



Detecting social transmission in networks

William Hoppitt^{a,*}, Neeltje J. Boogert^b, Kevin N. Laland^a

^a School of Biology, University of St Andrews, Bute Building, Westburn Lane, St Andrews, Fife KY16 9TS, United Kingdom

^b Department of Biology, McGill University, Montréal, Québec, Canada

ARTICLE INFO

Article history:

Received 26 August 2009

Received in revised form

4 January 2010

Accepted 5 January 2010

Keywords:

Social learning

Network-based diffusion analysis

Culture

Traditions

Order of acquisition

ABSTRACT

In recent years researchers have drawn attention to a need for new methods with which to identify the spread of behavioural innovations through social transmission in animal populations. Network-based analyses seek to recognise diffusions mediated by social learning by detecting a correspondence between patterns of association and the flow of information through groups. Here we introduce a new order of acquisition diffusion analysis (*OADA*) and develop established *time of acquisition diffusion analysis* (*TADA*) methods further. Through simulation we compare the merits of these and other approaches, demonstrating that *OADA* and *TADA* have greater power and lower Type I error rates than available alternatives, and specifying when each approach should be deployed. We illustrate the new methods by applying them to reanalyse an established dataset corresponding to the diffusion of foraging innovations in starlings, where *OADA* and *TADA* detect social transmission that hitherto had been missed. The methods are potentially widely applicable by researchers wishing to detect social learning in natural and captive populations of animals, and to facilitate this we provide code to implement *OADA* and *TADA* in the statistical package *R*.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

'Social learning' is broadly defined as learning that is influenced by observation of or interaction with a conspecific or its products (Heyes, 1994). Social learning can result in 'social transmission', which we define as occurring when the acquisition of information or a behavioural trait by one individual exerts a positive causal influence on the rate at which another acquires the same information or trait. Social learning appears widespread across both vertebrate and invertebrate taxa (Hoppitt and Laland, 2008; Leadbeater and Chittka, 2007), whilst experimental work has established that social transmission can result in the establishment of behavioural traditions (e.g. Galef and Allen, 1995; Whiten et al., 2005). This has led to claims of animal cultures in natural populations of apes (McGrew, 1998; Whiten et al., 1999; van Schaik et al., 2003), cetaceans (Rendell and Whitehead, 2001; Krützen et al., 2005) and monkeys (Perry and Manson, 2003). However, such claims remain controversial because studies fail to adequately rule out alternative explanations for local differences in behaviour, such as local environmental differences, or genetic differences between populations (Laland and Hoppitt, 2003; Laland and Janik, 2006). There is concern that the current 'ethnographic' method, which infers social transmission only where the alternatives of genetic or

environmental variation can be disregarded, will rule out genuine cases of social transmission that covary with these factors (Laland and Janik, 2006; Laland and Galef, 2009). Consequently, in recent years researchers have called for the development of quantitative methods for inferring social transmission from field and captive study data that can rule out alternative explanations for the observed effect (Laland and Janik, 2006; Laland and Galef, 2009, and chapters therein).

One type of data that has previously been used to infer social transmission in groups of animals is diffusion data, where researchers monitor the spread of a novel behavioural trait. For some time the shape of the 'diffusion curve' (the cumulative number of individuals seen to perform the novel behaviour plotted against time) was used to infer social learning (e.g. Lefebvre, 1995a, 1995b). The assumption was that if learning were asocial, the rate of learning would be the same for all individuals, resulting in an r-shaped diffusion curve. In contrast, if there were social transmission, the rate of learning would increase as the number of demonstrators increased, resulting in an s-shaped curve (Reader, 2004). However, this approach has been somewhat discredited, since there are a number of situations in which we expect to see an s-shaped diffusion curve in the absence of social transmission (Laland and Kendal, 2003; Reader, 2004), or an r-shaped curve in the presence of social transmission (Franz and Nunn, 2009).

An alternative method is to use the order in which individuals acquire a behavioural trait to infer social transmission from

* Corresponding author. Tel.: +44 1334 467230; fax: +44 1334 463600.
E-mail address: wjeh1@st-andrews.ac.uk (W. Hoppitt).

diffusion data, on the assumption that if social transmission is operating we might expect the spread to follow the patterns of associations between individuals (Boogert et al., 2008; Morrell et al., 2008). The reasoning here is that individuals that are closely associated are more likely to learn from each other (Coussi-Korbel and Fragaszy, 1995). A randomisation approach has already been applied to test for such a pattern (Boogert et al., 2008; see also Morrell et al., 2008), but below we demonstrate that this approach is vulnerable to both Type I and Type II errors.

Here we propose an alternative method, which we call *order of acquisition diffusion analysis*, or *OADA*, where a model of social learning is fitted to the data by maximum likelihood, and tested against a model with no social transmission.¹ Our approach is similar to a method recently proposed by Franz and Nunn (2009), which they term ‘network-based diffusion analysis’ (or *NBDA*). Franz and Nunn’s method exploits data on the time at which individuals acquire a behavioural trait, rather than the order in which they do so. However, as *OADA* and the randomisation approach of Boogert et al. (2008) are also network-based diffusion analyses, for clarity we rename Franz and Nunn’s approach *time of acquisition diffusion analysis* (or *TADA*), and retain *NBDA* as the more general term for network-based approaches. We see the *OADA* and *TADA* approaches as complementary, and in later sections of this paper we introduce the *OADA* model, extend Franz and Nunn’s *TADA* method, and provide a full comparison of *OADA* and *TADA* models. We end by illustrating the methods by applying them to a published data set: the diffusion of novel foraging traits in groups of starlings, *Sturnus vulgaris* (Boogert et al., 2008).

2. Boogert et al.’s (2008) randomisation method

First, we will describe Boogert et al.’s (2008) randomisation method and illustrate its limitations. To implement this method, for each group in which a diffusion is recorded, one needs a matrix containing an appropriate measure of association between individuals (the association matrix), and the order in which individuals acquired the behavioural trait (the ‘diffusion chain’). The test statistic is then simply the summed strength of associations between adjacent individuals in each diffusion chain, summed across groups. If social transmission were occurring preferentially between closely associated individuals, the test statistic is likely to be larger than if individuals were learning independently. To test this hypothesis, a null distribution is generated by randomisation (Manly, 2007): the diffusion chain is randomised for each group, and the test statistic calculated. If the diffusion of multiple behavioural traits has been observed, one can test the global null hypothesis of no social transmission by summing the test statistic across traits. Boogert et al. proposed a second test statistic where, instead of summing the associations between adjacent individuals on the diffusion chain, one sums the mean association between each individual and all individuals before it in the diffusion chain. The logic here is that an individual can learn from any informed individual, not just the preceding individual on the diffusion chain. These are referred to as the ‘linear’ and ‘averaging’ metrics, respectively.

Boogert et al.’s randomisation method is non-parametric, which has the obvious advantage that researchers need to make few assumptions about the way in which social transmission and asocial learning proceed in order to test the null hypothesis. A

disadvantage is that it does not allow inferences about the strength of social transmission to be made, which might be useful for testing hypotheses about the nature of the social learning strategy deployed (Laland, 2004). A more serious limitation is that it is susceptible to false positives if closely associated individuals happened to have a similar rate of acquisition through asocial learning. For instance, individuals of high social rank might have a higher rate of asocial acquisition due to increased access to the resources required for learning. If, in addition, individuals happened to associate with those of a similar social rank, this might result in a false positive for the detection of social transmission (see below). An alternative approach is to fit the data to a model that includes both variables representing the effects of social transmission and known variables that might influence asocial learning, thereby controlling statistically for the latter. Below we describe *OADA* and *TADA* methods that allow this to be done.

3. Order of acquisition diffusion analysis (OADA)

3.1. Modelling social transmission

Our starting model assumes that the rate at which social transmission occurs between a given dyad of informed and naïve individuals is linearly proportional to the association between them. This assumption is likely to be reasonable provided that (a) the probability a naïve individual observes, or is exposed to, the performance of the novel trait is proportional to its association with the demonstrator, and (b) all informed individuals are approximately equally likely to perform the trait. The rate of acquisition of the trait through social transmission for individual i at time t , or $\lambda_{S,i}(t)$, is given by

$$\lambda_{S,i}(t) \propto (1 - z_i(t)) \sum_{j=1}^N (a_{ij} z_j(t)), \quad (1)$$

where $z_i(t)$ is a binary indicator variable indicating whether i is naïve (0) or informed (1) at time t , and a_{ij} is the association between individuals i and j , in a population of size N .

