



Rethinking prestige bias

Azita Chellappoo¹

Received: 1 May 2019 / Accepted: 31 January 2020 / Published online: 8 February 2020
© The Author(s) 2020

Abstract

Some cultural evolution researchers have argued for the importance of prestige bias as a systematic and widespread social learning bias, that structures human social learning and cultural transmission patterns. Broadly speaking, prestige bias accounts understand it as a bias towards copying ‘prestigious’ individuals (which are typically described as high-status, due to a high level of skill or success in a socially valued domain, and so are treated by others with respect and deference). Prestige bias, along with other social learning biases, has been argued to play a crucial role in allowing cumulative cultural selection to take place, thereby generating adaptations that are key to our success as a species. However, I argue for skepticism about the plausibility and scope of a prestige bias account. I argue that although an account of prestige bias seems plausible or compelling on their face, it is committed to a particular view of the cognition underpinning the bias, and therefore to predictions regarding its flexibility and context-sensitivity. Given this, current empirical evidence gives us reason to doubt the explanatory value of a prestige bias account over a naive, goal-directed agent account. Additionally, the way that prestige is defined in empirical work is in tension with a general understanding of prestige, casting doubt upon its status as evidence of prestige bias. I examine two studies cited as evidence of prestige bias, arguing that in these cases we cannot clearly favour a prestige bias explanation over a goal-directed agent explanation.

Keywords Cultural evolution · Prestige · Bias · Cognition · Social learning

1 Introduction

Universal, or near-universal, social learning biases have been argued to be crucial in understanding how cultures evolve. According to the work of authors such as Robert Boyd, Peter Richerson, Joseph and Natalie Henrich, and Richard McElreath, there

✉ Azita Chellappoo
asc63@cam.ac.uk

¹ Department of History and Philosophy of Science, University of Cambridge, Free School Lane, Cambridge CB2 3RH, UK

are various kinds of systematic cognitive biases that structure human social learning and therefore the transmission of behaviours, ideas, and other components of culture between individuals. This can affect patterns of cultural change and stasis on a population level, and have been proposed as key drivers of the cumulative cultural evolution that has enabled humans to thrive in a vast range of environments (Henrich 2015).

Broadly, these biases are often understood as ‘rules of thumb’ regarding who to learn from, or what is learnt, which have evolved because of the high cost of individual learning (Boyd and Richerson 1985; Boyd et al. 2011; Henrich et al. 2008; Henrich and McElreath 2003). Models have been constructed that indicate that facultative social learning emerges as a dominant strategy when individual learning is costly, and social learning is cheap and accurate enough (Rendell et al. 2010). According to theoretical models, social learning biases can increase the benefits of social learning (Henrich and Boyd 1998, 2002). They do this by directing which variants are acquired or who the variants are acquired from, and thereby increasing the likelihood that the new cultural variant acquired will be adaptive for the learner.

Social learning biases are typically partitioned into three categories: content biases, frequency-dependent biases, and model biases (Boyd and Richerson 1985). Content biases (or direct biases) are those that favour the acquisition of one cultural variant (which could be a behaviour, idea, practice, or norm, amongst other things) over another by an individual because of characteristics of the variant itself (its content). Work done by Dan Sperber and colleagues on the existence of ‘cultural attractors’ can be said to fall within the category of content biases (Sperber 1996), although their understanding of what these biases constitute and how acquisition and transmission operates differs from work carried out by Boyd, Richerson and Henrich, and their students, who tend to focus on frequency-dependent and model biases.

Frequency-dependent biases, such as conformist or anti-conformist bias, are those that favour the acquisition of one cultural variant over another due to characteristics of the trait’s distribution in the population. The type of frequency-dependent bias that has received the most attention in the literature is conformist bias, where individuals are disproportionately more likely to copy variants that are present at a high frequency in the population (Henrich and Boyd 1998; Efferson et al. 2008). Models of conformist bias have claimed to explain a range of cultural phenomena, such as the diffusion of innovations (Henrich 2001).

Lastly, model biases are those that favour the acquisition of one cultural variant over another due to characteristics of the individual from whom the variant is acquired (the ‘model’). The type of model bias that has received the most attention and has been proposed as a key factor in cumulative cultural evolution, and is the focus of this paper, is prestige bias. This is where learners are more likely to copy variants from ‘prestigious’ individuals. Other model biases include a bias towards copying individuals based on age, familiarity, or similarity to the learner in respects such as gender or ethnicity.

These social learning biases also play an important role in the cultural selection literature. The kinds of social learning biases already described often form the foundation of approaches that focus on cultural selection as a key force for understanding cumulative cultural change. For some authors, it is the set of universal social learning biases (conformist bias, prestige bias, etc.) that has allowed cumulative cultural selec-

tion to take place, and therefore brought about adaptations that are key to our success as a species (e.g., Richerson and Boyd 2005). For example, conformist bias and prestige bias have been cited as important factors in minimising levels of within-group variation and maintaining levels of between-group variation, which are necessary for cultural group selection to take place (Henrich 2004).

Another key example comes from Henrich (2015), who argues that conformist bias forms a crucial part of the explanation of why the series of steps carried out to detoxify cassava emerged and spread throughout some South American populations. Some varieties of cassava contain dangerous levels of cyanide, which if eaten without being adequately processed to remove toxins, could result in chronic health problems, such as goiter and neurological issues. Communities in South America which rely on cassava as an important part of their diet have a multi-step processing technique that occurs over several days and results in detoxified and safe to eat cassava. Henrich argues that any one individual would have difficulty working out this technique by themselves, and therefore individual learning fails to explain this adaptive behaviour. This is because the poisoning is slow, and so it would be difficult to figure out the connection between eating cassava and experiencing health issues. Additionally, the cassava stops being bitter before it stops being toxic, and therefore individuals who are relying on the signal of bitterness would continue to ingest unsafe levels of cyanide. Here the steps in this procedure are “causally opaque”: individuals cannot easily infer the functions or importance of the behaviours they are carrying out. Therefore, invoking broad social learning biases could explain the spread of this adaptive behaviour when explanations in terms of individual learning fail. This feeds into a selective process because individuals who copy this behaviour will have fitness advantages over those who do not, and groups whose members perform this behaviour will outcompete those who do not.

However, some doubt has been cast on the existence or prevalence of these biases. Lewens (2015) offers a critique of the conformist bias literature, arguing that some of the key assumptions made in conformist bias models are not supported by empirical evidence, and that there may be some circularity in the way that empirical and theoretical evidence is used to construct explanations.

Similarly, I argue for skepticism about the existence and prevalence of prestige bias. Broadly speaking, prestige bias can be understood as a bias towards copying ‘prestigious’ individuals. What constitutes a prestigious individual is not always entirely clear, and not fully elaborated on in the literature. Intuitively, prestigious individuals are high-status (typically due to a high level of skill or success in a socially valued domain), are treated by others with respect and deference, and may receive tangible goods due to their status.

