1 **Movement and conformity interact to establish local behavioural traditions in**

2 **animal populations**

3

4 Marius Somveille[a,b]*, Josh A. Firth[a,c], Lucy M. Aplin[a,d], Damien R. Farine[a,e,f], Ben C.

5 Sheldon[a], Robin N. Thompson[g,h,i]

6

7 [a] Edward Grey Institute, Department of Zoology, University of Oxford, South Parks Road, Oxford

8 OX1 3PS, UK.

9 [b] Max Planck–Yale Center for Biodiversity Movement and Global Change, Yale University, 165

10 Prospect Street, New Haven, CT 06520-8105, USA.

11 [c] Merton College, Merton Street, University of Oxford, Oxford, OX1 4JD

12 [d] Cognitive and Cultural Ecology Research Group, Max Planck Institute for Ornithology,

13 Radolfzell, Germany.

14 [e] Department of Collective Behaviour, Max Planck Institute for Ornithology, Universitätsstrasse

15 10, 78457 Konstanz, Germany.

16 [f] Chair of Biodiversity and Collective Behaviour, Department of Biology, University of Konstanz,

17 Universitätsstrasse 10, 78457 Konstanz, Germany.

18 [g] Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

19 [h] Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road,

20 Oxford OX2 6GG, UK.

21 [i] Christ Church, University of Oxford, St. Aldates, Oxford OX1 1DP, UK.

22

23 * Corresponding author:

24 marius@somveille.com

25 **ABSTRACT**

26 The social transmission of information is critical to the emergence of animal culture.

27 Two processes are predicted to play key roles in how socially-transmitted information

28 spreads in animal populations: the movement of individuals across the landscape and

29 conformist social learning. We develop a model that, for the first time, explicitly

30 integrates these processes to investigate their impacts on the spread of behavioural

31 preferences. Our results reveal a strong interplay between movement and conformity

32 for determining whether local traditions establish across a landscape or whether a

33 single preference dominates the whole population. The model is able to replicate a real-

34 world cultural diffusion experiment in great tits *Parus major*, but also allows for a range

35 of predictions for the emergence of animal culture under various initial conditions,

36 habitat structure and strength of conformist bias to be made. Integrating social

37 behaviour with ecological variation will be important for understanding the stability and

38 diversity of culture in animals.

39

40 **INTRODUCTION**

41 The social transmission of information plays a central role in the lives of many animal

42 species [1–3]. Social learning via observation of, or interaction with, other individuals is

43 an efficient mechanism for acquiring information about the environment, leading to

44 adaptive adjustments of behavioural responses [4,5]. The transmission of information

45 through social networks can lead to the emergence of regional variations in behaviour

46 that are stable through time (called local cultures or traditions; [6–10]). However, we

47 still have little mechanistic understanding of the conditions under which local cultures

48 can emerge. Understanding how ecological, cognitive and social processes determine the

49 spread of information between individuals in wild populations is crucial if we want to

50 discern the conditions under which information spreads and local traditions emerge.

51 A key ecological process that is likely to affect the spread of information is movement.

52 First, movement of animals between discrete groups or sub-populations is expected to

53 accelerate information spread across the whole population. Second, moving individuals

54 can potentially import different behaviours into local groups or sub-population [11,12].

55 How individuals move in a landscape, itself influenced by a range of factors such as

56 habitat structure [13] and demography [14], is therefore likely to shape the dynamics of

57 behaviours in natural populations.

58   One of the main socio-cognitive factors thought to affect the emergence of culture is
59   conformity [15]. Conformist social learning is here defined as positive frequency
60   copying, where individuals are disproportionately likely to adopt the most common
61   behavioural trait [16]. Importantly, if individuals exhibit conformist learning, a single
62   socially learnt behavioural preference might fix in a group, remain stable over time and
63   be resistant to invasion by alternative variants, leading to the establishment and
64   persistence of group-specific traditions. Several experimental studies have therefore
65   suggested that conformity plays an important role in the establishment and stability of
66   local traditions in various animal species [11,12,17,18]. However, the interplay between
67   social learning biases such as conformity, and the ecological factors that determine the
68   context in which such learning takes place, has rarely been studied.

69   Here, we investigate how movement and conformity interact to shape the spread of
70   socially-transmitted information and the establishment of local traditions in animal
71   populations. We first develop a theoretical spatially-explicit model of the spread of
72   behavioural preference in which a puzzle (representing a novel foraging resource) can
73   be solved in one of two ways. The population is assumed to be composed of several
74   spatially distinct sub-populations, and each individual can either be unable to solve the
75   puzzle, or solve the puzzle and prefer one of the two solutions. Individuals can learn the
76   behaviour from each other, with a conformist bias guiding which preference they learn,
77   and move between sub-populations. We then use this model to investigate the
78   conditions under which local cultural traditions emerge. We consider simple scenarios
79   in which there are only two or three sub-populations. In addition, we test the model's
80   ability to replicate a real-world cultural diffusion experiment [11], in which alternative
81   novel foraging techniques – consisting of opening a bi-directional door puzzle-box by
82   sliding it either left or right to access food – were introduced in several wild sub-
83   populations of great tits *Parus major*. The spread of these foraging behaviours in the
84   population was monitored, revealing that the behaviour was socially transmitted with a
85   significant conformist bias, and that local foraging traditions established within the
86   population. As we show, our modelling approach can recreate the empirical results of
87   the experiment conducted by Aplin et al. [11], and predict the conditions under which
88   such local traditions are likely to establish and persist in the population.

89

90

RESULTS

The baseline model

Our model integrates the movement of individuals across the landscape and the social process of transmission of information between individuals, including a conformist bias. In the baseline version of the model, the spread of behavioural preferences for solutions to a puzzle (solutions s_1 and s_2) occurs in an environment with a metapopulation structure composed of two connected patches, each one hosting a sub-population (see details in Materials and Methods). These sub-populations are composed of naïve individuals (i.e. individuals who do not know how to perform the behaviour whose spread is being modelled) and knowledgeable individuals, or innovators (i.e. individuals who know how to perform one of the behaviours). At the start of each numerical solution of the model (hereafter, "simulation"), innovators are introduced in the system (initially only comprised of naïve individuals), with different behavioural preferences in each sub-population. The spread of behavioural preferences is then simulated for 150 time-steps.

We found that the emergence of contrasting local traditions (i.e. the situation where, at the end of the simulation, one sub-population is dominated by individuals with one behavioural preference while the other sub-population is dominated by individuals with the alternative behavioural preference) strongly depends on the relationship between the strength of conformity and the rate of movement of individuals between patches, determined by parameters λ and m respectively (see details in Materials and Methods). When conformity was not included in the model (i.e. $\lambda = 1$), the pattern that always emerged regardless of other parameter values (except when the difference in size between the two sub-populations was very large) was a mixture of behavioural preferences in both patches (Figs 1 and S1). When conformity was strong relative to the movement rate, local traditions established and remained stable (Fig 1). When conformity was weak relative to the movement rate, one of the two behavioural preferences dominated the whole system (Fig 1). In the latter case, which preference dominated depended on the relative sizes of the sub-populations. When the size of the sub-population (i.e. the number of naïve individuals in the patch) in which an innovator with a given behavioural preference was initially introduced was larger than the size of the other sub-population, then that preference came to dominate at the end of the

4

123   simulation (Fig S1). When the two sub-populations had exactly the same size, which
124   behavioural preference dominated at the end of the simulation appeared to be random
125   and very sensitive to the precise parameter values (Fig S1). Finally, the main effect of the
126   learning rate was to determine how much stronger/weaker conformity must be relative
127   to the magnitude of the movement rate to obtain these patterns: the lower the learning
128   rate, the stronger conformity had to be relative to the movement rate to see local
129   traditions establish and stabilise (Figs 1 and S1).