3.2. Inclusion of variables influencing asocial learning

At the same time we assume that it is possible that the individual may acquire the trait through trial and error or direct interaction with the environment, uninformed by the behaviour of others. The rate of asocial learning for i , $\lambda_{A,i}$ can be modelled as

$$\lambda_{A,i} \propto (1 - z_i(t)) \exp(\beta_1 x_{1,i} + \beta_2 x_{2,i} + \dots + \beta_V x_{V,i}), \quad (2)$$

where $x_{1,i}, x_{2,i}, \dots, x_{V,i}$ are the individual-level variables influencing asocial learning, and $\beta_1, \beta_2, \dots, \beta_V$ are the coefficients specifying the effect of each. Exponential transformation of the linear predictor ensures that the predicted rates are always positive, which is common practise in statistical modelling of rates (Therneau and Grambsch, 2001).

The question remains of how the effects of asocial learning and social transmission are combined in the model. Here we suggest two alternative approaches: (i) an additive model (Eq. (3)) and (ii) a multiplicative model (Eq. (4)). If social transmission occurs as an independent process by which individuals can acquire the trait, then the total rate of acquisition, $\lambda_i(t)$, will be the sum of the rates of asocial learning and social transmission, or

$$\lambda_i(t) = \lambda_0(t)(1 - z_i(t)) \left(s \sum_{j=1}^N a_{ij} z_j(t) + (1 - s) \exp\left(\sum_{k=1}^V \beta_k x_{k,i}\right) \right), \quad (3)$$

¹ In supplementary electronic material we provide code to run all analyses described in this paper in the statistical language R 2.8.1 (R Development Core Team 2008). Updated versions of the code will be made available on the authors’ website (<http://lalandlab.st-andrews.ac.uk/>).

where $\lambda_0(t)$ is a baseline rate of acquisition common to all individuals, and s is a parameter determining the strength of social transmission ($0 \leq s < 1$). Here $s=0$ indicates no social transmission and $s=1$ implies all learning is social. For a natural diffusion, $s \neq 1$ since the first individual must have acquired the behaviour through asocial learning. The additive model is likely to be appropriate if individuals can acquire the trait as a direct consequence of observation (Hoppitt and Laland, 2008), for instance, by imitation or some other form of observational learning.

Conversely, social transmission might often operate in an 'indirect' manner (Hoppitt and Laland, 2008), if the informed individual's behaviour influences the naïve individual's behaviour in a manner that leads indirectly to learning. For example, Leadbeater and Chittka (2007) found social transmission could speed the rate at which bumblebees (*Bombus terrestris*) learned to discriminate differently coloured artificial flowers, because they were attracted to rewarding flowers occupied by informed conspecifics, allowing them to learn by their own experience that these flowers are rewarding. Here the effect of social transmission is to increase the time spent in the area in which trait acquisition can occur (local enhancement, Thorpe, 1956), and so will weight the rate at which it occurs by otherwise asocial means. For these, and similar cases, we suggest that a multiplicative model is more appropriate, where

$$\lambda_i(t) = \lambda_0(t)(1-z_i(t)) \left(s \sum_{j=1}^N a_{ij}z_j(t) + (1-s) \right) \exp \left(\sum_{k=1}^V \beta_k x_{k,i} \right). \quad (4)$$

Here the $(1-s)$ term ensures that the effect of social transmission is weighted relative to the rate at which asocial learning occurs. The choice of model should not be seen as a nuisance. In cases where the experimenter has reasonable confidence in the likely social learning mechanism, the appropriate model can be selected. In other cases, both models may be used, and the model that best fits the data deployed. Indeed, this exercise could potentially be seen as providing information about the type of social transmission that is operating, although confidence in such inferences would be enhanced by experimental validation.

3.3. Model fitting

To implement an OADA we only need a relative measure of the rate at which individual i acquires the trait at time t (that is, relative to other naïve individuals), or $R_i(t) = \lambda_i(t)/\lambda_0(t)$. The probability that individual i is the next to learn can be written as

$$p_{\text{next},i}(t) = \frac{\lambda_i(t)}{\sum_{l=1}^N \lambda_l(t)} = \frac{\lambda_0(t)R_i(t)}{\lambda_0(t) \sum_{l=1}^N R_l(t)} = \frac{R_i(t)}{\sum_{l=1}^N R_l(t)} \quad (5)$$

and the probability that it will be the n th individual to acquire the trait, $p_{n,i}$, is given by

$$p_{n,i} = \frac{R_i(n)}{\sum_{l=1}^N R_l(n)}, \quad (6)$$

where $R_i(n)$ is i 's relative rate of acquisition immediately prior to the n th acquisition event. We can then write

$$R_i(n) = (1-z_i(n)) \left(s \sum_{j=1}^N (a_{ij}z_j(n)) + (1-s) \exp \left(\sum_{k=1}^V \beta_k x_{k,i} \right) \right) \quad (7a)$$

and

$$R_i(n) = (1-z_i(n)) \left(s \sum_{j=1}^N (a_{ij}z_j(n)) + (1-s) \right) \exp \left(\sum_{k=1}^V (\beta_k x_{k,i}) \right) \quad (7b)$$

for the additive (Eq. (3)) and multiplicative models (Eq. (4)), respectively, where $z_i(n)$ is the status of individual i prior to the n th acquisition event.

Eqs. (6) and (7) enable one to calculate the log-likelihood of the observed order of acquisition data for a given set of parameters, s and $\beta_1, \beta_2, \dots, \beta_n$ (e.g. see Morgan, 2009). The log-likelihood is easily calculated for multiple groups or multiple traits by adding together the log-likelihoods for each separate diffusion. The model is then fit by choosing the parameter values that minimise the log-likelihood, using a suitable numerical optimisation routine. In the supplementary material we provide R functions that fit both models (see ESM: "Additional Information" part C).

To fit the models, we find that the optimisation algorithms used are more likely to converge if we use the reparameterisation of $s' = s/(1-s)$ with $0 \leq s' < \infty$. This results in an additive model of

$$R_i(n) = (1-z_i(n)) \left(s' \sum_{j=1}^N a_{ij}z_j(n) + \exp \left(\sum_{k=1}^V \beta_k x_{k,i} \right) \right) \quad (8a)$$

and a multiplicative model of

$$R_i(n) = (1-z_i(n)) \left(s' \sum_{j=1}^N a_{ij}z_j(n) + 1 \right) \exp \left(\sum_{k=1}^V \beta_k x_{k,i} \right). \quad (8b)$$

To favour convergence of maximum likelihood estimation, we suggest use of Eq. (8) for model fitting, and transforming to the more intuitive parameterisation in Eq. (7) for interpretation.

3.4. Model selection and hypothesis testing

To test for social transmission, researchers can use a likelihood ratio test (LRT, see Morgan, 2009 for details) to compare the fitted model with a nested null model in which s is constrained to be zero. The significance of other parameters in the model can also be tested in this way, and the model reduced in a manner analogous to a multiple regression. Confidence intervals for parameters can be calculated using profile-likelihood techniques (ESM: Additional Information, part D; Morgan, 2009). Researchers can also use Akaike's Information Criterion (AIC) to compare alternative models with different degrees of freedom (Burnham and Anderson, 2002). This has the advantage that non-nested models can be compared, such as the best-fitting model containing social transmission and the best-fitting model without social transmission, when each contains different individual-level variables. Methods for dealing with tied data are given in the ESM (Additional Information, part F).

4. Comparison of OADA with TADA

Here we describe and extend Franz and Nunn's NBDA method, which we rename TADA, in the context of our OADA model, and using our notation. This facilitates a direct comparison between models reliant on order or time of acquisition.

TADA makes the same assumptions about social transmission as our model (Eq. (1)), but the models are fitted to time of acquisition data rather than to order of acquisition data, meaning the absolute rate of acquisition, $\lambda_i(t)$, is modelled, rather than the relative rate $R_i(t)$, and the baseline rate of acquisition is taken to be constant $\lambda_0(t) = \lambda_0$. Franz and Nunn suggest two approaches. The first involves fitting separate models for social transmission and asocial learning, with $\lambda_i(t) = \lambda_0$, and comparing the two models using AIC. However, this approach is only useful if the diffusion starts with informed individuals in the population, otherwise the likelihood of the

model for social transmission will always be zero, since the likelihood of the first individual's acquisition is zero. Similar to *OADA*, Franz and Nunn's second approach involves fitting a two-parameter model, which allows for both social transmission and a constant rate of asocial learning.

There are inherent strengths and weaknesses to both *TADA* and *OADA* methods. The fundamental difference is the type of data that is modelled, time or order. We demonstrate below that time of acquisition data typically possesses more power to detect a social transmission effect, which is the major advantage of *TADA*. However, *TADA* requires assumptions about the specific distribution of latencies: Franz and Nunn assume an exponential distribution, where the rate of acquisition at a given time is dependent only on the status of other individuals in the group. In contrast, *OADA* makes the less onerous assumption that the ratio of acquisition rates between two individuals is dependent only on the variables included in the model. The flexibility of this 'proportional hazards' assumption has led to the preference of the Cox proportional hazards model as the most widely used method for analysing time to event data (Therneau and Grambsch, 2001). The similarity of *OADA* to the Cox model is described in the ESM (Additional Information, part A). Below we show that the vulnerability of *TADA* and *OADA* to Type I error varies, and that each is more reliable than the other in some contexts.