Theoretical models have purported to show the adaptive benefit of prestige bias. Several studies have claimed to demonstrate the existence of prestige bias in both children and adults, and to show prestige-biased transmission networks in small-scale societies. As with conformist bias, prestige bias is an important component of cultural selection explanations. For example, Henrich and Henrich’s (2010) work on food taboos in Fiji characterises taboos surrounding eating certain marine species by pregnant and breastfeeding women as the result of prestige-biased learning. According to Henrich and Henrich, prestige-biased learning results in the spread of these beneficial

food taboos, leading to population-level patterns of adaptation. These social learning biases can be culturally selected for in themselves, as if they generally lead to the copying of adaptive behaviour, individuals who have them will outcompete individuals who do not. Additionally, groups with higher proportions of individuals copying the adaptive behaviour will do better than groups who do not.

Prestige bias models have also been invoked as a tool for understanding the existence of prestige in itself, and as explanations for a range of cultural phenomena, including the existence of celebrities, and of so-called ‘prestige goods’, that advertise the status of a prestigious individual (Henrich 2001; Plourde 2008; Mesoudi 2009).

I argue that although the accounts of prestige bias given by cultural evolution researchers seem plausible or compelling on their face, they in fact entail a set of commitments regarding the underlying cognition, how the bias operates, and how we should understand prestige. Once we relax these commitments, the prestige bias account loses explanatory value and struggles to offer anything distinctive over a naive, goal-directed agent account. In the following section I begin by summarising the main account of prestige bias developed in the literature (found in Henrich and Gil-White 2001). In Sect. 3 I argue that the characterisation of prestige under prestige bias accounts and in empirical tests of prestige bias is in tension with a general understanding of prestige. The definition of prestige in empirical work is pruned in a way which defers to the definition of prestige given by the prestige bias account, and this calls into question the status of this work as evidence of prestige bias. Additionally, the prestige bias account is committed to a particular view of the cognition underpinning the bias, and therefore to predictions regarding its flexibility and context-sensitivity. Both these considerations give us reason to question the explanatory value of a prestige bias account over a goal-directed agent account to explain patterns of apparent prestige-biased learning. I then make this comparison with a goal-directed agent account explicit in Sect. 4, drawing on two cases cited as evidence of prestige bias to argue that they do not clearly favour a prestige bias account over a goal-directed agent account.

2 An account of prestige bias

Henrich and Gil-White (2001) set out an extensive account of prestige bias. Although other authors have proposed similar accounts (e.g., Richerson and Boyd 2005), in addition to being the most comprehensive and detailed account, theirs has been explicitly cited and used as the basis for theoretical and empirical work (e.g., Bell 2013; Cheng and Tracy 2014; Chudek et al. 2012; Reyes-Garcia et al. 2008; Atkisson et al. 2012). Therefore, here I focus on their account.

They define prestige in contrast with dominance, and characterise it as unique to human societies. According to Henrich and Gil-White, dominance hierarchies are typical of chimpanzee societies (which do not display prestige). In dominance hierarchies, subordinates exhibit submissive behaviours towards superiors, avoid superiors when possible, and occasionally make aggressive challenges to the rank ordering. Dominant individuals in these hierarchies ‘grandstand’ to display their dominance, which is maintained primarily through physical force, or threat of physical force. In contrast,

in prestige hierarchies, subordinates will actively seek contact and interaction with superiors, and will spend sustained periods observing superiors. Prestigious individuals will not grandstand, and may even be self-deprecating and adopt non-dominant postures such as looking down. Henrich and Gil-White claim that human hierarchies are typically a mix of dominance and prestige.

The existence of prestige is supposedly puzzling on its face, in that individuals are paying ‘costs’ in terms of deference and perhaps material goods towards prestigious individuals, and therefore we would expect some kind of benefit for these individuals in order for this behaviour to be evolutionarily stable. Henrich and Gil-White argue for what they term the ‘information goods theory’. The key insight of this theory is to propose ‘information goods’ as the goods that deference-paying individuals are receiving, and which therefore stabilise prestige hierarchies in a population.

As they characterise it, prestige hierarchies will develop in populations with individuals that are already capable of ranking the relative skill of other members of the population, and of displaying deference selectively to certain chosen members. There are two key components to their story of the evolution of prestige bias: initially, individuals must have been able to determine the most adaptive models to copy from relatively reliably, and would have paid deference to those models in order to gain access so they could copy their behaviours. Once these two components are in place, rather than carrying out a potentially complex calculation to work out who is the best learning model, new entrants can use the low-cost cue of who is being paid the most deference to in order to determine who they should copy: this is prestige bias.

Firstly, in order for prestige hierarchies to emerge, individuals will have had the ability to rank potential learning models according to sets of cues. These cues include the model’s perceived competence, their perceived health, and their similarity to the learner in key respects such as age or gender (which may increase the likelihood that the cultural variants the model holds will be adaptive for the learner to acquire). Individuals that can pick up on these cues and successfully use them to selectively copy from certain models in the population will stand to benefit, as they will be more able to acquire adaptive cultural variants than through either individual learning (and at a lower cost than through individual learning) or ‘unguided variation’ (copying members of the population at random). In particular, individuals that can identify models with high levels of skill or knowledge in a valued domain (in the ancestral environment, this could have been hunting, foraging or knowledge of medicinal plants), and can copy the behaviour of these models, will have an adaptive advantage.

Secondly, Henrich and Gil-White argue that in ancestral environments, copying the behaviours of a chosen model requires perceptual access, and therefore requires the cooperation of the model themselves. For example, if novice hunters wish to learn from successful hunters, they would likely need to closely observe various behaviours and practices that successful hunters use, in order to accurately copy them. If it is the case that access to models is crucial, then successful or skilful individuals who many learners wish to copy would be able to control access and thereby copying opportunities (providing they cannot be coerced into supplying this information). Then, a system could develop whereby individuals seeking to learn from successful individuals (who it would be adaptive to copy from) pay them deference, accord them status, and provide

material benefits (such as exceptions from certain restrictive community norms), in return for access and the opportunity to observe their behaviours.

Eventually, in the population there will be a distribution of deference, where the most successful or skilful individuals will have many learners paying them deference (and therefore conferring ‘prestige’ upon them). When this happens, new individuals who are deciding who to learn from can decide based upon the level of deference (i.e., identifying the most prestigious individuals), rather than making assessments of skill or knowledge directly. This is supposedly adaptive because we would expect it to be significantly easier in many cases to identify the most prestigious individuals, rather than identifying them based on their skill or knowledge level. For example, who is the best hunter may be difficult to determine quickly or simply, given that there can be high levels of day-to-day variance in hunting returns, and even the best hunters can go through sustained periods without successful large kills. Rather than waiting to collect and compare data on hunting prowess, a novice hunter could simply use the ‘low-cost cue’ of who is the most prestigious hunter in the group (i.e., who is paid the most deference, and observed the most, by others), in order to determine who to copy.