130
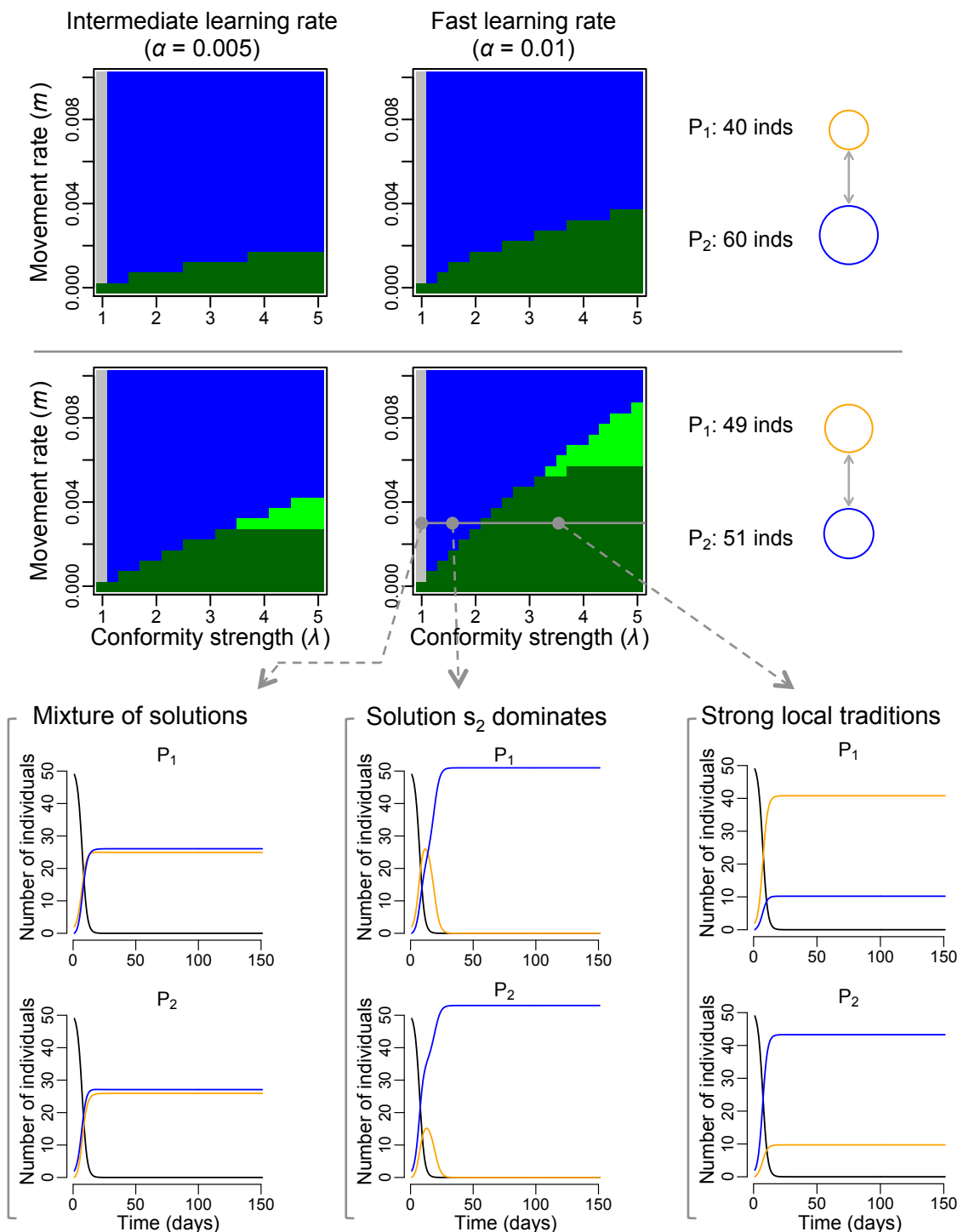
Fig 1. Local traditions emerge when conformity is strong relative to the movement rate. This panel shows the model outputs for the baseline model with two patches. At the start of every simulation, patch 1 contained two innovators using solution s_1 and contained P_1 naïve individuals, while patch 2 contained two innovators using solution s_2 and contained P_2 naïve individuals. Each pixel in the phase diagrams corresponds to a simulation run with the corresponding parameter values, and the colour of the pixel indicates the emerging pattern after 150 days: grey = mixture of solutions in every patch, blue = solution s_2 dominated the whole system, light green = weak local traditions, dark green = strong local traditions. The two columns of phase diagrams represent different learning rates: intermediate ($\alpha = 0.005$) and fast ($\alpha = 0.01$). The two rows of phase diagrams represent a different configuration of patch sizes, reflecting the differences in number of naïve individuals at the start of the simulation (U_0) occurring in each patch. Three examples of the evolution of the number of naïve individuals (black curve) and number of solvers using solution s_1 (orange curve) and solution s_2 (blue curve) in each patch, are shown for fixed movement rate ($m = 0.003$), learning rate ($\alpha = 0.01$) and patch sizes configuration, but with varying conformity strength. When no conformity bias was included ($\lambda = 1$), patches contained a mixture of solutions; when conformity was relatively weak ($\lambda = 1.5$), solution s_2 (that seeded in the larger population) ended up dominating in both patches; and when conformity as relatively strong ($\lambda = 3.5$), local traditions emerged.

Simple environmental setting

To examine the role of space in the spread of behavioural preferences, we extended the baseline model to a simple environmental setting with three spatially distinct patches, thereby effectively adding an extra patch containing no innovators at the start of the simulation (see Materials and Methods for details). Once again, when conformity was not included in the model ($\lambda = 1$), the pattern that always emerged regardless of other parameter values was a mixture of behavioural preferences in all patches (Fig 2). Similarly to the baseline model, there was a strong relationship between movement rate and conformity strength; when conformity was strong relative to the movement rate, local traditions established and were stable (i.e. two sub-populations were dominated by individuals with one behavioural preference while the third sub-population was dominated by individuals with the alternative behavioural preference; Fig 2), and when conformity was weak relative to the movement rate, one of the two behavioural preferences ultimately dominated the entire system (Fig 2). In the latter case, which preference dominated depended on both the relative sizes of the sub-populations as well as the spatial configuration (i.e. the distances separating the sub-populations, which determines the relative movement rates of individuals between pairs of sub-

6

populations). If the size of the sub-population in which innovators with a given behavioural preference were initially introduced was larger than the size of the sub-population in which innovators with the alternative behavioural preference were introduced, and all sub-populations were equidistant, then the former preference was predominant at the end of the simulation (Fig S2). Also, increasing the distance separating a sub-population in which innovators with a given behavioural preference were initially introduced from the two other sub-populations (i.e. creating unequal movement rates between patches) resulted in this preference not being able to dominate the system at the end of the simulation (Fig S2).