In its initial form, Franz and Nunn's *TADA* is also susceptible to the same problems of confounding variables as Boogert et al.'s randomisation method. Accordingly, here we extend *TADA* to include individual-level variables influencing rate of acquisition. By the above reasoning, the additive model can be written as

$$\lambda_i(t) = \lambda_0(1 - z_i(t)) \left(s \sum_{j=1}^N a_{ij} z_j(t) + (1-s) \exp \left(\sum_{k=1}^V \beta_k x_{k,i} \right) \right) \quad (9a)$$

and the multiplicative model as

$$\lambda_i(t) = \lambda_0(1 - z_i(t)) \left(s \sum_{j=1}^N a_{ij} z_j(t) + 1 - s \right) \exp \left(\sum_{k=1}^V \beta_k x_{k,i} \right) \quad (9b)$$

where λ_0 determines the overall rate of asocial acquisition, and s parameterises the social transmission effect relative to the rate of asocial acquisition. As for *OADA*, we find the reparameterisation $s' = s/(1-s)$ works better for maximum likelihood estimation. We have also found this reparameterisation preferable to independent parameters for the rate of social and asocial transmission, since in the latter case the estimators for each are highly negatively correlated (Morgan, 2009). Setting $\lambda_0 = 1/L_0$ can facilitate convergence of the optimisation routines. The model can either be fitted by treating time as a continuous variable or by splitting time into a number of discrete steps, depending on the way in which the data was collected (details are given in the ESM: Additional Information, part E). Functions to implement this extended version of *TADA* for the multiplicative and the additive models, using both discrete and continuous methods of fitting, are given in the ESM: Additional Information part E.

5. Simulation details

We compared how the *OADA*, *TADA* and randomisation models performed under different circumstances. All simulations considered the diffusion of a single learned behavioural trait through a single hypothetical group of animals of size N . Where the rate of acquisition of the trait was affected by an individual-level variable, this was generated by drawing a value for each individual from a normal distribution ($x \sim N(0,1)$). We simulated an association matrix for the population by first generating a matrix of associations that

was normally distributed with a specified correlation, c , with the magnitude of the differences in the individual-level variable. To make the matrix more realistic, we made the matrix symmetrical by setting $a'_{ij} = a'_{ji} = (a_{ij} + a_{ji})/2$. We then transformed the associations to vary between 0 and 1 by ranking the values and dividing each by the maximum rank. To explore the effect of different levels of connectedness within the group, we set associations less than a threshold value, T , to zero, and explored how the magnitude of T affected the utility of the models.

Order and time of acquisition data were simulated according to either the additive model (Eq. (8a)) or the multiplicative model (Eq. (8b)) for specified values of λ_0 , s and β . At each point in the diffusion chain, a value was drawn from an exponential distribution with an appropriate rate parameter for each naïve individual (determined by Eq. (9a) or (9b)). The individual with the lowest value was taken to be the next individual to solve the task, with the intervals between solving events determined by the value itself. The data was then analysed using the additive and multiplicative *OADA*, the additive and multiplicative *TADA*, randomisation tests using Boogert et al.'s linear metric (1000 randomisations) and averaging metric (100 randomisations only, due to larger computation time). The simulations were usually run 10,000 times for each combination of simulation parameter values. This was reduced to 1000 times when there were individual-level variables due to the increased computation time required to fit NBDA models.

Where there were no individual-level effects, we considered a variety of group sizes ($N=10, 20, 50$), and social transmission effect sizes ($s=0, 0.2, 0.4, 0.6, 0.8, 0.99$), and recorded the power of each technique to detect social transmission. Since there were no individual-level variables, the multiplicative and additive models are equivalent in this case.

We explored individual-level effects in simulations in which group size was fixed at 20, $\beta=10$, and there were a range of social transmission effect sizes ($s=0, 0.4, 0.8$) and levels of correlation between the association matrix and differences in the individual-level variable ($c=0, 0.4, 0.8$). We recorded the statistical power to detect social transmission at the 5% significance level and the *OADA* and *TADA* models preferred by AIC_c.

In another series of simulations, we allowed the baseline rate of acquisition, $\lambda_0(t)$, to vary within a diffusion, either (i) at random or (ii) systematically. For (i), to determine the initial baseline acquisition rate a number was drawn from a normal distribution with mean = $\log(0.0002)$ and a standard deviation of 0, 2, 4, 6 or 8 and then exponentially transformed. This process was repeated to generate a new baseline acquisition rate after each acquisition event. For (ii), the baseline hazard rate either increased or decreased with successive acquisition events, with $\lambda_0(t) = 0.0002 \exp(\varepsilon p(t))$, where $p(t)$ is the proportion of demonstrators in the population at time t , and ε determines the strength of the effect. We considered $\varepsilon = -4, -3, -2, -1, 0, 1, 2, 3, 4$, $s=0$ or 0.8 , and $\beta=0$.

To explore the effect of network connectedness, we altered the threshold value, T , under which simulated associations were set to zero ($T=0, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$), decreasing the number of non-zero associations in the network as T increases. Here we assumed $s=0.4$ or 0.8 , $N=20$ and $\beta=0$. Unless otherwise indicated, $T=0.8$, $\lambda_0=0.0002$.

6. Application of the models to Boogert et al. (2008)

We go on to illustrate the methods by applying *OADA* and *TADA* to a published dataset. Boogert et al. (2008) presented three captive groups of five starlings (*S. vulgaris*) with six different artificial foraging tasks. Each task was presented

separately for several sessions. The time (measured cumulatively over sessions) at which each individual first contacted each task and first solved each task was recorded. Associations between individuals were calculated as the proportion of discrete point samples a given dyad was within pecking distance. In addition, a number of individual-level variables were recorded: (a) a measure of asocial learning ability, (b) two measures of neophobia: (i) the latency to feed in a novel environment, and (ii) average latency to feed next to three novel objects, (c) two measures of social rank: (i) competitive rank: time spent dominating a limited resource and (ii) agonistic rank calculated as David's scores based on agonistic interactions (de Vries et al., 2006). The aims of the study were to investigate which individual-level variables predicted the diffusion dynamics, and whether the order of acquisition of task solution followed patterns of association. Boogert et al. pursued the former aim by fitting linear mixed models (LMMs) or generalised linear mixed models (GLMM) to data on the number of times an individual was first to solve a task within its group, and the latency to solve the task (excluding the first solver), each with the individual-level variables as predictors. The question of whether order of acquisition followed patterns of association was tested using the randomisation approach described above. Boogert et al. reported their analysis showed no evidence for social transmission.

Here we implement an alternative approach that uses *OADA* and our extended version of *TADA*, comparing the results of each with the original findings. The methods were applied to the data from all diffusions, across all groups and tasks, in a global analysis. To test for social transmission, we first identified the combination of individual-level variables best able to account for the data, in the absence of social transmission. We fitted models with all possible combinations of individual-level variables and recorded AIC_c in each case. We selected the two best models and used these as null models to test for social transmission, assuming additive and multiplicative functions. We then fit a model with separate social transmission parameters for each group. We used a LRT to test each of these against zero, and dropped those that were not significant at the 5% level. We quantified the significance of the terms left in the model by dropping each from the model and using a LRT. To assess whether the social transmission parameter differed between specific groups, we fitted a null model with the parameter constrained to be equal for each group, and used a LRT to compare this to a model where they were unconstrained. We also obtained approximate confidence intervals for each parameter using profile-likelihood techniques (see ESM: Additional Information, part D). The same approach was used to fit *TADA*. Individual-level variables representing an effect of 'group' and 'task' were considered alongside those considered in *OADA*.

7. Results

7.1. Comparison in the absence of individual-level effects

In the absence of individual-level effects, and for a given group and effect size, *TADA* typically had more statistical power to detect social transmission than did *OADA*, while both of these methods were more powerful than the averaging and linear randomisation methods (Fig. 1a and b). In the case of the randomisation methods, the averaging metric usually provided more power than the linear metric, especially for larger group sizes, where social transmission is less likely to occur between adjacent individuals in the diffusion chain. In most cases, power

increased with group size, except for the randomisation method with the linear metric. As expected, statistical power also increased with the strength of social transmission.

7.2. Effect of individual-level variables

When there was no correlation between the individual-level variable and association, the Type I error rates were appropriate (~5%) for all methods (Fig. 1d). However, the power to detect an effect using *OADA* or *TADA* was greatly increased by inclusion of the variable in the model (see Fig. 2).

As the correlation between the individual-level variable and association increased, Type I error rates were greatly inflated for all methods that did not include an individual-level variable (see Fig. 1d). However, inclusion of the individual-level variable in both *OADA* and *TADA* methods restored Type I error to an appropriate rate, for both multiplicative and additive models. When social transmission and asocial learning were additive, power to detect social transmission was little affected so long as the additive model was fitted to the data (see Fig. 2a and b). In contrast, when social transmission and asocial learning combined multiplicatively, power was markedly reduced, though again, there was more statistical power when the appropriate multiplicative model was used, rather than the additive model (see Fig. 2c and d). AIC_c was generally a successful criterion in selecting the appropriate model (additive versus multiplicative, see ESM: Additional Information, part B).