Henrich and Gil-White argue that this kind of scenario is typical in human history: individuals who were able to pick up on deference or prestige cues and use them to selectively copy models would have had an adaptive advantage within their group, and therefore prestige bias would have been selected for.

Henrich and Gil-White describe the development of prestige bias as potentially resulting in a kind of prestige ‘market’, arising from the differing interests of learners and models. Learners are seeking to copy the most successful or skilful models, whilst models have an interest in acquiring learners in order to receive deference benefits. This may mean that there is competition, where models compete to attract learners. If there is competition, models have an incentive to be ‘nicer’ in order to attract more learners. They use this to explain supposed ‘subdominant ethology’ found in some high-status individuals (such as self-deprecation, deep bows and lowered heads).

Although the details are not explicitly set out in the Henrich and Gil-White paper, it is possible this was followed by gene-culture coevolution, whereby, once prestige bias is prevalent and adaptive, genes that aid the reliable expression of prestige bias are selected for and spread. Additionally, Henrich and Gil-White (and other prestige bias accounts) do not dwell on the cognitive basis of these biases (for example, whether they must be implicit and automatic, or whether they can be available to or the product of conscious reflection).

3 Assessing prestige bias explanations

Although the evolutionary story that Henrich and Gil-White offer seems plausible on its face, particularly when brought to life with examples of modern-day prestigious individuals, when we dig deeper we see that in order for the prestige bias account to have explanatory force, it entails a set of particular commitments, for which the empirical evidence is mixed. Once we relax these commitments, it becomes difficult to make the case for the explanatory power of prestige bias accounts over, for example, a naive goal-directed agent explanation.

I will highlight issues that fall into two broad themes: firstly, that empirical evidence in support of prestige bias uses a concept of prestige that is pruned in deference to prestige bias theory, in a way which casts doubt upon the extent to which this work constitutes evidence in support of prestige bias, and secondly, that the prestige bias account entails particular commitments regarding the cognitive basis of prestige bias, and the extent to which it will be flexible and sensitive to context. Both these themes have a direct impact on the comparison between a prestige bias account and a goal-directed agent account. This comparison will be drawn explicitly in Sect. 4.

3.1 Characterising prestige

The first key problem with the prestige bias account concerns the vagueness regarding how the concept of prestige is understood. Precisely what prestige is comprised of is not fully characterised in Henrich and Gil-White's account, or indeed in other accounts of prestige bias. This is important to note, particularly when considering the evidence from empirical studies which use various proxies for prestige. There appears to be a disconnect between how prestige is understood within the Henrich and Gil-White account, how prestige is understood in general usage, and how prestige is operationalised in empirical work on prestige bias. In particular, the definition of prestige used in empirical work is often pruned in deference to the prestige bias account itself, calling into question the extent to which these studies can be used as evidence of prestige bias.

Henrich and Gil-White give a tentative ‘ethology’ of prestige, where they describe prestigious individuals as in control of the time spent with ‘subordinates’, having a free posture without grandstanding or violent gestures, and appearing confident yet self-deprecating. This is drawn in explicit contrast to dominance, where individuals rely on aggression and threat of physical force to maintain their rank. However, this is not comprehensive in that many individuals in a society may display these kinds of behaviours, without being regarded as ‘prestigious’. By this I mean that there are many individuals, at least in modern societies, that do not display either dominant (aggressive) behaviours, or subordinate (threatened or submissive) behaviours. They may project confidence in their interactions with others, and be in control of who they spend their time with, and yet not enjoy the status benefits we typically associate with prestige (and will not be named as prestigious by others). In fact, the observed behaviours of prestigious people may not differ from less prestigious people in any systematic way, other than that the behaviour of others may be modulated in the presence of people they consider prestigious.

Additionally, it is unclear to what extent their use of the term ‘prestigious’ lines up with the common usage of the term. They use the example of Stephen Hawking as a prestigious individual in modern-day society (which would line up with our common sense notions of prestige). However, the common-sense notion of a prestigious individual does not require that many people are looking to copy their behaviours or ideas.

In Henrich and Gil-White's evolutionary story, the most prestigious individuals are measured by the individuals that most learners are seeking to copy. Even if prestige bias

in modern-day societies operates differently than in ancestral environments, we would still minimally expect prestige bias to drive individuals to seek contact or opportunities for observation with prestigious models, and to be driven to preferentially copy them. It is therefore unclear whether the key component of prestige is deference from others and conferring of status onto prestigious individuals, or the seeking of prolonged access to and opportunities to copy from prestigious individuals from learners. We may expect these to often come apart.

The lack of a consistent conception or definition of prestige becomes a pressing issue when assessing the empirical evidence for the existence of prestige bias. In some lab-based studies prestige is measured by the number of people observing a model. In a 2012 study by Chudek et al. the ‘prestigious’ model was the one which was observed by two other individuals, while the other was not. Their measure of prestige is simply that other individuals are observing the model. Presumably this measure is chosen because of its relationship to the operation of prestige as described in the evolutionary origins story of Henrich and Gil-White: the number of people observing a model (and therefore paying deference benefits for access) was the low-cost cue that new learners could use. However, this appears to significantly differ from both common-sense notions of prestige, the identification of prestigious individuals in studies based in real world networks, and several examples or potential applications given in theoretical work.

A similar definition of prestige is used in a study by Atkisson et al. (2012), where participants were tasked to ‘design’ an arrowhead on a computer, with the opportunity to modify their design based on arrowheads which were presented alongside information regarding the ‘prestige’ of the individuals that designed them. Here prestige was represented by the time that four individuals spent examining the given arrowhead (generated randomly by the computer), where arrowheads designed by more prestigious individuals were examined by others for longer periods of time. In both the Chudek et al. and Atkisson et al. study, prestige is defined in a way that presupposes the concept of prestige given in Henrich and Gil-White’s account. To make use of an analogy, this is as if we are seeking to test the connection between depression and lower serotonin levels, and in our test we define depression as lower serotonin levels. We can therefore question the extent to which studies which define prestige in this way are in fact evidential support for prestige bias.

A study by Acerbi and Tehrani (2018) did use a definition of prestige which is congruent with our general notion of prestige, in their test of the relative role of content versus context biases in the selection of quotations. Quotations varied in their content, and also in their context (either by being associated with a popularity score, measuring conformist bias, or in the prestige of the individual they were attributed to, measuring prestige bias). Individuals were then asked to choose their preferred quote. They measured prestige by attributing quotes to either a famous individual (i.e., one we would generally consider prestigious), or to an unknown author. However, they did not find a statistically significant preference for quotes by more prestigious individuals.

Therefore, there appear to be, broadly, two definitions of prestige operating in the literature. There is the one found in much of the theoretical work, and in studies in small-scale societies, where prestige typically involves acknowledged success or skill in a given domain or domains, the payment of deference and of status-related goods.

