Reputation management and cultural evolution
Hugo Mercier

To cite this version:
Hugo Mercier. Reputation management and cultural evolution. Journal of Cognition and Culture, 2022, 22 (5), pp.485-498. 10.1163/15685373-12340147 . hal-03863402

HAL Id: hal-03863402
https://hal.science/hal-03863402
Submitted on 21 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Reputation management and cultural evolution
Hugo Mercier

Abstract. Sociologists and social psychologists have long seen reputation management as an important human motivation. More recently, evolutionary analyses have helped understand the function of reputation management, demonstrating the fitness consequences of being thought of as dominant, moral, or competent. Here, I argue that reputation management likely plays an important, but understudied, role in cultural evolution—whether one takes the perspective of dual inheritance theory or of cultural attraction theory. I illustrate the importance of reputation management through its role in the spread of non-actionable beliefs—beliefs which have few or no behavioral consequences, but which constitute a large part of culture.

Keywords. Cultural evolution; Reputation management; Dual inheritance theory; Cultural attraction theory; Non-actionable beliefs.

Sociologists and social psychologists have long recognized the importance of reputation (or impression) management (e.g., Goffman, 1959; Leary, 1995; Veblen, 1899). Reputation management consists in attempts to shape the opinions others have of us, on a
variety of dimensions.\(^1\) Evolutionary approaches have converged on the same realization, in different domains. In many species, sexual selection pushes individuals from at least one gender to advertise their traits—i.e. influence how other individuals perceive them—in order to increase their odds of mating. In group-living species, individuals (typically males) vie to appear dominant. In humans in particular, dominance takes diverse forms, and individuals also compete to be perceived as prestigious, as good leaders, and more generally as competent individuals (e.g., Boehm, 1999; Henrich & Gil-White, 2001). Finally, humans compete to appear fair, so that they are selected by others as cooperation partners (Baumard et al., 2013). In all of these cases, individuals have evolved dedicated traits in order to influence how others perceive them.

It is very plausible that much of our social behavior serves reputation management functions. This does not mean that we are constantly aware of the reputation management goals of our actions, but that some of the cognitive mechanisms shaping our behaviors evolved to serve reputation management functions. Even behaviors that do not directly serve such functions are influenced by reputation management concerns. For instance, if you want to hurt someone (not for reputation management purposes), you are still likely to take reputation management concerns into account (e.g. by doing it publicly to appear tough, or on the contrary discretely to preserve your moral reputation).

\(^1\) Technically, reputation is often understood as a public, shared opinion, by contrast with the more private opinions of individuals (Sperber & Baumard, 2012). As a result, it might be more correct to talk about ‘impression management.’ However, both terms are often used interchangeably, and reputation is often used in the evolutionary literature to merely mean the sum of others’ opinions, instead of specifically referring to its public dimension (e.g., Barclay, 2016; Nowak & Sigmund, 2005).
The importance of reputation management for cultural production and consumption has been highlighted by social scientists, from Veblen (1899) to Bourdieu (1979). Yet, in spite of the importance of reputation management in human behavior, it has played (to the best of my knowledge) only a minor role in studies of cultural evolution. My goal here is to make an argument for why reputation management should be taken seriously by the two main schools of cultural evolution research, dual inheritance theory and cultural attraction theory (on the differences between the two, see, Sterelny, 2017). I will then briefly illustrate the explanatory power of reputation management by sketching the following argument: the way we manage our reputation plays a huge role in the spread of non-actionable information.

Reputation management and dual inheritance theory

Dual inheritance theory sees culture as broadly adaptive (Henrich, 2015; Richerson & Boyd, 2005). In this framework, for culture to be adaptive, individuals have to be able to (i) faithfully reproduce the cultural elements they perceive (actions or beliefs) and (ii) select adaptive cultural variants to reproduce. The latter is achieved thanks to a series of heuristics regarding which cultural variants one should faithfully reproduce. The best studied of these heuristics are: content bias (be more likely to copy some contents over others), conformity bias (copy the majority), success bias (copy successful individuals), and prestige bias (copy individuals other people defer to).

Both the success and prestige biases are explicitly built upon reputation. As a result, the downstream effects of reputation—i.e. the effects of the success and prestige biases for cultural transmission—have been well studied within this tradition (e.g., Henrich & Gil-
What has been less studied are the effects of reputation management. Individuals gain from being thought of as successful or prestigious: they can influence others, and they often derive a series of other benefits from the deference displayed by others (see, Henrich & Gil-White, 2001; Jiménez & Mesoudi, 2019). People should thus strive to appear successful and to gain prestige.