These simulations demonstrate that the inclusion of individual-level variables in the analysis of diffusion data is highly desirable, both with respect to controlling Type I error rates, and maximising statistical power. This is an advantage that both *OADA* and our extension of *TADA* have over the randomisation techniques. Again, we see that *TADA* has more power than *OADA* in each case. Our analysis also lends confidence that the procedure we recommend will select a model (multiplicative or additive) appropriate to the data.

7.3. Varying baseline rate of acquisition

We manipulated the baseline rate of acquisition, both by increasing the variance of the underlying distribution (Fig. 3a) and by allowing it to increase or decrease as the diffusion proceeded (Fig. 3b). In all cases the power and Type I error rates remained approximately constant for the *OADA* method (see Fig. 3), as we would anticipate, since the baseline hazard function does not change the relative rate of acquisition. In contrast, *TADA* was very sensitive to changes in the baseline acquisition rate. When the baseline acquisition rate varied at random, statistical power dropped as the variance of the underlying distribution of rates increased (see Fig. 3a), whereas the Type I error rate increased. When the baseline acquisition rate decreased systematically throughout the diffusion, it obscured a social transmission effect from the *TADA* method (see Fig. 3b). Conversely, when the baseline acquisition rate increased, this resulted in an increase in Type I error for *TADA* (but see below).

These simulations illustrate the relative strengths and weaknesses of *OADA* and *TADA*. If there are fluctuating variables influencing the rate of acquisition that affect all individuals equally, then *OADA* is preferable to *TADA*. Likewise, if there is a factor that causes a systematic decrease in the baseline acquisition rate, *OADA* may be more likely to detect social transmission. This might occur if, for example, an increasing number of informed individuals depletes the resources necessary for trait acquisition, or increases the number of opportunities for scrounging, which might inhibit acquisition (Giraldeau and Lefebvre, 1987). The increase in

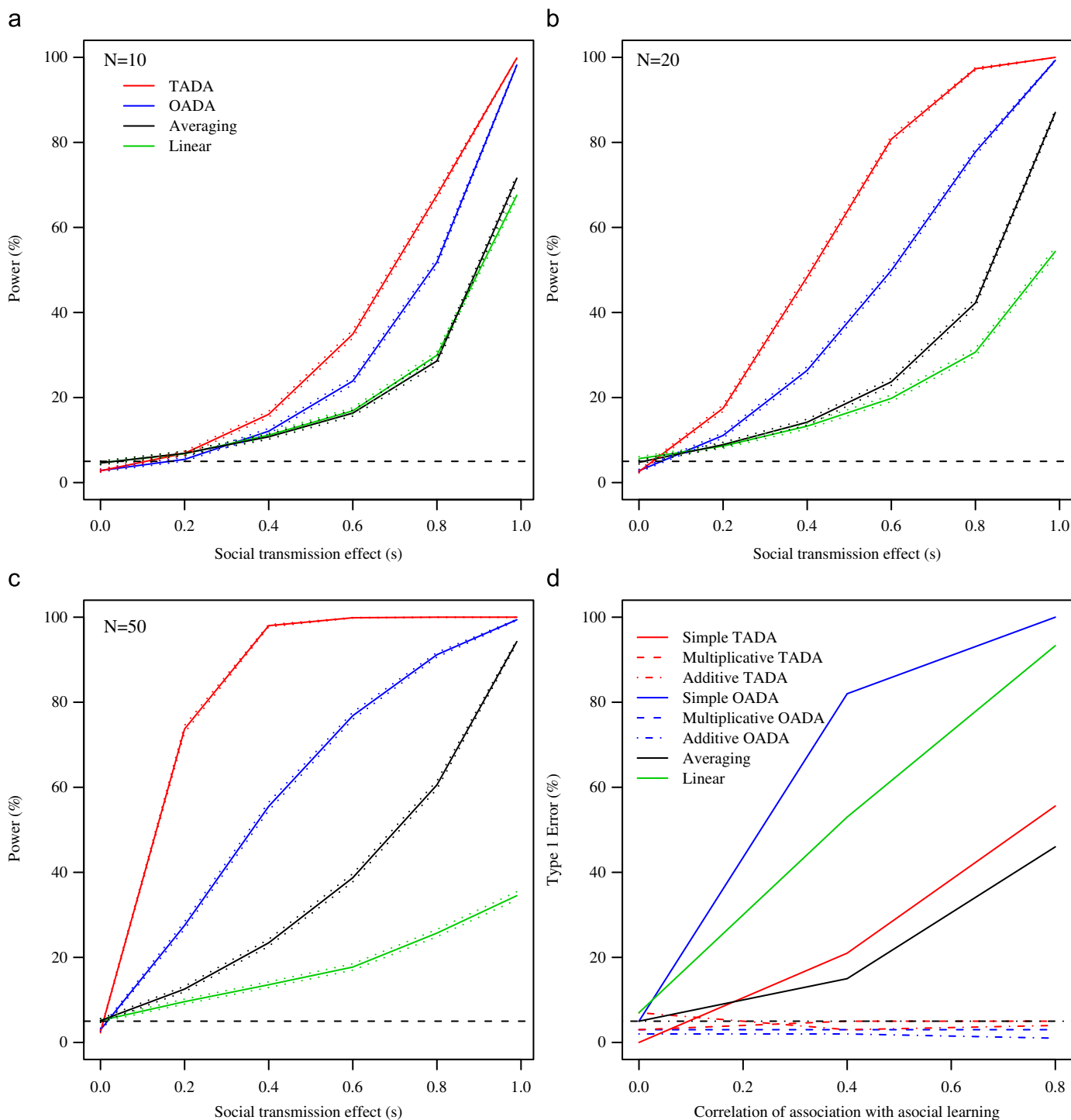


Fig. 1. Results of simulations for OADA, TADA and randomisation methods using linear and branching metrics. (a–c) Statistical power as a function of the strength of social transmission, for 10000 simulations of a single diffusion with differing group sizes. Dotted lines show Wilson's confidence intervals. (d) Type I error as a function of the level of correlation between association and asocial learning ability (see text for details), for 1000 simulations of a single diffusion with $N=20$. For OADA and TADA, the "Simple" model is one not including, and therefore not correcting for the individual-level variable influencing asocial learning. The type I error for the multiplicative and additive versions of OADA and TADA were always approximately 5% or less. Wilson's confidence intervals are excluded for clarity.

Type I error for an increasing baseline acquisition rate could be seen as a problem with TADA if there is reason to believe that a variable is influencing trait acquisition in this way. However, a systematic increase in baseline acquisition rate could be a direct result of the increased number of informed individuals, which would mean it is a case of social transmission by our definition. This shows that OADA is only sensitive to social transmission if it results in a difference in the relative rate of

acquisition by individuals, whereas TADA is also sensitive to absolute changes in the rate of acquisition (Fig. 3b).

7.4. Number of connections in the network

Network connectedness (the number of non-zero associations) had a different effect on OADA and TADA (see Fig. 4). For TADA,