This may or may not coincide with other individuals seeking to observe and spend time with prestigious individuals, and seeking to copy their behaviours. Alternatively, there is the operative definition of prestige used in lab-based studies, which focuses around the number of individuals that observe a model. This can occur without any of the characteristics associated with prestige in the first sense.

3.2 The cognitive basis of prestige bias

In addition to the tension between definitions of prestige which affect our assessment of the empirical evidence for prestige bias, I will argue that prestige bias as understood by Henrich and Gil-White is dependent upon a particular understanding of social learning biases as unconscious or not amenable to reflective consideration.

There are broadly two ways to understand what ‘prestige bias’ refers to. The first is on a purely populational level, where we see a pattern towards copying the behaviours of more prestigious people, with no commitment to how this bias is realised on a cognitive level. The bias could be the product of intelligent reasoning and conscious thought, it could be implicit and automatic, it could have an affective dimension, or not. The second way is to place constraints on what this bias entails: for example, understanding it as a subpersonal, automatic, implicit process (which has implications for how it operates). Here, I will argue that, while it is tempting to take a populational view that avoids specific cognitive commitments, the explanatory value of prestige bias explanations is in part determined by what kind of cognitive processes the bias is constituted by.

Cecilia Heyes, in her 2018 book *Cognitive Gadgets: the Cultural Evolution of Thinking*, outlines two ways to conceptualise the mechanisms that make social learning selective (i.e. that give rise to biases such as prestige bias): the ‘strategic’ approach and the ‘attentional’ approach. The strategic approach implies that the selectivity occurs at the output stage: in the case of prestige bias, if an observer is exposed to two models performing different actions (one more prestigious than the other), she will encode both inputs, and then when confronted with a situation where she has to make a decision about which action to choose, she is more likely to choose the action performed by the prestigious individual. In this approach, the agent “uses” a strategy, which Heyes believes implies that this depends on reportable, high-level processes, rather than low-level, automatic processes. In contrast, the attentional approach asserts that selection occurs at the point of information reception: if an observer is exposed to two models performing different actions (one more prestigious than the other), she will attend more closely to the more prestigious model, and therefore will learn more about this action than the other. In this case the bias is due to the modulation of learning by low-level or automatic attentional processes, rather than the application of an explicit rule.

Heyes argues that those who adopt the strategic approach tend to assume that these domain-specific, high-order selective social learning bias are genetically inherited, and see prestige bias as a “cognitive instinct” (Heyes 2018, p. 87). She argues that it is in fact preferable to understand these kinds of social learning biases as what she terms ‘cognitive gadgets’, or metacognitive rules or strategies. These metacognitive

strategies are likely to be culturally rather than genetically inherited, transmitted with fidelity and accuracy through training and socialisation, and may take highly domain-specific forms (such as the role ‘copy the boat builder with the largest fleet’). She claims that, while the attentional approach is plausible for a lot of social learning, especially in non-human species, the kind of social learning biases that make humans uniquely capable of cumulative culture involve this kind of metacognition.

If we understand prestige bias as implicit and automatic (or as a domain-general, attentional process), then the value of a ‘prestige bias’ explanation over a general ‘goal-directed agent’ explanation is evident. Not only do they differ in the cognitive details (prestige bias is subpersonal, not available to conscious deliberation and rational reflection), but they would also clearly differ in their predictions. For example, an individual with general prestige bias would show a tendency to copy the behaviours of a prestigious individual even in situations where skill did not correlate with prestige, and where copying the behaviour would not be an effective way of harnessing expertise. We would expect less flexibility, and less ability for individuals to reflect on and report on these biases. The differences in the explanatory capacity of these accounts will be elaborated on in the following section.

However, if we understand prestige bias using the strategic, or metacognitive approach, things become less clear. To take Heyes’ example of a metacognitive rule, ‘copy the boat builder with the largest fleet’, it is unclear how we would distinguish cognition and behaviour based on such a ‘rule’, compared to the intervention of general intelligence. For example, an individual may, upon conscious deliberation, decide that the copying the boat builder with the largest fleet is what is most likely to enable them to build the best boat. This may not be the same strategy they follow when building something else, and this strategy may change depending on circumstances. If this is the case, then it is not clear what the advantage is of conceptualising these behaviours as ‘rule’ or ‘strategy’-following. Heyes’ examples include the boat-builder, and ‘copy digital natives’ (i.e., copy those that we judge to have digital expertise) (2018, p. 82). We could think of similar rules for any kind of human decision-making, such as ‘copy the writing style of successful academics’. However, it is unclear that this constitutes a plausible reflection of human cognition. Heyes gives us no reason to identify certain behaviours as rule-following, and as examples of culturally evolved ‘cognitive gadgets’, over others.

Precisely what kinds of cognition underlie these biases matters. It matters because it will change the predictions of prestige bias explanations, which will either lie in contrast to or bear more similarities to goal-directed agent type explanations. An approach that remains agnostic to the cognitive underpinnings loses explanatory value. Imagine it were the case that we could understand individual behaviours just as well using a goal-directed agent type approach, understanding decisions through the goals and limitations of individual agents, but when seen from a populational level, a prestige bias pattern emerges (we find it is often effective to copy the most prestigious people). We could term this a prestige bias, but we would then lose what is distinctive from the existing accounts. We would lose the adaptive explanation for the emergence of prestige hierarchies, and an explanation for prestige that does not rely on general intelligence. It would be unclear why we should see prestige bias as an interesting explanation for behaviour, even if a pattern emerges at the populational level.

3.3 The context-sensitivity of prestige bias

Following from this, understanding prestige bias as an unconscious or implicit social learning rule results in certain predictions about how prestige bias operates. These predictions are crucial to the explanatory value of prestige bias. If prestige bias is highly context-sensitive and reliant on complex and shifting calculations carried out by individuals regarding who best to learn from, it becomes increasingly difficult to separate out its predictions from a goal-directed agent account. Lewens (2015), in his assessment of conformist bias, highlights the need for conformist bias to not simply be a tendency to copy the most common traits, but be a tendency for individuals to *disproportionately* copy the most common traits in a group. In the case of prestige bias, learners should be biased towards copying prestigious people over and above what we would expect from a general account of individuals as goal-directed agents. In an extreme case, if learners reason that in certain contexts prestige would be a useful cue, deploy it only when they believe it to be beneficial, and are able to articulate why they are relying on it in these cases, a prestige bias account would not offer anything distinctive over a goal-directed account. This would not support Henrich and Gil-White's particular evolutionary story, and would call into question why we should understand population-level patterns of prestige-biased learning, if they exist, in terms of individuals holding this specific social learning bias, rather than the result of rational, reflective, goal-directed action.

