Reputation for Competence: Social Learning Mechanisms Create an Incentive to Help Others

Douglas Schauer¹

Received: 21 June 2021 / Accepted: 24 January 2022 / Published online: 9 February 2022
© The Author(s) 2022

Abstract

Research on social learning has identified mechanisms that learners use to decide from whom to learn. Several of these mechanisms indicate that learners prefer to learn from more competent people over less competent people. This requires learners to measure the competence of other people. We use this article to analyze the incentives that this measure of competence creates. Learners measure the competence of models, people they would consider learning from, and share these judgments with other learners. This gives each model a reputation for competence within a community. Each model has a biological incentive to increase the magnitude of that reputation; at the very least, increasing the magnitude should make the model more attractive to potential mates. In this article, we present logic that indicates that one way for the model to increase the magnitude of that reputation is for the model to help people who think the model is competent. This reveals a novel evolutionary incentive for humans to help other humans.

Keywords Altruism · Competence · Cultural transmission · Helping · Prestige · Social learning

Introduction

Background

Much of what people do is not determined by genes, but rather is learned from other people. Customs, beliefs, practices, and methods of pursuing goals affect our biological fitness, but are not determined by genes alone. Many are learned from other people. Researchers have termed these customs, beliefs, practices, and methods “culture,” meaning some element of behavior that is passed from person to person by nongenetic means.

Prior work on cultural transmission has focused on two areas. One area is population dynamics. This work studies what happens when there are two or more cultural variants, for example two different customs, in a population, examining the prevalences of these cultural variants over generations. It asks whether one of the cultural variants is more adaptive than the others; and if so, whether the more adaptive one increases in prevalence over time,

decreases in prevalence over time, or if a stable equilibrium is reached (Boyd and Richerson 1985; Brown et al. 2011; Henrich 2016; Muthukrishna and Henrich 2016). The other area of work on cultural transmission studies the individual responses to culture. For example, it asks which cultural variant is adopted when a person is presented with two different options, and how that choice is made (Henrich and Gil-White 2001; Laland 2004; Mesoudi 2008; Henrich and Broesch 2011; Rendell et al. 2011; Jiménez and Mesoudi 2019).

Our work picks up from prior work on individual responses to culture, but rather than focusing on specific mechanisms that allow individuals to choose a cultural variant, we are going to take a more abstract approach. Our goal with this approach is to analyze the incentives that are created by decision mechanisms. As far as we know, no prior research has specifically focused on incentives that are created when an individual has to decide whether or not to adopt a cultural variant. In this article, we will present logic that implies that people have an incentive to help certain other people. If this logic is correct, and this incentive is real, it should be of interest to a wide variety of researchers in the sciences and humanities.

✉ Douglas Schauer
dougshauer1@gmail.com

¹ Frazier Park, CA, USA

Summary of Our Argument

When learners are presented with a cultural variant, there are many cues they can use to estimate whether it is adaptive or not, and thus choose to adopt it or not. Some of these cues are related to the cultural variant itself. For example, if they can evaluate the variant and determine that it is not adaptive, then they should not adopt it. However, much of the time, the learners will not be able to directly evaluate the cultural variant. Instead, indirect cues about its quality can be used. Many of these indirect cues come from the person who is modeling the cultural variant. Some of the indirect cues that prior work has identified include the success, skill, knowledge, age, health, and happiness of the model (Laland 2004; Mesoudi 2008; Rendell et al. 2011; Henrich 2016). We note that these cues seem to assess how good the model is, or how competent the model is. We abstract these cues to say that what the learners are really measuring is the competence of the model. If the model is more competent, learners have evidence that the behaviors the model displays are adaptive, so they will be more likely to adopt those behaviors. If the model is less competent, they are less likely to adopt the culture.

Before a learner can use the competence of the model to decide whether or not to adopt some culture, the learner must assess the competence of the model. One way to do this is by observing the model over time and estimating the model's competence based on those observations. Another way to estimate the model's competence is by getting information about it from an intermediate, third party. If the intermediary thinks that the model is competent, then the learner has some evidence that the model is competent. If the intermediary thinks the model is not competent, then the learner has evidence that the model is not competent. In this way the amount of competence that the learner should infer is a function of the intermediary's belief about the model.