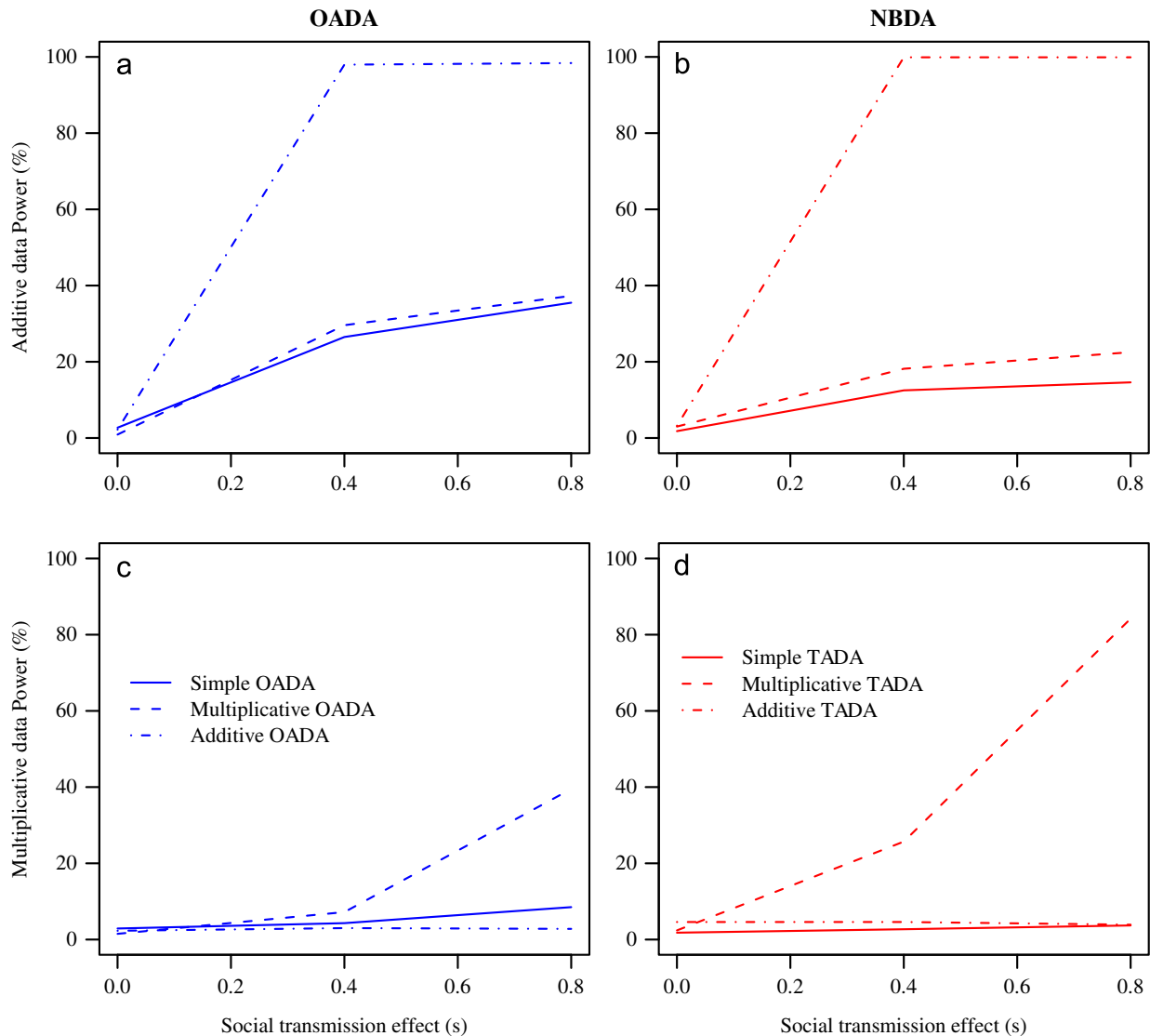


Fig. 2. Statistical power of *OADA* (left column) and *TADA* (right column) to detect different strengths of social transmission in additive (top row) or multiplicative (bottom row) data. “Simple” models are fitted without the individual-level variable influencing asocial learning. Solid lines show results for simulations in which the individual-level variable was independent of association, whereas dashed lines show results where the individual-level variable was correlated with association (level=0.8, see text for details). Number of simulated replications=10000; Group size=20. Wilson’s confidence intervals are excluded for clarity.

power either remained approximately constant ($s=0.8$) or declined ($s=0.4$) as connectedness went down (increasing T). In contrast, for *OADA*, the power increased in both cases, appearing to converge with the power for *TADA* when the proportion of zero associations was large. This is because *OADA* will detect social transmission when it results in large differences between the rates at which individuals acquire the trait, and works best when opportunities for social learning differ greatly between individuals at a given time. In contrast, *TADA* is also sensitive to the acceleration in the rate of acquisition which occurs as a result of an increased number of informed individuals. This effect will be more pronounced when there are many connections between individuals, offering many opportunities for social transmission.

7.5. Application of the models to Boogert et al. (2008)

Where the magnitude of the social transmission parameter was constant across groups, the best predictive *OADA* model included

object neophobia and asocial learning, and no social transmission (henceforth Model 1: $AIC_c=138.39$), but a model with latency to feed in a novel environment as sole predictor was almost as good (henceforth Model 2: $AIC_c=138.40$). Social transmission was not statistically significant when added to either model as an additive effect (Model 1: $LR=0$, $p=1$; Model 2: $LR=0.03$, $p=0.870$) or a multiplicative effect (Model 1: $LR=0.321$, $p=0.571$; Model 2: $LR=0.468$, $p=0.494$). However, when the social transmission parameter was allowed to vary between groups, we found a significant effect on group 1 in all models ($p < 0.05$, see Table 1), but no evidence for an effect on groups 2 or 3 ($p > 0.5$, see Table 1). For the additive model the social transmission effect on group 1 was also found to be significantly stronger than a putative effect on group 3 (Model 1: $LR=5.64$, $p=0.018$; Model 2: $LR=15.95$, $p < 0.001$) but not than that on group 2 (Model 1: $LR=0.65$, $p=0.420$; Model 2: $LR=1.02$, $p=0.312$). The same result was found for the multiplicative model: group 1 versus group 3: Model 1: $LR=5.15$, $p=0.023$; Model 2: $LR=4.30$, $p=0.038$; group 1 versus group 2: Model 1: $LR=0.91$, $p=0.340$; Model 2: $LR=0.25$, $p=0.614$. The best model, as judged by AIC_c included object neophobia and

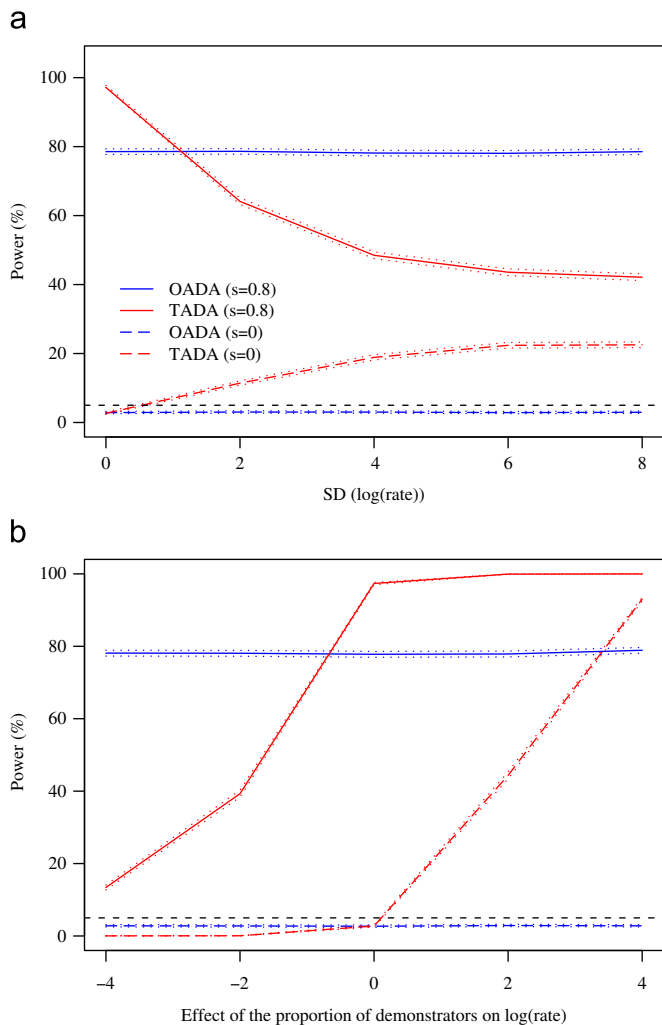


Fig. 3. Statistical power and type I error rate of *OADA* and *TADA* for a single simulated diffusion in which the baseline rate of acquisition varied. (a) Random variation: the baseline rate for each acquisition event was taken from a log-normal distribution with mean= $\log(0.0002)$ and shown on the x-axis. (b) Systematic variation: The baseline hazard rate varied systematically such that the log rate was a linear function of the proportion of demonstrators (informed individuals) in the population with intercept= $\log(0.0002)$ and coefficient shown on the x-axis. Positive values on the x-axis therefore indicate that the baseline rate of acquisition increased as the diffusions progressed, whereas negative values indicate that it decreased. In each case the solid lines show statistical power for a social transmission effect size of $s=0.8$ and the dashed lines shows type I error rate ($s=0$). Dotted lines show Wilson's confidence intervals. Number of simulated replications=10000; Group size=20.

asocial learning performance as individual-level variables, with an additive social transmission effect for group 1 only ($AIC_c=135.08$), although a multiplicative model worked almost as well ($AIC_c=135.14$). The AIC_c when all individual-level variables were dropped from the final model was 136.15, which is preferred to an additive model including latency to feed in a novel environment. None of the individual-level variables were significant at the 5% level when dropped from any of the final models. See Table 1 for full details of the best-fitting *OADA* models.