Therefore, the explanatory value of prestige bias accounts is dependent upon the extent to which the tendency for learners to copy prestigious individuals is the result of reflective deliberation or goal-directed intent. One indication of this is the extent to which prestige bias is cross-domain. By this I mean that when learners copy the behaviour of prestigious individuals, they copy their behaviour in many respects, not just a very limited subset of behaviours that are directly responsible for adaptive outcomes.

The evolution of prestige bias is based on prestige as a low-cost cue for who is best to copy. In this scenario, it would be too costly (in terms of time, material resources, or potential for error), or difficult, to directly discern which behaviours are adaptive and copy these, rather than gaining prolonged access to a model in order to copy many behaviours. If individuals were able to hone in on the behaviours of successful models that are responsible for their success, then presumably cooperation of the model in providing proximity and interaction would play a far smaller role. Additionally, if learners are copying the adaptive behaviours prestigious individuals only in cases when it would be adaptive to do so, this behaviour could be explained by a goal-directed agent account. It would therefore be difficult to see the explanatory gains from positing prestige bias in particular to explain these phenomena.

At least in modern-day societies, it does not seem to be the case that individuals copy suites of behaviours and practices from the most prestigious individuals. We can see this more clearly by looking at some of the examples used. In the prestige bias literature, we find a mix of informal and formal examples. One often invoked example is the existence of celebrities. Authors frequently gesture at the ability for prestige bias to explain why people in modern-day societies have a tendency to copy

the clothing choices of celebrities, amongst other things (e.g., Henrich 2001; Mesoudi 2009; Jiménez and Mesoudi 2019). However, prestige bias seems to be an ill-fitting explanation for this phenomenon in two ways. Firstly, many prestigious individuals are deferred to and given prestige by others who do not wish to copy them. As previously outlined, Henrich and Gil-White offer the example of Stephen Hawking as a prestigious individual. However, Stephen Hawking was regarded as ‘prestigious’ by a large number of people (who may well have paid him deference and other benefits of prestige), whilst we might imagine that only a very small number of people would be looking to copy his behaviours or ideas (most broadly, physicists). Secondly, celebrities who are copied are often strikingly *not* copied in domains that are responsible for their success. For example, many individuals might copy the style choices of Odell Beckham Jr., with no intention of trying to copy his American football skills, or the make-up of Kylie Jenner, without copying her business practices. If prestige bias is an unconscious or implicit bias towards copying prestigious people where learners cannot discern which behaviours are responsible for success, we would expect learners to copy all (or a large set of) the behaviours of a prestigious individual. In our evolutionary history, if prestige-biased learners selectively copied the non-adaptive traits of prestigious individuals whilst not copying the adaptive traits, prestige bias would not have been adaptive and would not have been culturally selected for.

Although it is outside the scope of this paper to defend a particular alternative explanation here, one possible explanation of the existence of prestigious individuals who are not generally copied from (such as Stephen Hawking) might be through a process of rational reflection. For example, individuals might value traits such as intelligence, and afford respect to those who they perceive as exemplifying these traits. Additionally, individuals may copy the style of someone like Kylie Jenner due to (in part) biologically and socially shaped aesthetic preferences. It is difficult to see how positing prestige bias explains features of the influence of celebrities in a way that intention-centred explanations cannot.

Additionally, empirical studies have found mixed evidence for the cross-domain action of prestige bias. Chudek et al. (2012) carried out studies where they claimed to find evidence of prestige-biased learning in children. In these studies they showed groups of children (around 4 years old) two adult female ‘models’ performing a series of actions. Firstly, they were shown the models making a choice between two objects. One model was being observed by two other individuals, who were not interacting with the model or each other, only gazing at the model (the ‘prestigious’ model). The other model was not being observed. Then, they were shown the same two models making a series of preferences: for example, picking toy A over toy B, or biscuit A over biscuit B. They were then given the opportunity to make choices between the two options themselves. They found that children were more likely to copy the choice of the ‘prestigious’ model, although this was significant for some domains and not others (in one study, for artifacts, but not for food or drink preferences, and in another, only for the domain that prestige was cued in).

Studies in small-scale societies have also found differences in the relevance of prestige in different domains, as well as the connection between prestige and age. Reyes-Garcia et al. (2008) studied connections between ethnomedicinal plant knowledge and prestige within an Amazonian group called the Tsimane’, and did not find a

clear association between such plant knowledge and level of prestige, and no association between prestige and age. Another study in Fijian villages examined success, knowledge, and model selection for three socially valued domains: fishing, yam-growing, and ethnomedicinal plant use (Henrich and Broesch 2011). Success in fishing and yam-growing were far better predictors of prestige than ethnomedicinal plant knowledge. Additionally, they only found age to be a predictor of model selection for some domains in some samples, and individuals were more likely to select models of the opposite gender to their own. Prestige bias models predict that older individuals are more likely to be copied and to have prestige, due to an average increase in skill and knowledge, and that same-gender individuals are more likely to be copied (in a society where what constitutes adaptive behaviours differs between genders).

Brand and Mesoudi (2019) found similar indications that prestige may be domain-specific. They investigated prestige and dominance hierarchies in groups of adults who had pre-existing relationships (they were members of established community groups, rather than participants placed into a new group for the purpose of a study). In their study, the prestige ratings of an individual given by their group members did not predict that individual's performance on the group task they were set, or who was selected to represent the group for a bonus task. The selection of the group representative was based not on perceived prestige or dominance, but correlated with the individual's actual performance in the group task.¹ The potential domain-specificity of prestige-biased learning, or who is accorded prestige, matters for the assessment of prestige bias as an explanation. If prestige-biased learning is highly domain-specific, this becomes easier to explain in terms of the goal-directedness of agents: individuals may be evaluating when it is appropriate to rely on prestige cues, and when it is not.

Jiménez and Mesoudi (2019), in their review of the prestige bias literature, note the mixed evidence for the specific predictions of the Henrich and Gil-White model. Given this, they suggest modifications to these predictions. These modifications include restricting the presence of prestige bias to cases where the domain of prestige is currently valued for a social group, where individuals carry out tasks which are domain-relevant and difficult, when the variation in knowledge and skill is large, and when prestige is highly correlated with success. Although these modifications bring the prestige bias account further into agreement with the empirical evidence, they do so at the expense of explanatory power. If prestige bias is no longer a general ‘rule-of-thumb’, deployed implicitly and unconsciously, biasing social learning in clear and systematic ways, but rather a highly context-dependent phenomenon that relies on individuals having access to large amounts of social information, it becomes more difficult to make the case that prestige bias offers explanatory benefits that a goal-directed agent account does not.

However, there is empirical work which does appear to suggest a prestige bias account over a goal-directed account. Firstly, evidence of overimitation, where people have a tendency to copy irrelevant actions carried out by prestigious individuals, is suggestive of a systematic bias that is not easily explained through intentional behaviour.

¹ Additionally, Jiménez and Mesoudi (preprint) tested the effects of prestige on the transmission of information. A prestige bias account would predict that information from more prestigious people would be more accurately recalled and transmitted. They did not find evidence for the effect of prestige on information transmission.