In this article, we will present logic that shows that the amount of competence that the learner should infer is not only a function of the intermediary's belief about the model. It is also a function of the competence of the intermediary. If a highly competent intermediary behaves as though a model is highly competent, then the model probably is highly competent, and the learner should infer that the model is highly competent. This will make the learner more likely to learn from the model. However, if an average intermediary behaves as though a model is highly competent, then the model probably is above average, but may not be extremely highly competent. In this case, the learner should infer that the model is above average. If an incompetent intermediary behaves as though a model is

highly competent, then the model is probably only average, and this is what the learner should infer. The learner is then less likely to learn from the model. In this way, the amount of competence that a learner infers is an increasing function of the intermediary's competence.

Learners have incentive to accurately assess the competence of models, so that they can learn from the best ones, and maximize their fitness. However, this is not the models' incentive. Models, as for any individual, have an incentive to convince people that they are competent. At the very least, this should increase their fitness by making them more attractive to potential mates. It might also have other fitness-enhancing benefits.

The amount of competence that the learner infers about the model is an increasing function of the intermediary's competence. If the model can increase the apparent competence of the intermediary, then the learner will infer a higher level of competence for the model, and the model's fitness will increase. This gives the model an incentive to increase the apparent competence of the intermediary. The competence of the intermediary measures how good the intermediary is at achieving valued goals. If the model can increase the apparent achievement of the intermediary, then the intermediary will appear more competent and the model's fitness will increase. One way for the model to increase the apparent achievement of the intermediary is to help the intermediary achieve those valued goals. In this way, the model has incentive to help the intermediary. This creates a biological incentive for the model to help other people.

In this article, we will explain this logic in more detail, as well as compare it to a competence-based status called "prestige."

Learning Logic

We present logic that we believe will help us understand the incentives that are created when one person learns from another. The logic is somewhat abstract. We believe that it will help us understand human behavior. The logic would have to be implemented by mechanisms in the human mind and body. This article does not discuss these implementation mechanisms.

What to Measure

We begin by assuming that each person has goals that he or she wants to achieve. In order to model these goals, we assume that each of these goals has a value, indicating how important it is. The goals and values are subjective, meaning that each person has his or her own set of goals and values for those goals. However, there are some basic human goals that are valued highly by most individuals. For this article, it

does not matter exactly what these goals are, but these could be things like having enough food to eat, having clothing and shelter, and having access to mating partners.

An individual's goals guide his or her actions. These goals can be achieved to different degrees. For the most part, higher degrees of achievement will have higher value. Individuals attempt to maximize the achievement of their goals by maximizing the total value of the goals achieved. We measure this as the sum of the values of the different goals that are achieved.

Individuals' achievements are not determined by their genes alone. They can learn to achieve their goals to higher degrees by learning from other people. In order to learn the behaviors that will result in the highest achievement of their goals, they should prioritize learning from other people who are good at achieving these goals. And in order to figure out from whom to learn, they need to measure or estimate how good other people are at achieving goals, to learn from the best people.

Learners have a variety of goals that they would like to achieve. They could try to figure out who is good at each goal separately and choose different models for different goals. However, there might be problems with this approach. The first problem is an opportunity cost problem. If a potential model is good at goal 1, but not good at goal 2, it might be that she uses methods that prioritize goal 1 over goal 2. If the learner adopts these methods, he will achieve goal 1 to a high degree but goal 2 to a low degree. The total of his goals will not be achieved to a high degree, so learning from this potential model will not benefit him. Even if this problem is solved and the learner is able to figure out who is efficient at achieving goal 1 but not at goal 2, there is another problem to solve, borrowed from Henrich and Gil-White (2001). The potential model will display many behaviors. It will be difficult for the learner to distinguish which behaviors cause goal 1 to be achieved well and which behaviors do not. Likewise, it will be difficult for the learner to distinguish which behaviors cause goal 2 to be achieved poorly and which do not. In this circumstance, the learner might choose the wrong behaviors, and again not learn to achieve the total of his goals to a high degree.

These are a priori arguments that a learner might not be able to tell which model is good at which goal or which of a model's behaviors make him good at this goal. However, if a learner can figure out which goals a model is good at and which actions make him good at this goal, then it might be beneficial for learners to perform domain-specific learning, in which they only copy behaviors that improve their achievement of a particular goal that the model is good at. They can then choose different models for different goals. Experimental research has shown that when learners are given a specific domain in which to learn and are presented with information that specifies which models are

knowledgeable in that domain, they prefer to learn from the models who are knowledgeable in that domain over models who are identified as having more general knowledge. This assumes that a learner is trying to learn in one specific domain and information about which models are knowledgeable in that domain is available to the learner. However, if information about which models are knowledgeable in the domain of interest is not available, the learners will prioritize learning from models who have general knowledge over models without general knowledge (Brand et al. 2021).