Increasing the distance between patches also affected how much stronger/weaker conformity must be in comparison to the movement rate to generate the emerging patterns: the smaller the distance, the stronger conformity had to be relative to the movement rate for local traditions to establish and stabilise (Figs 2 and S2). When both patch size and spatial configuration acted in opposite directions, then either behavioural preference could ultimately dominate depending on the relationship between the strength of conformity and the magnitude of the movement rate. This may arise when the size of the sub-population in which innovators with a given behavioural preference were initially introduced was larger than the size of the sub-population in which innovators with the alternative behavioural preference were initially introduced, but this sub-population was also further away form the two others (see illustration in Fig 2). When conformity was relatively strong, but not so strong as to generate local traditions, then the behavioural preference that was not released in the larger, further away sub-population ended up dominating the system, otherwise the alternative preference dominated at the end of the simulation (Fig 2).

Unexpected results were also observed when both sub-populations in which innovators were introduced at the start of the simulation had the same size, which is larger than the third sub-population, with one of these two sub-populations being slightly further away from the two others. In contrast to what might have been expected based on other results, when conformity was very weak compared to the movement rate, then the behavioural preference of the innovator that was initially introduced in the most distant sub-population ended up being the most prevalent in the system (Fig S2, second plot of fourth row, in blue).
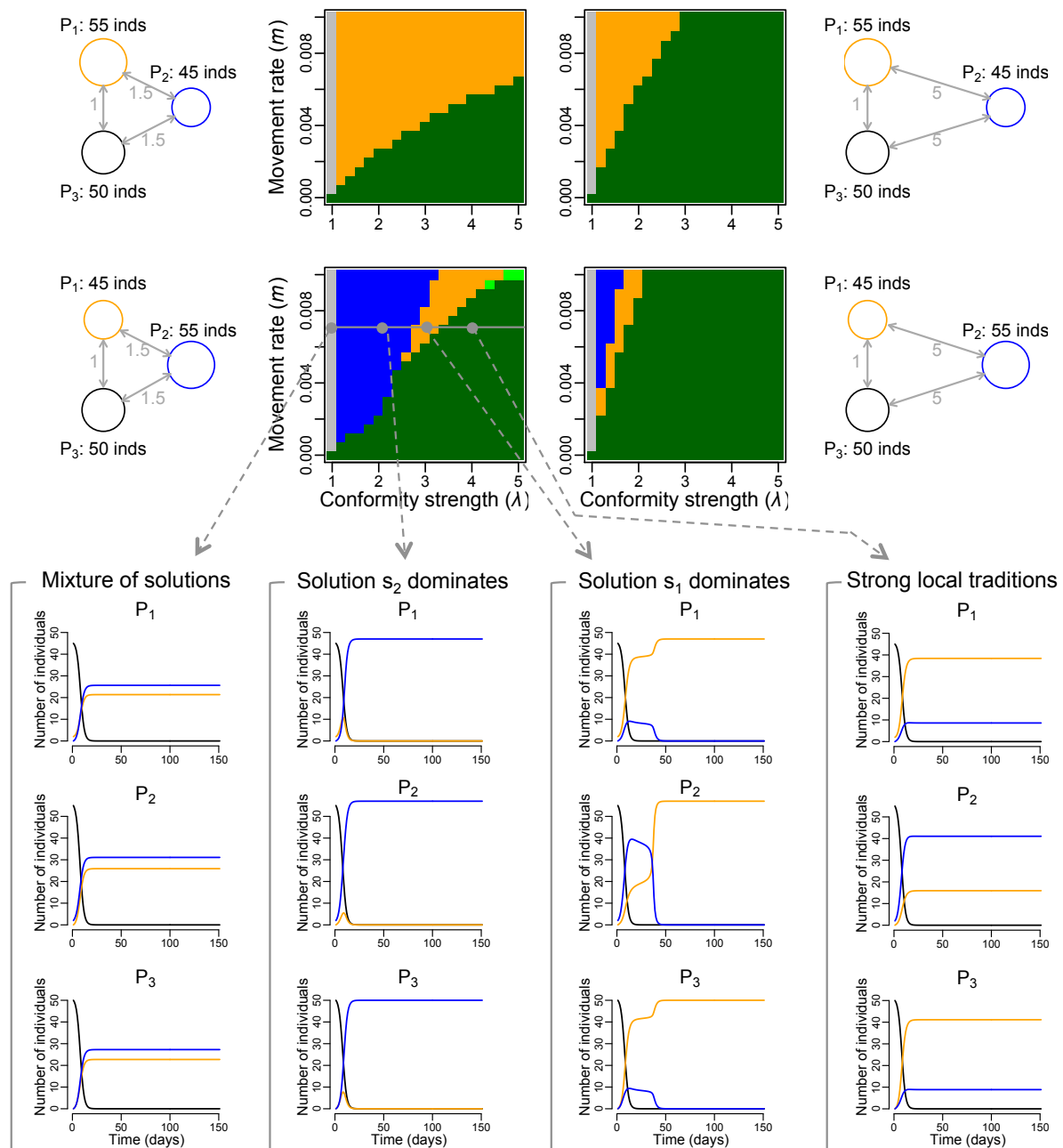
Fig 2. Space has a complex effect on how conformity and movement lead to either the emergence of local traditions or the domination of a single solution. This panel shows the model outputs in the three-patch case. At the start of every simulation, patch 1 contained two innovators using solution s_1 and contained P_1 naïve individuals, while patch 2 contained two innovators using solution s_2 and contained P_2 naïve individuals, and patch 3 contained only naïve individuals. Each pixel in the phase diagrams corresponds to a simulation run with the corresponding parameter values, and the colour of the pixel indicates the emerging pattern after 150 days: grey = mixture of solutions in every patch, blue = solution s_2 dominated the whole system, orange = solution s_1 dominated the whole system, light green = weak local traditions, dark green = strong local traditions. The two columns of phase diagrams represent different spatial configurations: a different distance between patch 2 and the other two patches (whose

distance separating them was set to 1), which results in different relative migration rates between pairs of patches. The two rows of phase diagrams represent a different configuration of patch sizes, reflecting differences in the number of naïve individuals (inds) at the start of the simulation (U_0) that occurred in patches P_1 and P_2. Four examples of the evolution of the number of naïve individuals (black curve) and number of solvers using solution s_1 (orange curve) and solution s_2 (blue curve) in each patch, are shown for fixed movement rate (m = 0.007), and spatial and patch sizes configuration, but with varying conformity strength. When no conformity bias was included ($\lambda = 1$), patches contained a mixture of solutions; when conformity was relatively weak ($\lambda = 2$), solution s_2 ended up dominating in every patch; when conformity was intermediate ($\lambda = 3$), solution s_1 ended up dominating in every patch; and when conformity was relatively strong ($\lambda = 4$), local traditions emerged.

