Author for correspondence:
Simon DeDeo
e-mail: sdedeo@andrew.cmu.edu

The cultural transmission of tacit knowledge

Helena Miton[1] and Simon DeDeo[1,2]

[1]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
[2]Social and Decision Sciences, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

(iD) HM, 0000-0003-0896-1167; SD, 0000-0002-5346-9393

A wide variety of cultural practices have a 'tacit' dimension, whose principles are neither obvious to an observer, nor known explicitly by experts. This poses a problem for cultural evolution: if beginners cannot spot the principles to imitate, and experts cannot say what they are doing, how can tacit knowledge pass from generation to generation? We present a domain-general model of 'tacit teaching', drawn from statistical physics, that shows how high-accuracy transmission of tacit knowledge is possible. It applies when the practice's underlying features are subject to interacting and competing constraints. Our model makes predictions for key features of the teaching process. It predicts a tell-tale distribution of teaching outcomes, with some students near-perfect performers while others receiving the same instruction are disastrously bad. This differs from standard cultural evolution models that rely on direct, high-fidelity copying, which lead to a much narrower distribution of mostly mediocre outcomes. The model also predicts generic features of the cultural evolution of tacit knowledge. The evolution of tacit knowledge is expected to be bursty, with long periods of stability interspersed with brief periods of dramatic change, and where tacit knowledge, once lost, becomes essentially impossible to recover.

1. Introduction

One of the hallmarks of human knowledge is its 'tacit' dimension; as Polanyi suggested, we start with 'the fact that *we can know more than we can tell*' [1]. A vast array of complex cultural practices have a significant tacit dimension, with key principles that cannot be verbalized. The tacit dimension goes by many names, including 'working knowledge' [2], 'practical' knowledge [3], 'know-how' [4] and 'knowing-how' [5,6]. A significant tacit dimension is found in everything from sports [7,8] and artistic performance [9] to architecture [10], medicine [11] and science itself [12], and has been studied in contexts ranging from traditional crafts [13] to the professions [14] and organizations [15,16] of the modern world.

For any practice, it is important to distinguish knowledge along the tacit dimension from both explicit knowledge (with which it is usually contrasted), and yet other forms. Explicit knowledge is usually defined as knowledge that can be readily codified, accessed and shared, but not all unverbalizable knowledge is tacit knowledge in the original sense. Recent work by Hoksbergen *et al.* [17], for example, draws attention to the 'dimension X' of knowledge: knowledge that is lost, not understood, or misappropriated, or even distorted or suppressed.

Ways in which knowledge can transform from one type to another have also drawn attention. For instance, explicit knowledge can become tacit once the code required to express it is lost, and tacit knowledge can become explicit in cases where it can be codified. The transmission question is crucial here, and the socialization, externalization, combination and internalization (SECI) model [16], for example, suggests that tacit knowledge is acquired through socialization and sharing direct experience and interactions with others. Tacit knowledge, in their model, can sometimes be made explicit through

externalization, when it is crystallized and made public. Conversely, explicit knowledge can become tacit when it is internalized. Work by Wheeler [18], meanwhile, suggests that while the transmission of the explicit dimension might be the province of deductive or inductive reasoning, the tacit dimension is often implicated in abductive forms of cognition [19].

Central to all these questions is the fact that the tacit dimension of any practice must, just as much as anything else, be transmitted from one generation to the next. The details of that transmission are expected to have a decisive impact on its evolution over time; for example, transmission needs to meet minimal levels of accuracy for evolution to be possible at all [20,21]. As we shall see, however, the very nature of tacit knowledge poses a direct challenge to the standard accounts of how transmission happens.

In cultural evolution, standard transmission mechanisms include teaching (where a teacher communicates their understanding to a learner), emulation (copying an end product) and imitation (copying the actions that produce the product) [22–25]. While these three mechanisms can account for part of how culture is transmitted, they can struggle to explain transmission of the tacit dimension. Three aspects, in particular, make the task challenging.

First, tacit knowledge is a mental representation. To be transmitted, that representation must be in some way made public [26]. One main way to do so is verbal instruction, and a great deal of culture is passed down by speech alone [25,27]. However, tacit knowledge cannot be transmitted in this fashion [28,29], because, by definition, even those who have the knowledge would not know what to say.

Second, tacit knowledge is combinatorially complex. It provides those who possess it with a set of contingently deployed, interconnected skills [30,31]. Key aspects of an expert's tacit knowledge may become relevant so rarely—say, 'under pressure', or in an exceptional context—that even the most diligent student may never encounter them through observation alone. This makes it difficult for a standard alternative to explicit instruction: the target goals, and their contingencies, are too various and mutable for straightforward imitation or emulation to work.

Third, tacit knowledge includes knowing which aspects of behaviour constitute the practice, and which are incidental. This can make imitation difficult: if a learner is to acquire skills through imitation, she needs to know, or be able to infer, what is relevant to imitate, including whether an action is understood as instrumental or not [32]; in purely observational situations, overimitation is common [33,34]. Knowledge of what is relevant, however, is itself tacit. I may be able to improve my technique by watching a skilled performer, but only after I have enough tacit knowledge to know the relevant from the incidental. A novice at the violin cannot learn by watching an orchestra perform. Similar challenges occur for emulation: when knowledge is tacit, a learner cannot determine which features of the end product matter.

This paper presents a domain-general model that shows how, despite these challenges, tacit knowledge may be faithfully transmitted. The solution we propose sees tacit knowledge as the emergent product of a network of interacting constraints, and transmission as a process of guiding a learner to a solution by the simultaneous, and mutually interfering, demands of both a teacher and the environment. The knowledge is tacit even in transmission because only an enigmatic fragment is ever present to the mind of either teacher or learner. The structure necessary to reconstruct the practice emerges from the interaction between the practitioner and the environment, and the teacher's task is to guide a learner towards the correct use of that structure. In particular, by careful intervention on a small fraction of the features, a teacher can guide the learner to discover the full structure of the culturally specific solution.

Our model shows how only around 10% of the task need be conveyed by a teacher's intervention. This helps make sense of a key feature of teaching seen across the anthropological record, where the most common forms of teaching in the cultural record are low-cost and involve significant underspecification; see, e.g. [35]. This is, of course, in contrast to the 'Western', or WEIRD [36], image of teaching as rationalized, explicit and high-cost.