The cues observers use to infer who is successful and whom to grant prestige to are imperfect. This creates opportunities for people to appear successful or to gain prestige even when they exert a negative influence on their audiences. For instance, in their review of work on the prestige bias, Jiménez and Mesoudi (2019) note that one’s expressed degree of confidence appears to be taken as a cue to success or prestige, and that people who express themselves more confidently tend to be granted higher status and to be believed more (Anderson & Kilduff, 2009). Since the correlation between confidence and accuracy is sometimes weak, people might end up granting prestige to the wrong individual, and therefore spreading potentially maladaptive cultural variants. However, it is not clear whether such mismatches can really have consequences that would affect the long term spread of culture. On the whole, people are able to adjust their judgments of others when better evidence becomes available (for review, see, Mercier, 2020). For instance, even if an individual can gain some traction over others by being confident, when it is revealed that they were in fact overconfident, trust in them plunges, and they can suffer serious reputational damage (Tenney et al., 2007; Vullioud et al., 2017).

A more plausible mechanism through which reputation management can foster the spread of maladaptive cultural variants is the interaction of the success and prestige bias on the one hand, and of content biases on the other. Most content biases should be adaptive, making us prefer, say, food that smells of vanilla than of vomit. However, combinations of
cognitive mechanisms can also create content biases that appear largely maladaptive. A striking example is provided by intuitive forms of therapies. Before modern medicine, human societies mostly relied on two forms of therapy: supernatural means, or attempting to physically remove some supposedly toxic substances from the body (through bloodletting, laxatives, emetics, or sudation) (Coury, 1967; Miton et al., 2015). Unfortunately, neither form of therapy is effective at least in purely medical terms (with some exceptions for the latter type). The success of these practices might stem from a combination of success or prestige biases and content biases. Ascertaining the success of a therapy is tremendously difficult without modern techniques of randomization and statistical inference. As a result, people might resort to content biases to evaluate medical practitioners, deeming more successful, and granting more prestige to those who rely on therapies that fit their content biases. This might help explain, for instance, why bloodletting is not only practiced in a large number of cultures, but also why it is typically practiced by a dedicated third party with some prestige in the society (Miton et al., 2015). The incentives to appear successful and prestigious might thus motivate people to not only rely on, but also defend practices that are intuitively compelling, but arguably maladaptive (see for instance how Galen accrued an enormous amount of prestige in part through his complex theoretical justifications for the use of bloodletting, Brain, 1986).

The last of the major heuristics that, in dual inheritance theory, guide our choices of cultural items to copy is the conformity bias (following what the majority does). At first, it might seem that reputation management only makes the conformity bias less reliable. One of the conditions necessary for the majority to be correct is that its members have formed their opinions largely independently of each other. However, people might join the majority not for epistemic reasons, but for normative reasons (on the distinction between
informational and normative conformity, see, e.g., Deutsch & Gerard, 1955). These normative reasons are related to reputation management: people do not want to suffer costs—reputational or worse—if they stand against the majority. Normative conformity is for instance what drives some participants to conform on some trials in the Asch conformity experiments (Asch, 1956). Normative conformity might thus artificially inflate the ranks of the majority, making informational conformity (the one required for the conformity bias to function) less reliable.

It has recently been argued, however, that reputation management can in fact save the conformity bias from itself (de Courson et al., 2021). A well-known issue with aggregation heuristics such as the conformity bias is that they are so efficient that everyone should use them. But if everyone looks at what other people are doing instead of gathering their own information, no one has reliable information anymore, and looking at what others are doing is worthless (which is also why normative conformity creates problems for informational conformity). This risk has been demonstrated in models (Curty & Marsili, 2006) and in experiments (Lorenz et al., 2011). A way out of this conundrum is to incentivize individuals to form opinions independently of the majority, even if these opinions are less likely to be accurate. De Courson et al. (2021) describe a model in which agents are rewarded not only for accuracy, but also for originality, through increased prestige. In this model, enough agents, incentivized by these reputational benefits, provide independent information, making it possible for the other agents to follow the majority.