Realistic environmental setting

To examine the role of habitat structure and the ecological process of movement in a realistic setting, we extended the baseline model to represent the great tit population of Wytham Woods, which has been the subject of a long-running study, and the site of a recent cultural diffusion experiment [11]. Running the model of spread of behavioural preference for this real-world animal population in its natural environment (see Materials and Methods for details) yielded results that were consistent with those for the baseline model and its extension to three patches. Three possible patterns emerged at the end of the simulation depending on parameter values: (1) a mixture of behavioural preferences in every sub-population when conformity was not included in the model (i.e. $\lambda = 1$), (2) domination of one behavioural preference across the population when conformity was weak relative to the magnitude of the movement rate, and (3) the establishment of local traditions when conformity was strong relative to the magnitude of the movement rate (i.e. some sub-populations were dominated by individuals with one behavioural preference while the rest were dominated by individuals with the alternative preference, Fig 3). Increasing the degree of fragmentation of the landscape (by only allowing individuals to move through contiguous forest, as opposed to moving along straight direct paths between patches) affected how much stronger/weaker conformity must be compared to the magnitude of the movement rate to generate the different emerging patterns (Fig 3). Increasing the learning rate had a similar effect to increasing the degree of landscape fragmentation (Fig 3). That is, for local traditions to establish and stabilise, conformity had to be

249 stronger relative to movement rates when either the habitat was less fragmented or
250 when the learning rate was slower (Fig 3).
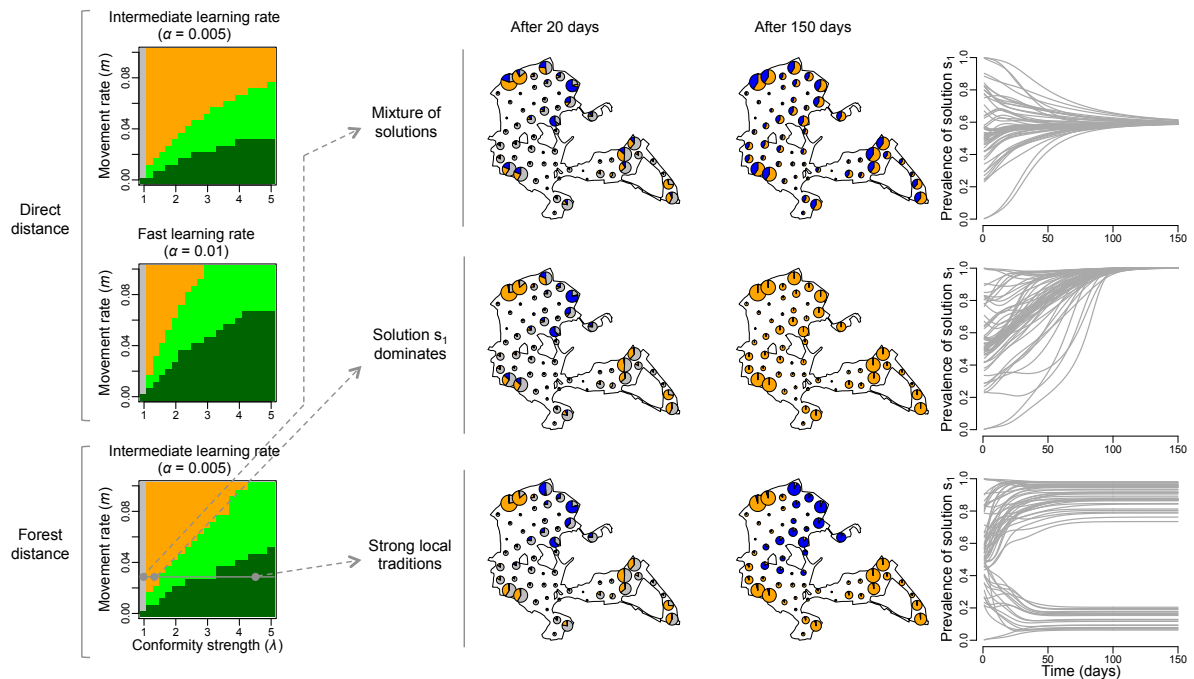
251

252 **Fig 3. Predictions for the spread of information in great tit cultural diffusion**
253 **experiment in Wytham Woods.** This analysis was designed to replicate the conditions
254 of the cultural diffusion experiment performed by Aplin et al. [11], in which two
255 alternative foraging techniques were introduced in the great tits population of Wytham
256 Woods in the United Kingdom, and their spread through time monitored. Each pixel in
257 the phase diagrams corresponds to a simulation run with the corresponding parameter
258 values, and the colour of the pixel indicates the emerging pattern after 150 days: grey =
259 mixture of solutions in every patch, blue = solution s_2 dominates the whole system,
260 orange = solution s_1 dominates the whole system, light green = weak local traditions,
261 light green = strong local traditions. The two top phase diagrams (with intermediate and
262 fast learning rates) correspond to model runs in which the distance separating patches
263 is the Euclidean distance (direct distance), while the phase diagram at the bottom
264 corresponds to model runs using the shortest route through the forest separating pairs
265 of patches (forest distance). Three examples of the evolution of the prevalence of
266 solution s_1 among solvers are shown for fixed movement rate (m = 0.03), learning rate
267 (α = 0.005) but with varying conformity strength. The maps represent the extent of
268 Wytham Woods and the location of the 60 feeders. Each feeder is represented by a pie
269 chart indicating the number of naïve individuals (grey), solvers using solution s_1
270 (orange) and solvers using solution s_2 (blue), and the size of the pie chart is proportional
271 to the total number individuals occurring around the feeder. When no conformity bias is
272 included (λ = 1) it results in a mixture of solutions in every patch; when conformity is
273 relatively weak (λ = 1.3) solution s_1 ends up dominating in every patch; and when
274 conformity is relatively strong (λ = 4.5) local traditions emerged.