McGuigan (2013) carried out a study in children that showed that they were more likely to overimitate when copying high social rank models than low social rank models. However, a study by Chudek et al. (2016) found no difference in the propensity to overimitate high versus low status models. Secondly, cases where individuals were not consciously aware of the effect of prestige in deciding who to learn from would speak in favour of prestige bias, as an unconscious or implicit bias, over a simple goal-directed account. Priestley and Mesoudi (2015) found that users of the social media website Reddit ranked ‘social influence’ low on the factors that drove them to ‘upvote’ or ‘downvote’ content. However, as they note, previous work has suggested that artificially adding ‘upvotes’ (which could be construed as a measure of social influence) significantly increased the chance of further ‘upvotes’ (Muchnik et al. 2013). If this is indicative of a broader pattern whereby individuals are not aware they are acting in a prestige-biased manner, this would support the prestige bias account over a goal-directed or intention-based explanation. Further work is needed to determine the extent to which prestige-biased patterns of learning are flexible and context-sensitive, and individuals are aware of the drivers of their decision-making.

4 Goal-directedness: an alternative explanation

In this section I will make comparison explicit between the prestige bias account and one proposed alternative, the ‘goal-directed agent account’. I argue that, upon close examination, it is unclear that the prestige bias model represents an increase in explanatory power over this naive, or common-sense model.

A goal-directed agent account treats individuals as intentional agents, acting in pursuit of their goals, where those actions are relative to their epistemic limitations and the social landscape. This does not necessarily mean that the preferences or choices of individuals will always be the adaptive optimum; there are many factors that may circumscribe or shape choices. This could also be characterised as a ‘common-sense’ view of individuals as agents, where individuals’ actions can be understood in terms of pursuit of their goals. Under this account, individuals would partake in a process of general evaluation of various relevant factors, that would be flexible and context-dependent.

The motivation for developing kinds of cultural evolution models frequently stems from the drive to explain phenomena that seem difficult to explain through a standard view of humans as goal-directed agents (such as the cassava case outlined by Henrich 2015). If some phenomena could in fact be adequately captured in this way, it would significantly diminish the explanatory potential of many kinds of cultural evolution explanations. In the context of a critique of the memetics project, Greenberg (2004) draws a contrast between memetics explanations and a ‘commonsensical’ account of cultural change that sees individuals as in “deliberate pursuit of [their] conscious goals”. He points out that meme theory needs to offer a distinctive explanation of why ideas spread, which competes with and contributes something beyond a common-sense goal-based account. A similar situation applies to other cultural evolutionary approaches. This is an issue because the onus is on cultural evolution approaches to show their distinctive explanatory value: ‘goal-directed agent’ type explanations are

so widespread and demonstrably effective in many domains, including our daily lives and common-sense understanding of others, that it is unclear why we should shift to a less intuitive framework that downplays human agency, if there is not a clear case made for its increased explanatory or predictive power.

It is for these reasons that I choose to contrast prestige bias explanations with ‘goal-directed agent’ explanations, to tease out their explanatory potential. This is not intended as an argument for a simplistic, goal-directed agent explanation as the best explanation for certain phenomena; other possible competing explanations could make reference to social or institutional structures, or genetically encoded aspects of cognition. However, if prestige bias explanations cannot offer anything distinctive over a common-sense, intuitive, everyday understanding of the behaviour of others, it would cast their usefulness into doubt.

To illustrate how these explanations would differ, here I elaborate on an example given by Atkisson et al. (2012) to represent prestige bias. This example is of a woman who marries into a new community, with different specialisations than the community she grew up in, and where the women are responsible for making stone arrowheads.

i. *Prestige bias explanation*

The new prestige-biased learner looks around and determines who the most prestigious arrowhead maker is in the group (defined by Atkisson et al. as the person that everybody pays the most attention to). She then pays the cost of access in terms of deference, and copies the behaviours of this arrowhead maker. If the costs are too high, she may choose a less prestigious model with lower costs of access.

ii. *Goal-directed agent explanation*

The new learner considers various factors in pursuit of her goal to make good arrowheads. She may consult friends in the group, explore through individual learning, or try to observe more successful arrowhead makers. If she pays deference to the most prestigious arrowhead maker, this could be (for example) because of a conscious calculation that deference is necessary in order to obtain the information she would like, or because she feels admiration for the prestigious arrowhead maker (which could occur whether or not she copies this individual), or because of the desire to conform to social norms.

The goal-directed agent model may give the same predictions as the prestige bias model in some cases, and may depart in others. However, even when the predictions are the same, the goal-directed agent model cleaves to how we typically understand individuals’ actions and has high *prima facie* plausibility, so the onus is on prestige bias models to provide explanations that improve on this in some way.

4.1 Comparing predictions

The main advantages of prestige bias models, as outlined by Henrich and Gil-White, appear to be in explaining the existence of prestige hierarchies in human societies in the first place, and in making predictions that individuals will seek contact with prestigious individuals, will copy them and pay deference to them, and that prestigious

individuals will be highly skilled and successful, and be able to maintain their status without aggression or dominant behaviour.

They may be correct in asserting that a tangible goods theory would struggle to explain why people show deference even when individuals stop producing tangible exchangeable goods, why they copy their behaviours and opinions, and why the dynamics of human social status rankings seem to differ from those of other social species. However, prestige bias models face similar explanatory challenges. As argued in the previous section, the explanatory power of prestige bias depends upon, at least in part, the extent to which it is unconscious, implicit and systematic, rather than flexible, context-sensitive, and amenable to rational deliberation. The empirical evidence seems to point towards the latter in some respects.

Is it possible to adequately explain the existence of prestige in itself without invoking prestige bias? Prestige bias models do give a compelling story for why prestige would have evolved in the first place. However, prestige could be explained through other means. For example, people generally accord prestige to those they perceive to be highly skilled or knowledgeable (and, as already argued, not necessarily those they wish to copy). This could be understood as valuing the skills or knowledge of the other individual, and translating that into admiration or respect. Additionally, the influence that prestigious people often have could be understood as others judging them to be a good source of information or advice based on direct assessment of their skill or knowledge, rather than a bias to copy based on prestige alone.

A possible response to this is to invoke proximate versus ultimate considerations here. Could it not be the case that emotions such as admiration are proximate mechanisms by which prestige bias is realised? This is where it becomes important to clarify what the cognitive basis of this bias is, and whether this matters to prestige bias explanations. However, my intention here is to note that an explanation that does not rely on an adaptive basis to prestige, and instead (perhaps) understands it as a side-effect of our general capacities for intelligence and reasoning, is possible and plausible. Here, it is not clear that a prestige bias account gives us new resources to explain the existence of prestige.