Our model can also, as we show, help explain a puzzling feature of cultural evolution: the fact that culture often appears to proceed in a bursty fashion, with long periods of stasis interspersed with short bursts of chaotic innovation leading to rapid and dramatic changes. Bursty evolution is common in cultural evolution (see, e.g. [37]). It is also a defining feature of prehistory: bursty changes in material culture, for example, provide the basis for how we divide prehistoric cultures into distinct periods (see [38] for a discussion). This alternation of periods of stasis and of short periods of rapid changes is seen in the evolution of technological systems [39,40] and is a focus for recent accounts of technological innovation [41] that emphasize analogies to biological evolution [42–44]. The theory of parasitism [45,46], where a technology's payoffs depend on the existence and spread of another technology, also predicts stepwise changes [47] which emerge because of interactions and constraints not within a system (as in our model), but between systems.

We present our work in three parts. We first present the model, showing how the mental representation of the practice is embedded in a network of embodied constraints, and how a teacher intervenes to help construct the representation for a learner. We then show how the fragmentary nature of these interventions combines with the constraints of the environment to allow for the accurate transmission of knowledge from generation to generation. Finally, we present our results on the cultural dynamics which arises spontaneously from the model, namely bursty evolution.

2. Model

Tacit knowledge can appear in a wide variety of domains. Our model attempts to capture the generalizable features of tacit knowledge by working at an abstract level. We present the model in two steps. First, we present the basic definitions, using a toy example to connect our work to broader questions in anthropology, archaeology and psychology. Then, we show how this qualitative account can be captured in a general mathematical framework, which allows us to predict how this process plays out in reality.

2.1. Constraint-based tacit knowledge

Our model shares a number of characteristics with that of [48]; in particular, both describe different forms of tacit cultural knowledge as systems of constrained and interacting choices. Formally, a particular case of tacit knowledge is

defined as a list of conditional behaviours. We refer to these as 'facets'. As an example, consider horse-riding. A particular style of riding corresponds to a tacit knowledge practice, and each style will involve a complex relationship between how, for example, the rider places their limbs in response to the movements of their mount.

In principle, each of these conditional behaviours can be specified by one of a set of symbols. If the body is in position so-and-so, and the horse does so-and-so, should the rider respond by lowering their hands ('option A for facet one', or A_1 for short), or, alternatively, by raising them (option B_1)? Should they relax their back (option A_2) or straighten it (option B_2)? Part of the facet specification for one style of riding might be A_1A_2, while another style might be B_1B_2, and so forth. The number of facets is potentially very high.

The second step of our model considers the interacting constraints between these different facets. What a rider does with one part of her body in a particular context will, because of the nature of the human or equine body, or because of the particular artefacts used for riding (the saddle, tack and so forth), be more or less consonant with what she does with another part of her body. For example, the combination of lowering one's hands and relaxing one's back may be a particularly consonant combination, while lowering one's hands and straightening one's back may not—i.e. an 'incorrect', or inexpert, response might be A_1B_2. A good combination will be something that, all other things being equal, the person can receive some sort of feedback on from the environment. For example, a consonant pairing may take less effort, or provide some other noticeable benefit such as fluency.

These consonance relationships, taken together, are called the constraint network. In our model, each facet in the network bears some relationship to the others. This can be a direct link, as in the example above, or it can be an indirect link, mediated by intermediate facets. For example, we might imagine a third conditional behaviour with two possibilities, A_3 and B_3, and that A_3 is more consonant with A_2, and B_3 is more consonant with B_2. The choice for the third conditional behaviour, in other words, is influenced by the choice for the second conditional behaviour. Because, however, the choice of A_1 versus B_1 is in turn influenced by the choice of A_2 versus B_2, the choice of A_3 versus B_3 has an impact as well. Such a network of interactions operationalizes intuitions of what makes a practice coherent.

The simplest version of such a model takes each facet to be a choice between one of two options (A or B; or more simply 0 or 1), and for the interactions between facets to be pairwise only. An example of such a network is shown in blue at the top of figure 1, with the very simple case of six facets. Each node corresponds to a facet, and lines between nodes reflect the two different types of consonance relationship. A solid line (e.g. the one connecting facets 1 and 2) says that the two facets in the 'same' state are preferred, while a dashed line (e.g. the one connecting facets 2 and 6) says that the two facets prefer to be in the opposite state. Thus, for example, the setting A_1A_2 is preferred to A_1B_2 (1 and 2 in the same state), all other things being equal, while A_2B_6 is preferred to A_2A_6. For simplicity, we can write out the full specification of the system as a binary string. One example of a string that satisfies many, though not all, of the constraints, is 111000; in this case, among other things, it satisfies the constraint that aligns facets 1 and 2, and that anti-aligns the facets 2 and 6.
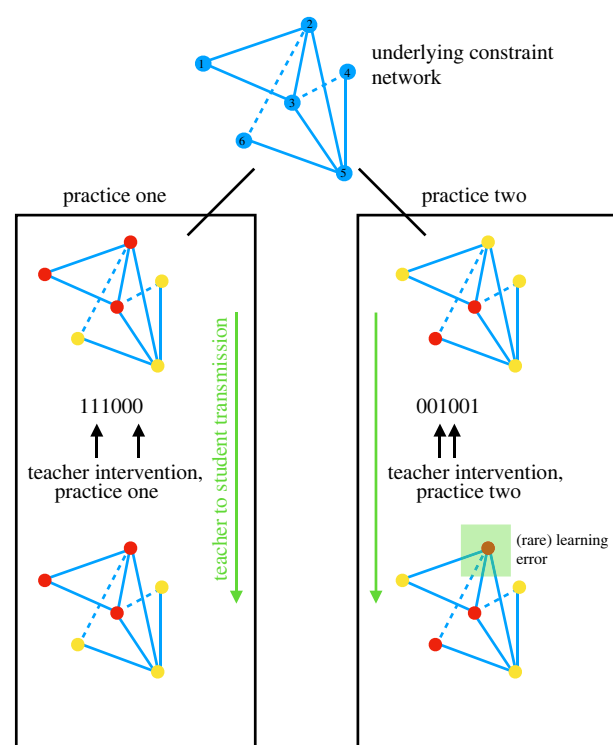

Figure 1. A constraint model of tacit knowledge transmission. A tacit cultural practice is a (usually partial) solution to a complex network of interacting constraints between aspects (facets) of the problem; we show a simple toy example here with only six facets, each of which takes on a binary value, i.e. an agent performs any particular aspect of the task in one of two mutually exclusive ways (red or yellow; written as one or zero, respectively). Constraints are pairwise, preferring either alignment (solid lines) or anti-alignment (dashed lines) of the facets.

Networks of interacting constraints like these, which include both preferences for alignment and anti-alignment, are often difficult, if impossible, to satisfy. In our simple example, facets 2, 5 and 6 cannot be set in a way that satisfies all three constraints simultaneously—as can be verified by trying the different combinations. Any particular specification for the facets, in other words, leads to difficulties.

