methods to detect indirect social transmission as a function of the relative effect of learning (relative to the naïve rate of performance) and the rate of learning (α). In the simulations observation resulted in a transient increase in trait performance, but did not result directly in learning. Power at the 5% significance level was calculated using 100 simulations of groups of 20 individuals (see main text for details). Error bars show Wilson's 95% confidence intervals. The results show *NBDA* can detect social transmission even when it operates by indirect means.

when the relative effect of learning (ω/μ) was low (c.f. Fig. 4), for both high and low rates of learning (α). Admittedly the transient effect modeled here is

strong, but this is based on an estimate made on real data [Hoppitt et al., in preparation].

The clear take-home message is that *NBDA* is capable of detecting social transmission, in spite of the fact that it assumes a very simplified form of learning. This conclusion holds even when social learning operates by indirect means (i.e. when observation of the trait does not directly result in trait acquisition).

SUMMARY

In summary, *NBDA* provides a promising method for detecting the social transmission of behavioral traits in primates, using data on the time or order of acquisition of the trait and a suitable social network. The simulations presented in this article suggest that *NBDA* works well even when the assumptions of the underlying model are violated, i.e. when individuals do not move directly from a naïve to an informed state, and when social transmission operates indirectly, by transiently influencing observers' behavior in a way that can lead to learning. *NBDA* can be carried out in the statistical environment R [R Development Core Team, 2008] using the code presented on our website (http://lalandlab.st-andrews.ac.uk/freeware.html), where technical instructions are also provided. We hope that the guidelines given in the first half of this article will aid researchers in selecting the appropriate variant of *NBDA*, and in fitting the model to their diffusion data.

ACKNOWLEDGMENTS

REFERENCES

Boogert NJ, Reader SM, Hoppitt W, Laland KN. 2008. The origin and spread of innovations in starlings. Animal Behaviour 75:1509–1518.

Burnham KP, Anderson DR. 2002. Model selection and multimodel inference: a practical information-theoretic approach. New York: Springer.

Byrne RW. 1994. The evolution of intelligence. In: Slater PJB, Halliday TR, editors. Behaviour and evolution. Cambridge: Cambridge University Press. p 223–265.

Campbell FM, Heyes CM. 2002. Rats smell: odour-mediated local enhancement, in a vertical movement two-action test. Animal Behaviour 63:1055–1063.

Coussi-Korbel S, Fragaszy D. 1995. On the relation between social dynamics and social learning. Animal Behaviour 50:1441–1453.

Croft DP, James R, Krause J. 2008. Exploring animal social networks. Oxford: Princeton University Press.

Dindo M, Stoinski T, Subiaul F, Whiten A. 2009. Social learning and diffusion of novel foraging tasks in orangutans (*Pongo pygmaeus* × *Pongo abelii*). American Journal of Primatology 71:48–48.

Franz M, Nunn CL. 2009. Network-based diffusion analysis: a new method for detecting social learning. Proceedings of the Royal Society B-Biological Sciences 276:1829–1836.

Franz M, Nunn CL. 2010. Investigating the impact of observation errors on the statistical performance of network-based diffusion analysis. Learning and Behavior 38:235–242.

Fritz J, Bisenberger A, Kotrschal K. 2000. Stimulus enhancement in greylag geese: socially mediated learning of an operant task. Animal Behaviour 59:1119–1125.

Galef BG, Laland KN. 2005. Social learning in animals: empirical studies and theoretical models. Bioscience 55:489–499.

Hauser MD. 1988. Invention and social transmission: new data from wild vervet monkeys. In: Byrne RW, Whiten A, editors. Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes and man. Oxford: Clarendon Press. p 327–343.

Henrich J. 2001. Cultural transmission and the diffusion of innovations: adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change. American Anthropologist 103:992–1013.

Hoppitt W, Laland KN. 2008. Social processes influencing learning in animals: a review of the evidence. Advances in the Study of Behavior 38:105–165.

Hoppitt W, Boogert NJ, Laland KN. 2010a. Detecting social transmission in networks. Journal of Theoretical Biology 263:544–555.

Hoppitt W, Kandler A, Kendal JR, Laland KN. 2010b. The effect of task structure on diffusion dynamics: implications for diffusion curve and network-based analyses. Learning and Behavior 38:243–251.

Horner V, Proctor D, Bonnie KE, Whiten A, de Waal FBM. 2010. Prestige affects cultural learning in chimpanzees. Plos One 5:e10625.

Kawai M. 1965. Newly acquired pre-cultural behaviour of the natural troop of Japanese monkeys on Koshima Islet. Primates 6:1–30.

Kendal JR. 2002. An investigation into social learning: mechanisms, diffusion dynamics, functions and evolutionary consequences. Unpublished doctoral thesis, University of Cambridge.

Kendal RL, Coolen I, Laland KN. 2004. The role of conformity in foraging when personal and social information conflict. Behavioral Ecology 15:269–277.

Kendal R, Kendal J, Hoppitt W, Lalalnd K. 2009. Identifying social learning in animal populations: a new "option-bias" method. PLOS ONE 4:e6541.

Kendal RL, Custance DM, Kendal JR, Vale G, Stoinski TS. 2010a. Evidence for social learning in wild lemurs. Learning and Behavior 38:220–234.