In this article, we will refer to being good at achieving the learner's goals as "competence." The people who are better at achieving the learner's goals will be more competent, and the people who are worse at achieving the learner's goals will be less competent. This could be domain-general competence, in which the learner measures competence across several or all goals or it could be domain-specific competence, in which the learner makes different measures of competence for different goals. Competence is subjective because it depends on the goals of the particular learner. However, since there are some goals and values shared across a wide range of people, there will be some consensus among learners about who is competent and who is not.

Lastly, we need to model competence for the purposes of this article. Competence is a measure of how good a person is at achieving a particular set of goals. We are going to model this as the expected value of his or her behavior, measured according to the goals and values of the learner. For any behavior, there is a range of possible outcomes. Each outcome has its own value, in terms of the learner's goals. The expected value of the model's behavior is the average of the values of the possible outcomes, weighted by the probability that each outcome occurs. This is an abstract model for this article, which we believe will guide human behavior. We do not require that each person actually model these scenarios or do this computation. A learner could use simpler techniques to approximate the expected value of a model's behavior.

How to Measure Competence

A learner wants to estimate how good other people are at achieving her goals. This means estimating the expected value of others' behaviors. The learner needs cost-effective techniques for making these estimates. The most obvious way is by observing other people for extended periods of time, seeing which goals they achieve and to what degrees, then valuing these achievements according to her own values. She can then combine the values and remember the total value of the goals that each person achieved. This serves as her measure of competence for each person. This might work for people who the learner has opportunity to observe frequently, for extended periods of time, but will not work

for many other potential models. The learner will not be able to observe everyone frequently. Even if someone can be observed frequently, some goals might take years to achieve and to measure, so this method of measuring will not be efficient.

Instead of observing people for long periods of time, the learner can use other cues that indicate how competent a person is. Examples of cues that have been proposed include the prior success of the model; the skill of the model; the health, happiness, or age of the model; and how much confidence a model displays (Laland 2004; Mesoudi 2008; Rendell et al. 2011; Henrich 2016). It is also easy to imagine other cues of competence, such as position in a company or other societal institution. For this article, the most important cue of competence is the opinion of other people. If other people think a person is competent, then he probably is, and if other people think he is not competent, then he probably is not. We will study this cue in more depth later.

A learner can combine evidence about a model's competence from several different cues with evidence from direct observation. This will allow the learner to estimate how competent a potential model is. If the model is competent, the learner might attempt to learn from the model.

How to Use the Measure of Competence to Learn

Each learner has estimates of the competence of potential models, one estimate per model. Each estimate of competence is an estimate of the expected value of the outcome of that person's behavior. Learners want to learn the behaviors that have the highest expected value, according to their own goals and values.

Learners will be exposed to a variety of behaviors, from a variety of people. In the complex, real world, they might try to combine two or more behaviors. For simplicity, we suppose that they are only trying to choose the best, single behavior to adopt. In order to do this, learners need to estimate the expected values of the behaviors to which they are exposed. Then they can choose the behavior with the highest expected value. If a learner has information about the expected value of each behavior, then this can be used to help make the decision. The learner can simply choose the behavior with the highest expected value.

If learners do not have information about the behaviors themselves, then other information can be used to figure out which behavior is best. This is where competence comes in. Each behavior is being modeled by a person. A learner has an estimate of the competence of each of these models. The measure of competence for the model is an estimate of the expected value of that model's behavior. The learner can use the measure of competence of the model as an estimate of the expected value of the behavior being modeled.

The learner would like to choose the behavior with the highest expected value. In the absence of information about the behaviors themselves, the measures of competence of the models can be used as the expected values of the modeled behaviors. In this simplified scenario, the learner can simply choose the behavior that is performed by the most competent model.

Transitivity

We mentioned previously that the opinions of others are important cues about how competent a potential model is. Our argument depends on understanding these cues better. In this section, we will dive into them more deeply.