Generically, there are different ways to satisfy these competing demands. For instance, similar outcomes can be obtained through distinct, yet functionally equivalent, movements while making pots [49]. Some specifications are better than others, and some are worse, but in general any particular tacit knowledge practice is a matter of how these difficulties are navigated; 'practice one' in our figure, for example, violates two constraints, while 'practice two' violates four. Others, not shown, are much worse; for example, the practice 101100 violates six constraints.

When a practice is a reasonably good solution to the constraint network, a practitioner who has learned the practice finds it easy to maintain. Deviations from the standard in many facets can be sensed and corrected. Consider, for example, someone implementing practice one. If she deviates by switching from the '0' state to the '1' state in facet four, she will experience an increased level of negative feedback from the environment, since she is now aligning with facet three (when it is more consonant to anti-align), and anti-aligning with facet five (when it is more consonant to align). This provides her with a signal that can be used to return to the standard. Even if she is unaware of which facet deviated,

she can make little (i.e. roughly single-facet) adjustments in her behaviour until consonance returns. When the practice is a reasonably good solution, in other words, the practitioner only needs to implement the solution. She does not need to understand it. Stable solutions like these are candidates for culturally transmitted tacit knowledge practices.

Transmission of the practice is now a matter of guidance. If the learner can be guided by a teacher close enough to the standard practice, the feedback from constraints will be sufficient to maintain her there. A very simple model of guidance is the intervention of the teacher to fix some of the facets into the culture's pattern. These may include physical interventions (to teach fly fishing, for example, a novice may be guided in proper form by literally tying his wrist to the rod), scaffolding (use of the barre in ballet), mnemonics ('eye on the ball', which maintains proper stance in tennis), or simple verbal guidance from the teacher (back straight!).

This way of considering teachers' intervention—as fixing some sub-part of students' practices—fits with a wide range of empirical results on teaching. It matches, in particular, real-life pedagogical interventions in which teachers reduce the degrees of freedom of their learners' movements. Downey [50] reports on this aspect of teaching in the case of the Brazilian martial art of capoeira, and [51] finds similar interventions from masters teaching kathak dance. Constraining interventions are also frequent in various sports involving athleticism, such as judo and swimming [7].

Careful interventions can do a great deal. In our toy example, practice one can be efficiently transmitted to the next generation by fixing only two critical nodes (nodes three and six). A learner who obeys her teacher's guidance in these two facets can learn the full pattern simply by remaining attentive to environmental feedback. She need only minimize the number of violated constraints, subject to the two instructional demands.

That effective subset of interventions (a kernel), when placed in an embodied context, reliably activates the characteristic and flexible behaviours of an expert. The nature of tacit knowledge means that the teacher is unaware of the exact nature of practice she exemplifies. However, the structure of the problem also can enable 'tacit teaching', which we define as the transmission of the tacit dimension of a practice through an attempt involving a teaching expert and a student. In our model, this is made possible when the teacher intervenes in only a fraction of the facets but nonetheless passes on the practice to some of the learners with near-perfect accuracy.

2.2. Mathematical framework

To study this model quantitatively, we need to specify how a learner responds to the constraint network. We adapt an approach used in a variety of cognitive models known as the maximum entropy principle [52–56]. The mathematics of our model is closely related to the Boltzmann machine in machine learning [57], Hopfield networks in neuroscience [58] and the 'spin glass' systems in physics [59]. Our model makes minimal assumptions about how facets interact. In the absence of teaching, the model fixes only the pairwise correlations between each pair of facets. The effect of teaching is to fix the average value of the particular facet being taught.

Once the constraint network is specified, this model has a free parameter, β, which governs the learner's sensitivity to constraints. When β is low, the learner pays little attention to the

constraints of her environment; when β is high, she is exceptionally rigorous. Assuming that the teacher is obeyed rigorously enough, our results are not particularly sensitive to the value of β, as long as it is past a critical point—essentially, the learner needs to be reasonably attentive to her environment (see electronic supplementary material for details).

All the relevant properties of the teaching process can be captured once we can compute the probability distribution of the learner over the different facet patterns. Following the discussion above, we assume that each facet for the learner, σ_i, can take on only one of two values; for simplicity, the two choices can be represented as +1 and −1. Then the probability distribution under the minimal model can be written [52] as

$$P(\{\sigma_i\}) = \frac{\exp\left(\beta \sum_{i,j} r_{ij}\sigma_i\sigma_j + \tau \sum_{i \in T} t_i\sigma_i\right)}{Z}, \qquad (2.1)$$

where r_{ij} is a matrix that describes the coupling between facets i and j; a positive value of r_{ij} indicates a preference for the two facets to be in the same state, and a negative value for them to be in opposite states (corresponding to the solid and dotted lines in figure 1, respectively). T is the set of nodes that are taught by intervention (the nodes marked with arrows in figure 1), t_i is the teacher's intervention (either +1, indicating a preference for the 'positive' practice, or −1, indicating a preference for the negative practice). The two constants β and τ govern the strength of the interaction between facets, and the influence of the teacher, respectively. Finally, Z is a normalization constant.

Equation (2.1) appears in many models in machine learning (Boltzmann machines [57]), neuroscience (Hopfield networks [58]) and physics (spin glass models [59]). The salient feature of all of these systems is the existence of multiple, distinct, 'metastable' (i.e. long-lived) patterns of activation. In the Hopfield case, these correspond to different 'memories'; for us, they correspond to different practices. The goal of the teacher is to guide the learner to her same solution.

We simulated different interaction configurations, where r_{ij} in any particular simulation is drawn from a random distribution, uniform between −1 and +1 (our qualitative conclusions are insensitive to the precise nature of this distribution, or the topology of the network; see electronic supplementary material). Once the interactions are fixed, the key parameters are β and τ. We set τ to be much larger than unity (in practice, between four and 10), and t_i to be zero for the facets i that the teacher does not intervene on, indicating that the teacher can make a strict intervention; i.e. that she can fix that small number of the student's facets with near perfection. Meanwhile, β indicates the extent to which the student is sensitive to the interacting constraints of the facets themselves. When β is zero, for example, the facets that are not being taught are completely free and uncorrelated. When β is high, the system has low tolerance for deviations from the patterns set by r_{ij}.