The best predictive *TADA* model excluding social transmission included latency to feed in a novel environment as a sole predictor ($AIC_c=1175.56$), so this was used as the null model to test for social transmission. When social transmission was added to the null model it was highly significant for both the additive ($LR=15.54$, $df=1$, $p < 0.001$) and multiplicative model ($LR=16.75$, $df=1$, $p < 0.001$). There was no evidence of a difference in the effect of social transmission between groups for either the

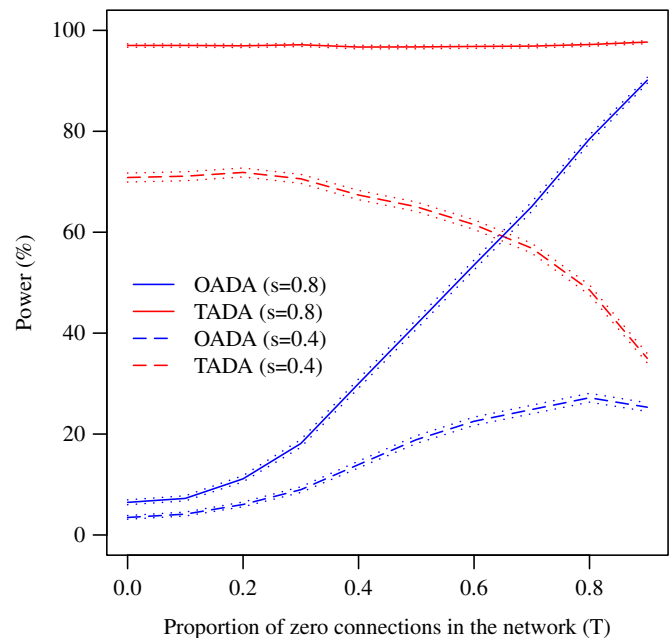


Fig. 4. Statistical power of *OADA* and *TADA* for a single simulated diffusion as a function of the proportion of zero connections in the network. Solid lines are for a social transmission effect size of $s=0.8$ and dashed lines $s=0.4$. Dotted lines show Wilson's confidence intervals. Number of simulated replications=10000; Group size=20.

additive ($LR=0.27$, $df=2$, $p=0.872$) or multiplicative model ($LR=1.30$, $df=2$, $p=0.523$). The best model, as judged by AIC_c , included latency to feed in a novel environment as an individual-level variable, with a common multiplicative social transmission effect for all groups ($AIC_c=1161.02$), although an additive model worked almost as well ($AIC_c=1162.24$). In contrast to the *OADA* model, there is a clear indication that latency to feed in a novel environment has a negative relationship with individuals' rates of acquisition (Additive model: $LR=5.04$, $df=1$, $p=0.025$; Multiplicative model: $LR=6.80$, $df=1$, $p=0.009$). See Table 2 for full details of the best-fitting *TADA* models.

Consistent with Boogert et al.'s original conclusions, when all groups were analysed together, there was no evidence of social transmission using the randomisation methods used by Boogert et al. (linear metric=206.5; $p=0.170$; averaging metric=204.2, $p=0.149$).² However, when groups were analysed separately (this was not done by Boogert et al.), both randomisation metrics provided evidence for social transmission in group 1 (linear metric=72, $p=0.013$; averaging metric=69.1, $p=0.012$; new metric: $G1=25.3$, $p=0.012$), but no evidence for groups 2 and 3 ($p > 0.15$ in all cases).

Whereas the randomisation tests used by Boogert et al. failed to find evidence of social transmission, based on the order of acquisition data, our *OADA* method found evidence for social transmission in group 1. When the randomisation methods were reapplied to the data from each group separately, the same results were found. However, unlike *OADA*, the randomisation methods do not enable us to construct confidence intervals on the effect of social transmission in each group. The 95% confidence intervals from *OADA* reveal that the data provide no resolution to distinguish social transmission from asocial learning in group 2, whereas in groups 1 and 3, the data are

² Note that the p values given here are one-tailed, whereas Boogert et al. calculated two-tailed p values by doubling the one-tailed p value, though there was a mistake in the calculation of the p value corresponding to the averaging statistic causing it to be reported as half its estimated value.

Table 1

Model used	Null model	Null model AIC _c	Estimated social transmission effect (s) 95% CI LRT (H ₀ :s'=0) ^a			Estimated effects for individual-level variables LRT (H ₀ :β=0) ^a			AIC _c for a model including social transmission for groups:		
			Group 1	Group 2	Group 3	ON	AL	LFNE	1	1 & 2	All
Additive OADA	Object neophobia+asocial learning	138.4	≈ 1 ^b [0.08,1]	0.06 [0,1]	0 [0,0.13]	-0.348 χ ₁ ² = 2.87, p=0.090	-0.350 χ ₁ ² = 3.18, p=0.075		135.08	137.26	139.63
	Latency to feed in a novel environment	138.4	χ ₁ ² = 5.53, p=0.018 ≈ 1 ^b [0.06,1]	χ ₁ ² = 0.12, p=0.734 0.12 [0,1]	χ ₁ ² = 0, p=1 0 [0,0.22]		-0.112 χ ₁ ² = 1.89, p=0.169	136.39	138.10	140.39	
Multiplicative OADA	Object neophobia+asocial learning	138.4	≈ 1 ^b [0.05,1]	0.02 [0,1]	0 [0,0.20]	-0.247 χ ₁ ² = 2.61, p=0.106	-0.322 χ ₁ ² = 3.18, p=0.075		135.14	137.42	139.79
	Latency to feed in a novel environment	138.4	χ ₁ ² = 5.47, p=0.019 ≈ 1 ^b [0.03,1]	χ ₁ ² = 0.02, p=0.898 0.18 [0,1]	χ ₁ ² = 0, p=1 0 [0,0.24]		-0.117 χ ₁ ² = 2.42, p=0.120	135.86	137.69	140.37	

ON=object neophobia; AL=asocial learning ability; LFNE=latency to feed in a novel environment.

^a LRTs are for significant parameters dropped from the final model, and for non-significant parameters when added to the final model. In all cases the final model included a social transmission effect for group 1 only. df=1 in all cases.

^b The likelihood continued to increase towards the upper bound 1, though s=1 has a likelihood of zero, since the first acquisition would be impossible (e.g. see Appendix D Fig. D.1).

Table 2

Model used	Null model	Null model AIC _c	Estimated social transmission effect (s) 95% CI LRT (H ₀ :s'=0) ^a	Estimated effects for individual-level variable (LFNE) 95% CI LRT (H ₀ :β=0) ^a	AIC _c including social transmission
Additive TADA	Latency to feed in a novel environment	1175.56	0.16 (0.08, 0.43) χ ₁ ² = 15.54, p < 0.001	-0.23 (-0.39, -0.04) χ ₁ ² = 5.04, p = 0.025	1162.24
Multiplicative TADA	Latency to feed in a novel environment	1175.56	0.17 (0.08, 0.33) χ ₁ ² = 16.75, p < 0.001	-0.16 (-0.36, -0.01) χ ₁ ² = 6.80, p = 0.009	1161.02

LFNE=latency to feed in a novel environment.

^a LRTs are for significant parameters dropped from the final model.

consistent with a lower and upper limit on social transmission, respectively (see Table 1).

In contrast to *OADA*, *TADA* provided evidence of social transmission in all groups, with no evidence of differences between them. This is probably the result of increased power resulting from inclusion of time of acquisition data, which is reflected in the narrower 95% confidence intervals for social transmission (see Table 2). The findings of *TADA* and *OADA* are less contradictory than they might appear at first. The *TADA* confidence intervals for social transmission are within the *OADA* confidence intervals for the effect for groups 1 and 2 and overlap with the *OADA* confidence intervals for group 3. The only real discrepancy is the finding from *OADA*, that social transmission was significantly stronger in group 1 than it was in group 3.

The results concerning individual-level variables are qualitatively similar for both *OADA* and *TADA*. *TADA* suggested that an individual's latency to feed in a novel environment was the best predictor of time of acquisition. In *OADA* this variable was also found to be a good predictor of the order of acquisition, though a model including object neophobia and asocial learning ability was approximately as good. However, when the model included social

transmission for group 1, none of these variables were significant at the 5% level. In Boogert et al.'s original analysis, significant differences in latency to solve were found between tasks (not significant in *TADA*), but no other variable was found to be significant (however, latency to feed in a novel environment was found to be correlated with the latency to contact the task). The critical differences between the *TADA* presented here and Boogert et al.'s analysis are: (a) social transmission was accounted for; (b) first-solvers were not excluded from the analysis; (c) individuals not solving the task were modelled as non-solvers rather than assigned a 'ceiling' value, which can distort an analysis of latencies (Crawley, 2002); and (d) we compared all possible subsets of variables, rather than using backward selection, which can be misleading when predictors are correlated (Weisberg, 1980).

The simulations presented above suggest that we should prefer *TADA* to *OADA* because of its greater power provided we are happy to assume that the baseline rate of acquisition is constant. We can think of no reason to reject this assumption in Boogert et al.'s diffusion experiment: the diffusions were conducted under laboratory conditions, reducing the possibility for external

influences on the birds' rate of acquisition. In addition, there were multiple versions of each task available and each was replenished as soon as it was solved, ensuring that informed individuals could not block naïve individuals from accessing the task. In any case, the blocking of naïve individuals would result in a decrease in the power of *TADA* to detect an effect (see above), whilst the failure of *OADA* to find a social transmission effect for groups 2 and 3 is likely to be a result of the reduced power of the analysis relative to *TADA*.