We suppose that we have three people: a learner, an intermediary, and a model. The learner does not know how competent the model is. The learner does know how competent the intermediary is, and the intermediary knows how competent the model is. If the learner can find out how competent the intermediary thinks the model is, then the learner can adopt the intermediary's estimate of the model's competence.

When the intermediary thinks the model is competent, the intermediary copies some of the model's behaviors. When the learner observes the intermediary copy the model, the learner knows that the model's competence, or rather the intermediary's estimate of it, is above some lower bound, because he knows that competence made the potential model good enough for the intermediary to copy. In order to estimate what this lower bound is, and how much competence the potential model has, given that it is greater than this lower bound, we need to use some mathematical modeling.

We are going to suppose that the competence of people in the total population is normally distributed with mean 100 and standard deviation 15. If the learner had no information about the potential model, he would have to guess that the competence of the model is 100, the average value in the population. In this scenario, the learner has additional information. He knows that the potential model's competence was high enough that the intermediary copied him. The intermediary would only copy the model if the expected value of the model's behavior, the model's competence, were higher than the expected value of what the intermediary would have done without copying. If the learner could find out what the intermediary would have done if he had not copied, and could find out the expected value of that action, then the learner would have a lower bound for the competence of the model. The learner's estimate of the competence of the model would then be the expected value of the competence of a random person from the population, given that the person's competence is greater than this lower bound. This allows us to express this conditional expected value as a function of the lower bound, borrowing a formula

Fig. 1 The amount of competence the learner infers about the model is an increasing function of the value of the action that the intermediary would have taken if he had not copied the model. If the value of the counterfactual, alternative action is low, then the model might be an average person in the population, with competence 100. If the value of the alternative action is high, then the model's competence is probably much above average, and that is what the learner will infer

from Panjer (2002). In other words, we express the learner's estimate of the model's competence as a function of the expected value of the action that the intermediary would have taken if he had not copied the model.

Figure 1 graphs the learner's estimate of the model's competence as a function of the expected value of the action that the intermediary would have taken, if he had not copied the model. See Eq. (A1) in Appendix A and the electronic supplementary material (Online Resource 1; contains the Python program modeling the transitive inference of competence) for more detail. If the expected value of this alternative action is very low, then knowing that the model's competence is higher than it does not give the learner much information, so he should still expect that the model's competence is roughly the average of the population, 100. If the expected value of the alternative action is high, then knowing that the model's competence is higher than that lower bound indicates that the model is very competent. This is intuitively correct; however, the learner probably cannot use this function to estimate the competence of the model because it is unlikely that the learner knows what the intermediary would have done had he not copied the model nor the expected value of that counterfactual action, so the learner does not actually know a lower bound for the competence of the model.

Even though the learner does not know how good the intermediary's counterfactual action would have been, he does know how good the intermediary's actions are, on average, because he knows how competent the intermediary is. The learner does not know the conditional expectation of the

particular alternative action. However, if the learner assumes that the possible alternative actions come from some probability distribution, he can find the expected value, or average, over all of the possible conditional expectations. This will be his estimate of the model's competence. This estimate of the model's competence can be written as a function of the intermediary's competence, the average value of the alternative actions. Figure 2 graphs the learner's estimate of the competence of the model as a function of the competence of the intermediary, using a few different assumptions for the probability distribution of the possible alternative actions. See Eq. (A2) in Appendix A and the electronic supplementary material for more detail.

We do not know exactly which distribution of the alternative actions is most accurate, and for this article, we do not need to. The only thing we need to notice is that for any reasonable distribution, the learner's estimate of the model's competence is an increasing function of the intermediary's competence. If the intermediary is not competent at all, and he copies the model, then the model might just be of average competence. If the intermediary is of average competence and chooses to copy the model, then the model is probably of above-average competence, but not necessarily very competent. If the intermediary is very competent and chooses to copy the model, then the model is probably very competent.

A somewhat overly simplified, but useful, way of stating this is that the measure of competence is transitive. If the learner thinks the intermediary is competent, and the intermediary thinks the model is competent, then the learner will

Fig. 2 The amount of competence the learner infers about the model is an increasing function of the known competence of the intermediary. This function depends on the assumed distribution of the alternative actions of the intermediary; however, the same pattern holds, regardless of assumption. If the intermediary's competence is low, then the model's competence is probably roughly average. If the intermediary's competence is high, then the model's competence is probably high, and that is what the learner will infer

think the model is competent. If the learner does not think the intermediary is competent, then this will not happen. Likewise, if the intermediary does not think the model is competent, then this will not happen.