Our qualitative results are insensitive to β as long as it is past the 'critical point' (around unity). Empirically, we find that the most accurate transmission is possible when τ is set to be large, and β is slowly increased from zero to a value comparable to τ (see electronic supplementary material). Once this process has finished, the learner, in turn, can act as a teacher for a new learner; the old learner/new teacher intervenes on this new learner in the same fashion.
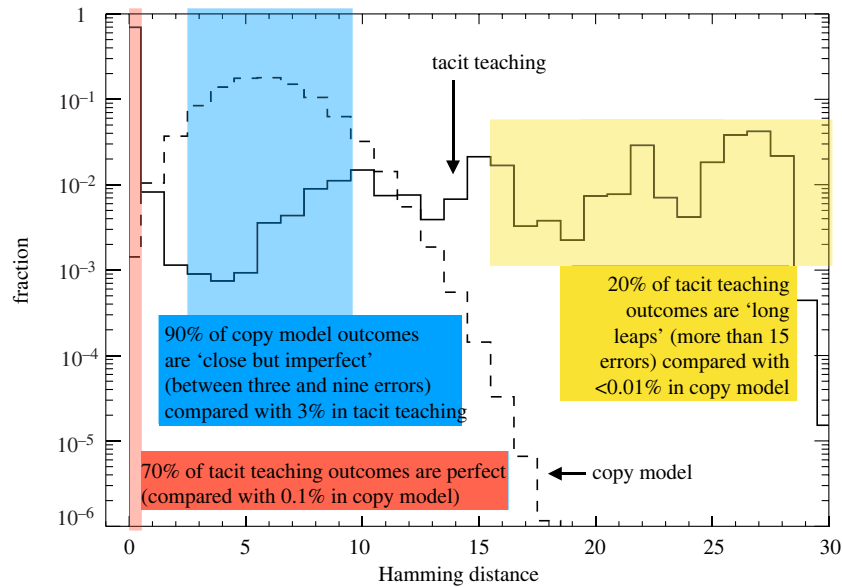
Figure 2. Tacit teaching leads to high-accuracy transmission from teacher to student. Shown here is the distribution of learning outcomes for the tacit teaching case (solid line) for the test case of 30 facets and (only) four teacher interventions. Even though the teacher intervenes in less than 15% of the facets, perfect reproduction occurs approximately 70% of the time (see the extreme left side of the distribution). The students who fail to learn properly, by contrast, often end up with very different solutions—the distribution of errors is non-normal (non-binomial) and long leaps are just as, if not more, common as small deviations. A null 'copy error' model (dashed line) has very different properties. If it is to achieve the same average accuracy as the tacit model, it must sacrifice any hope of accurate transmission.

All that remains is to determine a good candidate for T, the set of nodes the teacher intervenes on. We do this in a 'greedy' fashion. We first find the best facet for a single intervention (i.e. the facet that, if fixed by the teacher, allows the learner to best approximate the desired activity). We then iterate: we find the facet that, when fixed in conjunction with the first, produces the best outcome, and so forth. This heuristic can plausibly be followed by a beginning teacher, who experiments with enforcing different aspects, or an expert one who can adapt to a slightly changed culture. Over a decade range of network sizes (from 10 facets to 100), we find only between 10% and 15% of the facets need be fixed by teaching.

3. Results

We present our results on tacit teaching in two parts. First, we show how tacit teaching does, indeed, enable high-accuracy transmission of a practice from one generation to the next. We then show how the details of this process combine, at the population level, to leave traces on the long-term evolution of culture.

3.1. Tacit teaching

We first consider tacit teaching itself. Given a particular constraint network and cultural practice, we consider the effect of different kernel sizes on the accuracy of transmission. Accuracy is measured by Hamming distance, which counts the number of facets in which the student differs from the teacher; a Hamming distance of zero indicates perfect transmission (see electronic supplementary material for further details and alternative accuracy measures).

The results of our simulations suggest that under a variety of conditions, perfect transmission can be reliably obtained even when the number of interventions is significantly smaller than the number of facets. The kernel needs only be a small fraction of the whole, and a skilled teacher in

possession of that kernel would still be able to transmit the whole practice to the learner, even if only a small amount of information is conveyed between them.

An example of this phenomenon is shown in figure 2. In this case, a 30-facet practice can be transmitted with very high accuracy by intervening in only *four* facets. Despite the fact that, on the surface, only a tiny fraction of the total information is conveyed between teacher and student, perfect accuracy can be achieved nearly 70% of the time.

Neither teacher nor learner need know, in any conscious fashion, the correct pattern in all 30 facets—indeed, they need not even know how many facets there are. All that is needed for effective transmission is (i) that the teacher keep in mind four key features of the learner's behaviour and (ii) that the learner attend to the teacher's guidance while remaining attentive to the consonance demands of her environment. This is not only sufficient to guide the learner to the full 30-facet practice, but also to avoid other, potentially tempting—i.e. stable and similarly optimal—solutions that can be thought of as alternative cultural practices.

Two things are evident from figure 2. First, as noted, a majority of students learn the practice exactly. If teachers for the next generation are drawn from this subpopulation, the practice can persist with high levels of accuracy for multiple generations. Second, the distribution of errors is highly non-normal; of those who fail to learn, there are just as many who learn a practice (say) three Hamming units away as 20. Poor transmission is therefore expected to be far more noticeable.

This distribution arises because the underlying constraint network serves to correlate the errors made in learning: informally, a failed student learns 'bad habits' that connect together and re-enforce each other, driving the learner into a totally different part of solution space. This space is usually less optimal than the correct answer, but may have at least a modicum of stability. A simple example is in the teaching of juggling. A minority of learners find a satisfying, but in the end suboptimal, solution to the problem of juggling two

balls that involves passing, rather than tossing, one of the balls from one hand to the other.

Once a student has learned enough of these bad habits, further teaching may be in vain. Matching the teacher's practice would now require shifting a large number of facets simultaneously. The only other solution to this problem is if the student can start again—in our simple model, the necessary 'beginner's mind' is a random choice for each facet—and pay greater attention.

A number of consequences flow from this distribution of errors. First, it is easy to spot the majority of students who fail to imitate the practice: their overall behavioural pattern is generally very different from the cultural norm and (furthermore) the practice they do adopt is expected to be generally less effective in (for example) competition with learners who have correctly grasped the norm.

Second, however, not all errors are a combination of bad habits. It is entirely possible that a small fraction of students who fail to learn achieve, instead, 'true' alternative practices, meaning solutions to the constraint network that are, if not equally good as the standard practice of their culture, would be at least similarly stable. The rare discovery of unusual solutions, quite distant from the original practice, is a prediction of our model. It also appears to be found in the qualitative literature, with anecdotal records of sportsmen having unusual, yet high-performing, practices, like Oh Sadaharu in baseball [60] or Donald 'The Don' Bradman in cricket [61].