In summary, these new more powerful methods lead us to the conclusion that there is strong evidence for social transmission in all three groups of starlings, a finding starkly contrasting with that of Boogert et al. (2008).

8. Discussion

The above simulations bring home the desirability of including individual-level variables in an analysis to detect social transmission from diffusion data. The analyses establish that the inclusion of individual-level variables both increases statistical power and reduces Type I error rates. In addition, the sensitivity of the diffusion analyses to network structure prompts us to recommend that researchers use methods that can generate confidence intervals for the strength of social transmission, rather than relying on a rejection/acceptance procedure. For these reasons, our *OADA* and refined *TADA* methods are preferable to established randomisation approaches.

The simulations clearly show that *TADA* yields more statistical power than *OADA*. Consequently, in choosing which approach to utilise, we suggest that researchers use *TADA* unless there is good reason to suppose that the baseline rate of acquisition has changed over time. This might be the case, if, for example, the availability of a resource necessary to acquire the trait has varied over time, in which case the weaker assumption of proportional hazards is more appropriate, and the *OADA* method should be deployed (also see below). In principle, one could modify *TADA* to incorporate a non-constant baseline rate of acquisition. However, the success of this method would depend on the researcher choosing an appropriate baseline function. *OADA* has the advantage that it is insensitive to the shape of the baseline function.

The power of either method will depend critically on the association measure used in the analysis. Both models are built on the assumption that the rate of transmission between individuals is proportional to the association between them, and, if our interest is in testing for the presence or absence of social transmission, an association measure should be chosen for which this is likely to be true. We suggest that researchers utilise the association measure that is most relevant to the experimental context. For example, in the analysis of the diffusion of foraging task solutions (Boogert et al., 2008) presented above, a measure of association that reflects how often individuals feed together might have been preferable to the general proximity measure that was used. Note that the estimated effect of social transmission depends on the scaling of the association measures used, so if the effect of social transmission is compared between populations or species, either the same association measure needs to be used, or a case needs to be made that each association measure quantifies opportunities for social learning on a common scale. Franz and Nunn (2009) suggest an alternative approach: that different measures of association, reflecting different social and individual variables, can be used to fit separate NBDA's in order to identify which factors are important in determining diffusion dynamics. A third possible future application of NBDA is to use order or time of acquisition data to infer network structure. This could be of use in cases where it is known or assumed that behaviour is transmitted

socially, but social transmission-relevant association data is difficult to acquire. For example, in humpback whales (*Megaptera novaeangliae*) novel vocalisations are easily recorded, but association data is likely to be difficult to obtain in high latitudes (Noad et al., 2000).

There is clearly scope for a far more extensive investigation of how network structure influences both the overall rate of social transmission (Franz and Nunn, 2009) and the power of *OADA* and *TADA* to detect it. The simulations presented here must be viewed as a relatively crude first step. Nonetheless they are sufficient to show that network structures that promote social transmission (e.g. where all individuals are connected) are not necessarily the same as those that make it more likely to be detected, especially by *OADA*. Consequently, if researchers are to use these methods to make comparisons of the levels of social transmission between groups or species, which might have different network structures, we recommend that they obtain power estimates or (preferably) confidence intervals for the social transmission effect, rather than relying solely on presence/absence arguments based on hypothesis tests.

As discussed above, if all individuals have equal opportunity to learn from each other, *OADA* will have no power to detect social learning. In *TADA*, this situation can be modelled by setting all associations to 1, in which case *TADA* is effectively reduced to a diffusion curve analysis, since it is only sensitive to the acceleratory effect that an increasing number of informed individuals has on the rate of acquisition. However, in principle, our extended version of *TADA* may constitute an improved method for diffusion curve analysis (*DCA*), since it can statistically control for individual-level variables, which might otherwise obscure the underlying pattern.

The sensitivity of *TADA* to acceleration in the rate of acquisition could also be seen as a weakness. It has been noted that *DCA* is vulnerable to false positives if the latency to acquire a trait by social learning has a unimodal distribution (Reader, 2004), and *TADA* is also vulnerable under these circumstances. A unimodal distribution of latencies can arise if the process of trait acquisition has multiple steps, each of which is completed at a similar constant rate (Kendal, 2003). For example, to solve a foraging task an individual might first have to approach the task, and then interact with it in an appropriate way. If each of these component processes occurs at a similar constant rate, the overall latency to solve the task socially would have an approximately gamma distribution with shape parameter $k=2$, which would in turn result in an apparent acceleration in the rate of acquisition.

Though the models presented here, and the original *TADA* presented by Franz and Nunn (2009), assume a linear relationship between association and rate of social transmission, the methods could be adapted to accommodate other models of social transmission. For instance, the models could be refined to detect social transmission from the spatial spread of a behavioural trait through time (e.g. Fisher and Hinde, 1949). Here one merely needs to propose a relationship between the rate of transmission and the distance between individuals. If this is linear, or the distances can be transformed to linearise the relationship, researchers can use the above methods to fit the model.

The possibility that NBDA might allow us to infer something about the mechanism of social transmission is an issue worth pursuing. Given the fact that currently the ability to detect specific social learning mechanisms is restricted to the experimental laboratory, a method that could infer learning mechanisms from diffusion data could be extremely valuable. Above we suggested that if social transmission operates indirectly through social influences such as local enhancement, the multiplicative model is likely to provide a better fit to the data. In contrast, we suggest that if social transmission operates directly as an

independent learning process, such as imitation, the additive model might provide a better fit. However, such findings should only be taken as suggestive of mechanism at this stage, since there are a number of issues that might complicate this apparent dichotomy. For instance, if asocial and social learning ability covary between individuals, the multiplicative model might fit the data well even if the mechanism is additive in nature. Future extensions of NBDA could investigate these issues by including the effect of individual-level variables on the rate of social transmission (s). There is also the possibility that a number of social transmission processes, both direct and indirect, might operate in parallel. Either of these processes might result in a lack of resolution between multiplicative and additive models, as observed in our reanalysis of Boogert et al.'s data.

There are further improvements that can be made to the models in their current form. As it stands, if the models are fit to multiple diffusions involving the same individuals, they assume that the rate of acquisition by the same individual on different tasks is independent, conditional on the variables included in the model. In principle, this assumption could be dropped by incorporating a random effect for individuals. However, this is currently only implemented for the multiplicative *OADA* method, using our `multiCoxFit` function (see ESM: Additional Information, part A), which fits a Cox Proportional Hazards model with 'frailty' or 'cluster' terms (Therneau and Grambsch, 2001). A more general model would allow the user to specify a correlation structure between the rates of acquisition, for example, a spatial correlation structure (cf. Pinheiro and Bates, 2000). In the spatial analysis described above, this might allow us to control for the fact that two proximate individuals acquire the trait at a similar time because they have similar access to the resources necessary for trait acquisition.

NBDA appears to be a relatively novel approach to the statistical analysis of network data. Statistical methods have been developed to investigate properties of flow through networks, such as telecommunication interactions and traffic flow on roads and the internet (Kolaczyk, 2009). In contrast to NBDA, such models are more concerned with estimation of the strength of connections in the network, rather than testing for the presence of flow against an alternative hypothesis. In addition, such models assume that flow involves the continued transfer of material between nodes, rather than the switching of nodes from one state to another that is a feature of NBDA. In this respect, NBDA bears more resemblance with epidemiological models of the spread of a disease (e.g. Keeling, 1999) or the spread of rumours and fashions (Newman et al., 2006). However, such models assume that disease or information spreads through connections in the network, and usually aim to investigate theoretically the effect of network structure on the dynamics of spread (e.g. Meyers et al., 2006). In contrast, NBDA aims to test whether trait acquisition does spread through a given network, given real data on network connections and the pattern of trait acquisition. We are not aware of any equivalent epidemiological models that allow statistical inference about the transmission process based on an observed network (see Kolaczyk, 2009, p. 279).

Nonetheless, existing network models (e.g. Newman et al., 2006) could be used to investigate the effect of network structure on the spread of a behavioural trait as a result of social transmission. However, modifications might be necessary. For example, epidemiological models usually assume individuals move from 'susceptible' to 'infected' and then 'recovered' categories, sometimes then moving back to the 'susceptible' category (Watts, 1999). Whilst 'naïve' and 'informed' categories correspond closely to 'susceptible' and 'infected' categories, there is no obvious role for a 'recovered' category in the diffusion of many behavioural traits. A move back to the 'susceptible' category

is only applicable if individuals forget the behavioural trait. In addition, we have assumed that a behavioural trait can arise spontaneously in an individual through asocial learning, a feature which is absent from epidemiological models. Watts' (2002) model of information cascades in networks suggests another way in which NBDA could be formulated. In his model, individuals adopt a trait when they are connected to a threshold number of individuals displaying that trait. NBDA could be modified to investigate the factors that make an individual more likely to be an early adopter (low threshold), or one of the early (medium threshold) or late majority (high threshold).