Kendal RL, Galef BG, van Schaik CP. 2010b. Social learning research outside the laboratory: how and why? Learning and Behavior 38:187–194.

Laland K, Galef BG, editors. 2009. The question of animal culture. Cambridge, MA: Harvard University Press.

Laland KN, Hoppitt W. 2003. Do animals have culture? Evolutionary Anthropology 12:150–159.

Laland KN, Janik VM. 2006. The animal cultures debate. Trends in Ecology and Evolution 21:542–547.

Langergraber KE, Boesch C, Inoue E, Inoue-Murayama M, Mitani JC, Nishida T, Pusey A, Reynolds V, Schubert G, Wrangham RW, Wroblewski E, Vigilant L. 2011. Genetic and "cultural" similarity in wild chimpanzees. Proceedings of the Royal Society B-Biological Sciences 278:408–416.

Lefebvre L. 1995. Culturally-transmitted feeding-behavior in primates - evidence for accelerating learning rates. Primates 36:227–239.

Manly BFJ. 2007. Randomization, Bootstrap and Monte Carlo Methods in Biology. Boca Raton: Chapman & Hall/CRC.

McGrew WC. 1998. Culture in nonhuman primates? Annual Review of Anthropology 27:301–328.

McGrew WC, Tutin CEG, Baldwin PJ. 1979. Chimpanzees, tools, and termites—cross-cultural comparisons of Senegal, Tanzania, and Rio-Muni. Man 14:185–214.

Mitchell CJ, Heyes CM, Gardner MR, Dawson GR. 1999. Limitations of a bidirectional control procedure for the investigation of imitation in rats: odour cues on the manipulandum. Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology 52: 193–202.

Morrell LJ, Croft DP, Dyer JRG, Chapman BB, Kelley JL, Laland KN, Krause J. 2008. Association patterns and foraging behaviour in natural and artificial guppy shoals. Animal Behaviour 76:855–864.

Nahallage CAD, Huffman MA. 2007. Acquisition and development of stone handling behavior in infant Japanese macaques. Behaviour 144:1193–1215.

Perry S, Manson JH. 2003. Traditions in monkeys. Evolutionary Anthropology 12:71–81.

R Core Development Team. 2008. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Reader SM. 2004. Distinguishing social and asocial learning using diffusion dynamics. Learning and Behavior 32: 90–104.

Reader SM, Laland KN. 2003. Animal innovation: an introduction. In: Reader SM, Laland KN, editors. Animal innovation. Oxford: Oxford University Press.

Rogers E. 1995. Diffusion of innovations. New York: Free Press.

Roper TJ. 1986. Cultural-evolution of feeding-behavior in animals. Science Progress 70:571–583.

Sherry DF, Galef BG. 1984. Cultural transmission without imitation—milk bottle opening by birds. Animal Behaviour 32:937–938.

Sussmann RW. 1977. Socialization, social structure, and the ecology of two sympatric species of Lemur. In: Chevalier-Skolnikoff S, Poirier FE, editors. Primate biosocial development. New York: Garland. p 515–528.

Takahata Y, Hiraiwa-Hasegawa M, Takasaki H, Nyundo R. 1986. Newly-acquired feeding habits among the chimpanzees of the Mahale Mountains National Park, Tanzania. Human Evolution 1:277–284.

Takasaki H. 1983. Mahale chimpazees taste mangoes- toward acquisition of a new food item? Primates 24:273–275.

Terkel J. 1995. Cultural transmission in the black rat—pine-cone feeding. Advances in the Study of Behavior 24:119–154.

Thorpe WH. 1956. Learning and instinct in animals. London: Methuen.

van Schaik CP, Ancrenaz M, Borgen G, Galdikas B, Knott CD, Singleton I, Suzuki A, Utami SS, Merrill M. 2003. Orangutan cultures and the evolution of material culture. Science 299:102–105.

Watanabe K. 1989. Fish: a new addition to the diet of the Japanese macaques on Koshima Island. Folia Primatologica 52:124–131.

Whitehead H. 2008. Analyzing animal societies. London: University of Chicago Press.

Whiten A. 1998. Imitation of the sequential structure of actions by chimpanzees (Pan troglodytes). Journal of Comparative Psychology 112:270–281.

Whiten A. 2000. Primate culture and social learning. Cognitive Science 24:477–508.

Whiten A, Ham R. 1992. On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. Advances in the Study of Behavior 21:239–283.

Whiten A, Mesoudi A. 2008. Establishing an experimental science of culture: animal social diffusion experiments. Philosophical Transactions of the Royal Society B-Biological Sciences 363:3477–3488.

Whiten A, Goodall J, McGrew WC, Nishida T, Reynolds V, Sugiyama Y, Tutin CEG, Wrangham RW, Boesch C. 1999. Cultures in chimpanzees. Nature 399:682–685.

Whiten A, Horner V, de Waal FBM. 2005. Conformity to cultural norms of tool use in chimpanzees. Nature 437:737–740.

Whiten A, Spiteri A, Horner V, Bonnie KE, Lambeth SP, Schapiro SJ, de Waal FBM. 2007. Transmission of multiple traditions within and between chimpanzee groups. Current Biology 17:1038–1043.