This leads to an interesting paradox. On the one hand, tacit teaching is, despite the fragmentary nature of the teacher's interventions, extraordinarily reliable. A majority of students learn the practice accurately. On the other, however, tacit teaching is also highly evolvable. The deviations that do occur are often significantly different from the standard practice.

One way to understand this result is to compare it with a null model, an imitation-like 'copy error' model. This model assumes that all of the facets are observed by the student and copied independently with some level of error. To compare the copy error model with tacit teaching, we tune the error rate of copying so that the average Hamming distance matches that of the tacit model (see electronic supplementary material for details). The copy error model is shown in figure 2 by a dashed line. In this particular example, the derived error rate is roughly 15%, which is plausible given what we know for cultural imitation [62]; however, as can be seen in the figure, such an error rate for a complex practice makes it nearly impossible to preserve the practice.

When comparing the two error distributions, two things stand out. On the one hand, as mentioned, the copy error model achieves basically zero accuracy: it is essentially impossible for a learner to match the teacher's practice, despite the assumption that he is aware of and can attend to all of the facets in turn. Second, despite this high error rate, it is also very hard for the copy error model to make long leaps and discover viable alternative practices. In this particular case, the vast majority of outcomes for the copy-error model lead to 'close but imperfect' outcomes, with an error rate of 1/6; only around 0.1% of learners reproduce the practice perfectly, and less than 0.01% produce long leaps that modify more than half of the practice (see electronic supplementary material for more).

On the other hand, the tacit model produces a spectrum, with a large number of perfectly accurate students, and a small number of outlier eccentrics. Most of the outliers, of course, fail to create a new practice, but a small number may find novel, but stable and teachable, solutions. (By comparison, if we tuned the copy-error model to match the perfect accuracy rate of 70% of tacit model here, it would require an implausibly low error rate of 1% for the observer, for each of the 30 facets in turn—and would completely sacrifice evolvability.) This has suggestive consequences for cultural evolution dynamics, especially with regard to diversity and evolvability. We examine them in the next section.

3.2. Population level dynamics

No teaching method is perfect, and every culture needs to deal with the fact that some fraction of the students will fail to learn. While the tacit teaching model can achieve high accuracy, not everyone is successful. If transmission is solely a matter of learners who each become teachers to an independent group of their own in turn, the practice will soon decay.

One solution to this problem is institutional: learners participate in an educational situation, such as a classroom or dojo or gym, or in more informal and complex mentorship and feedback relationships that might emerge in a business organization. In a toy model based on the simple classroom case, learners in each generation agree on a consensus practice that is taught to the next. If there are 10 students, for example, in our 30-facet model above, roughly seven of them will learn the same practice. If consensus is simply a matter of voting on which pattern (or, rather, kernel) will be taught to the next generation, then error-free transmission can be sustained over many generations. This is robust for two reasons: because, on average, we expect the standard practice to dominate, but also because the deviations are often idiosyncratic. Even if the standard practice does not obtain a majority, it will usually retain a plurality.

Not always, however. This is in part because idiosyncratic fluctuations are not random: 'bad habits' tend to drive students to the same, suboptimal—or at least, different—parts of the solution space. This means that it is not that difficult or rare for a non-standard practice to obtain a plurality. When this does happen, two things follow. First, the initial, standard, culture's practice is lost. Second, it is replaced by something that is often worse compared with the original.

Suboptimal solutions, in turn, are more difficult to learn because there are more nearby solutions that are equally good. A learner who deviates in one or two facets may find that, rather than upset a fine balance, she has satisfied just as many constraints as she did before. Now there is no good signal to lead her back to the original pattern, and, unless the teacher makes more interventions, transmission will be unsuccessful.

Taken together, these effects predict that the cultural evolution of tacit knowledge is *bursty*. Long periods of stability, in which cultural practices change very little, are interspersed with chaotic periods. These chaotic periods begin with a long leap in the solution space, and the original tradition is completely lost. Communities of practice in these chaotic periods are then much worse at preserving their (new) traditions, and make long leaps in turn. This continues until a new, sufficiently stable, practice is discovered. A longer period of high-accuracy transmission commences, and the cycle repeats.

This is shown, first, in figure 3, with a sample simulation of seven learners conducting a majority vote. A vertical line
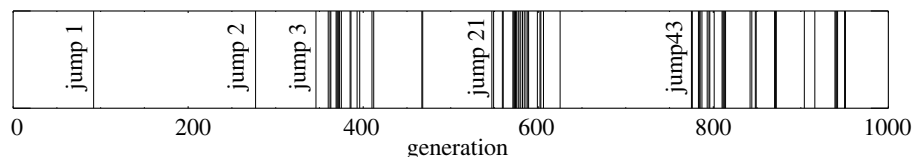
Figure 3. The distribution of jumps in a sample transmission sequence of 1000 generations, showing how jumps from one practice to another tend to be concentrated in time. The x-axis labels the generation number; a vertical line indicates a jump from one practice to another. Lines are clustered in groups indicate chaotic periods. Results of a simulation with 30 facets and four teaching nodes.

indicates a generation where the practice has switched. While the first two jumps are isolated events, the third jump leads to a chaotic cascade of jumps in the next 50 generations; similar turbulent periods appear every few hundred generations. The distribution of gaps between jumps is 'heavy-tailed', meaning that the majority of jumps are followed, one or two steps later, by another jump; once in a long while, however, these rapid jumps are interrupted by many hundreds of generations of stability. Standard statistical methods [63] show that the tail of this distribution is a power law, which means that there is no characteristic limit for how long this stability can last.

A final way to visualize this bursty behaviour is to track the evolution of the practice itself. Practices are high-dimensional objects; a 30-facet practice lies on one of the vertices of a 30-dimensional hypercube. This is, of course, impossible to visualize. However, we can use the fact that stable practices are sparsely distributed to our advantage. Since only a small fraction of the solution space corresponds to stable practices, we can use a dimensionality-reduction algorithm to map the shifts from generation to generation onto the two-dimensional page.

This is shown in figure 4. Each blue circle represents a point on the hypercube of tacit practices. The two-dimensional layout, provided by the multi-dimensional scaling (MDS) visualization algorithm, approximates Hamming distance: circles that are nearby each other on this plot have more facets set to the same value. Circle size is proportional to stability; larger circles indicate solutions that both satisfy more of the underlying constraints and are stable under perturbation; roughly a half-dozen such practices can be found.