I consider two studies whose results have been interpreted as evidence for the existence of prestige bias, in light of both the prestige bias and the goal-directed agent account. I argue that neither of these studies provide convincing reasons to favour the prestige bias account over the goal-directed agent account.

4.2 Case 1: building arrowheads

Atkisson et al. (2012) purport to demonstrate prestige-biased learning in adults. In this study they sat participants at computers and asked them to ‘design’ an arrowhead, that could vary in several dimensions, and then would be used to go on ‘hunts’. The closer their design was to the hidden optimal design, the higher their payoff would be. They had a series of hunts, with an opportunity each time to improve their design either through individual or social learning. If they chose to learn ‘socially’, they were shown a choice of five arrowheads, and if they clicked on one, they were given information about ‘prestige’. This information was the names of four other individuals and the

amount of time these individuals spent examining the given arrowhead (generated randomly by the computer). They found that individuals were more likely to select an arrowhead that was more ‘prestigious’ (those where they were given information that people had looked at it for longer).

Of course, methodological issues can be found in most studies in any areas. Additionally, with the necessarily atypical settings of any lab-based study there will always be questions about external validity. However, the issues that arise with these studies point to a deeper conceptual confusion in the prestige bias literature. As mentioned in the previous section, these studies often use proxies for prestige that may not adequately capture the concept as used in theoretical work.

The Atkisson et al. study in particular represents prestige in an unusual way. The participants do not even have access to individuals to observe, rather they can only observe the arrowheads produced, and only have information on the length of time people spent looking at the arrowhead. This is very far from prestige as described in theoretical work, or as used in daily language. In fact, prestige bias supposedly arose in part because of the difficulty of gaining useful adaptive information directly from sources such as artefacts. If this were possible, continued access to prestigious people would be much less important. Here there is no deference component (there is no indication that individuals regard the makers of the arrowhead as prestigious, or act as such), and no opportunity to copy other behaviours other than the arrowhead design itself. Although this study may point to something interesting about the way in which we learn, it does not appear to provide evidence for the prestige bias account.

If prestige bias is implicit, and automatic, we would expect it to be cued by signals such as the way that other individuals act around a prestigious person (i.e. deference). In the Atkisson et al. study participants had access to information that had to be interpreted: they would have had to connect the written information that others were looking at an arrowhead to some internal conception of prestige, where prestige is not directed towards an identified individual, but is rather routed through the product of (presumably) someone’s labour (the arrowhead). This seems be consistent with a deliberative, conscious process of reasoning, where these connections can be made.

It is therefore hard to separate a prestige bias explanation here from a goal-directed agent explanation, as the influence of conscious and deliberative processes seems clear. Alternatively, these results could even be construed to support another social learning bias: conformist bias. The information that others were looking at the arrowhead could have been interpreted as other individuals also choosing that arrowhead, and therefore we could tell a story of a bias towards choosing the arrowhead that others were choosing. There are likely to be issues with this interpretation also, however, I mention it to point out that the existence of a given social learning bias here is severely underdetermined by this empirical work.

4.3 Case 2: the evolution of food taboos

Henrich and Henrich (2010) carried out a study which they characterise as a case of prestige-biased transmission. They aim to show that social learning can give rise to ‘culturally evolved adaptations’, and they do this through examination of a population

of pregnant women in Fiji. The story they tell is that in Fiji, there are certain marine species that carry a high risk of toxins that cause food poisoning, with particularly dangerous effects for foetuses and nursing infants. In the Fijian population, there exist food taboos for pregnant and breast-feeding women surrounding these particular species, which cause them to remove those species completely from their diet. Pregnant and breast-feeding women ‘acquire’ these taboos both from close female relatives and from a small number of women in the population who are identified as being particular sources of knowledge on these taboos: ‘yalewa vuku’ or ‘wise women’.

According to Henrich and Henrich, close female relatives are low-cost, accessible learning models who share fitness incentives with the learner. As defined by the authors, yalewa vuku are “women who are well-respected and considered knowledgeable about traditional medicine, birthing and childcare” (p. 3720): these women are ‘preferred models’, and by selectively learning from them, learners can improve on the cultural variants acquired from their family members. Over time, this tendency stabilises a population at an adaptive equilibrium. The ‘adaptive’ behaviour of avoiding toxic marine species is explained through patterns of biased transmission, likely built through rounds of previous genetic and cultural selection.

In order to clarify the explanatory merit of the prestige bias model, I wish to introduce a comparative example. Let us consider the following case:

A woman living in Cambridge becomes pregnant for the first time. As she is concerned for the health of her foetus, she wishes to avoid any behaviours that would affect its health during her pregnancy. She consults her mother, her family GP, and the NHS webpage on ‘foods to avoid during pregnancy’. When asked why she chose these sources of information, she explains that she wants to learn from the experiences of her mother, who has been through several pregnancies, and that she places trust in the expertise of her family GP and the recommendations of the NHS. After consulting these sources, she decides to avoid soft cheeses with white rinds, such as brie and camembert. This decision is beneficial, as these cheeses are known to contain listeria bacteria.

This case follows the same structure as the case of Fijian food taboos. Both cases involve the expression of an ‘adaptive’ behaviour (in that it promotes the health of the foetus), which is acquired through social learning. The individuals in question selectively learn from close female relatives and ‘preferred models’.

However, in the second case, a prestige bias model seems intuitively implausible or unnecessary. This is because a goal-directed agent-type explanation jumps out at us as being the most plausible and most explanatory. We can explain the Cambridge woman’s behaviour through standard goal-directed means. She is motivated to maximise the health of her foetus, and synthesises information from various sources to achieve her goal. She chooses these sources based on a variety of reasons, all within conscious control (i.e., the decisions can be understood without reference to automatic, subpersonal, or domain-general processes).

Given the similar structure of the Cambridge case to the Fiji case, we can also apply this goal-directed agent explanation to the Fijian case. The Fijian women under study are motivated to maximise the health of their foetuses, and therefore consult a range of sources (family, and known experts) in order to determine the best course of

action. What this shows is that prestige bias models do not provide any explanatory advantage over common-sense goal-directed models in this case, with goal-directed models having higher *prima facie* plausibility.

These two cases are an indication that the evidence proposed in support of the existence of prestige bias can likely be equally as well explained by a simple goal-directed agent model. Therefore, it is difficult to make the case for understanding these aspects of social learning through this specific prestige bias account, which comes with theoretical commitments and entails particular assumptions.

5 Implications and conclusion

I have argued here that we have reasons to doubt the plausibility and explanatory power of a prestige bias account. Differences in the definitions of prestige used in theoretical and empirical work cast doubt upon the status of the empirical evidence for prestige bias. Additionally, I have highlighted that prestige bias as an explanation depends on a particular view of the cognition underpinning the bias. This is associated with particular predictions about the bias as implicit, systematic, and not available to conscious reflection. Empirical evidence suggesting prestige-biased learning is flexible and context-sensitive causes us to question whether we should favour the prestige bias interpretation over a goal-directed agent explanation. Furthermore, modifying the predictions of the prestige bias account to bring them in line with the empirical evidence weakens their explanatory value. I have demonstrated the way in which a goal-directed agent account could explain some of the empirical evidence for prestige bias through explicit consideration of two cases, demonstrating that in these cases a goal-directed explanation may do just as well.