On the whole, thus, it appears that within the dual inheritance theory framework, reputation management should play an important role, but that this role hasn’t received much attention yet.
Reputation management and cultural attraction theory

Dual inheritance theory is largely built on episodes of faithful cultural transmission which would enable a measure of cultural stability. By contrast, cultural attraction theory argues that such episodes are exceptional, and require explanation (Claidière et al., 2014; Sperber, 1996). Cultural attraction theory points to the immense degree of noise inherent in cultural transmission, meaning that each time a cultural (or potentially cultural) element is transmitted, it is transformed in a variety of ways (consider the difference between the thoughts I’m trying to convey and those you will form reading this paper, and those someone whom you would describe the paper to would form, etc.). In this context, cultural stability isn’t the outcome of faithful transmission, but of systematic biases in transmission pulling cultural elements in the same direction. These biases in transmission are factors of attraction, and the cultural variants they converge towards are attractors. Although factors of attractions can have different origins, I focus here on cognitive factors of attraction, since those are most closely related to reputation management. A convenient way of considering the manifold influences of cognitive factors of attraction is to consider the different steps that are necessary for cultural transmission to take place.

First, individuals must attend to a (potential) cultural element. Obviously, attentional mechanisms haven’t evolved chiefly for reputation management concerns, and they spontaneously guide our attention towards items that exhibit a range of intrinsic properties—for instance, our attention is naturally driven to attractive faces (Langlois et al., 2000), but also to disgusting stimuli (Charash & McKay, 2002). However, as soon as we direct our attention in public, reputation management concerns come into the picture—checking people out while being accompanied by one’s partner tends to be frowned upon. More importantly, reputation management concerns likely exert an important role in what
information we chose to consume. Even if our brains might be naturally inclined to pay attention to *Transformers 5*, with its beautiful faces and its explosions, we can choose to see an early Ozu movie instead, a choice that might be partly motivated by reputation management. Reputation management might also be one of the reasons people consume news, and in particular political news, as they seek information that might make them appear knowledgeable, or that would allow them to justify their political opinions (Mercier, 2020; Williams, 2022). It is thus plausible that reputation management plays a significant role in how we allocate our attention.

Second, individuals must remember the (potential) cultural element. Although this step is most crucial in oral cultures, it also plays an important role in written or even in digital cultures. People might not need to remember a meme to share it on social media, but if they remember it they are more likely to share it later, or to use it as a template to create a new meme. Like attention, memory systems were mostly shaped by selection pressures having nothing to do with reputation management. One potential exception is episodic memory. According to Mahr and Csibra (2017), the function of episodic memory would be communicative: to “determine when and how to assert epistemic authority in making claims about the past”. Episodic memory would thus have evolved in part to manage our reputation: when we make claims about the past, we are able to back them up with episodic memories of the events under discussion. For instance, participants led to view introversion as desirable were more likely to remember episodes in which they had been introverted (Sanitioso et al., 1990), possibly so they could convince others that they possessed a desirable trait.

Third, individuals must share the cultural element. There is a variety of reasons why people share cultural elements with others—to influence their behavior towards the world,
to create social bonds, etc. However, given that sharing is the most public of the three steps, it is also the one most likely to be influenced by reputation management. A dramatic illustration is provided by work showing massive discrepancies between what news people consume and what news people share. A study compared what articles people read and what articles people shared on the BBC News website (Bright, 2016). There was a negative correlation (-.14) between the topics that people were most likely to read, and those they were most likely to share. The most read topic (“accident and disaster”) was also the least shared one. An even more glaring example is porn, which is widely consumed (accounting for 20% of mobile searches in the US) but barely ever shared between consumers. Both of these examples reveal the importance of reputation management in sharing behaviors, as people do not want to appear morbidly curious, or reveal their sexual habits and preferences.

Given its impact on the different steps of cultural transmission—mostly but not only sharing information—reputation management also ought to play a significant role in cultural attraction theory.

**Reputation for competence and non-actionable information**

Some beliefs are ‘non-actionable’, by which I mean that their content does not affect how those who believe them interact with the world (for a related concept, see, Sperber, 1997). When I read in the news that a new exoplanet has been discovered, that Johnny Depp and Amber Heard are suing each other, or that there has been a volcanic eruption on the other side of the planet, this has no effect on my interactions with the world. The main effect encountering such news has is that I might want to share the information with others. Some beliefs that are non-actionable for the vast majority of people play a very important
role in the modern world. Consider relativity theory: some people have been able to act on it in order to design, say, precise GPS. However, for the vast majority of the people who believe in relativity theory, the belief is non-actionable—arguably, even for Einstein, who never considered practical applications of his theory. In spite of being useless for practical ends, non-actionable beliefs are pervasive in our cultural environments, and should therefore be an interesting object of study for cultural evolutionists.