Reputation for Competence

We have argued that the measure of competence is transitive. This means that a judgment of competence about a specific model will spread from person to person. If person A thinks a model is competent, this belief will spread to person B, who will pass it to person C, and so on. This means that the model will develop a reputation for competence. If the model does something to convince person A that she is more competent, eventually person B and person C will also believe she is more competent. Her reputation will increase. If a model does something to convince person A that she is less competent, her reputation will decrease.

When a learner assesses the competence of a model, the model's reputation will be one cue that is taken into account. The learner can combine the cue from the model's reputation with other cues to estimate how competent the model is.

Prestige as a Reputation for Competence

Our abstract framework has predicted that people should develop a reputation for competence. We now turn to the empirical literature to see if this reputation exists.

In our transitivity logic, learners realize that a model is competent when they see an intermediary copy, or learn from, the model. Our logic does not say exactly how learners observe that the intermediary learns from the model, just that they do. The research on the competence-based status called "prestige" says that when an intermediary attempts to learn from a model, the intermediary confers deference upon the model. This can take many forms, like not talking over the model, being especially trustworthy or helpful to the model, or other forms of generally "kissing up" to the model. The intermediary does these things in order to gain proximity to the model, so that he can learn more effectively. This pattern of deference can be observed by the learner. It signals to the learner that the intermediary is trying to learn from the model, which signals that the model's behavior is adaptive. When several intermediaries confer deference upon the same model, the model is said to have "prestige." This prestige signals to other learners that the model's behavior is adaptive (Henrich and Gil-White 2001; Henrich et al. 2015; Henrich 2016; Brand and Mesoudi 2019). Thus, prestige functions as a reputation for competence. When people learn from a model, other people notice this and use it as evidence that the model is competent, as our logic predicts.

There is lack of clarity in the literature about whether this pattern of people learning from prestigious individuals

should be explained by an implicit, unconscious bias in the minds of the learners or by a process of rational reflection that can be explained by general intelligence. It is possible that there is an unconscious bias in people's minds that causes them to preferentially learn from prestigious people. Alternatively, it is possible that people consciously seek out models who are skilled at certain goals and consciously copy their behaviors, resulting in a pattern of learning from prestigious people (Chellappoo 2021). Our logic is abstract enough to be mechanism-agnostic in this case. Either one of these explanations for the pattern of preferentially learning from prestigious people could create a reputation for competence, as our logic predicts.

Challappoo (2021) argues that, in the prestige literature, there is vagueness in the definition of prestige, and that different studies use different proxies for prestige, corresponding to different facets of the behavior that prestigious people or their learners supposedly exhibit. Some studies attempt to measure the acknowledged skill of models while others measure the amount of attention that learners pay to the models. Our logic suggests that prestige functions as a reputation for competence and that learners are more likely to learn from models with reputations for high competence over models with reputations for low competence. We hope this will provide additional clarity that will help prestige researchers align on how to measure the impact of prestige-biased social learning. We think that future research on prestige should attempt to assess the consequences of a model's reputation. If studies use the amount of attention that learners pay to a model as a proxy for the model's reputation, then the studies should evaluate the strength of that relationship before assuming it.

Because we built up the reputation for competence from the first principles of social learning, there is a clear distinction between the competence of a person and the reputation that results from that competence. This clarity about the relationship between competence and the resulting reputation will allow us to study the incentives created when people try to improve their reputations. Since prestige appears to function as a reputation for competence, predictions that we make about the reputation for competence should apply to prestige as well.

Resulting Incentives

Incentive to Appear Competent

A person's competence is a measure of how good that person is at achieving human goals. Likewise, a person's reputation for competence is a measure of how good that person is at achieving human goals. People can easily estimate how competent a person is by using his or her reputation. We

have focused on using this reputation for the purposes of learning, but this is not the only use of the reputation. People are likely to assess the competence of potential mates. Other things being equal, a potential mate with higher competence, who is better at achieving human goals, should be more desirable than a less competent one. A person can use the reputation of potential mates to assess their competence, and thus their attractiveness. Someone with a higher reputation for competence is likely to seem more attractive to potential mates, which should give access to either more mates or a better mate, increasing the individual's fitness. A person has a biological incentive to increase his or her reputation for competence because doing so will increase his or her fitness. This applies to all people, regardless of their reputations for competence and their roles as learners or models.