The yellow line shows a sample evolutionary trajectory through this space. The simulated population begins in the relatively stable and teachable practice A, which it maintains for a long time. After a hundred generations or so, it makes a long jump to practice B. Practice B is less teachable (it is rarely transmitted accurately from teachers to learners), and the culture enters a period of instability, making additional long jumps to practices near practice C, and spending tens of generations in the a hard-to-maintain cluster of practices near practice D. Eventually the system returns to, and settles down in, the highly stable practice C. Other practices (E, F, G, etc.) remain undiscovered by this culture even after many thousands of generations.

For simplicity of presentation, our discussion has focused on networks with 30 facets and with couplings between any two facets drawn from a uniform distribution between one and negative one (a 'random network' model). Simulations of both larger and smaller networks, and networks with different topologies, produce essentially identical qualitative features, at both the population and the individual level (see electronic supplementary material).

We find that larger networks can support more cultural practices, and tacit teaching requires more interventions as
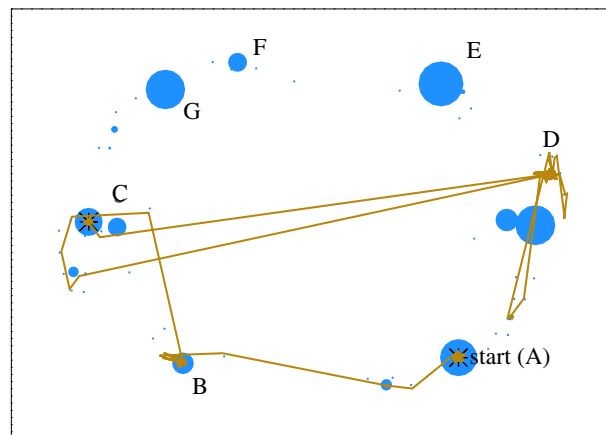

Figure 4. Exploration in cultural space, visualized in a sample simulation with the same parameters as figure 3. While any particular practice is a 30-dimensional binary vector, an approximate visualization is possible with the MDS algorithm. Circles correspond to different practices, with size proportional to stability. The yellow line follows the trajectory of a simulated culture, which begins in the practice labelled A, and wanders, in a characteristically bursty fashion, to land, finally, in practice C.

the complexity of the practice grows; for networks between 10 and 100 nodes, we find that tacit teaching with majority accuracy (i.e. at least half the time, a randomly chosen student matches the practice exactly) requires interventions on around 10% to 15% of the facets. This linear scaling is preserved for a variety of different distributions of edge weights.

Modelling the system of facet constraints as a randomly connected network has limitations. In many situations, we expect facets to organize themselves into roughly distinct 'modules' with tight interconnections within each module, and fewer, more disorganized connections, between modules. These modular organizations are expected under a range of circumstances. For example, when the facets concern material properties of the task, where spatial and temporal separations can generate nested topologies, as might happen for practices with clear stages of play (e.g. tennis). Or, when behavioural facets include the relative positions of different parts of the body, we expect there to be tighter constraints between groups of muscles that connect to the same joints, as might happen in dance. We also expect the emergence of modularity under generic tinkering and bricolage processes, as originally described by [48]; more recent work suggests that, if the underlying constraints are built up by combining and repurposing earlier practices, the resulting network will have high levels of modularity [64]. Under the assumption that each module has only two configurations, our results now apply at the module level. To teach a practice with 30 modules, for example, we expect tacit teaching to require interventions in (roughly) four of them.

Modules may be more complex, meaning that they may be able to support more than two internal configurations. A simple information-theoretic argument suggests that this will scale logarithmically in the number of additional interventions. If each module has N solutions, for example, then the demands on tacit teaching increase, albeit slowly, by a factor of $\log_2 N$.

4. Discussion

Cultural evolution, as a field, often opts to 'black box' how information is socially transmitted and learned [65]. Models such as [66–68] give us great insights into cultural dynamics at the population level. However, they have done so in part at the cost of ignoring the complexity of the cognitive agents who actually acquire and transmit culture.

Our model provides an explicit and quantitative account of the relationship between teacher and student in the commonly encountered case of tacit knowledge. It shows how high-accuracy 'tacit teaching' is possible, even in the case where both teacher and student lack conscious knowledge of up to 90% of the dimension. A small amount of guidance, well-presented, allows the majority of students to 'lock in' an efficient, culturally widespread practice. This is possible only when the features of underlying practice are subject to specific constraints, echoing observation of skill acquisition dynamic in ecological contexts [69,70].

Our results make empirical predictions for cognition at the individual level. One key feature of tacit teaching is the presence of an unusual and non-exponential distribution of learning errors: when tacit teaching is in place, we expect even diligent learners to, occasionally, learn something that diverges significantly from the correct performance. Conversely, we expect to find a deficit of near misses: students who are close to getting it, but miss only in one or two aspects. More generally, as seen in figure 2, we expect a characteristic pattern of error-making that looks very different from a model where the teacher teaches everything, and the student learns each piece independently. When most students do extremely well, but a small fraction, with otherwise equivalent abilities, do extremely poorly, it may be a sign that tacit teaching is at play.

These results have, in turn, implications for cultural evolution. They predict bursty, and sometimes very long-leap,

innovations, with a heavy-tailed power law distribution that makes it possible for a practice to change without going through a series of gradual mutations. These long leaps can enable, potentially, rapid adaptation to new condition (e.g. changes in the underlying constraint network). They come at a cost, however: once a leap has been made, it is very difficult to recover the prior practice, except by accident.

A cultural tradition is more than just a list of behavioural features. It is enabled by how those features fit together into a larger logic dictated by mental, material and environmental constraints. (Knowledge with a tacit dimension is, further, itself embedded in a socio-economic context which can modulate how it is expressed [71], and factors like emotional intelligence have been shown to positively impact the sharing of tacit knowledge [72].)

We have presented a minimal model that allows us to capture this higher-order logic, and to thereby go beyond accounts of cultural evolution that focus on the acquisition of individual traits considered in isolation. Attending to the cognitive aspects of transmission reveals how these interactions do more than channel culture: they make it possible to accurately transmit it, with a teacher's intervention serving as a seed for the learner's full practice. It also shows how this mechanism can drive the dynamics and long-leap changes that characterize the macroevolution of culture.

References

1. Polanyi M. 2009 The tacit dimension. Chicago, IL: University of Chicago Press.

2. Harper D. 1987 Working knowledge: skill and community in a small shop. Chicago, IL: University of Chicago Press.

3. Archer MS, Archer MS. 2000 Being human: the problem of agency. Cambridge, UK: Cambridge University Press.

4. Ryle G. 2009 The concept of mind. New York, NY: Routledge.

5. Harris M. 2007 Ways of knowing: new approaches in the anthropology of knowledge and learning, vol. 18. New York, NY: Berghahn Books.