275 The emergence of patterns depended on precisely where innovators with preferences
276 for solutions s_1 and s_2 occurred, particularly when conformity was weak relative to the
277 magnitude of the movement rate (Fig 4). When this was the case (i.e. $\lambda = 1.2$ and
278 $m = 0.02$), 62% of simulations in which initial conditions were randomised resulted in
279 the emergence of local traditions, and the rest of the simulations resulted in one
280 behavioural preference dominating the whole system (with half of those simulations
281 leading to solution s_1 to be predominant, and the other half of simulation with solution
282 s_2 dominating). Which behavioural preference ended up dominating was strongly
283 affected by the sizes of the pools of naïve individuals in contact with innovators
284 preferring each solution at the start of the simulation. If one behavioural preference
285 came to dominate the whole system, then it was likely to be the preference that was
286 initially added in the comparatively larger sub-population (Fig 4), consistent with
287 previous results. However, when conformity was strong relative to the magnitude of
288 movement rate (i.e. $\lambda = 4$ and $m = 0.02$), all the simulations resulted in the emergence
289 of local traditions. When no conformity was included (i.e. $\lambda = 1$ and $m = 0.02$), 79% of
290 simulations resulted in a mixture of behavioural preferences in every sub-population,
291 and the rest of simulations resulted in one behavioural preference dominating the whole
292 system (with half of simulations leading to solution s_1 throughout the landscape, and the
293 other half of simulations with solution s_2 predominant). Once again, the sizes of the
294 pools of naïve individuals in contact with each innovator at the start of the simulation
295 affected which pattern emerged, similarly to when a weak conformity was included (Fig
296 4).
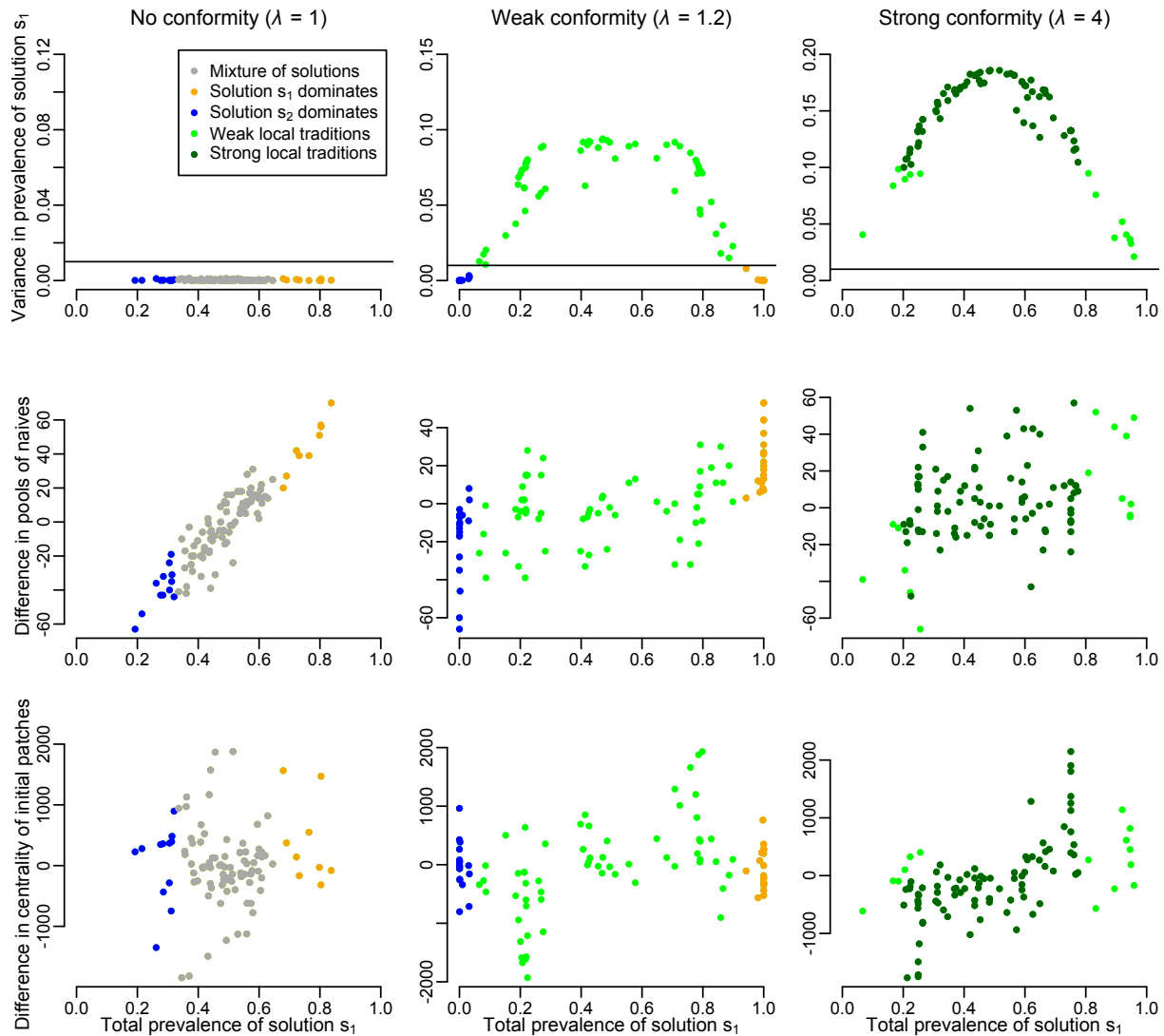
Fig 4. The outcome of the spread of information is sensitive to the initial conditions. This panel shows results for the randomisation of initial conditions for the spread of information in Wytham Woods. Each column of plots corresponds to a different conformity strength, for which 100 simulations with randomised initial conditions were run. The first row of plots indicates the values of simulations for the two summary statistics used to identify the emerging pattern. The horizontal line indicates the threshold above which local traditions were said to have emerged. The second row of plots investigates a relationship between the total prevalence of solution s_1 (one of the two summary statistics) and the difference between the pool of naïve individuals initially in contact with solution s_1 and s_2 (i.e. size of the sub-population in which innovators with solution s_1 were released at the start of the simulation minus the size of the sub-population in which innovators with solution s_2 were released). The third row of plots investigates a relationship between the total prevalence of solution s_1 and the difference in the centrality of the sub-populations in which innovators with solutions s_1 and s_2 were released at the start of the simulation. The centrality of a sub-population was computed as the median distance between itself and other sub-populations (the smaller the value the more central is the sub-population). The smaller the difference in

315 centrality, the more solution s_1 was released in a central location compared to solution
316 s_2.
317
318 **DISCUSSION**

319 Our results demonstrate the importance of the relationship between movement and
320 conformity for determining whether or not local traditions establish in animal
321 populations. First, our model indicates that a conformist bias in learning is key for the
322 emergence of local traditions, as none of our simulations in which a conformist bias was
323 not included led to the generation of local traditions (Figs 1–4; except in the two-patch
324 case when the difference in size between the two sub-populations is very large and the
325 learning rate is relatively fast, Fig S1). The importance of conformity in this scenario is in
326 line with previous hypotheses and indications from experimental results [11,12,16,18].
327 Second, we extended this finding to show that local traditions establish only when
328 conformity is relatively strong compared to the magnitude of the movement rate of
329 individuals between sub-populations. This was observed for the baseline model (two
330 patches; Fig 1) and its extension to three patches (Fig 2) as well as for the realistic
331 environmental setting of Wytham Woods (Fig 3). As highlighted for the baseline model,
332 when conformity was weak relative to the magnitude of the movement rate, moving
333 individuals could continuously invade a patch with alternative behavioural preferences
334 at a faster rate than which they could conform to the local behavioural preference in that
335 patch, thereby leading to the domination of a single behavioural preference across the
336 whole system by the end of the simulation (e.g. Figs 1–3). Since neither of the two
337 alternative behavioural preferences had a selective advantage, the solution that ended
338 up dominating was determined by the initial conditions: a given behavioural preference
339 that started in a larger pool of naïve individuals was more likely to dominate (Figs 1, 2
340 and 4). This is because it spread more quickly at the start of the simulation than the
341 alternative preference, and knowledgeable individuals moving out of that sub-
342 population therefore represented a relatively large proportion of knowledgeable
343 individuals in each sub-population that they arrived in.

344 Importantly, the spatial configuration of patches also influences the outcome of the
345 spread of information. Increasing habitat fragmentation led to more favourable
346 conditions for the establishment of local traditions by lowering the movement rate and
347 thus increasing the relative impact of conformist learning. This was observed for both

13

348 the three-patches setting (Fig 2) and in the realistic setting of Wytham Woods (Fig 3;
349 when we used *forest distance*, the fragmentation of the habitat was effectively larger
350 than when we used *Euclidean distance*). The spatial configuration of patches also affects
351 which of the two alternative behavioural preferences ultimately dominates when
352 conformity is weak relative to the magnitude of the movement rate. In the three-patch
353 setting, the behavioural preference that colonised the third patch first (in which no
354 innovators were introduced) was generally the one that ended up dominating at the end.
355 This effect was determined by how far the patches were from each other and the relative
356 sizes of the sub-populations. Interestingly, if the sub-population with the largest pool of
357 naïve individuals at the start of the simulation was also more distant from the other sub-
358 populations, which preference predominated at the end depended on the interplay
359 between conformity strength and the magnitude of movement rate (Fig 2 second row of
360 phase diagrams), as this affected which preference was better at colonising the patch
361 with no innovator.