Incentive to Help Others

If we analyze the transitivity logic more deeply, we will find that one way competent people can increase their reputations for competence is by helping others. Recall that when learners use the transitivity logic to assess the competence of a model, the amount of competence they estimate for the model is an increasing function of the competence of the intermediary. It seems then that if a model can increase the competence of the intermediary, the learner will estimate that the model has higher competence, and the model will have increased his reputation for competence. However, this will not work, because if the intermediary's competence increases, then the intermediary will not need to copy the model as often, and the learner should still be able to accurately assess the model's competence, on average.

If the model can make the learner think that the intermediary is more competent than he actually is, then the intermediary will still copy the model as frequently, but the learner will overestimate the competence of the model. This will effectively increase the model's reputation for competence. This creates a biological incentive for the model to make intermediaries seem more competent than they actually are. This means that the model has incentive to make it look like intermediaries achieve goals to higher degrees than is actually the case. One way for a model to do this is to artificially increase the achievement of the intermediaries' goals. For example, if one of the goals is finding food, then the model can give the intermediaries some food. As long as the model does this without other people knowing, then it will seem that the intermediaries are better at finding food than they actually are. This will increase the model's reputation for competence. The model has incentive to help the intermediaries without other people knowing that help is being provided.

A model is any person with at least some baseline level of competence so that people are interested in learning from

him. Anyone can be a learner. The intermediary is anyone who believes the model is competent, so that he will learn from the model and thus will spread the belief that the model is competent. This shows that any person who is competent enough to serve as a model has biological incentive to help the people who believe he is competent, as long as he does so without other people knowing.

Incentive to Behave as a Competent Person Would

Competent people have incentive to help others without other people knowing. They will not always be successful at keeping their actions hidden. Sometimes people will find out that a competent person helped someone. When this happens, the people who find out will not be tricked into thinking the model is more competent than is the case. However, knowing that the model helped someone serves as evidence that she is competent. This will either confirm or increase the amount of competence others estimate for the model.

People who are not competent have incentive to convince others that they are competent. In order to do this, they must behave as though they were competent. This is tricky because a competent person has incentive to help people without others noticing. If incompetent people help someone and no one notices, they don't get any benefit. No one will copy them, so the transitivity logic will not increase their reputations for competence; they don't have incentive to help people without others noticing. If they help someone and people notice, believing that the motivation is to help without being noticed, then people will think the incompetent individuals are behaving as competent people would. This serves as some evidence that they are competent, so their reputations will increase. If incompetent individuals help and people notice, but believe that their motivation is only to be seen helping, then the helpers will not convince others of their competence, because this is not how competent people behave. This will serve as evidence instead that they are not competent, so their reputations will remain low or become low.

Complex incentives now exist. Competent people benefit when they help people without others finding out. They also benefit when they help people and others do find out, as long as it looks like they had incentive to help without being noticed. A good strategy for competent people is to attempt to help people without being noticed. They benefit whether people find out or not.

Incompetent people do not benefit when they help people without others noticing. They also do not benefit if they overtly display that they are helping people. They only benefit when they convince people that they attempt to help without others noticing, but their attempts to keep the help hidden fail. In order to truly convince people that they are competent over the long term, they probably need to play

the same strategy that competent people play, which is to attempt to help people without being noticed. They could instead only help when they think people will notice, but if people realize they are playing this strategy, it will become apparent that they are not competent, and their fitness will decrease.

Reputation for Three Things

What started out as a measure of competence is no longer purely a measure of competence. It is now based on three things: one's ability to achieve goals, whether one helps other people or not, and the motivation behind helping those people. The reputation for the combination of these three things is what makes one an attractive mate and increases one's fitness.

Comparison to Prestige

We now return to the prestige literature to check the predictions of our logic. Since we believe that prestige functions as a reputation for competence, the main predictions for the reputation for competence apply to prestige as well. We made one main assumption and one main prediction for the reputation for competence. We assumed that people with a reputation for higher competence would be more attractive to potential mates. We predicted that this would incentivize people with a reputation for competence to help people who thought highly of them, via the transitivity mechanism. We can now test that assumption and prediction by looking at the prestige literature.