Social learning biases, and prestige bias in particular, form an important component of work into cultural selection processes. If there are reasons for skepticism about the existence of prestige bias, what consequences does this have for cultural selection models in general? I argue that, in fact, this does not have to have a significant negative impact on the plausibility of at least some cultural selection processes. To take the example of cultural group selection, prestige bias largely functions as a mechanism for minimising within-group variation and maintaining between-group variation (Henrich 2004). If individuals are all copying the most prestigious individuals in a group, that increases the likelihood that most individuals in the group are expressing the same cultural variant. In this case, as long as there is some mechanism for achieving this, it does not really matter whether it is prestige bias. For example, people could be learning in what looks like a prestige-biased pattern, without the existence of implicit, automatic biases towards prestigious individuals. In fact, everyone could be acting in a goal-directed way, explicable through a rational agent type explanation, and yet this still perpetuates enough within-group conformity for cultural group selection to take place. To return to the medical example outlined above, if everyone believes that their doctor is an important source of expertise, and doctors share the same opinions on, for example, what pregnant individuals should avoid, then we may well expect conformity within a group on tabooed foods during pregnancy.

However, there are still some important implications for cultural selectionists. Cultural selection explanations depend upon human intelligence or creativity factoring into these explanations in a limited way. This is not to say that cultural selection cannot accommodate human intelligence, but rather, that the explanatory power of cultural selection explanations derives from their ability to give us understanding of how cultural patterns or traits accumulated and emerged without resorting to human ingenuity as the locus of explanation. Cultural selection explanations have appeal when applied to cases that seem unexplainable or poorly explained through intelligent individual decision-making (Henrich 2015). This is why social learning biases have formed such a key part of the cultural selection research programme: we can construct an adaptive story about the emergence of these biases, which then themselves form the basis for selective processes that rely on automatic, domain-general ‘rules-of-thumb’, rather than domain-specific, intelligent, individual decisions. If the reason why fitness-enhancing variants spread is deliberate, conscious, reasoned adoption by individuals, rather than an automatic, evolved instinct or bias, then we would not expect selection dynamics to necessarily hold, or selection models to be useful tools.

Future work to further explore the plausibility and value of a prestige bias account in understanding prestige-biased learning should aim to tease apart and test the predictions made by prestige bias accounts compared to alternatives.

Acknowledgements I am deeply grateful to Tim Lewens, Andrew Buskell, and the two anonymous reviewers for their detailed and very helpful comments. This work was supported by the Arts and Humanities Research Council.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acerbi, A., & Tehrani, J. J. (2018). Did Einstein really say that? Testing content versus context in the cultural selection of quotations. *Journal of Cognition and Culture*, 18(3–4), 293–311.
- Atkisson, C., O’Brien, M. J., & Mesoudi, A. (2012). Adult learners in a novel environment use prestige-biased social learning. *Evolutionary Psychology*, 10(3), 147470491201000309.
- Bell, A. V. (2013). Evolutionary thinking in microeconomic models: prestige bias and market bubbles. *PLoS ONE*, 8(3), e59805.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(Supplement 2), 10918–10925.
- Brand, C. O., & Mesoudi, A. (2019). Prestige and dominance-based hierarchies exist in naturally occurring human groups, but are unrelated to task-specific knowledge. *Royal Society Open Science*, 6(5), 181621.

- Cheng, J. T., & Tracy, J. L. (2014). Toward a unified science of hierarchy: Dominance and prestige are two fundamental pathways to human social rank. In J. T. Cheng, J. L. Tracy, & C. Anderson (Eds.), *The psychology of social status* (pp. 3–27). New York: Springer.
- Chudek, M., Baron, A. S., & Birch, S. (2016). Unselective overimitators: The evolutionary implications of children's indiscriminate copying of successful and prestigious models. *Child Development*, 87(3), 782–794.
- Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: Bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, 33(1), 46–56.
- Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, 29(1), 56–64.
- Greenberg, M. (2004). Goals versus memes: Explanation in the theory of cultural evolution. In S. L. Hurley & N. Chater (Eds.), *Perspectives on Imitation* (pp. 339–353). Cambridge: MIT Press.
- Henrich, J. (2001). Cultural transmission and the diffusion of innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change. *American Anthropologist*, 103(4), 992–1013.
- Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, 53(1), 3–35.
- Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton: Princeton University Press.
- Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, 19(4), 215–241.
- Henrich, J., & Boyd, R. (2002). On modeling cognition and culture. *Journal of Cognition and Culture*, 2(2), 87–112.
- Henrich, J., Boyd, R., & Richerson, P. J. (2008). Five misunderstandings about cultural evolution. *Human Nature*, 19(2), 119–137.
- Henrich, J., & Broesch, J. (2011). On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1139–1148.
- Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22(3), 165–196.
- Henrich, J., & Henrich, N. (2010). The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proceedings of the Royal Society B: Biological Sciences*, 277(1701), 3715–3724.
- Henrich, J., & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, 12(3), 123–135.
- Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Cambridge: Harvard University Press.
- Jiménez, Á. V., & Mesoudi, A. (2019). Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications*, 5(1), 20.
- Jiménez, Á. V., & Mesoudi, A. (Preprint). Prestige does not affect the social transmission of controversial arguments. Preprint on PsyArXiv.
- Lewens, T. (2015). *Cultural evolution: Conceptual challenges*. Oxford: Oxford University Press.
- McGuigan, N. (2013). The influence of model status on the tendency of young children to over-imitate. *Journal of Experimental Child Psychology*, 116(4), 962–969.
- Mesoudi, A. (2009). How cultural evolutionary theory can inform social psychology and vice versa. *Psychological Review*, 116(4), 929.
- Muchnik, L., Aral, S., & Taylor, S. J. (2013). Social influence bias: A randomized experiment. *Science*, 331(6146), 647–651.
- Plourde, A. M. (2008). The origins of prestige goods as honest signals of skill and knowledge. *Human Nature*, 19(4), 374–388.
- Priestley, M., & Mesoudi, A. (2015). Do online voting patterns reflect evolved features of human cognition? An exploratory empirical investigation. *PloS One*, 10(6), e0129703.
- Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., et al. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, 328(5975), 208–213.

- Reyes-Garcia, V., Molina, J. L., Broesch, J., Calvet, L., Huanca, T., Saus, J., et al. (2008). Do the aged and knowledgeable men enjoy more prestige? A test of predictions from the prestige-bias model of cultural transmission. *Evolution and Human Behavior*, 29(4), 275–281.
- Richerson, P.J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.
- Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford: Blackwell.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.