362 Finally, a surprising effect of space was observed in a three-patch landscape in which
363 innovators were initially introduced in two large patches but where one of the two large
364 patches was located slightly further away from the other two patches. In this case, the
365 behavioural preference of the innovator in the most distant patch ended up dominating
366 the whole system when conformity was very weak relative to the magnitude of the
367 movement rate (Fig S2, second plot of fourth row, in blue). A possible explanation for
368 this result is that, with a very high movement rate relative to conformity strength, naïve
369 individuals from the patch without innovators moved *en masse* and slowed down the
370 initial spread of the behavioural preferences. This effect was less pronounced for the
371 preference introduced in the most distant patch (as movement was dependent on
372 distance) and so this preference could subsequently colonise the patch without
373 innovators faster than the alternative preference. Overall, these results highlight the
374 important effects of habitat configuration and fragmentation on the spread of culture in
375 animal populations (see also [19]), and allow for testable predictions to be made. It is
376 particularly relevant given the wide range of animal populations around the world that
377 are affected by habitat fragmentation [20,21].

378 Our model replicated the diffusion curves empirically observed in a cultural diffusion
379 experiment in great tits in Wytham Woods [11]. With the same initial conditions in our

model as in the field experiment (i.e. trained innovators released at the same locations in the landscape) and for an intermediate learning rate, the model predicted that sub-populations in which trained individuals were released should have reached a proportion of solvers approximately between 0.6–0.8 after 20 days (Fig S3; with one exception with a proportion of solvers of 0.35), and sub-populations in which no trained innovators were released to contain a proportion of solvers between 0.1–0.4 after 20 days (Fig S3). These values are very similar to the empirical results reported in the original study (Fig 1b in [11]). This supports the potential for this model to be used to make predictions about when novel behaviours could result in local traditions. These predictions could in turn be tested in cultural diffusion experiments.

Our model predicts that local traditions establish when the movement rate of individuals between sub-populations is low relative to the strength of conformity, based on a given learning rate. If movement rates are relatively high, the location where the different behaviours emerge has an important effect on the outcome. The model predicts that the centrality of location in the landscape largely does not affect the outcome but that the size of the pool of naïve individuals living there has a strong effect (Fig 4). If a conformist bias exists in the transmission of information, and for a given movement ability of the population, local traditions are more likely to establish and be well pronounced if two different behavioural preferences appear in sub-populations with similar sizes (Fig 4). These predictions have many implications for studying the emergence of behavioural traditions in animal populations in which social learning occurs. They highlight the key and often neglected role of movement, and particularly its interplay with conformist learning, as well as the importance of the initial conditions. It should therefore be interesting going forwards to test these model predictions for species with different levels of mobility – e.g. high mobility of fission-fusion bird populations [22] versus low inter-group movement rates by vervet monkeys [12] – and for various initial conditions.

In this study, we modelled a scenario in which two alternative behavioural preferences are introduced at the same time into a population of naïve individuals. This is consistent with cultural diffusion experiments. However, in natural settings, it is also likely that solutions to a foraging task might be discovered and rediscovered through repeated innovations [17]. Incorporating an asocial learning rate, whereby individuals can

spontaneously learn to solve the puzzle using a certain solution, would be an interesting, and relatively straightforward, addition to our model. However, it should have very little impact unless it is large relative to the social learning rate, or if individuals do not abandon personal preferences to conform. Future research could also extend our model to reflect other characteristics. For example, including demographic processes could be a fruitful avenue for making long-term predictions. We assumed that the total carrying capacity of the environment had been reached and that each sub-population had a constant number of individuals. However, including varying population sizes could be interesting for exploring whether or not local traditions remain stable across multiple generations. Furthermore, the model may also be useful for considering how individual-level differences interact with the emergence and spread of culture. For example, juveniles could potentially learn faster than adults, or conformity could vary across age classes [18]. Individual-level differences have recently been highlighted as being important in shaping the dynamics of collective behaviour in animal groups [23,24]. It is therefore likely that such differences could play a major role in shaping the spread of behaviours and the establishment of local traditions in natural populations. It would also be interesting to consider a stochastic version of our model, since random events soon after traditions arrive in a naïve population are likely to play an important role in determining the tradition that ends up dominating.

In summary, our results provide new insights into the interplay between the movement of individuals and conformist learning in the emergence of animal culture. By simply incorporating these two processes, our model is able to make predictions about the emergence and stability of local traditions, and allow the influence of quantities such as initial population conditions and the degree of habitat fragmentation to be tested. A major strength of the model is its generalisability. Future research could extend the model to explore the spread of animal culture for more than two behavioural preferences, other environmental settings and different time scales, and integrating individual differences and non-static sub-population demographics. Such exploration of the spread of socially-transmitted information in animal populations has the potential to provide additional insights into the conditions under which local traditions emerge and persist.

MATERIALS AND METHODS

The baseline model: The spatially-explicit model describing the spread of animal culture integrates two processes: (1) transmission of information between individuals with a conformity bias, and (2) movement of individuals between spatially distinct patches of habitat. A novel behaviour, which consists of two equally difficult, and equally rewarding, solutions to a novel foraging resource (s_1 and s_2) is introduced into a population of naïve individuals (who are unable to solve the task at the time of introduction) by adding innovators, which are individuals who know how to solve the task with a preference for either solution s_1 or s_2. The information about how to solve the novel task, along with the preference for either one of the two alternative solutions, can be socially transmitted to other individuals. The spread of the two behavioural preferences in the population is then modelled, with simulations being run for 150 days (with a daily time step). At any time, an individual is either naïve, a solver s_1, or a solver s_2. During encounters with other individuals, naïve individuals can learn from solvers, and in doing so copy their behavioural preference, with parameter α governing the magnitude of the learning rate. The rate at which naïve individuals acquire one of the two alternative solutions (s_1 and s_2) is a function of the proportion of solvers with this behavioural preference among all the solvers in the local sub-population. When individuals have a conformity bias (i.e. they are more likely to copy a specific behavioural preference than the prevalence of this preference among local solvers), which is given by the conformity parameter λ, then individuals use information about the behaviour of all other individuals in the patch when choosing which preference to acquire. The conformity parameter λ determines the strength of sigmoidality (i.e. S-shapedness) of the acquisition curve. An acquisition curve is the relationship between the prevalence of a preference for a solution in the local sub-population and the probability of adopting that preference (see equations below describing the conformist learning function L for learning s_1 and s_2). In this model, conformist learning (from naïve to solver) and conformist switching (from solving the puzzle using one solution to using the alternative solution) were modelled in the same way using the same parameters. By doing this, the likelihood of an individual learning from another is approximately independent of whether or not the individual already has a preference for either solution to the puzzle.