First, we investigate the assumption that prestigious people are attractive mates, since our incentive to help people is derived from this. Prestige should serve as a signal of their competence, enhancing their attractiveness. Prestige does seem to be related to being a more desirable mate, indicated by the observations that prestigious people have more children, have higher fitness, have more affairs, and sometimes marry younger (Von Reuden et al. 2011; Cheng and Tracy 2014; Henrich 2016).

Our main prediction is that reputation for competence, and therefore prestige status, is linked with helping behavior. According to our logic, people who are prestigious, and thus are competent, can further enhance their prestige by helping others, because helping others is a way of convincing people that they are more competent than they actually are. However, our logic specifies that in order to benefit from helping others, it must not appear that the motivation for helping is merely to advertise the fact that one helps. It must appear that there is motivation to help, even if no third party notices the help.

The prestige literature has noticed that prestigious people are generous and cooperative within their communities (Halevy et al. 2012; Cheng and Tracy 2014; Henrich et al. 2015; Henrich 2016; Jiménez and Mesoudi 2019). A few explanations for this have been proposed. Henrich and Gil-White (2001) proposed that prestigious people might act generously in order to advertise their skill. For example, sharing a large hunting kill advertises the fact that one is able to hunt successfully. This is similar to an explanation provided by costly signaling theory, which says that a person might share a large hunting kill in order to advertise the fact that he has enough meat to share, which might signal that he is able to successfully hunt or procure meat by some other means. This signal might benefit the generous individual by giving him more mating opportunities or other benefits (Smith and Bird 2000). Another explanation, espoused by Halevy et al. (2012), is that groups motivate people to act generously by rewarding them with status, such as prestige. This means that people will act generously toward their groups in order to receive benefits from those groups in return. A third possible explanation, modeled by Henrich et al. (2015), is that if prestigious people act cooperatively, their followers will copy them and will also act cooperatively. This creates an additional benefit for everyone in the group, and the prestigious people get part of that additional benefit. This gives the prestigious people incentive to motivate people to act cooperatively by acting cooperatively themselves.

Our logic offers an alternative, or perhaps complementary, explanation. Our explanation differs from those that have been provided because ours specifies that the motivation to help does not come from other people's knowledge of the help. In fact, in order to gain the benefit of helping, via this mechanism, one would have to convince others that one is not helping merely to be seen helping. The prestige literature has not noticed this subtlety in the helping behavior of prestigious people. The ostensible motivations for the generosity of prestigious people have not been studied. More research is needed in order to determine if our logic agrees with the behavior of prestigious people in this regard.

We have used the prestige literature as a reasonability check for the predictions of our model. The prestige literature confirms that there are people who are both competent and help others more than would otherwise be expected. It provides some evidence that there is a link between competence and helping others. However, we were not able to use it to validate some of the more specific predictions of our logic.

Conclusion

We have proposed two new things in this article. We have proposed new logic that predicts that humans will help each other. Helping signals that one is competent, which makes

one more attractive to potential mates, thereby increasing one's fitness. Over time, sexual selection could have increased the prevalence of helping behavior in humans.

We have also proposed a measure of a person, and a resulting reputation, that combines three different things. The measure combines competence, helping behavior, and the motivation for the helping behavior into one measure. We argued for and justified this measure because it increases individual fitness by helping an individual learn. Once this measure exists, it encourages people to help others.

The logic we proposed is more general than the conclusions we derived in this article. It has implications that reach beyond what we were able to explore. For example, this article explored a scenario in which people mostly share goals and values. In this scenario, people help each other. The logic might have different implications in a scenario in which people do not share goals and values or in a scenario in which different groups of people share different goals and values.

We hope that we have inspired readers to research in this area, both in order to determine if the logic we have presented is accurate and to explore its further implications.

Appendix A: Formulae

$$CTE(x_q) = E[X|X > x_q] = \mu + \frac{f(x_q)}{1 - F(x_q)} \sigma^2 \quad (A1)$$

where $f(\cdot)$ and $F(\cdot)$ are the probability density function and the cumulative distribution function of the normal distribution with mean μ and standard deviation σ (Panjer 2002).

$$M(\mu_l) = E[CTE(x_q)] = \int_{-\infty}^{\infty} g_{\mu_l}(x_q) CTE(x_q) dx_q \quad (A2)$$

where μ_l is the mean of the distribution of the values of an individual's actions and $g_{\mu_l}(\cdot)$ is the probability density function of that distribution.