6. Pavese C. 2021 Knowledge how. In The Stanford encyclopedia of philosophy (ed. EN Zalta).

Metaphysics Research Lab, Stanford University, Summer 2021 edition.

7. Jakubowska H. 2017 Skill transmission, sport and tacit knowledge: a sociological perspective, vol. 4. London, UK: Taylor & Francis.

8. Nyberg G. 2014 Exploring 'knowings' in human movement: the practical knowledge of pole-vaulters. Eur. Phys. Educ. Rev. 20, 72–89. (doi:10.1177/1356336X13496002)

9. Kaastra LT. 2016 Tacit knowledge in orchestral performance: an application of HH Clark's (1997) 'dogmas of understanding' to the analysis of performed music. In College Music Symp., vol. 56. Missoula, MT: College Music Society.

10. Alexander C. 1977 A pattern language: towns, buildings, construction. Oxford, UK: Oxford University Press.

11. Patel VL, Arocha JF, Kaufman DR. 1999 Expertise and tacit knowledge in medicine. In Tacit knowledge in professional practice: researcher and practitioner perspectives, pp. 75–99. New York, NY: Psychology Press.

12. Brock R. 2017 Tacit knowledge in science education, pp. 133–142. Rotterdam, The Netherlands: SensePublishers.

13. Marchand THJ. 2008 Muscles, morals and mind: craft apprenticeship and the formation of person. Br. J. Educ. Stud. 56, 245–271. (doi:10.1111/j.1467-8527.2008.00407.x)

14. Sternberg RJ, Horvath JA. 1999 *Tacit knowledge in professional practice: researcher and practitioner perspectives*. New York, NY: Psychology Press.

15. Baumard P. 1999 *Tacit knowledge in organizations*. Thousand Oaks, CA: Sage.

16. Nonaka I, Takeuchi H. 1995 *The knowledge-creating company: how Japanese companies create the dynamics of innovation*. New York, NY: Oxford University Press.

17. Hoksbergen M, Chan J, Peko G, Sundaram D. 2021 Illuminating and bridging the vortex between tacit and explicit knowledge: counterbalancing information asymmetry in high-value low-frequency transactions. *Decis. Support Syst.* **149**, 113605. (doi:10.1016/j.dss.2021.113605)

18. Wheeler W. 2016 'Do not block the path of inquiry!': Peircean abduction, the tacit dimension, and biosemiotic creativity in nature and culture. *Am. J. Semiot.* **24**, 171–187. (doi:10.5840/ajs2008241/312)

19. Sebeok TA, Danesi M. 2000 *The forms of meaning*. Berlin, Germany: De Gruyter Mouton.

20. Charbonneau M. 2020 Understanding cultural fidelity. *Br. J. Phil. Sci.* **71**, 1209–1233. (doi:10.1093/bjps/axy052)

21. Eigen M, Schuster P. 1977 The hypercycle: a principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften* **64**, 541–565. (doi:10.1007/BF00450633)

22. Tomasello M, Kruger AC, Ratner HH. 1993 Cultural learning. *Behav. Brain Sci.* **16**, 495–511. (doi:10.1017/S0140525X0003123X)

23. Caldwell CA, Millen AE. 2009 Social learning mechanisms and cumulative cultural evolution: is imitation necessary? *Psychol. Sci.* **20**, 1478–1483. (doi:10.1111/j.1467-9280.2009.02469.x)

24. Hoppitt W, Laland KN. 2013 *Social learning: an introduction to mechanisms, methods, and models*. Princeton, NJ: Princeton University Press.

25. Morgan TJH et al. 2015 Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat. Commun.* **6**, 1–8. (doi:10.1038/ncomms7029)

26. Sperber D, Hirschfeld L. 2007 Culture and modularity. In *The innate mind: culture and cognition*, vol. 2 (eds P Carruthers, S Stich, S Laurence), pp. 149–164. Oxford, UK: Oxford University Press.

27. Bietti LM, Bangerter A, Knutsen D, Mayor E. 2019 Cultural transmission in a food preparation task: the role of interactivity, innovation and storytelling. *PLoS ONE* **14**, e0221278. (doi:10.1371/journal.pone.0221278)

28. Neuweg GH. 2004 Tacit knowing and implicit learning. In *European perspectives on learning at work: the acquisition of work process knowledge*, pp. 130–147. Luxembourg: Office for Official Publications of the European Communities.

29. Collins H. 2010 *Tacit and explicit knowledge*. Chicago, IL: University of Chicago Press.

30. Stout D, Bril B, Roux V, DeBeaune S, Gowlett JAJ, Keller CM, Wynn T, Stout D. 2002 Skill and cognition in stone tool production: an ethnographic case study from Irian Jaya. *Curr. Anthropol.* **43**, 693–722. (doi:10.1086/342638)

31. Seifert L, Button C, Davids K. 2013 Key properties of expert movement systems in sport. *Sports Med.* **43**, 167–178. (doi:10.1007/s40279-012-0011-z)

32. Gergely G, Bekkering H, Király I. 2002 Rational imitation in preverbal infants. *Nature* **415**, 755–755. (doi:10.1038/415755a)

33. Lyons DE, Young AG, Keil FC. 2007 The hidden structure of overimitation. *Proc. Natl Acad. Sci. USA* **104**, 19 751–19 756. (doi:10.1073/pnas.0704452104)

34. Lyons DE, Damrosch DH, Lin JK, Macris DM, Keil FC. 2011 The scope and limits of overimitation in the transmission of artefact culture. *Phil. Trans. R. Soc. B* **366**, 1158–1167. (doi:10.1098/rstb.2010.0335)

35. Kline MA, Boyd R, Henrich J. 2013 Teaching and the life history of cultural transmission in Fijian villages. *Hum. Nat.* **24**, 351–374. (doi:10.1007/s12110-013-9180-1)

36. Henrich J, Heine SJ, Norenzayan A. 2010 The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83. (doi:10.1017/S0140525X0999152X)

37. Kuhn T, Helbing MD. 2014 Inheritance patterns in citation networks reveal scientific memes. *Phys. Rev. X* **4**, 041036. (doi:10.1103/PhysRevX.4.041036)

38. Barton CM, Clark GA. 2021 From artifacts to cultures: technology, society, and knowledge in the upper paleolithic. *J. Paleolit. Archaeol.* **4**, 1–21. (doi:10.1007/s41982-021-00091-8)