Given the lack of personal practical import of non-actionable beliefs, what can explain their cultural success?

The first element of the answer likely lies with the distinction between the proper and the actual domain of a cognitive mechanism (Sperber & Hirschfeld, 2004). The proper domain of a cognitive mechanism is the range of inputs that it evolved to process. For instance, our face recognition mechanism evolved to process the faces of humans we could see in front of us. However, most cognitive mechanisms respond to inputs that are similar enough to the input they evolved to process—in the case at hand, faces in pictures or movies, caricatures, masks, weird mountains on Mars, etc. This constitutes the actual domain of face recognition. As this example illustrates, the actual domains of cognitive mechanisms provides rich fodder for cultural evolution, as people are attracted to stimuli that will trigger their cognitive mechanisms (Sperber, 1996; Sperber & Hirschfeld, 2004).

Non-actionable beliefs fall within the actual domain of a range of cognitive mechanisms. Good theories like relativity theory trigger the mechanisms with which we evaluate explanations and arguments. Celebrity gossip satiates our appetite for information about important people in our community. News about disasters trigger our threat-detection mechanisms; and so on and so forth.
The exaggerated size of the actual domains of many cognitive mechanisms can explain why people consume a lot of non-actionable beliefs—reading gossip magazines, following the news, watching science videos on YouTube, etc. If someone is brilliant enough, and has time on their hands, this curiosity can even lead to significant new discoveries. A dramatic example is provided by Leonardo da Vinci’s endless curiosity, which led him (inter alia) to understand much of how the cardiovascular system functions through repeated observations.

Leonardo da Vinci is, obviously, exceptional: few people produce new and valuable non-actionable beliefs through their observations of nature; instead, they mostly consume non-actionable beliefs created by others. The question, then, is why do people create and share so many non-actionable beliefs? To answer this, curiosity is not enough. Indeed, we can see the limits of sheer curiosity even in Leonardo da Vinci’s case, since he never published his findings about blood circulation, which would have to be rediscovered centuries later.

Clearly, in the modern world, financial rewards constitute a large part of the answer, as intellectuals, scientists, journalists are paid to produce new non-actionable beliefs. However, for many producers of non-actionable beliefs, money is only a (possibly small) part of the motivation. For instance, studies have shown that scientists are willing to forgo a substantial increase in revenue (either in salary or in funding) in order to be able work on the topic of their choosing, and to publish their results (Myers, 2020; Stern, 2004). Moreover, many people contribute, if not to the production, but at least to the dissemination of non-actionable beliefs, without any financial incentives—by spreading celebrity gossip, discussing the news, sharing new scientific discoveries, etc.
The argument I would like to sketch here is that reputation management is one of the main drivers, if not the main driver of the search for and dissemination of non-actionable beliefs (although other mechanisms are possible, see, e.g. Dunbar, 1998). I will suggest three mechanisms through which people might gain in reputation for sharing non-actionable beliefs.

The first mechanism relies on the fact that some non-actionable beliefs are deemed relevant (because they fall within the actual domain of some cognitive mechanisms). Providing others with relevant information tends to be reputationally rewarded (Altay & Mercier, 2020b). People who share non-actionable but relevant beliefs should thus also be rewarded. For instance, people who share information about threats are deemed more competent (Boyer & Parren, 2015). This is true even if the threats do not directly affect the participants, and are thus essentially non-actionable (e.g. information about leeches in South-America for US participants). Non-actionable information might even have an advantage over actionable information: it is less likely to be rival. The value of information about a good mushroom spot decreases when it is shared, while my enjoyment of celebrity gossip or of a beautiful scientific theory doesn’t. As a result, acquiring a new piece of non-actionable information can allow us to improve our reputation in turn by passing it on to others, making the non-actionable information more relevant, and providing us with another reason to reward the source of the information (see, Mercier, 2020).