477 In the initial baseline version of the model, the environment is assumed to consist of two
478 patches, with each one hosting a sub-population. Within each sub-population, we
479 assume individuals mix entirely at random. The size of each sub-population is at
480 equilibrium throughout the simulation (i.e. no variation during the 150 days),
481 essentially assuming that each sub-population size is at the strict carrying capacity of
482 each patch. The movement rate of individuals between the patches decreases
483 exponentially with the distance d separating them, with a parameter m determining the
484 magnitude of the movement rate between the patches. In each patch j (where j = 1 or j =
485 2), the change of the numbers of individuals that are naïve ($U^{(j)}$), solvers with a
486 preference for solution s_1 ($S_1^{(j)}$) and solvers with a preference for solution s_2 ($S_2^{(j)}$)
487 through time is modelled using a system of differential equations. For example, the
488 change in the composition of individuals in patch j = 1 is given by:

489

$$
\begin{cases}
\dfrac{dU^{(1)}}{dt} = -\alpha\big(S_1^{(1)} + S_2^{(1)}\big)U^{(1)} - \dfrac{m\big(U^{(2)} + S_1^{(2)} + S_2^{(2)}\big)U^{(1)}}{d} + \dfrac{m\big(U^{(1)} + S_1^{(1)} + S_2^{(1)}\big)U^{(2)}}{d} \\[2em]
\dfrac{dS_1^{(1)}}{dt} = L_{S_1}^{(1)}\alpha\big(S_1^{(1)} + S_2^{(1)}\big)U^{(1)} - L_{S_2}^{(1)}\alpha\big(S_1^{(1)} + S_2^{(1)}\big)S_1^{(1)} + L_{S_1}^{(1)}\alpha\big(S_1^{(1)} + S_2^{(1)}\big)S_2^{(1)} - \dfrac{m\big(U^{(2)} + S_1^{(2)} + S_2^{(2)}\big)S_1^{(1)}}{d} + \dfrac{m\big(U^{(1)} + S_1^{(1)} + S_2^{(1)}\big)S_1^{(2)}}{d} \\[2em]
\dfrac{dS_2^{(1)}}{dt} = L_{S_2}^{(1)}\alpha\big(S_1^{(1)} + S_2^{(1)}\big)U^{(1)} - L_{S_1}^{(1)}\alpha\big(S_1^{(1)} + S_2^{(1)}\big)S_2^{(1)} + L_{S_2}^{(1)}\alpha\big(S_1^{(1)} + S_2^{(1)}\big)S_1^{(1)} - \dfrac{m\big(U^{(2)} + S_1^{(2)} + S_2^{(2)}\big)S_2^{(1)}}{d} + \dfrac{m\big(U^{(1)} + S_1^{(1)} + S_2^{(1)}\big)S_2^{(2)}}{d}
\end{cases}
$$

490

491 The equations for j = 2 are similar. In the equations above, the parameters $L_{S_1}^{(1)}$ and $L_{S_2}^{(1)}$
492 correspond to the conformist learning functions for learning solutions s_1 and s_2
493 respectively, which are functions of the prevalence of solution s_1 in the sub-population
494 ($P^{(1)}$; i.e. proportion of individuals in state s_1 among solvers in patch 1) and are defined
495 as follows:

$$
L_{S_1}^{(1)} = \frac{\left(\dfrac{P^{(1)}}{1 - P^{(1)}}\right)^{\lambda}}{1 + \left(\dfrac{P^{(1)}}{1 - P^{(1)}}\right)^{\lambda}}
$$

496

$$
L_{S_2}^{(1)} = 1 - L_{S_1}^{(1)}
$$

497 These conformist learning functions produce a sigmoidal relationship between a
498 solution's prevalence in the sub-population and the probability of adoption of that
499 behavioural preference (called acquisition curve; [25,26]; Fig S4). If λ = 1, there is no
500 conformity bias included in the model (i.e. straight 1:1 line; see Fig S4).

501 At the start of each simulation, two innovators (i.e. knowledgeable individuals) with
502 solution s_1 were added to one patch and two innovators with solution s_2 were added to
503 the other patch. We ran simulations for various conformity strengths: $\lambda \in [1, 5]$ sampled
504 every 0.1; and movement rate magnitudes: $m \in [0.0005, 0.01]$ sampled every 0.0005. To
505 investigate if results were affected by how quickly individuals learn, we ran the
506 simulations for different learning rate: $\alpha = 0.001$ (slow learning rate), $\alpha = 0.005$
507 (intermediate learning rate) and $\alpha = 0.01$ (fast learning rate). As there are only two
508 patches here, changing the distance between the patches is equivalent to changing the
509 movement rate m (see equations above), so we therefore set $d = 1$ for every simulation
510 run. To investigate the effect of patch size, we also ran simulations with different
511 numbers of naïve individuals in each patch at the start of the simulation (see Figs 1 and
512 S1).

513

514 **Simple environmental setting:** We extended the baseline model to three patches, each
515 containing a sub-population. In this case, the equations described above for the baseline
516 model were adapted for environmental settings with more than two patches (see
517 Supporting Methods). At the start of each simulation, two innovators with solution s_1
518 were added to one patch and two innovators with solution s_2 were added to another
519 patch (the third patch was assumed to initially consist of only naïve individuals). We ran
520 simulations for the same ranges of values of conformity strength (λ) and movement rate
521 (m) as for the baseline model. The learning rate was kept at the fast level ($\alpha = 0.01$) for
522 every simulation so that emerging patterns were more pronounced within the range of
523 values explored for λ and m. We also investigated the effect of patch size by running
524 simulations with the different numbers of naïve individuals in each patch at the start of
525 the simulation (see Figs 2 and S2). To investigate the effect of habitat fragmentation, we
526 varied the distance between the patch where two individuals trained to solve the puzzle
527 with solution s_2 were added and the two other patches, investigating distances 1, 1.5 and
528 5, while the distance separating the two other patches was maintained at 1 (see
529 schematics in Fig 2).

530

531 **Realistic environmental setting:** Wytham Woods, Oxfordshire, UK (51° 46' N, 01° 20'
532 W) is a 385ha broadleaf deciduous woodland surrounded by open farmland and covered
533 by an evenly-spaced grid of 60 feeders (see the map in Fig S5). This is the location where

534   Aplin et al. [11] performed the cultural diffusion experiment in great tits, introducing

535   alternative novel foraging techniques and monitoring their spread. We extended the

536   baseline model to this realistic setting, and used the adapted equations for more than

537   two patches (see Supporting Methods). We started simulations with the same initial

538   conditions as in the field study, releasing two innovators at targeted patches in a similar

539   fashion (i.e. at the same feeders; see Fig S5). We divided the landscape so that each patch

540   in our model contained one feeder. The total number of individuals across the woods

541   and relative patch size (i.e. the number of individuals in each patch around each feeder

542   in each time step) were derived from data described in [13]. We ran simulations for the

543   same ranges of values for conformity strength as described for the baseline model. We

544   investigated the following range of values for the movement rate: $m \in [0.005, 0.1]$. This

545   was different from the range of values explored for the two-patch and three-patch cases

546   for this parameter because the distances separating patches were in meters here rather

547   than in arbitrary spatial units. We modified the environment and the initial conditions to

548   investigate how these changes affected the model outcomes. First, we used two distance

549   measures between pairs of patches: *direct Euclidean distance* and *forest distance*, the

550   latter being computed as the length of the shortest route between the two patches

551   through the forest (without crossing open ground). This is known to be an ecologically

552   relevant measure of distance with regard to movement within this population [13].