Supplementary Information The online version of this article contains supplementary material available at <https://doi.org/10.1007/s13752-022-00396-z>.

Acknowledgments I would like to thank two anonymous reviewers for their helpful comments on an earlier version of this manuscript.

Funding No funding was received to assist with the preparation of this manuscript.

Availability of Data and Material Not applicable.

Code Availability Code used to create figures is published as supplementary material.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

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

- Boyd R, Richerson PJ (1985) Culture and the evolutionary process. University of Chicago Press, Chicago
- Brand CO, Mesoudi A (2019) Prestige and dominance-based hierarchies exist in naturally occurring human groups, but are unrelated to task-specific knowledge. *R Soc Open Sci* 6(5):181621. <https://doi.org/10.1098/rsos.181621>
- Brand CO, Mesoudi A, Morgan TJH (2021) Trusting the experts: the domain-specificity of prestige-biased social learning. *PLoS ONE* 16(8):e0255346. <https://doi.org/10.1371/journal.pone.0255346>
- Brown GR, Dickins TE, Sear R, Laland KN (2011) Evolutionary accounts of human behavioural diversity. *Philos Trans R Soc B* 366(1563):313–324. <https://doi.org/10.1098/rstb.2010.0267>
- Chellappoo A (2021) Rethinking prestige bias. *Synthese* 198:8191–8212. <https://doi.org/10.1007/s11229-020-02565-8>
- Cheng J, Tracy J (2014) Toward a unified science of hierarchy: dominance and prestige are two fundamental pathways to human social rank. In: Cheng J, Tracy J, Anderson C (eds) *The psychology of social status*. Springer, New York. https://doi.org/10.1007/978-1-4939-0867-7_1
- Halevy N, Chou EY, Cohen TR, Livingston RW (2012) Status conferred in intergroup social dilemmas: behavioral antecedents and consequences of prestige and dominance. *J Pers Soc Psychol* 102(2):351–366. <https://doi.org/10.1037/a0025515>
- Henrich J (2016) *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press, Princeton
- Henrich J, Broesch J (2011) On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philos Trans R Soc B* 366(1567):1139–1148. <https://doi.org/10.1098/rstb.2010.0323>
- Henrich J, Gil-White F (2001) The evolution of prestige: freely conferred status as a mechanism for enhancing the benefits of cultural transmission. *Evol Hum Behav* 22(3):165–196. [https://doi.org/10.1016/S1090-5138\(00\)00071-4](https://doi.org/10.1016/S1090-5138(00)00071-4)
- Henrich J, Chudek M, Boyd R (2015) The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. *Philos Trans R Soc B* 370(1683):20150013. <https://doi.org/10.1098/rstb.2015.0013>
- Jiménez ÁV, Mesoudi A (2019) Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Commun* 5(1):20. <https://doi.org/10.1057/s41599-019-0228-7>

- Laland KN (2004) Social learning strategies. *Anim Learn Behav* 32(1):4–14. <https://doi.org/10.3758/BF03196002>
- Mesoudi A (2008) An experimental simulation of the “copy-successful-individuals” cultural learning strategy: adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evol Hum Behav* 29(5):350–363. <https://doi.org/10.1016/j.evolhumbehav.2008.04.005>
- Muthukrishna M, Henrich J (2016) Innovation in the collective brain. *Philos Trans R Soc B* 371(1690):20150192. <https://doi.org/10.1098/rstb.2015.0192>
- Panjer HH (2002) Measurement of risk, solvency requirements and allocation of capital within financial conglomerates. https://www.soa.org/globalassets/assets/Files/Research/Projects/measurement_risk.pdf. Accessed 2 Mar 2021
- Rendell L, Fogarty L, Hoppitt WJE, Morgan TJH, Webster MM, Laland KN (2011) Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends Cogn Sci* 15(2):68–76. <https://doi.org/10.1016/j.tics.2010.12.002>
- Smith EA, Bird RLB (2000) Turtle hunting and tombstone opening: public generosity as costly signaling. *Evol Hum Behav* 21(4):245–261. [https://doi.org/10.1016/S1090-5138\(00\)00031-3](https://doi.org/10.1016/S1090-5138(00)00031-3)
- von Rueden C, Gurven M, Kaplan H (2011) Why do men seek status? Fitness payoffs to dominance and prestige. *Proc R Soc B* 278(1715):2223–2232. <https://doi.org/10.1098/rspb.2010.2145>

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