The second mechanism relies on people’s ability to use non-actionable beliefs to gauge the cognitive skills of those who share or generate them. This is an extension of a more general principle: in order to evaluate other people’s cognitive skills, we must be able to understand how well they solve their own problems. This means we must be able to recognize when they do something well, even if what they do is largely irrelevant to us. For
instance, if someone is able to solve their own computer problems, I might infer that they have good computer skills in general. Similarly, if someone is able to solve complicated mathematical problems, or to discover new scientific theories, even if these problems and theory are of little import for me, I might infer that they possess sound reasoning abilities. Much historical evidence shows that intellectuals can reap large reputational benefits from crafting sound arguments, advancing enlightening explanations, or gathering extensive knowledge—even if we largely deal with non-actionable information. That gaining reputation is one of the drivers of intellectual progress is suggested by the priority disputes that are the hallmark of academia (Merton, 1957). If intellectuals weren’t motivated by the reputational gains that can accrue from their ideas, they shouldn’t fight so hard to have their contributions recognized. Beyond intellectuals, experimental evidence suggests that adults and children are aware of the importance of having generated good ideas themselves (Shaw & Olson, 2015), and are able to manipulate how they present their ideas to maximize reputational gains (Altay et al., 2020).

Thirdly, non-actionable beliefs can also be used to signal a variety of traits that are unrelated to whether the belief is actionable or not. Many offensive beliefs are largely non-actionable (or at least not acted on)—for instance, people who believe in extreme conspiracy theories tend not to act in line with those beliefs (Kay, 2011; Mercier, 2020). However, offensive beliefs can serve reputation management purposes, in at least two ways: in an attempt to display dominance (‘look what I can get away with!’, see, de Araujo et al., 2020), or in order to ‘burn bridges’ (offending a given individual or coalition in order to honestly display one’s affiliation with another individual or coalition, see, Mercier, 2020; Williams, 2021). By contrast, other non-actionable beliefs are more positive—saying that “everyone has positive energy” for instance. Such beliefs can be used to signal one’s
niceness (Altay & Mercier, 2020a). These reputation management purposes might explain a substantial portion of sharing behavior on social media (see, e.g., Berriche & Altay, 2020; Brady & Crockett, 2019).

The existence of each of these mechanisms could be tested, as they make predictions about (i) the kind of non-actionable beliefs people will want to share; (ii) in what conditions people will want to share these non-actionable beliefs; (iii) the effects sharing these non-actionable beliefs should have. For instance, the second mechanism predicts that (i) people are more likely (everything else equal) to share non-actionable beliefs that they think will make them appear smart; (ii) they are more likely to share such beliefs when they are particularly motivated to appear smart (for some evidence, see Altay et al., 2020); (iii) people who share such beliefs are indeed thought to be smarter. Having demonstrated the psychological plausibility of these (and possibly other) mechanisms, one could then ask the question of which mechanisms are chiefly responsible for the spread of a given non-actionable beliefs. This would likely require more ecologically valid data and experiments. For instance, one could test whether a specific belief (say, a scientific theory) has the right properties (e.g., increasing the perceived competence of those who propound it) among those who actually share it (scientists, the interested lay public).

**Conclusion**

In this short paper, I have argued that reputation management likely plays an important role in cultural evolution. In dual-inheritance theory, people should strive to appear successful and prestigious; providing reputational rewards to independent thinkers might be what makes reliance on the conformity bias possible. In cultural attraction theory, reputation management likely affects all the steps of cultural transmission, exerting a
particularly strong influence on information sharing. As an illustration, I considered the case of non-actionable beliefs, briefly reviewing different reasons why people might share such beliefs to improve their reputation.

Since, by definition, most people do not act on them, non-actionable beliefs might appear inconsequential. However, beliefs that are non-actionable for most are often actionable for a few. Many scientific theories have applications, often unexpected, sometimes far-reaching. Political news might just be conversation pieces for most, but they will guide the attitudes and voting behavior of some, which in turn can have a sizeable impact on policy-making (e.g., Gelman & King, 1993; Stimson, 2004). Crucially, these beliefs would likely not be generated if it were only for the people who can act on them. Many scientific discoveries have been made without any thought of application, and journalism would be drastically reduced if only people who act on the news consumed them. As a result, it is important to understand why people are so keen on creating, consuming, and sharing beliefs even if they do not act on them.

Non-actionable beliefs constitute a sizable part of culture, and an increasingly large one. To take but one example, in the US citizens are increasingly following national instead of local news, even though national news is even less actionable than local news (e.g., Moskowitz, 2021). As a result, the role of reputation management in cultural evolution will also likely grow, and it should receive more attention.

References