Currently, methods for analysing diffusion data tend to assume that individuals fall into one of two binary categories, 'naïve' or 'informed', and that both social transmission and asocial learning result in a transition from the naïve to the informed state. Linked to this is the assumption that all informed individuals demonstrate the trait at the same rate once they are informed. *OADA* and *TADA* are no exceptions to these assumptions. In many cases the reduction to 'naïve' and 'informed' categories is a useful simplification that enables us to model social transmission in a relatively straightforward manner. However, it is worth noting that there may be some cases where this simplification is not appropriate, and that both *OADA* and *TADA* might fail to adequately model the underlying process. An individual's rate of performance of a trait is, in reality, a complex function of its own history of trait performance, observation and reward. Accordingly, we envisage that the process of acquisition may sometimes be better captured by a learning rule, such as Rescorla–Wagner (e.g. Kendal et al., 2009). However, this would make modelling a diffusion a much more challenging task, especially when data is limited.

Nonetheless, we envisage that the novel *OADA* method presented here, as well as Franz and Nunn's (2009) *Network-Based Diffusion Analysis* and our *TADA* extensions of it, will provide a useful toolkit for those wishing to detect and quantify social transmission in networks of animals, in captivity and the field. We hope that these methods will rejuvenate interest in collecting and analysing diffusion data, and add statistical rigour to the study of social transmission and culture be it in nonhuman animals or in humans.

Acknowledgements

W.H. was supported by a BBSRC grant (BB/D015812/1), N.B. by a McGill Milton Leong Fellowship and K.N.L. by grants from the BBSRC (BB/C005430/1 and BB/D015812/1) and an ERC Advanced Grant (EVOCULTURE, Ref. 232823). We would like to thank the members of the Laland Lab Journal Club for useful comments on an earlier draft, and to Laurel Fogarty, Tess Hanrahan and Joel Higgitt for 'test-driving' the R code.

Appendix. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jtbi.2010.01.004.

References

- Boogert, N.J., Reader, S.M., Hoppitt, W., Laland, K.N., 2008. The origin and spread of innovations in starlings. *Anim. Behav.* 75, 1509–1518, doi:10.1016/j.anbehav.2007.09.033.
- Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*, 2nd ed. Springer, New York.
- Coussi-Korbel, S., Fragaszy, D.M., 1995. On the relation between social dynamics and social learning. *Anim. Behav.* 50, 1441–1453.
- Crawley, M.J., 2002. *Statistical Computing: An Introduction to Data Analysis Using S-plus*. Wiley, Chichester, UK.

- de Vries, H., Stevens, J.M.G., Vervaecke, H., 2006. Measuring and testing the steepness of dominance hierarchies. *Anim. Behav.* 71, 585–592.
- Fisher, J., Hinde, R.A., 1949. The opening of milk bottles by birds. *Br. Birds* 42, 347–357.
- Franz, M., Nunn, C.L., 2009. Network-based diffusion analysis: a new method for detecting social learning. *Proc. R. Soc. B-Biol. Sci.* 276, 1829–1836, doi:10.1098/rspb.2008.1824.
- Galef Jr., B.G., Allen, C., 1995. A new model system for studying behavioral traditions in animals. *Anim. Behav.* 50, 705–717.
- Giraldeau, L.-A., Lefebvre, L., 1987. Scrounging prevents cultural transmission of food-finding behavior in pigeons. *Anim. Behav.* 35, 387–394.
- Heyes, C.M., 1994. Social-learning in animals: categories and mechanisms. *Biol. Rev.* 69, 207–231.
- Hoppitt, W., Laland, K.N., 2008. Social processes influencing learning in animals: a review of the evidence. *Adv. Study Behav.* 38, 105–165, doi:10.1016/S0065-3454(08)00003-X.
- Keeling, M.J., 1999. The effects of local spatial structure on epidemiological invasions. *Proc. R. Soc. B-Biol. Sci.* 266, 859–867.
- Kendal, J.R., 2003. Ph.D. Thesis, University of Cambridge.
- Kendal, R.L., Kendal, J.R., Hoppitt, W., Laland, K.N., 2009. Identifying social learning in animal populations: a new 'option-bias' method. *PLoS ONE* 4, e6541, doi:10.1371/journal.pone.0006541.
- Kolaczyk, E.D., 2009. *Statistical Analysis of Network Data: Methods and Models*. Springer, New York.
- Krützen, M., Mann, J., Heithaus, M.R., Connor, R.C., Bejder, L., Sherwin, W.B., 2005. Cultural transmission of tool use in bottlenose dolphins. *Proc. Natl. Acad. Sci. USA* 102, 8939–8943.
- Laland, K.N., 2004. Social learning strategies. *Learn. Behav.* 32, 4–14.
- Laland, K.N., Galef Jr., B.G., 2009. *The Question of Animal Culture*. Harvard University Press, Cambridge, MA, USA.
- Laland, K.N., Hoppitt, W., 2003. Do animals have culture? *Evol. Anthropol.* 12, 150–159, doi:10.1002/evan.10111.
- Laland, K.N., Janik, V.M., 2006. The animal cultures debate. *Trends Ecol. Evol.* 21, 542–547, doi:10.1016/j.tree.2006.06.005.
- Laland, K.N., Kendal, J.R., 2003. What the models say about social learning. In: Frigaszy, D., Perry, S. (Eds.), *The Biology of Traditions: Models and Evidence*. Cambridge University Press, Cambridge, UK, pp. 33–55.
- Leadbeater, E., Chittka, L., 2007. The dynamics of social learning in an insect model, the bumblebee (*Bombus terrestris*). *Behav. Ecol. Sociobiol.* 61, 1789–1796, doi:10.1007/s00265-007-0412-4.
- Lefebvre, L., 1995a. The opening of milk bottles by birds—evidence for accelerating learning rates, but against the wave-of-advance model of cultural transmission. *Behav. Process.* 34, 43–53.
- Lefebvre, L., 1995b. Culturally-transmitted feeding-behaviour in primates. *Primates* 36, 227–239.
- Manly, B.F.J., 2007. *Randomization, Bootstrap and Monte Carlo Methods in Biology*, third ed. Chapman and Hall/CRC, London.
- McGrew, W.C., 1998. Culture in nonhuman primates? *Annu. Rev. Anthropol.* 27, 301–328.
- Meyers, L.A., Newman, M.E.J., Pourbohloul, B., 2006. Predicting epidemics on directed contact networks. *J. Theoret. Biol.* 240, 400–418.
- Morgan, B.J.T., 2009. *Applied Stochastic Modelling*, second ed. CRC, London.
- Morrell, L.J., Croft, D.P., Dyer, J.R.G., Chapman, B.B., Kelley, J.L., Laland, K.N., Krause, J., 2008. Association patterns and foraging behaviour in natural and artificial guppy shoals. *Anim. Behav.* 76, 855–864, doi:10.1016/j.anbehav.2008.02.015.
- Newman, M., Barabasi, A.-L., Watts, D.J., 2006. *The Structure and Dynamics of Networks*. Princeton University Press, Oxford, UK.
- Noad, M.J., Cato, D.H., Bryden, M.M., Jenner, M.N., Jenner, K.C.S., 2000. Cultural revolution in whale songs. *Nature* 408 (6812), 537.
- Perry, S., Manson, J.H., 2003. Traditions in monkeys. *Evol. Anthropol.* 12, 71–81, doi:10.1002/evan.10105.
- Pinheiro, J.C., Bates, D.M., 2000. *Mixed-effects Models in S and S-Plus*. Springer, New York.
- Reader, S.M., 2004. Distinguishing social and asocial learning using diffusion dynamics. *Learn. Behav.* 32, 90–104.
- Rendell, L., Whitehead, H., 2001. Culture in whales and dolphins. *Behav. Brain Sci.* 24 (2), 309–324.
- Therneau, T.M., Grambsch, P.M., 2001. *Modeling Survival Data: Extending the Cox Model*. Springer, New York.
- Thorpe, W.H., 1956. *Learning and Instinct in Animals*. Methuen, London.
- van Schaik, C.P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C.D., Singleton, I., Suzuki, A., Utami, S.S., Merrill, M., 2003. Orangutan cultures and the evolution of material culture. *Science* 299 (5603), 102–105.
- Watts, D.J., 1999. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press, Oxford, UK.
- Watts, D.J., 2002. A simple model of global cascades on random networks. *PNAS* 99, 5766–5771, doi:10.1073/pnas.082090499.
- Weisberg, S., 1980. *Applied Linear Regression*. Wiley, Chichester, UK.
- Whiten, A., Goodall, J., McGrew, W.C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C.E.G., Wrangham, R.W., Boesch, C., 1999. Cultures in chimpanzees. *Nature* 399 (6737), 682–685.
- Whiten, A., Horner, V., de Waal, F.B.M., 2005. Conformity to cultural norms of tool use in chimpanzees. *Nature* 437 (7059), 737–740, doi:10.1038/nature04047.