553   Second, we randomised the locations of feeders where trained individuals were released

554   at the start of simulations. For three different values of conformity strength $\lambda = 1$ (no

555   conformity included), $\lambda = 1.2$ (weak conformity) and $\lambda = 4$ (strong conformity), and a

556   fixed movement rate magnitude ($m = 0.02$), we ran 100 simulations, each with a

557   random location (i.e. sub-population/patch) where two innovators with solution s_1, and

558   another random location where two innovators with solution s_2, were added at the start.

559

560   **Analysing emerging patterns:** In all model runs for every environmental setting (two

561   patches, three patches and Wytham Woods), we reported the total prevalence of

562   solution s_1 across the whole population at the end of the simulation (P_{tot}; i.e. proportion

563   of individuals with behavioural preference for solution s_1 among all solvers in all

564   patches) and the spatial variance of the final prevalence of solution s_1 in sub-populations

565   (P_{var}; i.e. variance in the proportion of individuals with behavioural preference for

solution s_1 among local solvers in each patch). These two summary statistics were used to identify the emerging patterns:

- if P_{var} > 0.1: *strong local traditions* established at the end of the simulation (i.e. some sub-populations are strongly dominated by one behavioural preference while the others are strongly dominated by the alternative preference)

- if 0.1 > P_{var} > 0.01: *weak local traditions* established at the end of the simulation (i.e. some sub-populations have a bit more of one behavioural preference while the others have a bit more of the alternative preference)

- if P_{var} < 0.01 and P_{tot} > 0.66: *solution s1 dominated* across the whole system at the end of the simulation

- if P_{var} < 0.01 and P_{tot} < 0.33: *solution s2 dominated* across the whole system at the end of the simulation

- if P_{var} < 0.01 and 0.33 > P_{tot} < 0.66: *mixture of solutions* in every sub-population

These criteria and thresholds were chosen in order to best reflect a visual identification of the emerging patterns (see examples in Fig 3).

REFERENCES

1. Dall SRX, Giraldeau L-A, Olsson O, McNamara JM, Stephens DW. Information and its use by animals in evolutionary ecology. Trends in Ecology & Evolution. 2005; 20(4): 180–193.

2. Aoki K, Feldman MW. Evolution of learning strategies in temporally and spatially variable environments: a review of theory. Theoretical Population Biology. 2014; 91: 3–19.

3. Jones TB, Aplin LM, Devost I, Morand-Ferron J. Individual and ecological determinants of social information transmission in the wild. Animal Behaviour. 2017; 129: 93–101.

4. Laland KN. Social learning strategies. Learning & Behavior. 2004; 32: 4–14.

5. Rendell L, Fogarty L, Hoppitt WJE, Morgan TJH, Webster MM, Laland KN. Cognitive culture: theoretical and empirical insights into social learning strategies. Trends in Cognitive Sciences. 2011; 15: 68–76.

6. Warner RR. Traditionality of mating-site preferences in a coral reef fish. Nature. 1988; 335: 719–721.

7. Rogers EM. Diffusion of innovations. New York: The Free Press; 1995.

599 8. Laland KN, Janik VM. The animal cultures debate. Trends in Ecology & Evolution.
600   2006; 21: 542–547.
601 9. Whiten A, Mesoudi A. Establishing an experimental science of culture: animal social
602   diffusion experiments. Philosophical Transactions of the Royal Society of London B.
603   2008; 363: 3477–3488.
604 10. Whiten A, Hinde RA, Laland KN, Stringer CB. Culture evolves. Philosophical
605   Transactions of the Royal Society of London B. 2011; 366: 938–948.
606 11. Aplin LM, Farine DR, Morand-Ferron J, Cockburn A, Thornton A, Sheldon BC.
607   Experimentally induced innovations lead to persistent culture via conformity in wild
608   birds. Nature. 2015; 518: 538–541.
609 12. Van de Waal E, Borgeaud C, Whiten A. Potent social learning and conformity shape a
610   wild primate's foraging decisions. Science. 2013; 340: 483–485.
611 13. Farine DR, Sheldon BC. Social ecology of a woodland songbird community: from
612   individual movements to the emergence of population social structure. BioRxiv.
613   2016; doi.org/10.1101/085944
614 14. Patterson TA, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J. State-space
615   models of individual animal movement. Trends in Ecology & Evolution. 2008; 23:
616   87–94.
617 15. Henrich J, Boyd R. The evolution of conformist transmission and the emergence of
618   between-group differences. Evolution and Human Behavior. 1998; 19: 215–241.
619 16. Boyd R, Richardson PJ. Culture and the evolutionary process. Chicago: University of
620   Chicago Press; 1985.
621 17. Whiten A, Horner V, de Waal FBM. Conformity to cultural norms of tool use in
622   chimpanzees. Nature. 2005; 437: 737–740.
623 18. Aplin LM, Sheldon BC, McElreath R. Conformity does not perpetuate suboptimal
624   traditions in a wild population of songbirds. PNAS. 2017; 114: 7830–7837.
625 19. He P, Maldonado-Chaparro AA, Farine DR. The role of habitat configuration in
626   shaping social structure: a gap in studies of animal social complexity. Behavioural
627   Ecology & Sociobiology. In revision.
628 20. Wilson MC, Chen X, Corlett RT, Didham RK, Ding P, Holt RD, Holyoak M, Hu G,
629   Hughes AC, Jiang L, Laurance WF, Liu J, Pimm SL, Robinson SK, Russo SE, Si X,
630   Wilcove DS, Wu J, Yu M. Habitat fragmentation and biodiversity conservation: key
631   findings and future challenges. Landscape Ecology. 2016; 31: 219–227.

632   21. Crooks KR, Brudett CL, Theobald DM, King SRB, Di Marco M, Rondinini C, Boitani L.
633   Quantification of habitat fragmentation reveals extinction risk in terrestrial
634   mammals. PNAS. 2017; 201705769.

635   22. Farine DR, Firth JA, Aplin LM, Crates RA, Culina A, Garroway CJ, Hinde CA, Kidd LR,
636   Milligan ND, Psorakis I, Radersma R, Verhelst B, Voelkl B, Sheldon BC. The role of
637   social and ecological processes in structuring animal populations: a case study from
638   automated tracking of wild birds. Royal Society Open Science. 2015; 2: 150057.

639   23. Aplin LM, Farine DR, Mann RP, Sheldon BC. Individual-level personality influences
640   social foraging and collective behaviour in wild birds. Proceedings of the Royals
641   Society B. 2014; 218: 20141016.

642   24. Jolles JW, Boogert NJ, Sridhar VH, Couzin ID, Manica A. Consistent individual
643   differences drives collective behaviour and group functioning of schooling fish.
644   Current Biology. 2017; 27: 2862-2868.

645   25. Morgan TJH, Laland KN. The biological bases of conformity. Frontiers in
646   Neuroscience. 2012; 6: 87.

647   26. Smaldino PE, Aplin LM, Farine DR. Do sigmoidal acquisition curves indicate
648   conformity? BioRxiv. 2017; doi.org/10.1101/159038.