39. Gorshenev AA, Pis'mak YM. 2004 Punctuated equilibrium in software evolution. *Phys. Rev. E* **70**, 067103. (doi:10.1103/PhysRevE.70.067103)

40. Valverde S, Solé RV. 2015 Punctuated equilibrium in the large-scale evolution of programming languages. *J. R. Soc. Interface* **12**, 20150249. (doi:10.1098/rsif.2015.0249)

41. McNerney J, Farmer JD, Redner S, Trancik JE. 2011 Role of design complexity in technology improvement. *Proc. Natl Acad. Sci. USA* **108**, 9008–9013. (doi:10.1073/pnas.1017298108)

42. Kauffman S, Macready W. 1995 Technological evolution and adaptive organizations. *Complexity* **1**, 26–43. (doi:10.1002/cplx.6130010208)

43. Solé RV, Valverde S, Casals MR, Kauffman SA, Farmer D, Eldredge N. 2013 The evolutionary ecology of technological innovations. *Complexity* **18**, 15–27. (doi:10.1002/cplx.21436)

44. Valverde S. 2016 Major transitions in information technology. *Phil. Trans. R. Soc. B* **371**, 20150450. (doi:10.1098/rstb.2015.0450)

45. Coccia M. 2019 Comparative theories of the evolution of technology. In *Global encyclopedia of public administration, public policy, and governance*. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-31816-5_3841-1)

46. Coccia M. 2019 A theory of classification and evolution of technologies within a generalised Darwinism. *Technol. Anal. Strat. Manage.* **31**, 517–531. (doi:10.1080/09537325.2018.1523385)

47. Coccia M. 2021 Technological innovation. *Innovations* **11**, I12.

48. Alexander C. 1964 *Notes on the synthesis of form*. Cambridge, MA: Harvard University Press.

49. Gandon E, Bootsma RJ, Endler JA, Grosman L. 2013 How can ten fingers shape a pot? Evidence for equivalent function in culturally distinct motor skills. *PLoS ONE* **8**, e81614. (doi:10.1371/journal.pone.0081614)

50. Downey G. 2008 Scaffolding imitation in capoeira: physical education and enculturation in an Afro-Brazilian art. *Am. Anthropol.* **110**, 204–213. (doi:10.1111/j.1548-1433.2008.00026.x)

51. Dalidowicz M. 2015 Crafting fidelity: pedagogical creativity in kathak dance. *J. R. Anthropol. Inst.* **21**, 838–854. (doi:10.1111/1467-9655.12290)

52. Schneidman E, Berry II MJ, Segev R, Bialek W. 2006 Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007–1012. (doi:10.1038/nature04701)

53. Lezon TR, Banavar JR, Cieplak M, Maritan A, Fedoroff NV. 2006 Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. *Proc. Natl Acad. Sci. USA* **103**, 19 033–19 038. (doi:10.1073/pnas.0609152103)

54. Seno F, Trovato A, Banavar JR, Maritan A. 2008 Maximum entropy approach for deducing amino acid interactions in proteins. *Phys. Rev. Lett.* **100**, 078102. (doi:10.1103/PhysRevLett.100.078102)

55. Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M, Walczak AM. 2012 Statistical mechanics for natural flocks of birds. *Proc. Natl Acad. Sci. USA* **109**, 4786–4791. (doi:10.1073/pnas.1118633109)

56. Lee ED, Broedersz CP, Bialek W. 2015 Statistical mechanics of the US Supreme Court. *J. Stat. Phys.* **160**, 275–301. (doi:10.1007/s10955-015-1253-6)

57. Ackley DH, Hinton GE, Sejnowski TJ. 1985 A learning algorithm for Boltzmann machines. *Cogn. Sci.* **9**, 147–169. (doi:10.1207/s15516709cog0901_7)

58. Hopfield JJ. 1982 Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl Acad. Sci. USA* **79**, 2554–2558. (doi:10.1073/pnas.79.8.2554)

59. Sherrington D, Kirkpatrick S. 1975 Solvable model of a spin-glass. *Phys. Rev. Lett.* **35**, 1792–1796. (doi:10.1103/PhysRevLett.35.1792)

60. Kelly WW. 2008 Learning to swing: Oh Sadaharu and the pedagogy and practice of Japanese baseball. In *Learning in likely places: varieties of apprenticeship in Japan*, p. 265. Cambridge, UK: Cambridge University Press.

61. Renshaw I, Glazier P, Davids K, Button C. 2005 Uncovering the secrets of the Don: Bradman reassessed. *Sport Health* **22**, 16–21.

62. Claidière N, Sperber D. 2010 Imitation explains the propagation, not the stability of animal culture. *Proc. R. Soc. B* **277**, 651–659. (doi:10.1098/rspb.2009.1615)

63. Clauset A, Shalizi CR, Newman MEJ. 2009 Power-law distributions in empirical data. *SIAM Rev.* **51**, 661–703. (doi:10.1137/070710111)

64. Solé R, Valverde S. 2020 Evolving complexity: how tinkering shapes cells, software and ecological

networks. *Phil. Trans. R. Soc. B* **375**, 20190325. (doi:10.1098/rstb.2019.0325)

65. Heyes C. 2016 Blackboxing: social learning strategies and cultural evolution. *Phil. Trans. R. Soc. B* **371**, 20150369. (doi:10.1098/rstb.2015.0369)

66. Henrich J. 2004 Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *Am. Antiq.* **69**, 197–214. (doi:10.2307/4128416)

67. Boyd R, Richerson PJ. 1988 *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

68. Mesoudi A. 2011 Variable cultural acquisition costs constrain cumulative cultural evolution. *PLoS ONE* **6**, e18239. (doi:10.1371/journal.pone.0018239)

69. Button C, Seifert L, Chow JY, Davids K, Araujo D. 2020 *Dynamics of skill acquisition: an ecological dynamics approach*. Champaign, IL: Human Kinetics Publishers.

70. Hristovski R, Davids K, Araújo D, Button C. 2006 How boxers decide to punch a target: emergent behaviour in nonlinear dynamical movement systems. *J. Sports Sci. Med.* **5**, 60.

71. Hou C, Liu Z. 2021 Tacit knowledge mediates the effect of family socioeconomic status on career adaptability. *Soc. Behav. Pers.: Int. J.* **49**, 1–9. (doi:10.2224/sbp.9974)

72. Malik S. 2021 The nexus between emotional intelligence and types of knowledge sharing: does work experience matter? *J. Workplace Learn.* **33**. 619–634. (doi:10.1108/JWL-10-2020-0170)

73. Miton H, DeDeo S. 2022 The cultural transmission of tacit knowledge. Figshare. (doi:10.6084/m9.figshare.c.6238482)